การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป เมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์

ปริญญานิพนธ์

ของ

สุทธิวรรณ พีรศักดิ์โสภณ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
เมษายน 2546

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

370.284
ส๋๋ส
๋.๋

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป เมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

บทคัดย่อ

ของ

สุทธิวรรณ พีรศักดิ์โสภณ

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษาดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
เมษายน 2546

ฝ่ายวิจัย

สุทธิวรรณ พีรศักดิ์โสภณ. (2546). การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปเมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์. ปริญญานิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ
คณะกรรมการควบคุม : รองศาสตราจารย์ ดร. บุญเชิด ภิญโญอนันตพงษ์,
อาจารย์ ดร.ชูศักดิ์ ขัมภลิขิต และผู้ช่วยศาสตราจารย์ ดร. นิคม ดังคะพิภพ.

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาและหาความแม่นยำของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยตรวจสอบจากความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย ด้วยการแปลงคะแนนแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ เมื่อใช้วิธีการประมาณค่าสี่วิธี คือ วิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายจำนวนสามฉบับ คือ ฉบับด้านตัวเลข ด้านภาษา และด้านเหตุผล กลุ่มตัวอย่างที่ใช้ในการวิจัยได้แก่ นักเรียนชั้นมัธยมศึกษาตอนปลาย ปีการศึกษา 2544 จำนวน 2,000 คน และสุ่มมาเป็นตัวอย่างจำลองสามขนาดคือ 300, 500 และ 800 คน แต่ละขนาดได้มาโดยการสุ่มซ้ำ 30 ครั้ง

ผลการวิจัยสรุปได้ดังนี้

1. ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ เมื่อใช้วิธีประมาณค่าแบบทวินาม ทวินามประกอบ วิธีของเฟลด์ - ควอลส์ ให้ค่าสูงตรงกลางของคะแนน PR ให้ค่าต่ำตรงด้านปลายของ PR รูปร่างของกราฟคล้ายรูปตัวยูหัวกลับ แต่เมื่อคะแนนแปลงเป็นคะแนนปกติเมื่อใช้วิธีประมาณค่าแต่ละวิธีให้ค่าต่ำตรงช่วงกลางของคะแนนปกติ รูปร่างของกราฟเป็นเส้นหยักมีลักษณะเป็นโค้งสองยอด และวิธีทฤษฎีการตอบข้อสอบ ให้ค่าใกล้เคียงกันกับเมื่อใช้คะแนนแปลงแบบคะแนนปกติ แต่รูปร่างของกราฟคล้ายตัวยูหัวกลับแต่ไม่ราบเรียบ ดังนั้นเมื่อใช้คะแนนแปลงต่างกัน รูปร่างของกราฟต่างกัน วิธีการแปลงแบบคะแนนปกติเป็นการแปลงที่ทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดเหมาะสมกับการประมาณค่าความสามารถที่แท้จริง

2. ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสำหรับตัวอย่างจำลองที่มีขนาดต่างกันในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ เมื่อใช้วิธีประมาณค่าแบบเดียวกันมีค่าสูงกว่าในคะแนนแปลงแบบคะแนนปกติ แต่ไม่สามารถสรุปได้แน่ชัดว่าสำหรับตัวอย่างขนาดใดจะทำให้ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยสูงกว่ากัน และวิธีทฤษฎีการตอบข้อสอบให้ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยใกล้เคียงกับแต่ละวิธีที่ใช้ในคะแนนแปลงแบบคะแนนปกติ

ESTIMATING CONDITIONAL SCALE – SCORE STANDARD ERRORS OF MEASUREMENT USING ITEM RESPONSE THEORY, BINOMIAL, COMPOUND BINOMIAL, AND FELDT – QUALLS METHOD.

AN ABSTRACT

BY

SUTHIWAN PIRASAKSOPON

Presented in partial fulfillment of the requirements

For the Doctor of Education degree in Testing and Measurement

at Srinakharinwirot University

April 2003

Suthiwan Pirasaksopon. (2003). Estimating Conditional Scale - Score Standard Errors of Measurement Using Item Response Theory, Binomial, Compound Binomial, and Feldt - Qualls Method. Doctor Thesis, (Testing and Educational Measurement). Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee : Assoc. Prof. Dr. Boonchird Pinyoanuntapong, Dr. Choosak Khampalikit, and Assist. Prof. Dr. Nikom Tangkapipop.

This study aimed to estimate and evaluate the accuracy of the estimated conditional standard errors of measurement : average bias (abias) and average error variance (aev), for the scale score (CSEMs) using Item response theory, Binomial, Compound Binomial and Feldt – Qualls procedure. The instruments used were the scholastic aptitude tests ; Numeric, Verbal and Reasoining. The data obtained from 2,000 high school students enrolled in the academic classes of 2,001. The sampling model sizes : 300, 500 and 800 respectively, were replicated 30 times were analyzed and tabulated. The findings were as follows :

1. The conditional standard errors of measurement for percentile ranks used by Binomial, Compound Binomial and Feldt - Qualls method, were varied across scale score and the same curve. The conditional percentile ranks standard errors of measurement were nearly consistent in the middle of percentile ranks and lower in the extreme scale. The curve looked like an inverted U shape. The conditional standard errors of measurement for normalized score were low in the interval of middle scale score and jagged CSEMs as bimodal shape. The CSEMs used item response theory method was similar to CSEMs that of normalized score. In summary, the conditional normalized standard errors of measurement were useful for implication.

2. Using the Binomial, Compound Binomial and Feldt – Qualls procedures with 300, 500, and 800 sample size, the average bias and average error variance of percentile ranks and normalized score of each test were low and similar to each other, but the curve - shape was different. The average bias and average error variance of percentile rank were higher than that of normalized score. The item response theory procedure produced average bias and average error variance close to that of normalized score. Therefore, it could not obviously be concluded which method provided a high average error variance when using different sample sizes.

ปริญญานิพนธ์ฉบับนี้ ได้รับทุนสนับสนุนการวิจัยจาก ทุนทบวงมหาวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ ประจำปีการศึกษา 2545

ปริญญานิพนธ์
เรื่อง

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป
เมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม
วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์

ของ
นางสาวสุทธิวรรณ พีรศักดิ์โสภณ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษาดุษฎีบัณฑิต สาขาการทดสอบและวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

..... คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.นภาภรณ์ หะวานนท์)
วันที่ 11 เดือน กุมภาพันธ์ พ.ศ. 2546

คณะกรรมการสอบปริญญานิพนธ์

.....ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์)

.....กรรมการ
(อาจารย์ ดร.ชูศักดิ์ ขัมภลิขิต)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.นิคม ดังคะพิภพ)

.....กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ ดร. วิรัช วรรณรัตน์)

.....กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ดร.บุญลือ ทองอยู่)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์และความกรุณาจาก รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์ ดร.ชูศักดิ์ ขัมภลิขิต และ ผู้ช่วยศาสตราจารย์ ดร.นิคม ตังคะพิภพ ที่กรุณาให้คำปรึกษา ข้อคิดเห็น และแก้ไขข้อบกพร่อง จนทำให้งานวิจัยเรื่องนี้สำเร็จลงได้ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณกรรมการที่แต่งตั้งเพิ่มเติมคือ รองศาสตราจารย์ ดร.วิรัช วรรณรัตน์ และ ดร.บุญลือ ทองอยู่ ที่ให้ความกรุณาเป็นกรรมการในการสอบปริญญานิพนธ์และให้ข้อเสนอแนะเพิ่มเติมที่เป็นประโยชน์ ทำให้ปริญญานิพนธ์นี้มีความสมบูรณ์ยิ่งขึ้น

ขอขอบคุณ ดร.วรรณี ตปนียากร ที่ช่วยกรุณาติดต่อสำเนาโปรแกรม CBINSS และขอขอบคุณ Prof. Robert L. Brennan แห่งมหาวิทยาลัยไอโอวา ประเทศสหรัฐอเมริกา ที่ให้คำแนะนำในการวิเคราะห์ข้อมูลจากโปรแกรมนี้จนทำให้สำเร็จลุล่วงไปด้วยดี

ขอขอบคุณ อาจารย์สมกิจ กิจพูนวงษ์ สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรี นครินทรวิโรฒ อาจารย์ธีระพล ศิระบูชา คณะบัญชีและการจัดการ อาจารย์วัชรชัย วีรกุลเกษตร กลุ่มงานระบบเครือข่าย มหาวิทยาลัยมหาสารคาม และ อาจารย์ภาณุวัชร ปุรณะศิริ สำนักงานสามัญศึกษาจังหวัดบุรีรัมย์ ที่ให้ความอนุเคราะห์และช่วยเหลือเป็นอย่างดีในเรื่องของเทคนิคในการวิเคราะห์ข้อมูล

สุดท้าย ผู้วิจัยขอขอบพระคุณทุกท่านที่ไม่ได้กล่าวนามมา ณ ที่นี้ที่สนับสนุนให้ความช่วยเหลือในการทำปริญญานิพนธ์ฉบับนี้จนสำเร็จลุล่วงไปด้วยดี คุณค่าและประโยชน์ของปริญญานิพนธ์ฉบับนี้ ขอมอบเป็นเครื่องบูชาแด่ บิดา - มารดา ครู - อาจารย์ ตลอดจนผู้มีพระคุณของผู้วิจัยทุกท่านที่ช่วยให้ผู้วิจัยประสบความสำเร็จในการศึกษา

สุทธิวรรณ พีรศักดิ์โสภณ

สารบัญ

บทที่		หน้า
1	บทนำ	1
	ภูมิหลัง	1
	จุดมุ่งหมายของการวิจัย	4
	ความสำคัญของการวิจัย	5
	ขอบเขตของการวิจัย	5
	นิยามศัพท์เฉพาะ	6
2	เอกสารและงานวิจัยที่เกี่ยวข้อง	9
	การแปลงคะแนนและประเภทของการแปลงคะแนน	9
	คะแนนดิบ	9
	คะแนนแปลงรูปหรือคะแนนที่มีมาตรวัดเดียวกัน	9
	การแปลงแบบเชิงเส้นตรง	10
	การแปลงแบบไม่เป็นเชิงเส้นตรง	11
	ความคลาดเคลื่อนมาตรฐานในการวัด	13
	การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนดิบ	14
	การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงรูป	20
	วิธีที่ใช้โมเดลคะแนนจริงอย่างเข้มงวด (Strong True Score)	20
	การประมาณค่า CSEM โดยใช้ทฤษฎีการตอบข้อสอบ	25
	วิธีพหุนาม (Polynomial Procedure)	28
	วิธีเฟลด์ - ควอลส์ (Feldt - Qualls)	29
	วิธีแบบทวินาม	30
	วิธีแบบทวินามประกอบ (Compound Binomial)	32
	ตัวประมาณที่ดี	35
	งานวิจัยที่เกี่ยวข้อง	35
	งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนมาตรฐานในการวัด	35
	งานวิจัยที่เกี่ยวข้องกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ	37
	งานวิจัยที่เกี่ยวข้องกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป	41
	สมมติฐานการวิจัย	46

สารบัญ (ต่อ)

บทที่		หน้า
3	วิธีดำเนินการวิจัย	47
	การกำหนดประชากรและกลุ่มตัวอย่าง	47
	เครื่องมือที่ใช้ในการวิจัย	49
	การคัดเลือกข้อสอบ	52
	การทดลองเครื่องมือ	52
	การเก็บรวบรวมข้อมูล	53
	การจัดกระทำข้อมูล	54
	สถิติและการวิเคราะห์ข้อมูล	55
4	ผลการวิเคราะห์ข้อมูล	60
	สัญลักษณ์และอักษรย่อที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล	60
	ผลการวิเคราะห์ข้อมูล	61
5	สรุป อภิปรายผล และข้อเสนอแนะ	117
	จุดมุ่งหมายการวิจัย	117
	ข้อมูลที่นำมาศึกษา	117
	การเก็บรวบรวมข้อมูลและจัดกระทำข้อมูล	118
	เครื่องมือที่ใช้ในการวิจัย	119
	สรุปผลการวิจัย	120
	อภิปรายผล	126
	ข้อเสนอแนะ	133
บรรณานุกรม		135
ภาคผนวก		141
	ภาคผนวก ก	142
	ภาคผนวก ข	153
	ภาคผนวก ค	157
	ภาคผนวก ง	161
	ภาคผนวก จ	171
ประวัติย่อของผู้วิจัย		190

บัญชีตาราง

ตาราง หน้า

1 กลุ่มตัวอย่างที่ใช้ในการวิจัย 48

2 ค่าพิสัยความยาก (p) อำนาจจำแนก (r) และค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน เมื่อใช้สูตร KR – 20, KR – 21และสูตร r_B 53

3 ค่าสถิติพื้นฐานจากตัวอย่างขนาดใหญ่ 62

4 ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่ 62

5 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างจำลองขนาด 300 500 และ 800 คน..... 63

6 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างจำลองขนาด 300 500 และ 800 คน..... 65

7 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 300 500 และ 800 คน..... 67

8 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์..... 70

9 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ (Z) จากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีทฤษฎีการตอบข้อสอบ.....71

10 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์..... 75

11 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์และวิธีทฤษฎีการตอบข้อสอบ..... 77

12 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์..... 81

บัญชีตาราง (ต่อ)

ตาราง หน้า

13 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์ และวิธีทฤษฎีการตอบข้อสอบ 82

14 พิสัยค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งที่เปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน 88

15 พิสัยค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน 89

16 พิสัยค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ 90

17 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองแต่ละขนาดที่ใช้คะแนนตำแหน่งเปอร์เซ็นไทล์ (PR) และคะแนนปกติ (Z) เมื่อใช้วิธีประมาณค่าต่างกัน.... 91

18 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างแต่ละขนาดที่ใช้คะแนนวิธีทฤษฎีการตอบข้อสอบ 92

19 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากกลุ่มตัวอย่างขนาดใหญ่ 93

20 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข 96

21 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา 98

บัญชีตาราง (ต่อ)

ตาราง หน้า

22 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล 100

23 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน เมื่อใช้วิธีทฤษฎีการตอบข้อสอบ 102

24 สรุปค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่ 120

25 สรุปค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน 121

26 สรุปค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง (CSEM) จากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่ 122

27 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข 143

28 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา 145

29 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล 148

30 ค่า p_i,(1 - p_i), ΣB_{ji} และ λ_iรายข้อของแบบทดสอบวัดความถนัดด้านตัวเลข 150

31 ค่า p_i,(1 - p_i), ΣB_{ji} และ λ_iรายข้อของแบบทดสอบวัดความถนัดด้านภาษา 151

32 ค่า p_i,(1 - p_i), ΣB_{ji} และ λ_iรายข้อของแบบทดสอบวัดความถนัดด้านเหตุผล 152

33 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข 154

34 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา 155

35 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล 157

36 ค่าพารามิเตอร์ข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและเหตุผล 158

37 ค่าพารามิเตอร์ข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา 159

38 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 300 คน 162

บัญชีตาราง (ต่อ)

ตาราง หน้า

39 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 500 คน 163

40 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 800 คน 164

41 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองขนาด 300 คน 165

42 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองขนาด 500 คน 166

43 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองขนาด 800 คน 167

44 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างจำลองขนาด 300 คน 168

45 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างจำลองขนาด 500 คน 169

46 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างจำลองขนาด 800 คน 170

47 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านตัวเลข จากตัวอย่างจำลองขนาด 300 คน 172

48 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านตัวเลข จากตัวอย่างจำลองขนาด 500 คน 174

49 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านตัวเลช จากตัวอย่างจำลองขนาด 800 คน 176

50 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านภาษา จากตัวอย่างจำลองขนาด 300 คน 178

51 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านภาษา จากตัวอย่างจำลองขนาด 500 คน 180

