

ประสิทธิภาพของวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ
และวิธีอันซายแมนเทล-แฮนส์เซล ในการตรวจสอบการทำหน้าที่ต่างกันของ
ข้อสอบแบบอเนกรูป

*The Efficiency of Variation of the Mantel-Haenszel and Unsigned
Mantel-Haenszel Procedures in Detecting the Non-Uniform
Differential Item Functioning*

ลินจง บัวซ้อน*

lintm@hotmail.com

กิตติมา พฤกษณ**

ศุภวรรณ พรหมเพรา***

พงศ์เทพ จิระโร****

บทคัดย่อ

การวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อเปรียบเทียบผลการตรวจพบข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูป ระหว่างวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันซายแมนเทล-แฮนส์เซลโดยใช้ข้อมูลเชิงประจักษ์ ตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2553 ของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษากาญจนบุรี จำนวน 1,400 คน ได้มาโดยการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบสอบวิชาคณิตศาสตร์ ที่มีค่าความยากง่ายตั้งแต่ 0.23-0.80 ค่าอำนาจจำแนกตั้งแต่ 0.20-0.92 และค่าความเชื่อมั่น เท่ากับ 0.90 ผลการวิจัยพบว่า วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันซายแมนเทล-แฮนส์เซล ตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยวิธีอันซายแมนเทล-แฮนส์เซล มีประสิทธิภาพในการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปได้ดีกว่าวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ

คำสำคัญ : การทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ วิธีอันซายแมนเทล-แฮนส์เซล

*นิสิตระดับมหาบัณฑิต สาขาวิจัยและประเมินผลการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏกาญจนบุรี

**อาจารย์ ดร. คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏกาญจนบุรี

***ผู้ช่วยศาสตราจารย์ ดร. คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครศรีธรรมราช

****นางสาวตรี ดร. ภาควิชาวิจัยและจิตวิทยาประยุกต์ คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา

Abstract

The purpose of this research was to compare the results in detecting the non-uniform differential item functioning (DIF) test items between variation of the Mantel-Haenszel (VMH) and unsigned Mantel-Haenszel (UMH) procedures with empirical data. The sample consisted of 1,400 Mattayomsuksa 1 students in school under Kanchanaburi Educational Service Area Office, academic year 2010, selected by stratified random sampling. The research instrument was a mathematics achievement test with difficulty of 0.23-0.81, discrimination of 0.20-0.92 and reliability of 0.90. The results revealed that both procedures, VMH and UMH, detected different number of non-uniform DIF, with a statistical significance at .01 level. The UMH procedure was found to be more powerful than the VMH in detecting non-uniform DIF.

Keywords : non-uniform DIF, variation of the Mantel-Haenszel, unsigned Mantel-Haenszel

บทนำ

การทำหน้าที่ต่างกันของข้อสอบ (Differential Item Functioning: DIF) หรือเดิมเรียกว่า ความลำเอียงข้อสอบ (Item bias) เกิดขึ้นเมื่อผู้สอบที่มีความสามารถเท่าเทียมกันจากกลุ่มผู้สอบย่อย ต่างกัน มีความน่าจะเป็นในการตอบข้อสอบข้อนั้นถูก แตกต่างกัน (Mazor, Clauser, & Hambleton, 1994) สาเหตุดังกล่าวอาจเนื่องมาจากแบบทดสอบไม่ได้วัดความสามารถ เป้าหมายที่ต้องการวัด (Target Ability) เพียงอย่างเดียว แต่ยังวัดความสามารถแทรกซ้อนที่ไม่ต้องการวัด (Nuisance Ability) การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบเป็นการเปรียบเทียบผลการตอบข้อสอบระหว่างกลุ่มผู้สอบอย่างน้อย 2 กลุ่มขึ้นไป ปกตินิยมทำการเปรียบเทียบ 2 กลุ่ม ที่มีความสามารถระดับเดียวกัน โดยผู้สอบกลุ่มแรกเรียกว่า กลุ่มเปรียบเทียบ (Focal Group: F) เป็นกลุ่มที่สนใจศึกษาและคาดว่าจะจะเป็นกลุ่มที่เสียเปรียบในการตอบข้อสอบ และกลุ่มที่สอง เรียกว่า กลุ่มอ้างอิง (Reference Group: R) เป็นกลุ่มที่คาดว่าจะได้เปรียบในการตอบข้อสอบได้ถูกต้อง (Angoff, 1993, อ้างถึงใน วลีมาศ แซ่อึ้ง: 2543)

การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ มี 2 ประเภท คือ การทำหน้าที่ต่างกันของข้อสอบแบบเอกรูป (Uniform DIF) ซึ่งหมายถึง ข้อสอบที่ทำให้ผู้สอบกลุ่มอ้างอิงมีโอกาสในการตอบข้อสอบถูกมากกว่าผู้สอบกลุ่มเปรียบเทียบอย่างสม่ำเสมอในทุกระดับความสามารถ และการทำหน้าที่ต่างกันของข้อสอบแบบอนเอกรูป (Non-uniform DIF) ซึ่ง หมายถึง ข้อสอบที่ทำให้โอกาสในการตอบข้อสอบถูกของผู้สอบระหว่างกลุ่มแตกต่างกันอย่างไม่สม่ำเสมอ ในทุกระดับความสามารถ (Mellenbergh, 1982 อ้างถึงใน ศิริชัย กาญจนวาสี, 2550)

วิธีการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบมีหลายวิธี และวิธีหนึ่งที่มีความนิยมใช้มาอย่างต่อเนื่องทั้งนักการศึกษาไทยและนักการศึกษาต่างประเทศ ในช่วงหลายปีที่ผ่านมา คือ วิธีแมนเทล-แฮนส์เซล ที่เสนอโดยฮอลล์แลนด์และเทเยอร์ ในปี 1985 วิธีแมนเทล-แฮนส์เซล มีขั้นตอนในการวิเคราะห์ดังนี้ (Holland, & Thayer, 1985; อ้างถึงใน Holland, & Thayer, 1986)

1. คำนวณสัดส่วนการตอบข้อสอบของผู้สอบ 2 กลุ่ม คือ กลุ่มอ้างอิง (R) และกลุ่มเปรียบเทียบ (F) ในแต่ละช่วงคะแนน ดังนี้

กลุ่มผู้สอบ	ช่วงคะแนนที่ j		รวม
	ตอบถูก ได้ 1	ตอบผิด ได้ 0	
R	P_{Rj}	Q_{Rj}	1
F	P_{Fj}	Q_{Fj}	1

เมื่อ P_{Rj} แทน สัดส่วนของกลุ่มอ้างอิงที่อยู่ในช่วงคะแนน j ที่ตอบข้อสอบถูก

P_{Fj} แทน สัดส่วนของกลุ่มเปรียบเทียบที่อยู่ในช่วงคะแนน j ที่ตอบข้อสอบถูก

Q_{Rj} แทน 1 - P_{Rj}

Q_{Fj} แทน 1 - P_{Fj}

2. คำนวณค่าความน่าจะเป็นในรูปของสัดส่วน การตอบข้อสอบถูกและผิด ระหว่างกลุ่มในทุกช่วงคะแนน จากสูตร

$$\hat{\alpha}_{MH} = \frac{\sum A_j D_j / N_j}{\sum B_j C_j / N_j}$$

เมื่อ $\hat{\alpha}_{MH}$ เป็นตัวประมาณค่าขนาดอิทธิพลของ ข้อสอบที่ทำหน้าที่ต่างกัน (DIF effect size) มีค่าตั้งแต่ 0 ถึง μ

3. ทดสอบนัยสำคัญด้วยสถิติทดสอบไคกำลัง สอง (X^2) เพื่อทดสอบว่าค่า $\hat{\alpha}_{MH}$ ที่คำนวณได้มีค่า แตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้าพบว่า ค่า $\hat{\alpha}_{MH}$ แตกต่างจาก 1 อย่างมีนัยสำคัญทางสถิติ แสดงว่า ข้อสอบข้อนั้นทำหน้าที่ต่างกัน โดยจะทำหน้าที่ต่างกัน ของข้อสอบไปทางใด พิจารณาจากเกณฑ์ดังนี้

ถ้า $\hat{\alpha}_{MH} = 1.0$ แสดงว่า ข้อสอบทำหน้าที่ไม่ ต่างกัน

$\hat{\alpha}_{MH} < 1.0$ แสดงว่า ข้อสอบง่ายสำหรับ ผู้สอบกลุ่มเปรียบเทียบ

$\hat{\alpha}_{MH} > 1.0$ แสดงว่า ข้อสอบง่ายสำหรับ ผู้สอบกลุ่มอ้างอิง

วิธีแมนเทิล-แฮนส์เซล เป็นวิธีตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบที่มีความสะดวกในทางปฏิบัติ เสียค่าใช้จ่ายไม่แพง และสามารถใช้ได้กับตัวอย่างขนาด ไม่ใหญ่นัก (เสรี ชัดเข้ม, 2540: 10) มีขั้นตอนการคำนวณ

ที่ไม่สลับซับซ้อน โดยใช้หลักการของตารางการณัจจร (Contingency Table) แบบประเพณีนิยม (Traditional) ไม่มีการคำนวณทวนซ้ำ (Noniterative) มีการทดสอบ ทางสถิติแบบไม่ใช้พารามิเตอร์ (Nonparametric) จึง ไม่จำเป็นต้องใช้โมเดลประมาณค่า (ศิริชัย กาญจนวาสี, 2550: 127) นอกจากนี้ยังพบว่า ผลการตรวจสอบการทำ หน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล-แฮนส์เซล กับ วิธีที่ใช้ทฤษฎีการตอบสนองของข้อสอบ มีความคล้ายคลึง กันอีกด้วย (Camilli, & Shepard, 1994)