บัญชีตาราง (ต่อ)

ตาราง หน้า

52 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านภาษา จากตัวอย่างจำลองขนาด 800 คน 182

53 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านเหตุผล จากตัวอย่างจำลองขนาด 300 คน..... 184

54 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านเหตุผล จากตัวอย่างจำลองขนาด 500 คน..... 186

55 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดด้านเหตุผล จากตัวอย่างจำลองขนาด 800 คน..... 188

บัญชีภาพประกอบ

ภาพประกอบ หน้า

1 คะแนนดิบที่สอบได้กับ Scale Score 10

2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ ตำแหน่งเปอร์เซ็นไทล์ (PR) ในแบบทดสอบวัดความถนัดทางการเรียน ด้านตัวเลขจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 73

3 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ คะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 73

4 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ ตำแหน่งเปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียน ด้านภาษาจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 79

5 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ คะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 79

6 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ อันดับที่เปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียน ด้านเหตุผลจากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 84

7 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ .... คะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 84

8 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัด ความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผลจากตัวอย่าง ขนาดใหญ่ เมื่อใช้วิธีทฤษฎีการตอบข้อสอบ 86

9 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ ตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษาและด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 94

บัญชีภาพประกอบ (ต่อ)

ภาพประกอบ หน้า

10 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ 95

11 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์... 103

12 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 104

13 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์... 105

14 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 106

15 ค่าความลำเอียงเฉลี่ยของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 107

16 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 108

17 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 109

บัญชีภาพประกอบ (ต่อ)

ภาพประกอบ หน้า

18 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 110

19 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ แบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 111

20 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแบบคะแนนปกติในแบบทดสอบ วัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 112

21 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 113

22 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติในแบบ ทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์..... 114

23 ความลำเอียงเฉลี่ยเมื่อใช้วิธีทฤษฎีการตอบข้อสอบของแบบทดสอบวัดความถนัด ทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน..... 115

24 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อใช้วิธีทฤษฎีการตอบข้อสอบของแบบ ทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน 116

25 ฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน..... 160

บทที่ 1
บทนำ

ภูมิหลัง

การวัดผลการศึกษาเป็นองค์ประกอบหนึ่งของการจัดการศึกษา ที่ต้องดำเนินการให้สอดคล้องกับจุดประสงค์ของหลักสูตรและกระบวนการเรียนการสอน โดยใช้วิธีการที่หลากหลายเพื่อวัดคุณลักษณะของผู้เรียน เช่น ความรู้ความสามารถ เจตคติ ความถนัด หรือคุณธรรม เป็นต้น คุณลักษณะเหล่านี้ไม่สามารถวัดได้โดยตรงเหมือนกับการวัดทางกายภาพที่สามารถวัดซ้ำได้หลายครั้งและยังคงให้ผลการวัดคงเดิม แต่การวัดคุณลักษณะภายในซึ่งเป็นการวัดทางจิตภาพนั้นเราไม่สามารถวัดซ้ำได้หลายครั้งส่วนใหญ่เป็นการวัดเพียงครั้งหรือสองครั้งเท่านั้นและในการวัดซ้ำนั้นผลที่ได้จากการวัดก็อาจคลาดเคลื่อนไปจากเดิมได้อาจเนื่องมาจากความเหนื่อยล้าของผู้สอบหรืออาจเป็นผลมาจากการฝึกฝน (Lord & Novick. 1968 : 13) ในทฤษฎีการวัดแบบมาตรฐานเดิม (Classical Test Theory) คะแนนที่ได้จากการวัด (Observed Score) ประกอบด้วยคะแนนจริง (True Score) และคะแนนความคลาดเคลื่อน (Error Score) (Allen & Yen. 1979 : 57)

คะแนนจริงเป็นคะแนนเฉลี่ยที่ได้จากการวัดซ้ำหลายครั้ง (Feld & Brennan. 1989 : 106) และในการสอบแต่ละครั้งคะแนนที่ได้จากการสอบจะแสดงถึงความรู้ความสามารถหรือพฤติกรรมของบุคคลนั้นว่ามีอยู่เท่าใด คะแนนนี้จะบ่งบอกถึงคะแนนจริงได้มากน้อยเพียงใดย่อมขึ้นอยู่กับความคลาดเคลื่อนในการวัด (Error of measurement) ถ้าความคลาดเคลื่อนในการวัดมีมาก แสดงว่าคะแนนที่ได้จากการวัด จะแตกต่างจากคะแนนจริงมาก

ในทางปฏิบัติจะไม่ทราบคะแนนจริงที่ได้จากการสอบด้วยแบบทดสอบชุดเดียวหลาย ๆ ครั้ง และในแต่ละครั้งที่สอบวัดก็มีความคลาดเคลื่อนเกิดขึ้นซึ่งเป็นผลมาจากองค์ประกอบหลายอย่าง เช่น สภาวะและเงื่อนไขของการทดสอบ การดำเนินการสอบ เครื่องมือที่ใช้ทดสอบ ผู้เข้าสอบ ซึ่งความคลาดเคลื่อนหรือความไม่คงที่ของคะแนนนี้ แม้จะรู้ว่ามีอยู่จริงแต่ก็ไม่สามารถบอกได้ว่าคะแนนที่สอบได้ของแต่ละคนมีความคลาดเคลื่อนเท่าใด แต่ก็สามารถอาศัยวิธีการทางสถิติเพื่อประมาณความคลาดเคลื่อนโดยเฉลี่ยของบุคคลนั้นว่าเป็นเท่าใด ซึ่งเรียกว่า ความคลาดเคลื่อนมาตรฐานในการวัด (Standard error of measurement) และค่านี้มีประโยชน์ที่จะทำให้เราสามารถประมาณคะแนนจริงของผู้สอบได้ อีกทั้งช่วยทำให้การแปลความหมายของคะแนนที่แต่ละคนสอบได้มีความแม่นยำและชัดเจนขึ้น

การแปลความหมายของความคลาดเคลื่อนมาตรฐานในการวัดที่ผ่านมา กำหนดให้ความคลาดเคลื่อนมาตรฐานในการวัดมีค่าคงที่ตลอดทุกช่วงคะแนน นั่นคือไม่ว่าผู้เข้าสอบคนใดได้คะแนนต่างกันอย่างไรจะกำหนดให้มีคะแนนความคลาดเคลื่อนในการวัดเท่ากัน แต่จากการศึกษาของโมเลนคอปฟ์ (Mollenkopf. 1949 :189) พบว่า ความคลาดเคลื่อนมาตรฐานในการวัดจะมีค่าคงที่เมื่อการแจกแจงของคะแนนที่ได้จากการวัดเป็นการแจกแจงแบบสมมาตร (Symmetrical or

Mesokurtic Distribution) นอกจากนั้นแล้วมาตรฐานในการทดสอบทางการศึกษาและจิตวิทยาข้อที่ 2.10 กล่าวว่าในการรายงานผลการสอบควรที่จะรายงานค่าความคลาดเคลื่อนมาตรฐานในการวัดแต่ละระดับคะแนน (American Educational Research Association, American Psychological Association & National Council on Measurement in Education. 1985 : 22) ซึ่งจะเป็นประโยชน์และมีคุณค่าต่อผู้ใช้ผลการทดสอบของทุกคนที่สอบเพื่อสื่อให้ทราบถึงดัชนีของความคลาดเคลื่อนในการวัดและการสร้างช่วงแห่งความเชื่อมั่นของคะแนน (Kolen, Hanson & Brennan. 1992 : 286)

ในการหาความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนที่ได้จากการทำแบบทดสอบถูกที่ระดับความสามารถเฉพาะ หรือที่แต่ละระดับคะแนนเราเรียกว่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข (Conditional Standard Error of Measurement : CSEM) (Feldt & Qualls. 1998 : 159 -160) ความคลาดเคลื่อนมาตรฐานในการวัดที่แต่ละระดับคะแนนขึ้นอยู่กับฟังก์ชันของคะแนนจริงไม่ใช่คะแนนที่สังเกตได้ (Lord. 1984 : 239) ซึ่งเป็นค่าที่มีความสำคัญและมีคุณค่ามากต่อการแปลความหมายของคะแนนที่ได้จากการวัดเพราะจะทำให้การแปลความหมายคะแนนของผู้สอบมีความแม่นยำมากขึ้น ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขนั้นมีการพัฒนาวิธีการไว้อย่างหลากหลายวิธี สำหรับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนดิบที่เป็นคะแนนที่ได้จากการตอบถูก แรกเริ่มทีเดียวในปี 1949 ก็มีวิธีของโมเลนคอปฟ (Mollenkopf) ได้เสนอไว้ ต่อมาเป็นวิธีของลอร์ด (Lord) ได้เสนอไว้ในปี 1951 1957 1980 และ 1984 นอกจากวิธีที่บุคคลเหล่านี้ได้นำเสนอแล้วก็ยังมีวิธีของเฟลด์ (Feldt) วิธีของจาร์จัวรา (Jajoura) วิธีของวูดรัฟฟ์ (Woodruff) วิธีของ เฟลด์และควอลส์ (Feldt & Qualls) และวิธีของเบรนนัน (Brennan) ที่ได้นำเสนอกันมาตามลำดับ (Lee, Brennan & Kolen. 2000 : 1 - 2) แต่เนื่องจากวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเหล่านี้ล้วนแต่เป็นวิธีการประมาณค่าในคะแนนดิบ ซึ่งโดยปกติแล้วมักไม่นำคะแนนดิบที่สอบได้ไปใช้ตัดสินผลใด ๆ เพราะคะแนนดิบไม่ให้ความหมายอะไรมากนัก และจะไม่นำไปเปรียบเทียบกันเพราะคนที่สอบได้คะแนนเท่ากันในแบบทดสอบต่างฉบับกันก็ไม่สามารถเปรียบเทียบกันได้ ดังนั้นเมื่อต้องการให้คะแนนของผู้สอบมีความหมายยิ่งขึ้นเราจึงมักแปลงคะแนนดิบเหล่านั้นให้เป็นคะแนนที่มีมาตรวัดเหมือนกัน (Scale Score) จะเห็นได้จากในการสอบคัดเลือกเข้ามหาวิทยาลัยจะแปลงคะแนนของผู้สอบไปเป็นคะแนนตำแหน่งเปอร์เซ็นไทล์เพื่อนำไปเป็นส่วนหนึ่งสำหรับการตัดสินใจคัดเลือกนักเรียนเข้าเรียนต่อในมหาวิทยาลัย สำหรับการแปลงคะแนนดิบเป็นคะแนนแปลงรูปที่มีมาตรวัดเหมือนกันมีวิธีการแปลงจำแนกออกเป็น 2 แบบ คือการแปลงคะแนนแบบเป็นเส้นตรง (Linear Transformation) และการแปลงคะแนนแบบไม่เป็นเชิงเส้นตรง (Nonlinear Transformation)

การแปลงคะแนนดิบเป็นคะแนนมาตรฐานแบบเส้นตรงนั้น ค่าความเชื่อมั่นของคะแนนที่แปลงจะไม่แตกต่างไปจากค่าความเชื่อมั่นของคะแนนดิบ และความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของคะแนนแปลงรูปจะเท่ากับผลคูณของค่าคงที่ค่าหนึ่งกับความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของคะแนนดิบนั่นเอง (Kolen, Hanson & Brennan. 1992 : 285) แต่สำหรับการแปลงคะแนนดิบแบบไม่เป็นเส้นตรงนั้นค่าความเชื่อมั่นของคะแนนที่แปลงรูปนั้นจะเปลี่ยนแปลงไปและมีผลต่อขนาดสัมพัทธ์ของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข จึงทำให้ไม่ค่อยมีการศึกษาและวิจัย

เกี่ยวกับค่าความเชื่อมั่นและค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบเป็นเส้นตรง แต่จะไปศึกษาจากคะแนนแปลงรูปแบบไม่เป็นเส้นตรงแทน ในการศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบไม่เป็นเส้นตรงนี้มีผู้ศึกษาไว้หลายคน เช่น โคเลน แฮนสัน และ เบรนนอน (Kolen, Hanson & Brennan. 1992 : 288-259) ได้ศึกษาวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปที่ต่างกัน จำนวนสเกลที่ใช้ และการเทียบคะแนน โดยใช้ข้อมูลจากการทดสอบในโครงการทดสอบ ACT Assessment โดยใช้โมเดลคะแนนจริงแบบเข้มงวด (Strong True Score) ต่อมา โคเลน เซง และแฮนสัน (Kolen, Zeng & Hanson. 1996 : 129 - 140) ก็ได้ศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยใช้วิธีทฤษฎีการตอบข้อสอบ (Item Response Theory) ซึ่งเป็นวิธีที่เน้นเกี่ยวกับความเป็นมิติเดียวในการวัด (Unidimensional) และนำผลที่ได้ไปเปรียบเทียบกับวิธีของโคเลน แฮนสัน และเบรนนอน ซึ่งพบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยวิธีที่ใช้โมเดลคะแนนจริงแบบเข้มงวดกับวิธีที่ใช้ทฤษฎีการตอบข้อสอบให้ผลคล้ายกัน นอกจากนี้ เบรนนัน และลี (Brennan & Lee. 1997, 1999) ยังได้ศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยใช้วิธีทวินามและวิธีทวินามประกอบ ซึ่งเป็นวิธีที่ใช้รูปแบบของโมเดลแบบทวินาม (Binomial Model) ซึ่งก็คือวิธีการประมาณค่าความคลาดเคลื่อนอย่างมีเงื่อนไขของคะแนนตามวิธีของลอร์ด (Lord. 1955, 1957) และโมเดลแบบแบบทวินามประกอบ (Compound Binomial Model) ซึ่งก็คือวิธีการประมาณค่าความคลาดเคลื่อนอย่างมีเงื่อนไขของคะแนนดิบตามวิธีของเฟลด์ (1984) นั่นเอง จากการศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่กล่าวมาเป็นวิธีการที่ต้องคำนวณยุ่งยากพอสมควร ดังนั้นเฟลด์ – ควอลส์ (Feldt – Qualls. 1998) จึงได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบคะแนนเทียบชั้นเรียน (Grade Equivalents) สองวิธีคือวิธีโพลิโนเมียล (Polynomial Method) และวิธีการกะประมาณ (Approximation Method) จากคะแนนเทียบชั้นที่ได้ โดยใช้ข้อมูลจากโครงการทดสอบทักษะพื้นฐานของไอโอวา (ITBS) ทั้งสองวิธีจะอาศัยความคลาดเคลื่อนมาตรฐานในการวัดจากคะแนนดิบด้วย ผลการวิจัยพบว่าทั้งสองวิธีให้ค่าใกล้เคียงกัน นั่นคือจะให้ค่าสูงตรงคะแนนเทียบชั้นตรงกลางและค่าที่ได้ไม่คงที่

จากผลการศึกษาที่กล่าวมาจะเห็นได้ว่าในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปมีอยู่หลายวิธี แต่ละวิธีต่างก็มีข้อตกลงเบื้องต้นแตกต่างกันและเป็นการศึกษาในต่างประเทศ สำหรับประเทศไทยการศึกษาเกี่ยวกับความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขแทบจะไม่ปรากฏให้เห็นในประเทศไทยเลย เท่าที่พบ ในปี พ.ศ. 2542 มีเพียงงานวิจัยของ อำนวยพร คำพันธ์ (อำนวยพร คำพันธ์. 2542 : 54-58) ได้ศึกษาผลของการจัดเรียงข้อสอบวัดความสามารถด้านตัวเลขแบบอนุกรมธรรมดาสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 ที่มีต่อค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะ โดยใช้วิธีการประมาณค่าสามวิธี ได้แก่ วิธีของธอร์นไดค์ วิธีของลอร์ด และวิธีของคีตส์ ผลการวิจัยพบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะของแบบทดสอบที่จัดเรียงจากง่ายไปยากมีค่าต่ำกว่าแบบทดสอบที่จัดเรียงแบบสุ่ม และค่าความคลาดเคลื่อนมาตรฐานในการวัดที่คำนวณโดยวิธีทั้งสามมีค่าแตกต่างกัน

ต่อมาปี พ.ศ. 2544 วิชุดา กิจธรธรรม (วิชุดา กิจธรธรรม. 2544 : 56 - 57) ได้ศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะของแบบทดสอบความถนัดทางวิชาชีพพยาบาลตามแนวคิดของฟลานาแกน ด้วยวิธีการวิเคราะห์ความแปรปรวน วิธีทฤษฎีการตอบข้อสอบ และวิธีการของคีตส์ ผลการวิจัยพบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเมื่อคำนวณด้วยวิธีการทั้งสามวิธี มีค่าไม่คงที่โดยเฉพาะอย่างยิ่งในช่วงปลายของคะแนน นั่นคือ มีค่าต่ำในช่วงปลายคะแนนต่ำและช่วงปลายคะแนนสูง และสูงสุดในช่วงคะแนนตรงกลาง งานวิจัยทั้งสองเรื่องนี้เป็นวิธีที่ใช้ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดสำหรับคะแนนดิบ แต่การศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสำหรับคะแนนแปลงรูปนั้นยังไม่ปรากฏว่ามีผู้ใดศึกษาเลย

ดังนั้น จึงทำให้ผู้วิจัยสนใจที่จะนำคะแนนที่ได้จากการสอบวัดมาทำให้เป็นคะแนนที่มีมาตรเดียวกัน (Scale Score) โดยใช้คะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (Percentile Rank Score) และวิธีคะแนนแบบปกติ (Normalized Score) เพราะวิธีการแปลงทั้งสองนี้มีรูปแบบการแปลงต่างกันโดยที่คะแนนตำแหน่งเปอร์เซ็นไทล์เป็นการแปลงโดยยึดตำแหน่งของบุคคลภายในกลุ่มเป็นหลักเพื่อดูว่าคะแนนนั้นอยู่ตำแหน่งใดของกลุ่มและอยู่เหนือคนอื่นกี่คนจาก 100 คน แต่คะแนนแปลงแบบคะแนนปกติเป็นการแปลงคะแนนดิบให้เข้าสู่คะแนนที่มีรูปแบบการกระจายปกติที่มีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนมาตรฐานเป็นหนึ่ง และเป็นคะแนนแปลงที่สามารถพัฒนาให้เป็นการแปลงคะแนนมาตรฐานแบบอื่นได้อีกด้วย โดยใช้วิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์ ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเพราะวิธีการประมาณค่าเหล่านี้มีข้อตกลงเบื้องต้นต่างกัน จึงทำให้ผู้วิจัยสนใจว่าเมื่อวิธีการประมาณค่าจากวิธีที่มีข้อตกลงเบื้องต้นต่างกันนี้จะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเป็นอย่างไร นอกจากนี้ยังต้องการศึกษาอีกว่าเมื่อใช้ตัวอย่างขนาด 300 500 และ 800 คน จะทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงมีค่าแตกต่างไปจากกรณีที่มีตัวอย่างขนาดใหญ่หรือไม่ ซึ่งในการศึกษาลักษณะเช่นนี้เท่าที่พบยังไม่มีผู้ใดศึกษาได้ผลสรุปเป็นที่แน่ชัด จึงเป็นเหตุให้ผู้วิจัยได้ทำการศึกษาในเรื่องนี้ ผลที่ได้จะทำให้ผู้ใช้แบบทดสอบตระหนักถึงความจำเป็นของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขและเลือกใช้วิธีที่เหมาะสม เพื่อนำผลไปใช้ในการแปลความหมายของความสามารถที่แท้จริงของผู้สอบแต่ละคนได้อย่างชัดเจน นั่นคือจะช่วยทำให้ทราบว่าคะแนนจริงของผู้สอบว่าอยู่ในช่วงใดด้วยความเชื่อมั่นที่น่าจะเป็นไปได้ และทำให้การตัดสินใจเกี่ยวกับผลการสอบของผู้สอบแต่ละคนเป็นไปด้วยความแม่นยำมากขึ้น

จุดมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง โดยกำหนดจุดมุ่งหมายเฉพาะ ดังนี้

1. เพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ จากคะแนนแบบทดสอบวัดความถนัดทางการเรียน

2. เพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้การประมาณค่าด้วยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์ จากคะแนนแบบทดสอบวัดความถนัดทางการเรียน

3. เพื่อศึกษาความแม่นยำในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด ได้แก่ ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย จากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ โดยใช้วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขวิธีทฤษฎีการตอบข้อสอบ วิธีแบบทวินาม วิธีแบบทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

ความสำคัญของการวิจัย

ผลที่ได้จากการศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขครั้งนี้ทำให้สามารถนำไปใช้ในการตัดสินใจเกี่ยวกับการสอบหรือการคัดเลือกบุคคลได้ถูกต้องมากขึ้นว่าการแปลงคะแนนแบบใดด้วยวิธีการใด ที่ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างเหมาะสม เพราะจะทำให้ผู้ใช้แบบทดสอบสามารถนำไปประมาณค่าคะแนนจริงของผู้สอบได้ถูกต้องแม่นยำด้วยความเชื่อมั่นในระดับที่ต้องการ และยังทำให้เกิดแนวทางการศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในสถานการณ์ของการสอบที่หลากหลาย อันทำให้เกิดประโยชน์ในการเสริมสร้างองค์ความรู้ทางด้านการวัดผลมากยิ่งขึ้น

ขอบเขตของการวิจัย

ประชากรที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย ของโรงเรียนสังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม ปีการศึกษา 2544 จำนวน 21,534 คน

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาตอนปลายที่สุ่มมาจากประชากร โดยการสุ่มแบบสองขั้นตอน (Two - Stage Random Sampling) จำนวน 1,500 คน และกลุ่มตัวอย่างที่ได้มาอย่างเจาะจงจากโรงเรียนขนาดใหญ่/ใหญ่พิเศษในจังหวัดร้อยเอ็ด ขอนแก่น นครพนม และยโสธร จำนวน 500 คน รวมเป็นกลุ่มตัวอย่างทั้งหมด 2,000 คน

เครื่องมือที่ใช้ในการวิจัยเป็นแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล ที่ได้จากโครงการวิจัยการสร้างและพัฒนาแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ปีงบประมาณ 2539 - 2541 คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

ตัวแปรอิสระ ได้แก่

1. ประเภทของคะแนนแปลงรูป ได้แก่
   1.1 ตำแหน่งเปอร์เซ็นไทล์
   1.2 คะแนนปกติ
2. วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ได้แก่
   2.1 วิธีทฤษฎีการตอบข้อสอบ
   2.2 วิธีแบบทวินาม

2.3 วิธีแบบทวินามประกอบ

2.4 วิธีเฟลด์ - ควอลส์

3. ขนาดตัวอย่าง ได้แก่ ขนาด 300 500 และ 800 คน

ตัวแปรตาม ได้แก่ ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข

นิยามศัพท์เฉพาะ

1. ความคลาดเคลื่อนมาตรฐานในการวัด หมายถึงค่าเบี่ยงเบนมาตรฐานของคะแนนความคลาดเคลื่อนซึ่งเป็นผลต่างระหว่างคะแนนที่สอบได้กับคะแนนจริง ในการวิจัยครั้งนี้เป็นการหาค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ได้จากการดำเนินการสอบเพียงครั้งเดียว

2. ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข หมายถึงความคลาดเคลื่อนมาตรฐานในการวัด ณ แต่ละระดับความสามารถซึ่งแปลงมาจากคะแนนดิบแบบไม่เป็นเส้นตรง

3. วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข หมายถึงการประมาณค่าความคลาดเคลื่อนมาตรฐานของคะแนนที่แปลงมาจากคะแนนดิบแบบไม่เป็นเส้นตรงที่ได้จากการทดสอบด้วยแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล ในการวิจัยครั้งนี้ใช้วิธีการประมาณค่า ดังนี้

3.1 วิธีทฤษฎีการตอบข้อสอบ หมายถึง วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยใช้รูปแบบทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับการตอบข้อคำถาม นั่นคือ โอกาสที่แต่ละคน i ที่มีความสามารถ θ จะตอบข้อคำถามถูกด้วยฟังก์ชัน $P_j(\theta)$

3.2 วิธีทวินาม หมายถึง วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ใช้ข้อตกลงเบื้องต้นของโมเดลแบบทวินามที่กล่าวว่าคะแนนที่สังเกตได้คือผลรวมของคะแนนสำหรับข้อคำถามที่ตอบถูก ซึ่งมีการให้คะแนน 1 เมื่อตอบถูก และให้ 0 เมื่อตอบผิด และการตอบข้อคำถามถูกเป็นอิสระจากกัน ข้อคำถามมีความยากเท่ากัน ความคลาดเคลื่อนมาตรฐานในการวัดแปรเปลี่ยนไปตามระดับของคะแนนจริงและไม่ขึ้นอยู่กับการแจกแจงคะแนนของกลุ่มตัวอย่างที่ทดสอบ

3.3 วิธีทวินามประกอบ หมายถึง วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ใช้ข้อตกลงเบื้องต้นของโมเดลแบบทวินามประกอบซึ่งเป็นโมเดลทั่วไปของโมเดลแบบทวินาม ซึ่งข้อคำถามแต่ละข้อมีความยากไม่เท่ากัน ความคลาดเคลื่อนมาตรฐานจะแปรเปลี่ยนไปตามระดับของคะแนนจริงและไม่ขึ้นอยู่กับการแจกแจงคะแนนของกลุ่มตัวอย่างที่ทดสอบ

3.4 วิธีเฟลด์ - ควอลส์ หมายถึง วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ต้องอาศัยค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ได้จากคะแนนดิบก่อนและหาอนุพันธ์ของการแปลงคะแนนโดยพิจารณาถึงจำนวนข้อคำถามในแบบทดสอบ

4. ความแม่นยำในการประมาณค่า หมายถึง ความสามารถของค่าสถิติที่ประมาณค่าได้ใกล้เคียงกับค่าที่ได้จากประชากร ในที่นี้ผู้วิจัยใช้ค่าสถิติสองค่าเพื่อแสดงถึงความแม่นยำในการประมาณค่าคือ

4.1 ความลำเอียงเฉลี่ย (Average Bias) หมายถึง ค่าสถิติที่แสดงถึงขนาดและทิศทางของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากตัวอย่างขนาดเล็กที่เบี่ยงเบนไปจากค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากตัวอย่างขนาดใหญ่

4.2 ความแปรปรวนคลาดเคลื่อนเฉลี่ย (Average Error Variance) หมายถึงค่าเฉลี่ยของกำลังสองของค่าเบี่ยงเบนที่ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากตัวอย่างจำลองเบี่ยงเบนไปจากค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากตัวอย่างขนาดใหญ่

5. การแปลงคะแนนดิบแบบไม่เป็นเส้นตรง หมายถึง การแปลงคะแนนดิบที่รูปแบบการแจกเดิมอาจเป็นรูปแบบใดรูปแบบหนึ่งไปสู่การแจกแจงของคะแนนที่ปรับให้เป็นโค้งปกติเพื่อทำให้คะแนนที่แปลงมีหน่วยเหมือนกันสามารถนำมาเปรียบเทียบกันได้

6. คะแนนแปลงรูป (Scale Score) หมายถึง คะแนนระบบหนึ่งที่ได้จากการแปลงรูปคะแนนดิบ ซึ่งมีจุดมุ่งหมายเพื่อทำให้คะแนนดิบนั้นมีหน่วยวัดเดียวกันและสามารถตีความหมายระดับความสามารถของผู้สอบจากคะแนนที่สอบได้ชัดเจนขึ้น ซึ่งในการวิจัยครั้งนี้ใช้การแปลงรูปแบบไม่เป็นเชิงเส้นตรง ดังนี้

6.1 ตำแหน่งเปอร์เซ็นไทล์ (Percentile Rank) หมายถึง ระบบการเทียบคะแนนที่ใช้หลักว่า ถ้ามีผู้เข้าสอบ 100 คน แต่ละคนได้คะแนนต่างกัน นำคะแนนนั้นมาจัดอันดับแต่ละตำแหน่งแล้วคะแนน ณ ตำแหน่งนั้นจะเป็นคะแนนที่อยู่เหนือคนอื่นกี่คนจาก 100 คน เช่น นาย ก สอบได้ 48 คะแนน จากคะแนนเต็ม 60 คะแนน คิดเป็นตำแหน่งเปอร์เซ็นไทล์ 70 นั่นคือ นาย ก สอบได้คะแนนสูงกว่าคะแนนของนักเรียนคนอื่นอยู่ 70 คนใน 100 คน

6.2 คะแนนปกติ (Normalized Score) หมายถึงการแปลงคะแนนดิบให้เข้าสู่คะแนนที่มีรูปแบบการกระจายปกติ ซึ่งมีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนมาตรฐานเป็นหนึ่ง

7. แบบทดสอบวัดความถนัดทางการเรียน (Scholastic Aptitude Tests) หมายถึง แบบทดสอบที่ใช้วัดความสามารถทางสมองของบุคคลแต่ละบุคคลที่เป็นผลมาจากการได้ศึกษาเล่าเรียนหรือจากประสบการณ์ทั้งปวง ที่สร้างขึ้นตามแนวคิดของเทอร์สโตน (Thurstone) 3 ด้าน ดังนี้

7.1 ความสามารถด้านตัวเลข (Numerical Ability) เป็นการวัดความเข้าใจทางด้านปริมาณจำนวน ในเชิงความสัมพันธ์ของปริมาณตัวเลข ความคิดรวบยอด หลักการทางคณิตศาสตร์ ทักษะการคิดคำนวณ ความสามารถด้านเหตุผลเชิงปริมาณ ในการวิจัยครั้งนี้ใช้แบบทดสอบวัดความถนัดด้านตัวเลขที่ประกอบด้วยเรื่องทักษะการคำนวณ คณิตศาสตร์เหตุผล และอนุกรมธรรมดา

7.2 ความสามารถด้านภาษา (Verbal Ability) เป็นการวัดความสามารถในการเข้าใจความหมายของคำ กลุ่มคำ ประโยค และสามารถอ่านจับใจความสำคัญ แปลความหมาย และตีความหมายของเรื่องราวที่เกี่ยวข้องกับภาษาได้เป็นอย่างดี ในการวิจัยครั้งนี้ใช้แบบทดสอบวัดความถนัดด้านภาษาที่ประกอบด้วยเรื่องการหาคำตรงข้าม และความเข้าใจภาษา

7.3 ความสามารถด้านเหตุผล (Reasoning Ability) เป็นการวัดความสามารถในการคิดอย่างมีวิจารณญาณ หาเหตุผล ค้นคว้าหาความสำคัญ ความสัมพันธ์ และหลักการทั้งหลายที่สร้างกฎหรือทฤษฎี หลักวินิจฉัย และสามารถลงสรุปได้อย่างถูกต้อง ในการวิจัยครั้งนี้ใช้แบบทดสอบวัดความถนัดด้านเหตุผลแบบภาพที่ประกอบด้วยเรื่องการจำแนกประเภท อุปมาอุปไมย และอนุกรมภาพ

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป ผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องในเรื่องต่อไปนี้

1. การแปลงคะแนนและประเภทของการแปลงคะแนน
2. ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขและวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบและคะแนนแปลงรูป
3. ตัวประมาณที่ดี
4. งานวิจัยที่เกี่ยวข้อง
   - 4.1 งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนมาตรฐานในการวัด
   - 4.2 งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ
   - 4.3 งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป

การแปลงคะแนนและประเภทของการแปลงคะแนน

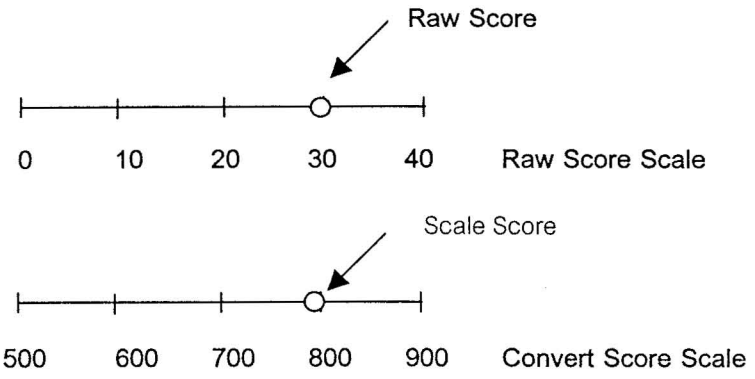
คะแนนดิบ (Raw Score Scale)

คะแนนที่ได้จากการสอบวัดเรามักได้จากการนับจำนวนข้อที่ตอบแบบทดสอบถูกเมื่อมีการให้คะแนนข้อถูกเป็นหนึ่งและข้อผิดเป็นศูนย์ โดยไม่คำนึงถึงว่าจะมีการแก้การเดาหรือไม่ก็ตาม คะแนนที่ได้จากการทดสอบเรียกว่าคะแนนดิบ และให้ความหมายของคะแนนที่ได้แต่ละคะแนนว่าแทนปริมาณความสามารถที่มีอยู่จากการสอบวัดที่เท่ากัน แต่แท้ที่จริงแล้วคะแนนดิบที่ได้จากการสอบนี้มีความหมายน้อยมากสำหรับการตีความหมายและการนำไปใช้ เพราะว่าแบบทดสอบแต่ละฉบับมีความยากแตกต่างกัน และความสามารถของคนภายในกลุ่มที่สอบเป็นอย่างไรเราก็ไม่ทราบได้ ดังนั้นในการที่จะทำให้คะแนนที่ได้จากการสอบมีความหมายและสามารถแปลความได้อย่างถูกต้องยิ่งขึ้นจำเป็นที่จะต้องมีการแปลงคะแนนจากแบบทดสอบให้อยู่ในรูปคะแนนชนิดเดียวกันหรือมีมาตรวัดเดียวกัน (Scale Score)

คะแนนแปลงรูปหรือคะแนนที่มาตรวัดเดียวกัน (Derived Score or Scale Score)

เป็นคะแนนระบบหนึ่งที่ได้จากการแปลงรูปคะแนนดิบที่สามารถสะท้อนระดับความสามารถหรือผลสัมฤทธิ์ของผู้สอบได้ เช่น ผู้สอบคนหนึ่งทำแบบทดสอบ 40 ข้อ ที่ให้คะแนน 1 ถ้าตอบถูก และให้ 0 ถ้าตอบผิด ทำได้ถูกทั้งหมด 30 ข้อ ดังนั้น คะแนนที่เป็นไปได้ทั้งหมดจะมีค่าพิสัยตั้งแต่ 0 ถึง 40 แต่คะแนนที่สอบได้นี้มักไม่ค่อยนิยมนำมาใช้ในการรายงานผลการสอบ ส่วนใหญ่รายงานผลการ

สอบในรูปของคะแนนที่มีหน่วยวัดเดียวกัน โดยทำคะแนนดิบให้เป็นคะแนนที่แปลงรูปขึ้นมาใหม่ (Popham. 1990 : 158) ดังภาพประกอบ 1 เป็นการแปลงคะแนนดิบเป็นคะแนนแปลงรูปแบบหนึ่ง



ภาพประกอบ 1 คะแนนดิบที่สอบได้กับคะแนนแปลงรูป (Popham. 1990 : 158)

สำหรับการแปลงคะแนนดิบให้เป็นคะแนนที่มีมาตรวัดเดียวกัน จำแนกได้ 2 ลักษณะคือ

1. การแปลงแบบเชิงเส้นตรง (Linear Transformation) เช่น คะแนนที (T-Score) คะแนนมาตรฐาน (Z-Score) เป็นต้น

2. การแปลงแบบไม่เป็นเชิงเส้น (Nonlinear Transformation) เช่น คะแนนมาตรฐานปกติ (Normalized standard scores) คะแนนเทียบชั้น (Grade equivalents) คะแนนเทียบอายุ (Age equivalents) และคะแนนตำแหน่งเปอร์เซ็นไทล์ (Percentile ranks)

การแปลงแบบเชิงเส้นตรง เป็นการแปลงรูปคะแนนโดยใช้สมการเส้นตรง คะแนนที่แปลงได้จะยังคงมีการแจกแจงคล้ายกับการแจกแจงของคะแนนเดิม

ถ้าให้ f เป็นการแปลงคะแนนดิบ X ไปยังคะแนนมาตรฐาน S แบบเชิงเส้น จะได้ว่า

$$S = f(X) = B + A(X)$$
$$= B + A(T_X + E_X)$$
$$= [B + A(T_X)] + A(E_X) \quad \text{.....(1)}$$
$$= T_S + E_S \quad \text{.....(2)}$$

จากสมการ (1) และ (2) ได้ $E_S = A(E_X)$ และ $\sigma_{E_S} = A\sigma_{E_X}$ ดังนั้นจะได้

$$\sigma_{E(S)} = A\sigma_{E(X)} \quad \text{.....(3)}$$

สมการ (3) นำไปประยุกต์ใช้กับทั้งค่าเฉลี่ย SEM และ SEMs ที่คะแนนของแต่ละบุคคล $X = x$ ดังนั้นสามารถเขียนสมการ (3) ได้เป็น

$$\sigma_{E(S|x)} = A\sigma_{E(X|x)} \quad \text{.....(4)}$$

ตัวอย่างคะแนนแปลงแบบเส้นตรง

คะแนนมาตรฐาน (Z - Scores) เป็นคะแนนที่บอกให้เราทราบว่าคะแนนดิบนั้นอยู่เหนือกว่าหรือต่ำกว่าคะแนนเฉลี่ยเป็นกี่หน่วยเบี่ยงเบนมาตรฐาน คะแนนประเภทนี้มีค่าเฉลี่ยเป็นศูนย์และค่าเบี่ยงเบนมาตรฐานเป็น 1 มีลักษณะการกระจายเป็นโค้งปกติ สูตรที่ใช้แปลงคือ

$$Z = \frac{X - \bar{X}}{S.D.}$$

เมื่อ X แทน คะแนนดิบ

$\bar{X}$ แทน คะแนนเฉลี่ย

และ S.D. แทน ค่าเบี่ยงเบนมาตรฐาน

คะแนนที (T - Score) เป็นคะแนนมาตรฐานหนึ่งที่ได้ปรับมาจาก Z - Score เพื่อมิให้ Z-Score ติดลบ T- Score มีค่าเฉลี่ยเป็น 50 และส่วนเบี่ยงเบนมาตรฐานเป็น 10 สูตรที่ใช้แปลงคือ

$$T = 50 + 10Z$$

หรือ

$$T = 50 + 10\left(\frac{X - \bar{X}}{S.D.}\right)$$

คะแนนสแตนไนน์ (Stannine - Score) เป็นคะแนนแปลงชนิดหนึ่งที่มีค่าเฉลี่ยเป็น 5 และส่วนเบี่ยงเบนมาตรฐานเป็น 2 เป็นการแบ่งคะแนนให้กระจายอยู่ในรูปโค้งปกติที่แบ่งออกเป็น 9 ช่วง นั่นคือจากสแตนไนน์ที่ 1 ถึงสแตนไนน์ที่ 9 แต่ละช่วงของ Stannine จะห่างกันประมาณ 0 - 5 ช่วงของคะแนนมาตรฐาน Z

การแปลงคะแนนแบบไม่เป็นเส้นตรงนี้ไม่มีผลกระทบต่อการแจกแจงของข้อมูลแต่อย่างใด แต่จะมีผลต่อค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงเท่านั้น (Stockburger. 2000)

การแปลงแบบไม่เป็นเชิงเส้นตรง

เป็นการแปลงคะแนนดิบให้เป็นคะแนนอีกระบบหนึ่ง โดยที่รูปร่างการแจกแจงของคะแนนที่แปลงแบบนี้จะแตกต่างจากรูปร่างการแจกแจงของคะแนนเดิม คะแนนที่แปลงแล้วจะมีการแจกแจงเป็นโค้งปกติ สมการที่ใช้แปลงไม่สามารถเขียนได้ในรูปสมการเชิงเส้นตรง แอลเลนและเยน (Allen & Yen. 1979 : 149) ได้จำแนกไว้เป็น 5 ประเภท คือ 1) Percentiles 2) Age and Grade Scores 3) Expectancy Table 4) Normalized scores และ 5) Equal-interval Scales

ตัวอย่างคะแนนที่แปลงแบบไม่เป็นเชิงเส้นตรง

คะแนนตำแหน่งเปอร์เซ็นไทล์ (Percentile Rank Scale)

เป็นคะแนนแปลงรูปชนิดหนึ่งที่นิยมใช้กันมาก นั่นคือเป็นคะแนนที่บอกให้ทราบว่ามีผู้เข้าสอบด้วยกันกี่เปอร์เซ็นต์ของผู้เข้าสอบทั้งหมดที่ได้คะแนนต่ำกว่าคะแนนนั้น การคำนวณหาคะแนนเปอร์เซ็นไทล์นั้นเป็นการนำเอาจำนวนความถี่ของผู้ที่ได้คะแนนต่ำกว่าคะแนนนั้นทั้งหมดมา

รวมกันและบวกด้วยครึ่งหนึ่งของจำนวนความถี่ของผู้ที่ได้คะแนนนั้น แล้วหารด้วยจำนวนความถี่ทั้งหมดหรือจำนวนคนผู้เข้าสอบทั้งหมดนั่นเอง

ขั้นตอนในการแปลงคะแนน มีดังนี้

1. เรียงลำดับคะแนนดิบที่ได้ (X) จากสูงไปหาต่ำ
2. หาความถี่ (f) ของคะแนนในแต่ละคะแนน
3. หาความถี่สะสม (cf) โดยการรวมความถี่ของคะแนนจากต่ำสุดขึ้นมาจนถึงความถี่ของชั้นคะแนนนั้น
4. หาความถี่สะสมที่แท้จริง (cf + ½f) โดยเอาความถี่สะสมของคะแนนในชั้นที่ต่ำลงไปรวมกับครึ่งหนึ่งของความถี่ในชั้นคะแนนนั้น
5. คำนวณค่าเปอร์เซ็นไทล์ซึ่งเท่ากับ $\frac{(cf + \frac{1}{2}f)x100}{N}$ เมื่อ N เป็นจำนวนผู้เข้าสอบทั้งหมด

คะแนนปกติ

เป็นการทำให้รูปแบบการกระจายความถี่ของคะแนนดิบเข้าสู่แบบการกระจายปกติ ซึ่งมีขั้นตอนการคำนวณต่อจากการคำนวณค่าเปอร์เซ็นไทล์ โดยอ่านคะแนนปกติจากการเปิดตารางโค้งปกติ ค่าที่อยู่บนแกนนอนเป็นค่า Z ซึ่งจะตรงกับแต่ละเปอร์เซ็นไทล์ เช่น เปอร์เซ็นไทล์ที่ 95 ตรงกับคะแนน Z เท่ากับ 1.64

คะแนนเทียบอายุและคะแนนเทียบชั้นเรียน

คะแนนเทียบอายุ (Age Equivalent Scale) เป็นการดูว่าคะแนนที่ผู้สอบทำได้นั้นเทียบเท่ากับอายุเท่าใด คะแนนเทียบอายุของคะแนนดิบใดก็คืออายุที่ตรงกับคะแนนเฉลี่ยของเด็กระดับอายุนั้น การคำนวณหาคะแนนเทียบอายุมีดังนี้

1. นำแบบทดสอบที่จะสร้างเกณฑ์เทียบอายุไปสอบกับกลุ่มตัวอย่างที่มีอายุหลายระดับ ข้อคำถามที่ถามควรครอบคลุมทั้งข้อที่ง่ายสำหรับเด็กอายุน้อยหรือเด็กที่อ่อน และข้อที่ยากสำหรับเด็กที่อายุมากหรือเด็กที่มีสติปัญญาดี
2. หาค่าเฉลี่ยหรือค่ามัธยฐาน ของคะแนนในแต่ละช่วงอายุ และนำไปเขียนกราฟ
3. ลากเส้นเชื่อมต่อจุดที่เป็นค่าเฉลี่ยหรือค่ามัธยฐานในแต่ละช่วงอายุและปรับโค้งให้เรียบ ปรับขยายเส้นกราฟออกไปให้ยาวทั้งหัวและท้ายเพื่อให้ครอบคลุมทุกช่วงอายุ
4. ค่าเฉลี่ยที่ได้จากการปรับนั้นจะแทนคะแนนเทียบอายุนั้น

การแปลความหมายผลที่ได้จากการเทียบอายุไม่ว่าจะได้สูงหรือต่ำกว่ากลุ่มคนในระดับอายุนั้นมากน้อยเท่าใด มิได้หมายความว่าผู้นั้นจะเป็นไปดั่งที่กลุ่มอายุนั้นเทียบได้

คะแนนเทียบชั้นเรียน (Grade Equivalent Scale) เป็นการเทียบคะแนนโดยใช้ค่าเฉลี่ยหรือค่ามัธยฐานของกลุ่มนักเรียนในระดับชั้นต่าง ๆ ที่เป็นกลุ่มมาตรฐาน วิธีการคำนวณหาคะแนน

เทียบชั้นมีขั้นตอนทำนองเดียวกับคะแนนเทียบอายุ เช่น ถ้านักเรียนคนหนึ่งได้คะแนนเท่ากับคะแนนเฉลี่ยของนักเรียนที่เริ่มเรียนชั้น ป. 6 ในกลุ่มปกติ ก็เรียกว่านักเรียนคนนั้นมีคะแนนเทียบชั้นเท่ากับ 6.0

ในการแปลงคะแนนแบบที่เป็นเชิงเส้นตรงและไม่เป็นเชิงเส้นตรงนี้ มีผลต่อการหาคุณภาพของแบบทดสอบ ซึ่งโคเลน แฮนสัน และเบรนนัน (Kolen, Hanson และ Brennan.1992) กล่าวว่าในการแปลงคะแนนดิบเป็นคะแนนแปลงรูปแบบเชิงเส้นตรง ค่าความเชื่อมั่นของคะแนนแปลงรูปจะเท่ากับค่าความเชื่อมั่นของคะแนนดิบ และความคลาดเคลื่อนอย่างมีเงื่อนไขของคะแนนแปลงรูปจะเท่ากับผลคูณของค่าคงที่กับความคลาดเคลื่อนอย่างมีเงื่อนไขของคะแนนดิบ และถ้าการแปลงคะแนนดิบไปเป็น คะแนนแปลงรูปแบบไม่เป็นเชิงเส้น จะทำให้ความเชื่อมั่นเปลี่ยนแปลงและมีผลต่อขนาดสัมพัทธ์ของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ด้วยเหตุผลนี้จึงทำให้ผู้วิจัยเลือกใช้การแปลงคะแนนแบบไม่เป็นเชิงเส้นตรงสำหรับในการวิจัยครั้งนี้

ความคลาดเคลื่อนมาตรฐานในการวัด (Standard Error of Measurement : SEM)