จากการศึกษาของนักการศึกษาหลายท่าน พบว่า วิธีแมนเทิล-แฮนส์เซล มีประสิทธิภาพสูงใน การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบ เกรุป (Swaminathan, & Rogers, 1990, Roger, & Swaminathan, 1993, Narayanan, & Swaminathan, 1994) อย่างไรก็ตาม เมื่อนำวิธีแมนเทิล-แฮนส์เซล ไป ใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ แบบอนิกรุปพบว่ามีประสิทธิภาพในการตรวจสอบไม่ ดีนัก (Swaminathan, & Rogers, 1990, Rogers, & Swaminathan, 1993, Narayanan, & Swaminathan, 1996, วลีมาศ แซ่อึ้ง, 2543) ซึ่งการทำหน้าที่ต่างกันของ ข้อสอบแบบอนิกรุปมักเกิดขึ้นได้ในสถานการณ์จริง (Mazor, Clauser, & Hambleton, 1994: 286)

ด้วยเหตุผลดังกล่าวข้างต้นจึงมีนักการศึกษา หลายท่าน พยายามที่จะพัฒนาวิธีแมนเทิล-แฮนส์เซล ให้ ใช้ได้ดีกับการทำหน้าที่ต่างกันของข้อสอบแบบอนิกรุป

ด้วย เช่น มาซอร์ คลาสเซอร์ และแฮมเบิลตัน (Mazor, Clauser, & Hambleton, 1994) ได้เสนอวิธีแมนเทล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบ (Variation of the Mantel-Haenszel Methods: VMH) ด้วยการจำลองกลุ่มผู้สอบขนาด 1,000 คน ต่อกลุ่ม ผลการศึกษาพบว่า วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ตรวจพบการทำหน้าที่ต่างกันของข้อสอบแบบอนุรูปได้ดีโดยไม่ทำให้อัตราความคลาดเคลื่อนประเภทที่ 1 สูงขึ้น และยังคงเป็นวิธีที่มีความสะดวกในทางปฏิบัติเช่นเดียวกับวิธีแมนเทล-แฮนส์เซลแบบดั้งเดิม ต่อมา เสรี ชัดเข้ม (2540) เสนอวิธีแมนเทล-แฮนส์เซลเพื่อใช้ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุรูปโดยใช้หลักการเช่นเดียวกับวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบที่เสนอโดยมาซอร์ คลาสเซอร์ และแฮมเบิลตันในปี ค.ศ. 1994 ด้วยการเพิ่มเงื่อนไขเกี่ยวกับค่าความยากของข้อสอบลงไป เรียกวิธีนี้ว่า วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ และเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุรูประหว่างวิธีแมนเทล-แฮนส์เซลแบบปกติ กับวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ โดยใช้วิธี IRT เป็นเกณฑ์ในการตัดสินความถูกต้องในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ โดยศึกษาจากข้อมูลเชิงประจักษ์พบว่าวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ สามารถตรวจพบข้อสอบที่ทำหน้าที่ต่างกันแบบอนุรูปได้สอดคล้องกับวิธี IRT มากกว่าวิธีแมนเทล-แฮนส์เซลแบบปกติ

นอกจากนี้ ควาก ดาวิสสัน และดาเวนพอร์ต (Kwak, Davison, & Davenport, 1997) เสนอวิธีอันซายแมนเทล-แฮนส์เซล (Unsigned Mantel-Haenszel Methods: UMH) ในการตรวจสอบการ

ทำหน้าที่ต่างกันของข้อสอบแบบอนุรูป โดยอาศัยหลักการวิธีแมนเทล-แฮนส์เซลแบบปกติ ต่างกันตรงวิธีจำแนกกลุ่มอ้างอิงและกลุ่มเปรียบเทียบ โดยวิธีแมนเทล-แฮนส์เซลแบบปกติ จะกำหนดให้กลุ่มที่คาดว่าจะได้เปรียบในการตอบข้อสอบได้ถูกต้องเป็นกลุ่มอ้างอิงเสมอในทุกข้อของข้อสอบและทุกกลุ่มคะแนนรวมที่วิเคราะห์ ในขณะที่วิธีอันซายแมนเทล-แฮนส์เซลในแต่ละข้อและแต่ละกลุ่มคะแนนรวมที่วิเคราะห์ กลุ่มอ้างอิงจะเปลี่ยนไปตามสัดส่วนของผู้ที่ตอบถูก ผลการศึกษาด้วยข้อมูลจำลองพบว่าวิธีอันซายแมนเทล-แฮนส์เซล สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุรูปได้ดีเท่ากับข้อสอบแบบอนุรูป และสามารถตรวจพบข้อสอบ ทำหน้าที่ต่างกันได้ดีกว่าวิธี Absolute Mean Deviation (AMD)