จากทฤษฎีการวัดแบบมาตรฐานเดิมกล่าวว่าคะแนนที่สังเกตได้ประกอบด้วยคะแนนจริงและคะแนนความคลาดเคลื่อน ถ้าในการสอบวัดครั้งใดไม่มีความคลาดเคลื่อนเกิดขึ้นคะแนนจริงจะเท่ากับคะแนนที่สังเกตได้ ดังนั้นคะแนนความคลาดเคลื่อนในการวัดจึงเป็นความแตกต่างหรือคะแนนที่คลาดเคลื่อนไปจากคะแนนจริงของผู้สอบในลักษณะแบบสุ่ม (เยาวดี วิบูลย์ศรี. 2540 : 91) คะแนนคลาดเคลื่อนที่เกิดจากการวัดแต่ละครั้งมีค่าแตกต่างกันไปอาจมีค่าบวกหรือค่าลบก็ได้

ความคลาดเคลื่อนมาตรฐานในการวัดเป็นค่าที่สะท้อนให้เห็นถึงความแปรปรวนของคะแนนแต่ละบุคคลที่ได้ทำแบบทดสอบซ้ำหลายครั้ง แต่ในการปฏิบัติไม่สามารถทำการทดสอบซ้ำบ่อยครั้งได้ (Popham. 1990 : 137) คะแนนความคลาดเคลื่อนนี้ไม่มีความสัมพันธ์กับคะแนนจริง (มีความสัมพันธ์กันเป็นศูนย์) นั่นคือผู้สอบที่ได้คะแนนสูงไม่ได้แสดงว่ามีคะแนนความคลาดเคลื่อนสูงหรือผู้ที่ได้คะแนนผลสอบต่ำก็ไม่ได้แสดงว่าคะแนนความคลาดเคลื่อนจะต่ำด้วย การคำนวณหาค่าความคลาดเคลื่อนมาตรฐานในการวัดสามารถหาได้จากสูตร

$$S_e = S_x\sqrt{1 - r_{xx}}$$

เมื่อ s_e เป็น ความคลาดเคลื่อนมาตรฐานในการวัด

r_{xx} เป็นค่าความเชื่อมั่นของแบบทดสอบ

s_x เป็นค่าเบี่ยงเบนมาตรฐานของแบบทดสอบ

ประโยชน์ของความคลาดเคลื่อนมาตรฐานในการวัด

ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าที่ได้นี้สามารถนำไปสร้างช่วงแห่งความเชื่อมั่นเพื่อประมาณค่าคะแนนจริงของผู้เข้าสอบได้ เช่น นักเรียนคนหนึ่งทำแบบทดสอบฉบับหนึ่งได้ 37 คะแนน ซึ่งค่าความเชื่อมั่นและค่าเบี่ยงเบนมาตรฐานของแบบทดสอบเป็น .91 และ 7.5 ตามลำดับ ดังนั้นประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดได้เป็น

$$S_e = 7.5\sqrt{1-.91}$$

$$S_e = 7.5\sqrt{.09} \quad = 7.5(.3)$$

$$= 2.25$$

ฉะนั้น ถ้าจะให้มีความเชื่อมั่นได้ 95 % นักเรียนคนนี้จะมีคะแนนจริงอยู่ในช่วง $37 \pm 2(2.25)$

จะเห็นได้ว่าในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดตามทฤษฎีการวัดแบบมาตรฐานเดิม จะกำหนดให้ความคลาดเคลื่อนมาตรฐานในการวัดที่หาได้มีค่าคงที่ตลอดทุกช่วงคะแนน นั่นคือไม่ว่าผู้สอบคนใดได้คะแนนเท่าใดก็มีความคลาดเคลื่อนมาตรฐานในการวัดเท่ากัน แต่แท้จริงแล้วความคลาดเคลื่อนมาตรฐานในการวัดจะไม่เป็นค่าคงที่ไม่เท่ากันตลอดทุกช่วงคะแนน ดังนั้นในการหาความคลาดเคลื่อนมาตรฐานในการวัดจึงควรประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนที่ได้จากการทำแบบทดสอบถูกที่ระดับความสามารถเฉพาะ หรือที่แต่ละระดับคะแนน ซึ่งเรียกว่าเป็นความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข (Conditional Standard Error of Measurement : CSEM)

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ (CSEM)

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ ได้มีการศึกษากันมาแล้วกว่า 50 ปี ผู้ที่นำเสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในระยะแรก ก็คือ โมเลนคอปฟ (Mollenkopf. 1949) และ ธอร์นไดค์ (Thorndike. 1951) ต่อมา ลอร์ด (Lord. 1955, 1957) ก็ได้พัฒนาวิธีการประมาณค่าขึ้นมาโดยใช้โมเดลความคลาดเคลื่อนในการวัดแบบทวินาม (Binomial Error Model) ซึ่งเป็นวิธีที่รู้จักกันดี หลังจากนั้น เฟลด์ (Feldt. 1984) ก็ได้เสนอวิธีการประมาณค่าโดยใช้โมเดลความคลาดเคลื่อนในการวัดแบบทวินามประกอบ (Compound Binomial Error Model) โดยกำหนดไว้ล่วงหน้าว่าแบบทดสอบเป็นแบบคู่ขนานที่สุ่มมาจากข้อสอบที่แบ่งเป็นชั้นหรือเป็นกลุ่มตามเนื้อหาหรือตามความยาก จากนั้นในปี 1980 ลอร์ด ก็ได้ใช้วิธีการประมาณค่าโดยใช้ทฤษฎีการตอบข้อสอบ (IRT) และในปี 1998 เบรนนัน (Brennan) ก็นำเสนอวิธีการประมาณค่าโดยใช้ทฤษฎีสัมประสิทธิ์การสรุปอ้างอิง (G - Theory) ในแต่ละวิธีเหล่านี้ต่างก็มีข้อตกลงเบื้องต้นในการใช้ที่แตกต่างกันไป และในที่สุด เฟลด์ และเบรนนัน (Feldt & Brennan. 1989) และ เฟลด์และควอลส์ (Feldt & Qualls. 1996) ก็ได้นำวิธีการทั้งหลายเหล่านั้นมาประยุกต์ใช้ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข

สำหรับแบบทดสอบที่นำมาใช้เพื่อการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขนั้นส่วนใหญ่เป็นการดำเนินการสอบครั้งเดียวและลักษณะแบบทดสอบเป็นแบบคำถามโดด (Item-based) นั่นคือการตอบแต่ละข้อจะมีคำถามมาให้เพียงคำถามเดียว การตอบจะไม่เกี่ยวกับเงื่อนไขหรือสถานการณ์ใดที่กำหนดให้ ต่อมา ลี (Lee. 2000) ได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบที่ประกอบด้วยข้อคำถามเป็นชุด แต่ละชุดเป็น

สถานการณ์ บทความ โคลง กลอน กราฟ และนำไปใช้ตอบคำถามหลายข้อ โดยใช้วิธีการประมาณค่าตามวิธี G - Theory และวิธี IRT แบบสามพารามิเตอร์

ดังนั้นวิธีการประมาณค่าแบบ IRT และ G - Theory ใช้ได้กับแบบทดสอบที่ประกอบด้วยข้อคำถามโดด และ ข้อคำถามที่เป็นชุด ส่วนวิธีที่ใช้โมเดลความคลาดเคลื่อนทวินาม (Binomial Error Model) ของ ลอร์ด (Lord. 1955, 1957) และวิธีที่ใช้โมเดลความคลาดเคลื่อนทวินามประกอบ (Compound Binomial Error) ของ เฟลด์ เป็นกรณีเฉพาะของวิธี G - Theory นั่นคือทั้งวิธี IRT และ G - Theory เป็นวิธีที่ได้รวบรวมเอาความคิดพื้นฐานเกือบทั้งหมดของหลายวิธี เข้าไว้ด้วยกัน ยกเว้นวิธีของ ธอร์นไดค์ (Thorndike. 1951) และ วิธีของ โมเลนคอป์ฟ (Mollenkopf. 1949)

สำหรับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ แต่ละคะแนนดิบ สามารถประมาณค่าได้หลายวิธี พอสรุปได้ ดังนี้

วิธีที่ใช้ทฤษฎีการทดสอบแบบมาตรฐานเดิม ได้แก่ วิธีของโมเลนคอป์ฟ (1949) และวิธีของธอร์นไดค์ (1951)

วิธีของโมเลนคอป์ฟ

โมเลนคอป์ฟ (1949) ได้นิยามความคลาดเคลื่อนมาตรฐานในการวัด ดังนี้

$$\sigma_{meas_x} = \sigma_x \sqrt{1 - r_{xx}} \qquad \text{.....(5)}$$

เมื่อ r_{xx} เป็นความเชื่อมั่น และ σ_x เป็นค่าเบี่ยงเบนมาตรฐานของคะแนนแบบทดสอบที่ได้จากกลุ่มผู้สอบทั้งหมดซึ่งแปรเปลี่ยนไปตามคะแนนจริงถ้ากลุ่มผู้สอบเหล่านี้ได้ทำข้อสอบที่คู่ขนานจำนวนมาก ดังนั้นเมื่อผู้เข้าสอบทั้งหมด N คน ทำข้อสอบที่คู่ขนาน K แล้วค่าความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละคน สามารถนิยามได้ดังนี้

$$\sigma_{meas_x} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{K}(x_{ij} - t_i)^2}{KN}} \qquad \text{.....(6)}$$

เมื่อ x_{ij} แทน คะแนนที่สอบได้ของแต่ละคนที่ตอบแบบทดสอบ j

t_i แทน คะแนนจริงของแต่ละคน

N แทน จำนวนคนที่สอบ

K แทน จำนวนแบบทดสอบที่คู่ขนานกัน

ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ได้จากสมการ (6) จะมีค่าใกล้เคียงกับค่าที่ได้จากสมการ (5) ก็ต่อเมื่อคะแนนจริงของแต่ละคนเป็นค่าเฉลี่ยของคะแนนสอบที่ได้จากการตอบข้อสอบคู่ขนานกันจำนวนมาก นั่นคือ

$$t_i = \frac{\sum_{j=1}^{K} x_{ij}}{K}$$

วิธีของธอร์นไดค์ (1951)

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของธอร์นไดค์นี้ ได้กำหนดให้ X เป็นคะแนนรวมที่ได้จากคะแนนแต่ละครึ่งฉบับ (X_1 และ X_2) ที่คู่ขนานกัน นั่นคือ $X = X_1 + X_2$ จากโมเดลทฤษฎีการทดสอบแบบมาตรฐานเดิมนั้น คะแนนของแบบทดสอบแต่ละครึ่งฉบับและคะแนนรวมทั้งหมดสามารถเขียนอยู่ในรูปส่วนประกอบของคะแนนจริงและคะแนนความคลาดเคลื่อนได้ดังนี้

$$X_1 = T_1 + E_1$$
$$X_2 = T_2 + E_2$$
$$X = X_1 + X_2$$
$$= (T_1 + T_2) + (E_1 + E_2)$$
$$= T + E$$

จากทฤษฎีการทดสอบแบบมาตรฐานเดิม ค่าสหสัมพันธ์ระหว่าง E_1 และ E_2 เป็นศูนย์ จึงทำให้

$$\sigma_E = (\sigma_{E_1}^2 + \sigma_{E_2}^2)^{1/2}$$

เพราะว่า $X_1 - X_2 = (T_1 - T_2) + (E_1 - E_2)$

สำหรับแบบทดสอบคู่ขนานถ้า $T_1 = T_2$ จะได้ว่าค่าเบี่ยงเบนมาตรฐานของความแตกต่างระหว่างคะแนนแต่ละครึ่งฉบับเท่ากับความคลาดเคลื่อนมาตรฐานของคะแนนแบบทดสอบทั้งฉบับ

$$\sigma_{X_1 - X_2} = \sigma_{E_1} + \sigma_{E_2}$$
$$= (\sigma_{E_1}^2 + \sigma_{E_2}^2)^{1/2}$$
$$= \sigma_E$$

วิธีนี้มีข้อจำกัดตรงจำนวนคนในกลุ่มตัวอย่างที่ได้คะแนนในแต่ละระดับคะแนน ถ้ามีผู้ได้คะแนน ณ ระดับคะแนนที่สนใจเป็นจำนวนน้อย ควรรวมจำนวนคนที่ได้คะแนนทั้งฉบับเป็นช่วงคะแนนดีกว่าที่จะใช้แต่ละคะแนน

วิธีที่ใช้โมเดลความคลาดเคลื่อนแบบทวินาม (Binomial Error Model : BEM) เช่น วิธีของลอร์ด (1955) คีตส์ (1957) และวิธีของเฟลด์ (1984)

วิธีของลอร์ด

ลอร์ด (1955) ได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐาน ณ แต่ละระดับคะแนนภายใต้โมเดลแบบทวินามซึ่งกล่าวว่าข้อคำถามจำนวน k ข้อ เป็นข้อที่สุ่มมาจากประชากรข้อสอบทั้งหมดที่คู่ขนานกัน และ X_i เป็นคะแนนที่ผู้สอบคนที่ i ตอบข้อคำถามได้ถูกเมื่อมีการให้คะแนนแบบ 0 - 1 สูตรที่ใช้ในการประมาณค่าความแปรปรวนคลาดเคลื่อนของผู้เข้าสอบแต่ละคนสามารถประมาณค่าได้ดังนี้

$$\tilde{\sigma}_{e_i}^2 = \frac{X_i(k - X_i)}{k - 1} \qquad \text{.....(7)}$$

ความแปรปรวนคลาดเคลื่อนเฉลี่ยของผู้เข้าสอบ N คน คือ

$$\text{ค่าเฉลี่ย } \tilde{\sigma}_e^2 = \frac{[\bar{X}\ (k-\bar{X}) - S_X^2]}{k-1} \qquad \text{.....(8)}$$

เมื่อ $\bar{X}$ และ S_X^2 เป็นค่าเฉลี่ยและความแปรปรวนของคะแนนที่สอบได้ เมื่อนำค่าเฉลี่ยความแปรปรวนคลาดเคลื่อนนี้ไปแทนค่าในสูตรความเชื่อมั่น จะได้ว่า

$$\text{reliabilty} = 1 - \frac{\sigma_e^2}{\sigma_X^2} \qquad \text{.....(9)}$$

ค่าที่ได้จากสูตรนี้จะเท่ากับค่าที่ได้จากสูตรของคูเดอร์ - ริชาร์ดสัน 21

ให้ $p_1, p_2,\ldots,p_n$ เป็นสัดส่วนของการตอบถูกใน k ข้อ ที่สุ่มมาจากประชากรข้อสอบทั้งหมด ดังนั้น

$$\sum_h^k p_h = k\bar{p} = \bar{X}_A$$

ความแปรปรวนที่ถูกประมาณค่าเฉลี่ยของแบบทดสอบที่เป็นฟอร์มคู่ขนานที่ได้จากการทำข้อสอบซ้ำ k ข้อ จากประชากรทั้งหมดคือ

$$\bar{\sigma}_X^2 = \bar{\sigma}_{kp}^2 = k^2\bar{\sigma}_p^2 = k^2\left(\frac{S_p^2}{k-1}\right) \qquad \text{.....(10)}$$

เมื่อ S_p^2 เป็นความแปรปรวนของ $p_1, p_2,\ldots,p_k$ ของข้อคำถามในฟอร์ม A นำสมการ (10) ลบออกจากสมการ (8) จะได้

$$\begin{aligned}
\bar{\sigma}_e^2 &= \frac{\bar{X}(k-\bar{X}) - S_X^2}{k-1} - k^2\left(\frac{S_p^2}{k-1}\right) \\
&= \frac{\bar{X}(k-\bar{X}) - S_X^2}{k-1} - \left(\frac{k^2}{k-1}\right)\left(\frac{\sum p_h^2}{k} - \bar{p}^2\right) \\
&= \frac{\bar{X}(k-\bar{X}) - S_X^2}{k-1} - \left(\frac{k^2}{k-1}\right)\left(\frac{\sum p_h^2}{k} - \frac{\bar{X}^2}{k^2}\right) \\
&= \frac{k\bar{X} - \bar{X}^2 - S_X^2 - k\sum p_h^2 + \bar{X}^2}{k-1} \\
&= \frac{k\sum p_h - k\sum p_h^2 - S_X^2}{k-1} \\
&= \frac{k\sum p_h(1-p_h) - S_X^2}{k-1} \qquad \text{.....(11)}
\end{aligned}$$

ให้ $q_h = 1 - p_h$ แทนค่าความแปรปรวนคลาดเคลื่อนในสมการ (11) ลงในสมการ (9) จะได้สูตรการหาความเชื่อมั่นดังนี้

$$r_{XX} = 1 - \frac{k\sum p_h q_h - S_X^2}{(k-1)S_X^2}$$

$$= \frac{k}{k-1}\left(\frac{S_X^2 - \sum p_h q_h}{S_X^2}\right) \qquad \text{.....(12)}$$

สมการ (12) ก็คือสูตรการหาความเชื่อมั่นของคูเดอร์ - ริชาร์ดสัน นั่นเอง นั่นแสดงว่าทฤษฎีความคลาดเคลื่อนแบบทวินามได้ถูกปรับมาเพื่อให้ได้มีการยอมรับสูตร KR20 มากกว่า KR21

วิธีของคีตส์

จากการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขวิธีของลอร์ดในสมการ (7) นั้น นับเป็นวิธีการที่มีจุดอ่อนมาก เพราะไม่ได้คำนึงถึงเรื่องของการจับคู่ฟอร์มของแบบทดสอบ ความยากของแบบทดสอบหรือคุณลักษณะอื่นใดของแบบทดสอบเลย จึงทำให้การประมาณค่าที่ได้เกินความเป็นจริง ถ้าพิจารณาให้ดีแล้วค่ากำลังสองของสมการนี้ก็คือสัมประสิทธิ์ความเชื่อมั่นที่เป็นสูตรของคูเดอร์ - ริชาร์ดสัน 21 นั่นเอง ดังนั้นคีตส์ (1957) จึงเสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขขึ้นมาโดยนำค่าคงที่ไปคูณทางด้านขวาของสมการที่ลอร์ดเสนอเรียกว่าเป็นวิธีปรับแก้ของคีตส์ บางครั้งเขียนเป็นสัญลักษณ์ว่า $_{KL}S_E^2$ สูตรที่ใช้คือ

$$S_E = \left[\left(\frac{X(k-X)}{k-1}\right)\left(\frac{1-r_{xx'}}{1-r_{21}}\right)\right]^{1/2} \qquad \text{.....(13)}$$

เมื่อ r_{21} เป็นค่าความเชื่อมั่นของแบบทดสอบที่หาโดยสูตรของคูเดอร์ - ริชาร์ดสัน 21 และ $r_{xx'}$ เป็นค่าประมาณความเชื่อมั่นของแบบทดสอบซึ่งคีตส์เสนอว่าควรเป็นสัมประสิทธิ์ความเชื่อมั่นที่ได้จากฟอร์มคู่ขนาน แต่ในทางปฏิบัติอาจใช้สัมประสิทธิ์ความเชื่อมั่นแบบแบ่งครึ่งหรือสัมประสิทธิ์แอลฟ่าของครอนบัคก็ได้ แต่เฟลต์ - ควอลส์ (1998) ได้ใช้สัมประสิทธิ์ความเชื่อมั่นสูตรของคูเดอร์ - ริชาร์ดสัน 20

สำหรับการวิจัยครั้งนี้ผู้วิจัยได้ใช้ค่าสัมประสิทธิ์ความเชื่อมั่น r_B ตามรูปแบบคะแนนจริงสัมพันธ์ที่ บุญเชิด ภิญโญอนันตพงษ์ (2542 : 59 - 72) ได้พัฒนาขึ้น เพื่อเป็นการผ่อนปรนข้อตกลงเกี่ยวกับคะแนนของแบบทดสอบที่ใช้สูตรสัมประสิทธิ์แอลฟ่ามาเป็นข้อตกลงของคะแนนจริงสัมพันธ์ (Congeneric) ซึ่งมีแบบจำลองคะแนน ดังนี้

$$T_i = \lambda_{ij} T_j + C_{ij}$$

จากข้อตกลงของคะแนนจริงสัมพันธ์นี้ คะแนนจริงแต่ละส่วนของคู่ใด ๆ จะแตกต่างกันไปตามค่าคงที่ C_{ij} และแตกต่างกันไปตามลักษณะเฉพาะของคะแนนแต่ละส่วนคู่นั้นคือ λ_{ij}

สูตรที่ใช้ในการหาค่าสัมประสิทธิ์ความเชื่อมั่น r_B คือ

$$r_B = \frac{1}{1-\sum\lambda_i^2}\left[1-\frac{\sum P_i(1-P_i)}{S_x^2}\right] \qquad \text{.....(14)}$$

ขั้นตอนการหาค่าสัมประสิทธิ์ r_B

การหาค่าสัมประสิทธิ์ r_B จากแบบทดสอบที่มี k ข้อ ซึ่งให้คะแนนแบบ 0 - 1 มีขั้นตอนดังนี้

1. สร้างเมตริกซ์คะแนนที่ได้ของแต่ละข้อและคะแนนรวมของแต่ละคน โดยแต่ละคอลัมน์เป็นผลของคะแนนที่ได้แต่ละข้อ และในแนวแถวเป็นผลการตอบของแต่ละคน

2. หาค่า P_i ของแต่ละข้อ ซึ่งเป็นสัดส่วนของจำนวนผู้ตอบถูกแต่ละข้อนั่นเอง และหา $\sum P_i(1-P_i)$ (i = 1, 2 , ..., k)

3. หาค่า $\bar{X} = \sum P_i$ และ $S_x^2 = \frac{N\sum X^2 - (\sum X)^2}{N(N-1)}$

4. สร้างเมตริกซ์ของสัดส่วนจำนวนผู้ตอบข้อสอบถูกทั้งสองข้อแต่ละคู่ทุกคู่ (B_{ji} ; $j \neq i$) เช่น ถ้ามีข้อสอบ 7 ข้อ จะได้คู่ของ B_{ji} ทั้งหมด 21 คู่

5. คำนวณค่า λ_i จากสูตร

$$\lambda_i = [\sum B_{ji} - P_i(\bar{X}-1)]/S_x^2$$

6. แทนค่า P_i , λ_i และ S_x^2 ลงในสมการ (14) เพื่อประมาณค่าสัมประสิทธิ์ความเชื่อมั่น r_B

วิธีแบบทวินามประกอบ

เฟลด์ (1984) ได้ใช้โมเดลความคลาดเคลื่อนแบบทวินามของลอร์ด (1965) ที่ใช้การสุ่มข้อคำถามแบบแบ่งชั้นที่เป็นการสร้างแบบทดสอบฟอร์มคู่ขนาน โดยใช้วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ แต่ละระดับคะแนนจากการทำแบบทดสอบที่แบ่งเป็นชั้นเป็นกลุ่ม โดยใช้สูตรดังนี้

$$S_{E(i)} = \left[\sum_{h=1}^{c}\frac{X_{ih}(k_h - X_{ih})}{k_h - 1}\right]^{1/2} \qquad \text{.....(15)}$$

เมื่อ $S_{E(i)}$ เป็นความคลาดเคลื่อนมาตรฐานของแต่ละบุคคล i

X_{ih} เป็นคะแนนของบุคคล i ในแต่ละชั้น

c เป็นจำนวนข้อของแต่ละชั้น

k_h เป็นจำนวนข้อในชั้น h ซึ่งต้องมากกว่า 1

วิธีที่ใช้โมเดลทฤษฎีการตอบข้อสอบ

ทฤษฎีการตอบข้อสอบเป็นทฤษฎีที่แสดงความสัมพันธ์ระหว่างการตอบข้อสอบแต่ละข้อกับคุณลักษณะที่ข้อสอบนั้นจะวัด ความสัมพันธ์นี้แทนได้ด้วยฟังก์ชันการตอบข้อสอบ $P_j(\theta)$ หรือเส้นแนวทาง (trace line) และสามารถใช้ประมาณค่าโอกาสที่แต่ละบุคคลหนึ่งที่มีความสามารถ θ จะตอบ

ข้อสอบข้อใดได้ถูกต้อง ในการประมาณค่าความแปรปรวนคลาดเคลื่อนในการวัดของแต่ละบุคคลมีสองขั้นตอน คือ

ขั้นแรก ประมาณค่าความสามารถ θ ของแต่ละบุคคลซึ่งอาจใช้โมเดลที่มีหนึ่งพารามิเตอร์ สองพารามิเตอร์ หรือสามพารามิเตอร์ ก็ได้

ขั้นที่สอง หาค่าฟังก์ชันต่อไปนี้ ซึ่งเป็นการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละบุคคล

$$S^2_{E(i)} = \left\{ \sum_j^k [P_j(\theta_i)][1 - P_j(\theta_i)] \right\}^{1/2} \quad \text{.....(16)}$$

เมื่อ $P_j(\theta_i)$ เป็นค่าฟังก์ชันของข้อคำถาม j ณ ระดับความสามารถ θ ในแต่ละบุคคล i

วิธีนี้คล้ายกับวิธีแบบทวินามประกอบที่ให้การประมาณค่าในผู้สอบแต่ละคน ในการใช้วิธีนี้ ผู้เข้าสอบจะถูกจัดกลุ่มตามระดับคะแนนที่ได้ และคำนวณ $S^2_{E(i)}$

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป (CSEMs)

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงรูป นับว่าเป็นวิธีที่ค่อนข้างใหม่ทีเดียว ซึ่งมีผู้เสนอได้ศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดไว้หลายท่าน นับแต่ปี ค.ศ. 1992 โคเลน แฮนสัน และเบรนนัน (Kolen, Hanson & Brennan) ได้นำเสนอวิธีการประมาณค่า CSEMs โดยการประยุกต์ใช้โมเดลของคะแนนจริงอย่างเข้มงวด ซึ่งเปรียบได้กับวิธี Four-Parameter Beta ของลอร์ด และวิธีที่ใช้โมเดลความคลาดเคลื่อนแบบทวินามประกอบ ต่อมาในปี ค.ศ. 1996 โคเลน เซง และแฮนสัน (Kolen, Zeng & Hanson) และ วัง โคเลน และแฮริส (Wang, Kolen & Harris) ได้ใช้วิธีการประมาณค่าโดยทฤษฎีการตอบข้อสอบ นอกจากนี้แล้วในปีเดียวกัน คอลตัน เกา และโคเลน (Colton, Gao & Kolen) ก็ได้ใช้วิธี Bootstrap ในการประมาณค่า และในปี ค.ศ. 1998 เฟลต์และควอลส์ (Feldt & Qualls) ได้เสนอวิธีการประมาณค่าแบบพหุนาม (Polynomial) และวิธีการกะประมาณ (Approximation) นอกจากนี้แล้ว ในปี ค.ศ.1999 เบรนนัน และลี (Brennan & Lee) ก็ได้ใช้วิธีประมาณค่าแบบทวินาม และแบบทวินามประกอบ

ในการประมาณค่าแต่ละวิธีต่างก็มีลักษณะเฉพาะและข้อตกลงในการประมาณค่าต่างกัน ดังต่อไปนี้

วิธีที่ใช้โมเดลคะแนนจริงอย่างเข้มงวด (Strong True Score Model)

โคเลน แฮนสัน และเบรนนัน (Kolen, Hanson & Brennan. 1992) ได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนอย่างมีเงื่อนไขของคะแนนแปลงรูป และประมาณค่าความเชื่อมั่นโดยใช้โมเดลคะแนนจริงอย่างเข้มงวด ดังนี้

ในการแปลงคะแนนดิบเป็นคะแนนแปลงรูปนั้นโดยทั่วไปคะแนนดิบจะเป็นคะแนนที่เป็นเลขจำนวนเต็มและแปลงเป็นคะแนนที่มีหน่วยวัดเดียวกันที่มีค่าเป็นเลขจำนวนเต็มด้วย เมื่อแปลงคะแนนดิบแล้วจะประมาณค่าความคลาดเคลื่อนมาตรฐาน โดยใช้รูปแบบที่เกี่ยวกับการแจกแจงของคะแนนจริงและการแจกแจงของความคลาดเคลื่อนในการวัดเมื่อกำหนดคะแนนจริงให้ โดยใช้โมเดลของคะแนนจริงอย่างเข้มงวดหรือโมเดลแบบ IRT

คุณสมบัติของคะแนนแปลงรูป

ในการศึกษาคุณสมบัติของคะแนนแปลงรูป ให้คะแนนดิบของแบบทดสอบฉบับหนึ่งเป็นผลรวมของข้อคำถาม K ข้อที่ให้คะแนนแบบ 0 – 1 ดังนั้น ความน่าจะเป็นที่ตัวแปรเชิงสุ่มของคะแนน X จะเท่ากับ i (i = 0,...,K) ในประชากรของผู้เข้าสอบ คือ

$$\Pr(X = i) = \int_0^1 \Pr(X = i \mid \tau) g(\tau) d\tau \qquad \text{.....(17)}$$

เมื่อ τ เป็นสัดส่วนของคะแนนจริงที่ตอบถูกด้วยฟังก์ชันความหนาแน่น $g(\tau)$ ในประชากร เทอมของ $\Pr(X = i \mid \tau)$ เป็นการแจกแจงอย่างมีเงื่อนไขของคะแนนดิบสำหรับผู้เข้าสอบเมื่อกำหนด τ ให้ เพื่อใช้สมการ (17) นี้จำเป็นที่จะต้องกำหนดและประมาณค่ารูปแบบในการแจกแจงของ $g(\tau)$ และ $\Pr(X = i \mid \tau)$

จากสมการ (17) เมื่อกำหนดคะแนนจริงที่เป็นสัดส่วนในการตอบถูกของผู้สอบแล้ว ค่าคาดหวังของคะแนนดิบนิยามให้เป็น K เท่าของสัดส่วนในการตอบถูกของคะแนนจริง ดังนั้นตัวแปรเชิงสุ่มของคะแนนดิบ X ก็คือ

$$K\tau = \varepsilon(X \mid \tau) = \sum_{i=0}^{K} i \Pr(X = i \mid \tau) \qquad \text{.....(18)}$$

เมื่อ ε เป็นค่าคาดหวัง ความแปรปรวนคลาดเคลื่อนของผู้สอบเมื่อกำหนดคะแนนจริงที่เป็นสัดส่วนของการตอบถูก นิยามได้ดังนี้

$$\sigma^2(X \mid \tau) = \varepsilon[(X - K\tau)^2 \mid \tau = \sum_{i=0}^{K} (i - K\tau)^2 \Pr(X = i \mid \tau) \qquad \text{.....(19)}$$

ความแปรปรวนคลาดเคลื่อนเฉลี่ยในการวัด (E) ของผู้สอบทั้งหมด คือ

$$\sigma^2(E) = \int_0^1 \sigma^2(X \mid \tau) g(\tau) d\tau \qquad \text{.....(20)}$$

นิยามให้ s(i) ว่าเป็นคะแนนดิบที่ถูกแปลง (i = 0,...,K) ในกรณีนี้ ค่าเฉลี่ยของประชากรผู้เข้าสอบ s(X) ก็คือ

$$\mu[s(X)] = \sum_{i=0}^{K} s(i) \Pr(X = i), \qquad \text{.....(21)}$$

เมื่อ $\Pr(X = i)$ แทน การแจกแจงความถี่สัมพัทธ์ของคะแนนดิบ ที่ได้นิยามไว้ในสมการ (17) ดังนั้นความแปรปรวนก็คือ

$$\sigma^2[s(X)] = \sum_{i=0}^{K}\{s(i) - \mu[s(X)]\}^2 \Pr(X = i) \qquad \text{.....(22)}$$