จากการศึกษาเอกสารต่าง ๆ ข้างต้น สรุปได้ว่าการทำหน้าที่ต่างกันของข้อสอบที่มักพบในสถานการณ์จริงเป็นแบบอนุรูปมากกว่าแบบอนุรูป และยังไม่พบการเปรียบเทียบประสิทธิภาพการตรวจสอบการทำหน้าที่ต่างของข้อสอบแบบอนุรูปด้วยวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบกับวิธีอันซายแมนเทล-แฮนส์เซล ผู้วิจัยจึงมีความสนใจที่จะศึกษาเปรียบเทียบในประเด็นดังกล่าวกับข้อมูลเชิงประจักษ์ โดยกำหนดให้ผู้สอบจากโรงเรียนขนาดใหญ่เป็นกลุ่มอ้างอิงและผู้สอบจากโรงเรียนขนาดเล็กเป็นกลุ่มเปรียบเทียบ เลือกใช้จำนวนข้อสอบที่ทำหน้าที่ต่างกันแบบอนุรูปที่ตรวจพบโดยวิธีดถอยโลจิสติกเป็นเกณฑ์ เนื่องจากวิธีดถอย โลจิสติกเป็นวิธีการที่มีโมเดลพื้นฐานทางสถิติ สามารถตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนุรูปและแบบอนุรูปได้ดี ถูกต้องใกล้เคียงกับวิธี IRT ผลการตรวจสอบการทำหน้าที่ต่างกัน มีความถูกต้องสูง (Hollad and Thayer, 1988)

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบผลการตรวจพบข้อสอบที่ทำหน้าที่ต่างกันแบบอนุกรม ระหว่างวิธี แมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันชายแมนเทิล-แฮนส์เซล โดยใช้จำนวนข้อสอบทำหน้าที่ต่างกันแบบอนุกรม ที่ตรวจพบตามวิธีถดถอยโลจิสติก เป็นเกณฑ์

สมมติฐานของการวิจัย

จำนวนข้อสอบทำหน้าที่ต่างกันแบบอนุกรม ที่ตรวจพบระหว่างวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันชายแมนเทิล-แฮนส์เซล แตกต่างกัน

วิธีดำเนินการวิจัย

1. ประชากรที่ใช้ในการวิจัย ได้แก่ นักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2553 ของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษากาญจนบุรี จำนวน 11,192 คน โดยกำหนดให้นักเรียนจากโรงเรียนขนาดใหญ่ เป็นกลุ่มอ้างอิงและนักเรียนจากโรงเรียนขนาดเล็ก เป็นกลุ่มเปรียบเทียบ

2. ตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2553 ของโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษากาญจนบุรี จำนวน 1,400 คน จำแนกเป็นนักเรียนจากโรงเรียนขนาดใหญ่ 726 คน และจากโรงเรียนขนาดเล็ก 674 คน ซึ่งได้มาโดยวิธีสุ่มตัวอย่างแบบแบ่งชั้น (Stratified Random Sampling) กำหนดให้ขนาดโรงเรียนเป็นชั้น (Strata) และมีโรงเรียนเป็นหน่วยสุ่ม (Sampling Unit)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในวิจัยครั้งนี้ เป็นแบบสอบถามด้านพุทธิพิสัย กลุ่มสาระการเรียนรู้วิทยาศาสตร์ เรื่อง จำนวน สำหรับนักเรียน ชั้นมัธยมศึกษาปีที่ 1 ชนิด 4

ตัวเลือก จำนวน 40 ข้อ มีค่าความยากง่าย ตั้งแต่ 0.23-0.80 มีค่าอำนาจจำแนก ตั้งแต่ 0.20-0.92 มีค่าความเชื่อมั่น เท่ากับ 0.90 และเป็นข้อสอบที่มีความเป็นเอกมิติ โดยพิจารณาจากผลการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) พบว่าค่าไอเกน (eigenvalue) ขององค์ประกอบแรกมีค่าเท่ากับ 4.732 และองค์ประกอบที่ 2 เท่ากับ 1.634 แสดงให้เห็นว่าอัตราส่วนของค่าไอเกนขององค์ประกอบแรก และองค์ประกอบถัดไปมีค่าสูง แสดงว่าแบบสอบชุดนี้วัดคุณลักษณะเด่นเพียงลักษณะเดียว

การเก็บรวบรวมข้อมูล

1. ผู้วิจัยขอหนังสือจากสำนักงานบัณฑิตศึกษามหาวิทยาลัยราชภัฏกาญจนบุรี ถึงผู้บริหารสถานศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษากาญจนบุรี ที่ถูกเลือกเป็นตัวอย่าง เพื่อขอความอนุเคราะห์ในการจัดสอบนักเรียน

2. ผู้วิจัยรวบรวมกระดาษคำตอบ และตรวจให้คะแนน

3. นำข้อมูลที่ได้มาทำการวิเคราะห์ทางสถิติโดยใช้โปรแกรมสำเร็จรูป SPSS for Windows

การวิเคราะห์ข้อมูล

1. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอนุกรม ด้วยวิธีถดถอยโลจิสติก โดยใช้สมการมาตรฐานของโมเดลการถดถอยโลจิสติกคำนวณผลการตอบข้อสอบถูก ดังนี้ (Swaminathan & Rogers, 1990)

$$P(U_{ij} = 1 / \theta_{ij}) = \frac{\exp(\beta_0 + \beta_1 \theta_{ij})}{1 + \exp(\beta_0 + \beta_1 \theta_{ij})}, i = 1, 2, \dots, n; j = 1, 2$$

เมื่อ U_{ij} แทน ผลการตอบข้อสอบของผู้เข้าสอบคนที่ i ในกลุ่ม j

θ_{ij} แทน ค่าความสามารถที่สังเกตได้ของผู้เข้าสอบคนที่ i ในกลุ่ม j

β_{0j} แทน ค่าพารามิเตอร์จุดตัด (intercept Parameter)