และโมเมนต์อันดับที่สูงกว่านี้ก็สามารถนิยามได้ทำนองเดียวกัน

ค่าคาดหวังของคะแนนแปลงในผู้สอบเมื่อกำหนดคะแนนจริงที่เป็นสัดส่วนในการตอบถูก คือ

$$\varepsilon(\tau) = \varepsilon[s(X) \mid \tau] = \sum_{i=0}^{K} s(i) \Pr(X = i \mid \tau) \qquad \text{.....(23)}$$

ซึ่งเป็นคะแนนแปลงของคะแนนจริงในผู้สอบคนนั้น ความแปรปรวนคลาดเคลื่อนในการวัดของคะแนนแปลงในผู้สอบเมื่อกำหนดสัดส่วนคะแนนจริงของการตอบถูกในคะแนนแปลงมาให้ ก็คือ

$$\sigma^2[s(X) \mid \tau] = \varepsilon\{[s(X) - \xi(\tau)]^2 \mid \tau\} = \sum_{i=0}^{K}[s(i) - \xi(\tau)]^2 \Pr(X = i \mid \tau) \qquad \text{.....(24)}$$

นิยามให้ E_s เป็นความควาดเคลื่อนในการวัดของคะแนนแปลง ดังนั้นค่าเฉลี่ยของความแปรปรวนคลาดเคลื่อนของคะแนนแปลงของผู้สอบทั้งหมด ก็คือ

$$\sigma^2(E_s) = \int_0^1 \sigma^2[s(X) \mid \tau] g(\tau) d\tau \qquad \text{.....(25)}$$

ความแปรปรวนของคะแนนจริง $\sigma^2(\xi)$ ได้มาจากการลบสมการ (25) ออกจากสมการ (22)

ดังนั้นถ้าสมมติให้คะแนนจริงและความคลาดเคลื่อนของคะแนนแปลง ไม่มีความสัมพันธ์กัน แล้วสัมประสิทธิ์ความเชื่อมั่นของคะแนนแปลง ก็คือ

$$\frac{\sigma^2(\xi)}{\sigma^2(\xi) + \sigma^2(E_s)} = 1 - \frac{\sigma^2(E_s)}{\sigma^2[s(X)]} \qquad \text{.....(26)}$$

ในสมการ (26) นี้อาจจะไม่สอดคล้องกับการหาความเชื่อมั่นแบบสหสัมพันธ์ของแบบทดสอบคู่ขนาน

โมเดลคะแนนจริงอย่างเข้มงวด (Strong True Score Models)

โมเดลของคะแนนจริงอย่างเข้มงวดเหมาะกับแบบทดสอบเลือกตอบที่ให้คะแนนตามจำนวนข้อที่ตอบถูก โมเดลนี้เริ่มด้วยโมเดลแบบทั่วไปมากกว่าโมเดลที่แสดงในสมการ (17) ในโมเดลทั่วไปนี้ข้อคำถามแต่ละข้อถูกจัดลงไปในกลุ่มเนื้อหาหนึ่งในตารางวิเคราะห์หลักสูตรที่มี M ชั้น ในแต่ละชั้น M มีอยู่ K_m ข้อ ในแต่ละชั้นความน่าจะเป็นที่ตัวแปรเชิงสุ่มของคะแนนดิบ X_m ของชั้นนั้นเท่ากับ i ($i_m = 0,\ldots, K_m$) ในประชากรของผู้สอบคือ

$$\Pr(X_m = i_m) = \int_0^1 \Pr(X_m = i_m \mid \tau_m) g(\tau_m) d\tau_m \qquad \text{.....(27)}$$

เมื่อ τ_m เป็นคะแนนจริงที่เป็นสัดส่วนในการตอบถูกในชั้น m ด้วยฟังก์ชันความหนาแน่น $g(\tau_m)$ ในประชากร สมการ (27) เท่ากับสมการ (17) ยกเว้นโมเดลที่มีการกระจายในแต่ละชั้นมากกว่าในแบบทดสอบทั้งหมด

สมมติว่าการแจกแจงความคลาดเคลื่อนอย่างมีเงื่อนไข $Pr(X_m = i_m \mid \tau_m)$ เป็นแบบทวินาม ให้คะแนนดิบเป็นอิสระกัน เมื่อกำหนดคะแนนจริงที่เป็นสัดส่วนการตอบถูก โอกาสอย่างมีเงื่อนไขของตัวแปรเชิงสุ่มของคะแนนดิบ X ทั้งหมด ก็จะเท่ากับ i ($i = 0,\ldots,K$ เมื่อ $K = \sum_m K_m$) เมื่อกำหนดคะแนนจริงของชั้น τ_m ($m = 1,\ldots,M$) มาให้ เขียนสัญลักษณ์ได้เป็น

$$Pr(X = i \mid \tau_1, \tau_2, \ldots, \tau_M) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(28)$$

สมการ (28) ได้จากการแจกแจงแบบทวินามประกอบที่มีพารามิเตอร์ $\tau_1, \tau_2, \ldots, \tau_m$ ในข้อตกลงที่กำหนดไว้ให้คะแนนดิบแต่ละชั้นเป็นอิสระเมื่อกำหนดสัดส่วนการตอบถูก จะเหมือนกับข้อตกลงของความเป็นอิสระของข้อสอบในโมเดล IRT ที่กล่าวว่าการตอบข้อสอบทั้งหมดเป็นอิสระจากกันเมื่อกำหนดพารามิเตอร์ความสามารถมาให้ ซึ่งเป็นข้อตกลงที่แข็งแกร่งกว่าข้อตกลงที่กำหนดในโมเดลคะแนนจริงแบบเข้มงวด

ลอร์ด (Lord. 1965, 1969) เสนอการประมาณค่าแบบสองเทอมของการแจกแจงทวินามประกอบ ดังนี้

$$Pr(X = i \mid \tau_1, \tau_2, \ldots, \tau_M) \approx p_K(i) + \frac{K}{2} V_2 C_2(i) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(29)$$

เมื่อ

$$p_n(i) = \binom{n}{i} \tau^i (1-\tau)^{n-i}$$

$$C_2(i) = \sum_{j=0}^{2} (-1)^{j+1} \binom{2}{j} p_{K-2}(i-j)$$

$$V_2 = \frac{1}{K} \sum_{m=1}^{M} K_m (\tau_m - \tau)^2, \text{and}$$

$$\tau = \frac{1}{K} \sum_{m=1}^{M} K_m \tau_m$$

เทอม n ใน $p_n(i)$ เป็น Placeholder τ เป็นคะแนนจริงที่เป็นสัดส่วนการตอบถูกเฉลี่ยในทุกกลุ่มที่กำหนดค่าน้ำหนักโดยจำนวนข้อในแต่ละกลุ่ม การแจกแจงนี้ยังขึ้นอยู่กับตัวพารามิเตอร์ M ที่ไม่ทราบค่า อย่างไรก็ตาม V_2 ที่เป็นฟังก์ชันของ τ ใช้แทน τ_m ($m = 1,\ldots,M$) ฟังก์ชันที่ลอร์ด (1965) ใช้ประมาณค่า V_2 คือ $\frac{2k}{K}\tau(1-\tau)$ เมื่อ k เป็นพารามิเตอร์ตัวหนึ่ง ฉะนั้นความน่าจะเป็นที่ตัวแปรเชิงสุ่ม X จะเท่ากับ i ในประชากรของผู้สอบสามารถประมาณค่าโดยสมการ (17) เมื่อ $Pr(X = i \mid \tau)$ เป็นการประมาณค่าแบบสองเทอมของการแจกแจงทวินามประกอบที่เน้นเพียงพารามิเตอร์ τ ของคะแนนจริงที่เป็นสัดส่วนการตอบถูกเพียงตัวเดียว

ค่า k หาได้โดยการกำหนดค่าทางทฤษฎีของความแปรปรวนคลาดเคลื่อนเฉลี่ยภายใต้การประมาณค่าการแจกแจงแบบสองเทอมแบบทวินามประกอบ (ซึ่งเป็นฟังก์ชันของ k) ซึ่งเท่ากับการประมาณค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยภายใต้การประมาณค่าการแจกแจงแบบทวินามประกอบแบบสองเทอม ก็คือ

$$\frac{\sigma^2(X)[K(K-1)+2k-K^2]+(K-2k)\mu(X)[K-\mu(X)]}{K(K-1)+2k} \qquad \text{.....(30)}$$

เมื่อ $\mu(X)$ และ $\sigma^2(X)$ เป็นค่าเฉลี่ยและความแปรปรวนของคะแนนดิบ ให้สมการ (30) เท่ากับความแปรปรวนคลาดเคลื่อนเฉลี่ย $\sigma^2(E)$ และแก้สมการหา k ดังนี้

$$k=\frac{K\{(I-1)[\sigma^2(X)-\sigma^2(E)]-K\sigma^2(X)+\mu(X)+\mu(X)[K-\mu(X)]\}}{2\{\mu(X)[K-\mu(X)]-[\sigma^2(X)-\sigma^2(E)]\}} \qquad \text{.....(31)}$$

ฉะนั้นค่าของ k จะประมาณค่าได้เมื่อกำหนดค่า $\sigma^2(E)$, $\mu(X)$ และ $\sigma^2(X)$ มาให้

การประมาณค่าของ $\mu(X)$ และ $\sigma^2(X)$ ประมาณได้โดยใช้ค่าเฉลี่ยและความแปรปรวนของกลุ่มตัวอย่าง ในการประมาณค่า $\sigma^2(E)$ ตามแนวคิดของลอร์ด (1965) แล้วข้อคำถามแต่ละข้อก็คือข้อสอบในแต่ละชั้นนั่นเอง (นั่นคือ 1 ข้อก็คือ 1 ชั้น) ดังนั้น เมื่อต้องการประมาณค่าของ $\sigma^2(E)$ จะเน้นการประมาณค่าความแปรปรวนคลาดเคลื่อนแต่ละคนที่กำหนดให้ชั้นเป็นกลุ่มของเนื้อหานั่นเอง ความแปรปรวนคลาดเคลื่อนของแต่ละบุคคลตามวิธีของเฟลต์ จะประมาณค่าของผู้สอบที่ตอบจำนวนข้อสอบถูก p (ไม่ใช่สัดส่วนการตอบถูก) ดังนี้

$$\hat{\sigma}^2(E_p)=\sum_{m=1}^{M}\frac{x_{pm}(K_m-x_{pm})}{K_m-1} \qquad \text{.....(32)}$$

เมื่อ x_{pm} เป็นคะแนนของแบบทดสอบที่สังเกตได้ในผู้สอบ p ในชั้น m ค่าเฉลี่ยของความแปรปรวนคลาดเคลื่อนที่ถูกประมาณค่าในกลุ่มผู้สอบทั้ง N คน คือ

$$\hat{\sigma}^2(E_p)=\sum_{p=1}^{N}\frac{\hat{\sigma}^2(E_p)}{N} \qquad \text{.....(33)}$$

ในการประยุกต์ใช้ค่า k ที่อยู่ในการประมาณค่าสองเทอมในการแจกแจงแบบทวินามประกอบจะได้จากการใช้ $\hat{\sigma}^2(E)$ จากสมการ (33) ว่าเป็นการประมาณค่าของ $\sigma^2(E)$ ในสมการ (31)

โมเดลของลอร์ด สมมติให้การแจกแจงสัดส่วนที่ตอบถูกของคะแนนจริงเป็นค่าหนึ่งของตระกูลการแจกแจงเบต้า 4 พารามิเตอร์ ซึ่งการแจกแจงแบบนี้มีประโยชน์ที่เหมาะกับการแจกแจงคะแนนจริงที่เป็นสัดส่วนในการตอบถูก เพราะว่าเป็นค่าต่อเนื่อง มีขอบเขตที่จุดปลายทั้งสอง และสามารถที่จะให้ความแปรเปลี่ยนอย่างกว้างขวางของรูปร่าง พารามิเตอร์ 2 ตัวของการแจกแจงนี้ให้สัญลักษณ์เป็น α และ β พารามิเตอร์อีก 2 ตัว ให้สัญลักษณ์ว่า l และ u เป็นขอบเขตจำกัดล่างและบนของการแจกแจงคะแนนจริงที่เป็นสัดส่วนการตอบถูก (นั่นคือ $g(\tau) = 0$ เมื่อ $\tau < 1$ และ $\tau > u$) l และ u

จะถูกจำกัดให้อยู่ระหว่าง 0 และ 1 สำหรับความเหมาะเจาะพอดี (fitting) ของการแจกแจงคะแนนจริงที่เป็นสัดส่วนการตอบถูก

ดังนั้นการประยุกต์ใช้สำหรับวิธีนี้ ขั้นตอนแรกต้องประมาณค่าพารามิเตอร์ของโมเดลคะแนนจริงแบบเข้มงวด และนำค่าประมาณเหล่านี้ไปใช้แทนค่าพารามิเตอร์ในสมการ (18) ถึง สมการ (26) เมื่อกำหนดโมเดลของคะแนนที่เหมาะเจาะแล้ว ก็ประมาณค่าคุณสมบัติของคะแนนแปลง ที่ได้นำเสนอไว้ในสมการ (21) ถึงสมการ (26) เป็นสมการที่สามารถคำนวณได้อย่างตรงไปตรงมา ยกเว้น $\sigma^2(E_s)$ ในสมการ (25) ซึ่งต้องใช้การอินทิเกรดในการคำนวณ

การประมาณค่า CSEM โดยใช้ทฤษฎีการตอบข้อสอบ

ในปี ค.ศ. 1982 ลอร์ด ได้เสนอวิธีการประมาณค่า CSEM เมื่อความสามารถแฝงถูกแปลงเป็นฟังก์ชั่นแบบไม่เป็นเชิงเส้นตรง (ที่เป็นแบบต่อเนื่อง) โดยใช้ทฤษฎีการตอบข้อสอบ แต่ลอร์ดไม่ได้เน้นถึงประเด็นที่คะแนนแปลงถูกตัดทอนออกไป หรือกรณีที่คะแนนดิบหลายค่าถูกแปลงไปเป็นคะแนนแปลงรูปได้เพียงค่าเดียว แต่ลอร์ดเน้นถึงการแปลงคะแนนในสเกลของความสามารถแฝงมากกว่าที่จะใช้คะแนนดิบ

ต่อมาโคเลน และแฮนสัน (Kolen & Hanson. 1989 ; citing Kolen, Hanson & Brennan. 1992 : 286) ได้เสนอวิธีที่เป็นทางเลือกอื่นอีกสำหรับการประมาณค่า CSEM และความเชื่อมั่นของคะแนนแปลงรูปอีกด้วย แต่วิธีของ โคเลน และแฮนสัน จะไม่เหมือนกับของลอร์ด ตรงที่มีการรวมการแปลงคะแนนดิบแบบไม่ต่อเนื่องไปยังคะแนนแปลงรูป และคำนึงถึงประเด็นที่คะแนนแปลงรูปถูกตัดออกไปบางส่วน หรือกรณีที่มีคะแนนดิบมากกว่า 1 ค่าที่ถูกแปลงเป็นคะแนนแปลงได้เพียงค่าเดียว

ในวิธีการทั้งของลอร์ด และวิธีของโคเลนและแฮนสัน ใช้โมเดลความคลาดเคลื่อนในการวัดที่ไม่ใช้กับแบบทดสอบที่เน้นความเร็วในการสอบ