β_{1j} แทน ค่าพารามิเตอร์ความชันสำหรับกลุ่ม j (Slope Parameter)

จากโมเดลดังกล่าว ถ้า $01 = 02$ และ $11 = 12$ แล้ว ฟังก์ชันถดถอยโลจิสติกของ ผู้เข้าสอบสองกลุ่มเหมือนกัน แสดงว่าข้อสอบทำหน้าที่ไม่ต่างกัน ถ้า $11 = 12$ แต่ $01 \neq 02$ แล้ว ฟังก์ชันถดถอยโลจิสติกของ ผู้เข้าสอบสองกลุ่มขนานกันแต่ไม่ทับกัน แสดงว่า ข้อสอบทำหน้าที่ต่างกันแบบเอกรูป (Uniform DIF) และถ้า $01 \neq 02$ แต่ $11 \neq 12$ แล้ว ฟังก์ชันถดถอยโลจิสติกของ ผู้เข้าสอบไม่ขนานกัน แสดงว่าข้อสอบทำหน้าที่ต่างกันแบบอนเอกรูป (Non-uniform DIF)

ผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอนเอกรูปจากวิธีถดถอยโลจิสติกจะถูกนำไปใช้เป็นเกณฑ์ สำหรับการตรวจสอบการทำหน้าที่ต่างกันของ วิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันซายแมนเทิล-แฮนส์เซล

2. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอนเอกรูป ด้วยวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ที่เสนอโดย มาซอร์ และคณะ (Mazor, Clauser, & Hambleton, 1994) ตามขั้นตอน ดังนี้

2.1 ใช้กลุ่มผู้สอบทั้งหมดในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีแมนเทิล-แฮนส์เซล แบบปกติ โดยกำหนดกลุ่มอ้างอิงและกลุ่มเปรียบเทียบไว้ในเบื้องต้น ซึ่งในงานวิจัยนี้ กำหนดให้นักเรียนจากโรงเรียนขนาดใหญ่ เป็นกลุ่มอ้างอิง นักเรียนจากโรงเรียนขนาดเล็ก เป็นกลุ่มเปรียบเทียบ

2.2 หาค่าเฉลี่ยของคะแนนทั้งหมด เพื่อใช้เป็นเกณฑ์ในการแบ่งกลุ่มความสามารถของผู้สอบ ออกเป็น 2 กลุ่ม เรียกกลุ่มที่มีคะแนนสูงกว่าค่าเฉลี่ย ว่า กลุ่มที่มีความสามารถสูง และกลุ่มที่มีคะแนนต่ำกว่าค่าเฉลี่ยว่า กลุ่มที่มีความสามารถต่ำ

2.3 วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธี แมนเทิล-แฮนส์เซล แบบปกติ เฉพาะกลุ่มที่มีความสามารถสูง

2.4 วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบด้วยวิธี แมนเทิล-แฮนส์เซล แบบปกติ เฉพาะกลุ่มที่มีความสามารถต่ำ

3. วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบแบบอนเอกรูป ด้วยวิธีอันซายแมนเทิล-แฮนส์เซล ที่เสนอโดย ควาก ดาวิสสัน และดาเวนพอร์ท (Kwak, Davison, & Davenport, 1997) ตามขั้นตอน ดังนี้

3.1 จำแนกกลุ่มอ้างอิงซึ่งจะเปลี่ยนไปตามผลการสอบ ตามหลักการของวิธีแมนเทิล-แฮนส์เซล กลุ่มอ้างอิงจะเป็นกลุ่มที่ตอบได้ถูกต้องเสมอ แต่วิธีอันซายแมนเทิล-แฮนส์เซล แต่ละข้อ ของข้อสอบ กลุ่มอ้างอิงจะเปลี่ยนไปตามสัดส่วนของผู้ที่ตอบ

3.2 วิเคราะห์การทำหน้าที่ต่างกันของข้อสอบ ด้วยวิธีแมนเทิล-แฮนส์เซล แบบปกติ

4. เปรียบเทียบสัดส่วนของข้อสอบทำหน้าที่ต่างกันของข้อสอบแบบอนเอกรูป ระหว่างวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและวิธีอันซายแมนเทิล-แฮนส์เซล

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการหาคุณภาพของแบบสอบ ได้แก่ ความเที่ยงของคะแนนแบบสอบ จากการประมาณค่าสัมประสิทธิ์ความคงที่ภายใน โดยวิธี Kuder Richardson 20 (KR 20) (Crocker, & Algina, 1986) และความเป็นเอกมิติของแบบสอบ โดยใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA)

2. สถิติที่ใช้ในการทดสอบสัดส่วนของข้อสอบทำหน้าที่ต่างกันของข้อสอบแบบอนเอกรูประหว่างวิธีแมนเทิล-แฮนส์เซล แบบแบ่งกลุ่มความสามารถของผู้สอบและวิธีอันซายแมนเทิล-แฮนส์เซล ใช้สถิติทดสอบ Z (Z-test)

ผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูล พบว่า วิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันซายแมนเทิล-แฮนส์เซล ตรวจสอบจำนวนข้อสอบทำหน้าที่ต่างกันแบบอนุกรม ได้สอดคล้องกับวิธีถดถอยโลจิสติกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยพบว่าวิธีอันซายแมนเทิล-แฮนส์เซล ตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้สอดคล้องกับวิธีถดถอย โลจิสติกมากกว่าวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ดังตารางที่ 1 และ 2

ตารางที่ 1 การเปรียบเทียบข้อสอบทำหน้าที่ต่างกันแบบอนุกรม ตามวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ (VMH) และ ตามวิธีอันซายแมนเทิล-แฮนส์เซล (UMH) โดยใช้วิธีถดถอยโลจิสติก (LR) เป็นเกณฑ์

ข้อที่	ค่าความยาก	LR		VMH		UMH	
		B	Sig.	$\hat{a}_{MH}$	Sig.	$\hat{a}_{MH}$	Sig.
1	0.50	-0.073	0.025*	-	-	-	-
3	0.23	0.101	0.000**	-	-	0.623	0.001**
4	0.46	-0.116	0.000**	1.534	0.001*	1.304	0.044*
5	0.54	-0.082	0.002*	-	-	0.717	0.007*
6	0.38	0.067	0.008*	1.734	0.000**	0.577	0.000**
7	0.46	-0.060	0.016*	0.694	0.0031*	0.698	0.004*
13	0.50	-0.051	0.023*	-	-	0.773	0.032*
17	0.46	-0.106	0.000**	-	-	0.732	0.010**
20	0.54	-0.086	0.000**	1.326	0.017*	0.695	0.002*
25	0.38	0.053	0.018*	-	-	0.713	0.007*
26	0.50	0.084	0.000**	0.698	0.045*	0.588	0.000**
27	0.46	0.050	0.037*	-	-	0.752	0.041*
29	0.35	0.054	0.019*	1.808	0.000**	0.570	0.000**
32	0.46	0.050	0.022*	-	-	0.657	0.001**
36	0.46	0.053	0.012*	-	-	0.726	0.007*
39	0.50	0.067	0.003*	-	-	0.710	0.006*
40	0.65	0.056	0.013*	1.746	0.000**	0.583	0.000**

* ที่ระดับนัยสำคัญ .05 **ที่ระดับนัยสำคัญ .01

เกณฑ์พิจารณาระดับค่าความยากง่าย (พรณี สীগิจวัฒน์ 2554)

0.81 – 1.00 ง่ายมาก

0.61 – 0.80 ง่าย

0.40 – 0.60 ปานกลาง

0.20 – 0.39 ยาก

0.00 – 0.19 ยากมาก

ตารางที่ 2 การเปรียบเทียบอัตราการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบข้อสอบแบบอเนกรูป ตามวิธี
แมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ (VMH) และวิธีอันชายแมนเทิล-แฮนส์เซล
(UMH)

วิธี	จำนวนข้อ	%	Z
VMH	7	41.18	- 3.30a
UMH	16	94.12	

a แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ($\pm Z.005 \approx \pm 2.575$)

จากตารางที่ 1 และ 2 พบว่า วิธีถดถอย โลจิสติกตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป 17 ข้อ วิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันได้สอดคล้องกับวิธีถดถอยโลจิสติก จำนวน 7 ข้อ คิดเป็นร้อยละ 41.18 โดยข้อสอบข้อ 40 เป็นข้อสอบที่มีค่าความยากง่ายระดับง่าย ข้อ 4, 7, 20 และ 26 เป็นข้อสอบที่มีค่าความยากง่ายระดับปานกลาง และข้อ 6 และ 29 เป็นข้อสอบที่มีค่าความยากง่ายระดับยาก

วิธีอันชายแมนเทิล-แฮนส์เซล ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอเนกรูปได้สอดคล้องกับวิธีถดถอยโลจิสติก จำนวน 16 ข้อ คิดเป็นร้อยละ 94.12 โดยข้อสอบข้อ 40 เป็นข้อสอบที่ค่าความยากง่ายระดับง่าย ข้อสอบข้อ 4, 5, 7, 13, 17, 20, 26, 27, 32, 36, 39 เป็นข้อสอบที่ค่าความยากง่ายระดับปานกลาง และข้อสอบข้อ 3, 6, 25, 26 เป็นข้อสอบที่ค่าความยากง่ายระดับยาก

เมื่อทดสอบสัดส่วนข้อสอบทำหน้าที่ต่างกันแบบอเนกรูป ระหว่างวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและวิธี

อันชายแมนเทิล-แฮนส์เซล เมื่อใช้จำนวนข้อสอบ ทำหน้าที่ต่างกันแบบอเนกรูป ที่ตรวจพบโดยวิธีถดถอยโลจิสติก เป็นเกณฑ์ ดังแสดงในตารางที่ 2 พบว่า วิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันชายแมนเทิล-แฮนส์เซล ตรวจพบจำนวนข้อสอบที่ทำหน้าที่ต่างกันแบบอเนกรูปแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยพบว่า วิธีอันชายแมนเทิล-แฮนส์เซล ตรวจพบข้อสอบทำหน้าที่ต่างกันได้สอดคล้องกับวิธีถดถอย โลจิสติก มากกว่าวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และเพื่อเป็นการยืนยันผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูป โดยวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและวิธีอันชายแมนเทิล-แฮนส์เซล ผู้วิจัยจึงจำลองข้อมูลจากข้อมูลเชิงประจักษ์ (1,400 คน) จำนวน 10 แฟ้มข้อมูล โดยแต่ละแฟ้มข้อมูลมีจำนวนตัวอย่าง 800 คน จำแนกเป็น ตัวอย่างจากโรงเรียนขนาดใหญ่ 400 คนและตัวอย่างจากโรงเรียนขนาดเล็ก 400 คนและดำเนินการวิเคราะห์ข้อมูลตามรูปแบบเดิม ได้ผลแสดงดังตารางที่ 3

ตารางที่ 3 ผลการตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกัน ตามวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่ม ความสามารถของผู้สอบ (VMH) และตามวิธีอันซายแมนเทิล-แฮนส์เซล (UMH) โดยการวิเคราะห์ ข้อมูลจากการจำลองตัวอย่าง จากข้อมูลเชิงประจักษ์ จำนวน 10 แฟ้มข้อมูล

แฟ้มข้อมูล	LR	UMH		VMH		Z
	จำนวนข้อ	จำนวนข้อ	%	จำนวนข้อ	%	
1	7	6	85.71	3	42.86	1.67 ^b
2	10	7	70.00	4	40.00	1.35 ^c
3	12	10	83.33	5	41.67	2.11 ^b
4	12	10	83.33	6	50.00	1.73 ^b
5	7	7	100.00	4	57.14	1.95 ^b
6	14	10	71.43	7	50.00	1.16
7	10	9	90.00	8	80.00	0.63
8	11	8	72.73	9	81.82	-0.51
9	14	12	85.71	10	71.43	0.92
10	12	10	83.33	7	58.33	1.35 ^c

^b มากกว่าอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($Z_{.05} \approx 1.645$)

^c มากกว่าอย่างมีนัยสำคัญทางสถิติที่ระดับ .10 ($Z_{.10} \approx 1.282$)

จากตารางที่ 3 พบว่า จากการจำลองข้อมูลตัวอย่าง จากข้อมูลเชิงประจักษ์ จำนวน 10 ครั้ง พบว่า วิธีอันซายแมนเทิล-แฮนส์เซล ตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันของข้อสอบแบบอนุกรม ได้สอดคล้องกับวิธีถดถอยโลจิสติก ได้มากกว่าวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 4 ครั้ง ได้แก่ ครั้งที่ 1, 3, 4 และ 5 และวิธีอันซายแมนเทิล-แฮนส์เซล ตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้สอดคล้องกับวิธีถดถอยโลจิสติก ได้มากกว่าวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ .10 จำนวน 2 ครั้ง ได้แก่ ครั้งที่ 2 และ 10 ซึ่งผลการวิจัยเป็นไปในทิศทางเดียวกับการตรวจสอบในตัวอย่าง 1,400 คน

สรุปผลการวิจัย

วิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันซายแมนเทิล-แฮนส์เซล ตรวจพบจำนวนข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้สอดคล้องกับวิธีถดถอยโลจิสติก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยพบว่าวิธีอันซายแมนเทิล-แฮนส์เซล ตรวจพบข้อสอบ ทำหน้าที่ต่างกันแบบอนุกรมได้สอดคล้องกับวิธีถดถอยโลจิสติก จำนวน 16 ข้อ (ส่วนใหญ่เป็นข้อสอบที่มีค่าความยากง่ายอยู่ในระดับปานกลาง) ซึ่งมากกว่าวิธีแมนเทิล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ที่ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบอนุกรมได้สอดคล้องกับวิธีถดถอยโลจิสติกเพียง 7 ข้อ (ส่วนใหญ่เป็นข้อสอบที่อยู่ในระดับปานกลาง)

อภิปรายผลการวิจัย

การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบบอเนกรูป ตามวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ส่วนใหญ่ตรวจพบข้อสอบทำหน้าที่ต่างกันเมื่อข้อสอบเป็นข้อสอบที่มีค่าความยากง่ายอยู่ในระดับปานกลาง ชัดแจ้งกับการศึกษาของมาซอร์และคณะ (Mazor, et al., 1994) ที่พบว่า วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบบอเนกรูป ส่วนใหญ่เป็นข้อสอบที่มีค่าความยากง่ายอยู่ในระดับง่าย แต่ไม่พบในข้อสอบที่มีค่าความยากง่ายอยู่ในระดับยาก ในขณะที่ เสรี ชัดแจ้ง, (2540) พบว่า วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและค่าความยากของข้อสอบ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบบอเนกรูปในข้อสอบที่มีค่าความยากง่ายอยู่ในระดับง่ายถึงปานกลาง และเมื่อพิจารณาวิธีอันชายแมนเทล-แฮนส์เซล พบว่าวิธีนี้ ตรวจพบข้อสอบทำหน้าที่ต่างกันแบบบอเนกรูปได้ในข้อสอบที่มีค่าความยากง่ายอยู่ในระดับปานกลาง จากผลการศึกษาที่กล่าวมาข้างต้นอาจพอสรุปได้สอดคล้องกับโรเจอร์และ สวานิมาทาน (Rogers, & Swaminathan, 1993) และมาซอร์ และคณะ (Mazor, et al., 1994) ที่กล่าวว่าค่าความยากง่ายของข้อสอบมีผลต่อการตรวจพบการทำหน้าที่ต่างกันของข้อสอบแบบบอเนกรูป

ข้อเสนอแนะ

ข้อเสนอแนะที่ได้จากการวิจัย

วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและวิธีอันชายแมนเทล-แฮนส์เซลสามารถตรวจสอบข้อสอบทำหน้าที่ต่างกันแบบบอเนกรูปได้ในทางปฏิบัติจริง ถ้าเป็นข้อสอบที่มีค่าความยากง่ายในระดับง่าย สามารถใช้ได้ทั้งสองวิธี หากเป็นข้อสอบที่มีค่าความยากง่ายในระดับปานกลางถึงยาก วิธีอันชายแมนเทล-แฮนส์เซลดีกว่าวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ

ข้อเสนอแนะสำหรับการทำวิจัยครั้งต่อไป

1. ควรมีการศึกษการเปรียบเทียบประสิทธิภาพการทำหน้าที่ต่างกันของข้อสอบแบบบอเนกรูป วิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและวิธีอันชายแมนเทล-แฮนส์เซล กับวิธีวิเคราะห์อื่นๆ
2. ควรมีการศึกษผลการตรวจพบข้อสอบทำหน้าที่ต่างกันแบบบอเนกรูป ด้วยวิธีแมนเทล-แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบ และวิธีอันชายแมนเทล-แฮนส์เซล กับกลุ่มตัวอย่างขนาดต่างๆ กัน และจำนวนข้อสอบที่ต่างกัน

เอกสารอ้างอิง

- วลีมาศ แซ่อึ้ง. (2543). *การเปรียบเทียบอำนาจการทดสอบและอัตราความคลาดเคลื่อนชนิดที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบบอเนกรูประหว่างวิธีซิปเทสท์ปรับปรุงใหม่ วิธีซิป เทสท์ วิธีแมนเทล-แฮนเซลและวิธีถดถอยโลจิสติก*. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา, คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย.
- พรณี ลีกิจวัฒนา. (2554). *วิธีวิจัยทางการศึกษา*. (พิมพ์ครั้งที่ 7 ปรับปรุงแก้ไข). ห้างหุ้นส่วนจำกัด มิน เซอร์วิส ซัพพลาย, คณะครุศาสตร์อุตสาหกรรม, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ศิริชัย กาญจนวาสี. (2550). *ทฤษฎีการทดสอบแนวใหม่*. กรุงเทพมหานคร, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- เสรี ชัดแจ้ง. (2540). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบระหว่างวิธีแมนเทิล-เฮนส์เซลส์แบบปกติ กับวิธีแมนเทิล - เฮนส์เซลส์แบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ. วิทยานิพนธ์ดุษฎีบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา, คณะครุศาสตร์, จุฬาลงกรณ์มหาวิทยาลัย.
- Camilli, G. and Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. California : Sage Publication. 64-70.
- Crocker, Linda & Algina, James. (1986). *Introduction to classical and modern test theory*. Orland, Florida: Harcourt Brace Jovanovich College.
- Holland, P.W. & Thayer, D. T. (1986). Differential item functioning and the Mantel-Haenszel procedure. (Technical Rep. No.86-96.) Princeton, NJ: Educational testing Service.
- Holland, P. W. & Thayer, D. T. (1988). *Differential item functioning and the Mantel-Haenszel procedure*. In Wainer, P.W. and Braun, H. T.(Eds.) *Test validity* . Hillsdale, NJ: Lawrence Erlbaum Associates. 129-145.
- Kwak, Davison, & Davenport, (1997). *Unsigned Mantel-Heanszel Statistic for Detecting Uniform and Nonuniform DIF*. 199 : 3.
- Mazor, K. M.,B.E. Clauser., & R.K. Hambleton. (1994). *Identification of non uniform differential item functioning using a variation of the Mantel-Haenszel procedure*. Educational and Psychological Measurement. 54(2): 284-291.
- Narayanan, P., & H. Swaminathan. (1994). *Performance of the Mantel-Haenzel and imultaneous Item Bias Precedures for Detecting Differential Item Function*. Applied Phychological Measurement.(December) 315-328.
- Narayanan, P., & H. Swaminathan. (1996). *Identification of item that show non uniform DIF*. Applied Psychological Measurement. (September) 257-274.
- Roger H. J. & Swaminathan. (1993). *A comparison of logistic regression and Mantel-Haenszel procedures for detecting difference item functioning*. Applied Psychological Measurement.17(2) : 105-116.
- Swaminathan, H. & Rogers H. J. (1990). Detecting Differential Item Functioning using Logistic Regression Precedures. *Journal of Education Measurement*. 27(4) : 361-370 ; Winter.