ต่อมาปี ค.ศ. 1992 โคเลน เซง และแฮนสัน (Kolen, Zeng & Hanson. 1996 ; citing Kolen, Hanson & Brennan. 1992 : 286) ก็ได้นำเสนอวิธีการประมาณค่าความคลาดเคลื่อนอย่างมีเงื่อนไขโดยใช้วิธีการตอบข้อคำถาม วิธีนี้จะทำให้ได้การแจกแจงของจำนวนข้อที่ตอบถูกที่เป็นคะแนนดิบ เมื่อกำหนดความสามารถ (θ) และพารามิเตอร์ของข้อคำถามมาให้ ผลที่ได้จากการแจกแจงของคะแนนดิบจะถูกแปลงไปเป็นคะแนนแปลงรูปโดยใช้การแปลงรูปวิธีหนึ่ง ค่าเบี่ยงเบนมาตรฐานของการแจกแจงของคะแนนแปลงรูป เมื่อกำหนดระดับความสามารถมาให้จะถูกทำให้เป็นค่า CSEM ณ ระดับความสามารถนั้น การใช้ทฤษฎีการตอบข้อสอบ นอกจากจะใช้ประมาณค่า CSEM แล้วยังสามารถนำมาประมาณค่าความเชื่อมั่นของแบบทดสอบได้อีก

การแจกแจงอย่างมีเงื่อนไขของจำนวนข้อที่ตอบถูก

(Conditional Distribution of Number-Correct Raw Scores)

ความน่าจะเป็นที่ตัวแปรเชิงสุ่ม X เท่ากับคะแนนดิบ i ในการทำแบบทดสอบ K ข้อ เมื่อกำหนดระดับความสามารถมาให้ เขียนเป็นสัญลักษณ์ได้ดังนี้

$$Pr(X = i \mid \theta)$$

ซึ่งเป็นการแสดงการแจกแจงแบบมีเงื่อนไขของคะแนนดิบจากจำนวนข้อที่ตอบถูก (Number – Correct Raw Score) เมื่อกำหนดความสามารถมาให้ ซึ่งลอร์ด (Lord. 1980 ; citing Kolen, Zeng & Hanson. 1996 : 30) ได้ชี้ให้เห็นว่าการแจกแจงคะแนนดิบอย่างมีเงื่อนไขสามารถเขียนการแจกแจงได้เหมือนกับการแจกแจงแบบทวินามประกอบ เมื่อกำหนดพารามิเตอร์ของข้อคำถามสำหรับแบบทดสอบที่มี K ข้อมาให้

ลอร์ดและวิงเกอร์สกี้ (Lord & Wingersky. 1984) ได้นำเสนอสูตรที่จะนำมาคำนวณความน่าจะเป็นเหล่านี้ และต่อมา โคเลนและเบรนนัน (Kolen & Brennan. 1995) ก็ได้นำเสนอโดยนิยาม X_r ว่าเป็นตัวแปรเชิงสุ่ม r ข้อคำถามแรกของแบบทดสอบ ($X_r = 0,1,\ldots,r \leq K$) ดังนั้น $Pr(X_r = i \mid \theta)$ แทนโอกาสที่ X_r เท่ากับ i เมื่อกำหนด θ ให้ ทำนองเดียวกันนิยามโค้งคุณลักษณะของข้อคำถามข้อ j เป็น $p_j(\theta)$ ว่าคือโอกาสของการตอบข้อสอบข้อที่ j ถูกต้อง ณ ระดับความสามารถ θ สำหรับข้อคำถามแรก

$$Pr(X_1 = 1 \mid \theta) = p_1(\theta)$$
$$Pr(X_1 = 0 \mid \theta) = [1 - p_1(\theta)] \quad \text{สำหรับ } r > 1$$

สูตรที่นำมาพิจารณาใหม่คือ

$$\begin{aligned} Pr(X_r = i \mid \theta) &= Pr(X_{r-1} = i \mid \theta)[1 - p_r(\theta)] && i = 0 \\ &= Pr(X_{r-1} = i \mid \theta)[1 - p_r(\theta)] + Pr(X_{r-1} = i-1 \mid \theta)p_r(\theta) && 0<1< r \\ &= Pr(X_{r-1} = i-1 \mid \theta)p_r(\theta) && i = r \end{aligned} \qquad \ldots\ldots\ldots\ldots\ldots(34)$$

ในการใช้สูตรนี้ จะเริ่มที่ r = 1 และหาค่าความน่าจะเป็นซ้ำเมื่อเพิ่ม r ในแต่ละครั้ง และหยุดทำหลังจากที่ r = K นั่นคือ $Pr(X = i \mid \theta) = Pr(X_K = i \mid \theta)$ ความแปรปรวนของผลของการแจกแจงก็คือ ความแปรปรวนอย่างมีเงื่อนไขของคะแนนดิบที่เป็นจำนวนข้อที่ตอบถูก ณ ที่ความสามารถ θ เขียนได้เป็น $\sigma^2(X \mid \theta)$ รากที่สองของ $\sigma^2(X \mid \theta)$ ก็คือความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของการวัดจากจำนวนข้อที่ตอบถูก ณ ที่ความสามารถ θ นั่นเอง

การแจกแจงอย่างมีเงื่อนไขของคะแนนแปลงรูป

(Conditional Distribution of Scale Score)

สมมติให้การแปลงคะแนนดิบไปเป็นคะแนนแปลงรูป ใช้สัญลักษณ์ 's การแปลงคะแนนนี้จะถูกนำไปประยุกต์ใช้กับการแจกแจงแบบมีเงื่อนไขของคะแนนดิบจากจำนวนข้อที่ตอบถูก ณ ระดับสามารถ θ เพื่อให้ได้การแจกแจงความน่าจะเป็นอย่างมีเงื่อนไขของคะแนนแปลงรูป เมื่อกำหนด θ

ค่าเฉลี่ยของการแจกแจงอย่างมีเงื่อนไข คือ

$$\xi(\theta) = \varepsilon[s(X) \mid \theta) = \sum_{i=0}^{K} s(i) Pr(X = i \mid \theta) \qquad \ldots\ldots\ldots\ldots\ldots(35)$$

ซึ่งเป็นคะแนนแปลงรูปของคะแนนจริง ณ ที่ θ ในสมการ (35) ε หมายถึงค่าที่คาดหวัง ความแปรปรวนคลาดเคลื่อนของการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป ณ ที่ θ คือ

$$\sigma^2[s(X) \mid \theta = \varepsilon\{[s(X) - \xi(\theta)]^2 \mid \theta) = \sum_{i=0}^{K}[s(i) - \xi(\theta)]^2 \Pr(X = i \mid \theta) \quad \text{.....(36)}$$

เมื่อถอดรากที่สองของสมการ (36) ก็จะเป็นความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข นั่นเอง

ความแปรปรวนมาตรฐานเฉลี่ยและความเชื่อมั่น

(Average Standard Errors and Reliability)

โคเลนและคณะ (Kolen et al. 1992) ให้เหตุผลว่า ความแปรปรวนคลาดเคลื่อนเฉลี่ยและดัชนีความเชื่อมั่นของคะแนนดิบจากจำนวนข้อที่ตอบถูกและคะแนนแปลงรูป จะถูกกำหนดได้อย่างเฉพาะเจาะจงถ้ากำหนดการแจกแจงของ θ ให้

ให้ $\psi(\theta)$ เป็นฟังก์ชันความหนาแน่นของความสามารถในประชากร ดังนั้นความแปรปรวนความคลาดเคลื่อนเฉลี่ย (E) ของคะแนนที่ทำถูก สามารถเขียนได้เป็น

$$\sigma^2(E) = \int_\theta \sigma^2(X \mid \theta)\psi(\theta)\, d\theta \quad \text{.....(37)}$$

การแจกแจงของคะแนนที่เป็นจำนวนข้อที่ตอบถูกในประชากรทั้งหมดคือ

$$\Pr(X = i) = \int_\theta \Pr(X = i \mid \theta)\psi(\theta)\, d\theta \quad \text{.....(38)}$$

และให้ $\sigma^2(X)$ เป็นความแปรปรวนของการแจกแจงคะแนนของจำนวนข้อที่ตอบถูก ดังนั้นความเชื่อมั่นของคะแนนจากจำนวนข้อที่ตอบถูกนิยามว่าเท่ากับ 1 ลบด้วยอัตราส่วนระหว่างความแปรปรวนคลาดเคลื่อนและความแปรปรวนทั้งหมด นั่นคือ

$$rel_{raw} = 1 - \frac{\sigma^2(E)}{\sigma^2(X)} \quad \text{.....(39)}$$

ดังนั้นความแปรปรวนความคลาดเคลื่อนเฉลี่ยในการวัดของคะแนนแปลงรูป (E_S) ก็คือ

$$\sigma^2(E_S) = \int_\theta \sigma^2[s(X) \mid \theta]\psi(\theta)\, d\theta \quad \text{.....(40)}$$

ค่าเฉลี่ยของคะแนนแปลงรูปในประชากร เขียนได้ดังนี้

$$\mu[s(X)] = \sum_{i=0}^{K} s(i) \Pr(X = i) \quad \text{.....(41)}$$

และความแปรปรวนคือ

$$\sigma^2[s(X)] = \sum_{i=0}^{K} \{s(i) - \mu[s(X)]\}^2 \Pr(X = i) \quad \text{.....(42)}$$

ความเชื่อมั่นของคะแนนแปลงรูป คือ

$$rel_{scale} = 1 - \frac{\sigma^2(E_s)}{\sigma^2[s(X)]} \qquad \text{.....(43)}$$

ข้อตกลงเบื้องต้นที่สำคัญของทฤษฎีการตอบข้อสอบ

อย่างไรก็ตามในการใช้ทฤษฎีการตอบข้อคำถามเพื่อประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ต้องคำนึงถึงข้อตกลงเบื้องต้นเกี่ยวกับข้อมูลที่จะนำมาประมาณค่าที่สำคัญ (Hambleton & Swaminathan.1985 : 16 - 30) ดังนี้

ประการที่หนึ่ง มิติของลาเทนท์สเปส (Dimensionality of Latent Space) กล่าวถึง ลักษณะหรือความสามารถที่เป็นตัวกำหนดพฤติกรรมการตอบข้อสอบจะมีเพียงลักษณะเดียว เรียกว่า Unidimensional

ประการที่สอง ได้แก่ ความเป็นอิสระในการตอบข้อสอบ (Local Independence) ข้อตกลงเบื้องต้นนี้หมายถึง พฤติกรรมการตอบข้อสอบต่าง ๆ ในแบบทดสอบของแต่ละคนมีความเป็นอิสระในเชิงสถิติ นั่นคือ การตอบข้อสอบข้อหนึ่งไม่มีผลต่อการตอบข้อสอบข้ออื่น ในแบบทดสอบฉบับเดียวสามารถเขียนเป็นสมการได้ดังนี้

$$Prob[U_1 = u_1, U_2 = u_2, \ldots, U_n = u_n \mid \theta] = \prod_{i=1}^{n} P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i}$$

เมื่อ U_i = 1, 2, 3, ..., n แทน ผลการตอบข้อสอบ n ข้อ เมื่อตอบถูกให้ 1
ตอบผิดให้ 0

P_i แทน โอกาสที่ผู้สอบคนหนึ่งทำข้อสอบข้อที่ i ถูก

$Q_i = 1 - P_i$

ประการที่สาม โค้งลักษณะข้อสอบ (Item Characteristic Curve) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นถูกกับระดับความสามารถของผู้สอบ

วิธีพหุนาม (Polynomial Procedure)

เฟลต์และควอลส์ (Feldt & Qualls. 1998) ได้เสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง ($\sigma_{E(S|x)}$) ที่แปลง f(X) แบบไม่เป็นเชิงเส้นตรงโดยวิธีพหุนาม ดังนี้

จากสมการ (3) และ (4) $\sigma_{E(S)} = A\sigma_{E(X)}$ .....(3)

$\sigma_{E(S|x)} = A\sigma_{E(X|x)}$ .....(4)

A เป็นค่าคงที่สำหรับทุกค่าที่ X = x ในนิยามของแคลคูลัสเกี่ยวกับการดิฟเฟอร์เรนเชียล ค่า A จะเป็นอนุพันธ์อันดับที่หนึ่งของ f(X) ซึ่งเขียนเป็น f'(X) นั่นคือในสมการ (3) สามารถเขียนได้เป็น

$$\sigma_{E(S)} = f'(X)\sigma_{E(X)} \quad \text{.....(44)}$$

สำหรับความสัมพันธ์แบบไม่เป็นเชิงเส้น เราจำเป็นต้องหาอนุพันธ์อันดับที่หนึ่งของ f(X) (ซึ่งก็คือความชันนั่นเอง) ณ ทุกค่าที่ X = x ดังนั้น SEM ณ ที่แต่ละคะแนน x สามารถถูกประมาณค่าได้เป็น

$$\sigma_{E(S|x)} = f'(X)\sigma_{E(X|x)} \quad \text{.....(45)}$$

โคเลนและเบรนนัน (Kolen & Brennan. 1995) กล่าวว่า สมการ (45) ของการประมาณค่า $\sigma_{E(S|X)}$ ก็คือการประมาณค่าแบบ "delta-method" มี 2 ขั้นตอน คือ

ขั้นที่ 1 ประมาณค่า $\sigma_{E(X)}$ โดยวิธีการหาค่า SEM ของลอร์ดหรือของเฟลด์หรือสูตรของเบรนนัน

ขั้นที่ 2 ประมาณค่า $f'(X)$ โดยใช้โพลิโนเมียลอันดับสูง การปรับเส้นโค้งที่มีอันดับสาม หรือใช้วิธีแบบง่าย ที่เสนอโดยเฟลด์และควอลส์

สำหรับการใช้วิธีการแบบพหุนาม เพื่อประมาณ $f'(x)$ นั้นเริ่มด้วยการหาดีกรี n ที่เหมาะกับ f(X)

$$F(X) = c_0 + c_1X + c_2X^2 + \ldots + c_nX^n \quad \text{.....(46)}$$

อนุพันธ์อันดับแรกของ f(X) ณ ที่ X = x คือ

$$f'(x) = c_1 + 2c_2x + \ldots + nc_nx^{n-1} \quad \text{.....(47)}$$

จะเห็นว่าวิธีแบบพหุนามนี้ดูเหมือนจะง่ายแต่มีความยุ่งยากในทางปฏิบัติ เพราะประการแรกจะต้องเลือกอันดับ n ซึ่งเป็นวิธีหนึ่งที่อาศัยการพิจาณาเกี่ยวกับจำนวนการวิเคราะห์ข้อมูลแบบสำรวจ และประการที่สองเป็นวิธีการที่ต้องใช้โปรแกรมคอมพิวเตอร์ช่วยในการคำนวณ

วิธีเฟลด์ - ควอลส์ (Feldt – Qualls Procedure)

ในการประมาณค่า CSEM โดยวิธีพหุนามมีความยุ่งยากในทางปฏิบัติซึ่งต้องหา $f'(x)$ เสียก่อน ดังนั้น เฟลด์และควอลส์ (Feldt & Qualls. 1998 ; citing Brennan and Lee. 1999 : 8-9)) จึงเสนอวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด เพื่อสามารถนำมาใช้ได้ทางปฏิบัติซึ่งมีการคำนวณที่ง่าย แต่ต้องไม่ทำให้ส่วนใดของการแจกแจงของคะแนนที่แปลงอยู่นอกพิสัย f(x – C) ถึง f(x + C) สูตรที่ใช้ในการประมาณค่าคือ

$$f'(X) = \begin{cases} \dfrac{f(x+C)-f(x-C)}{2C} & \text{เมื่อ } x-C \geq 0 \text{ และ } x+C \leq k \\ \dfrac{f(x+C)-f(0)}{x+C} & \text{เมื่อ } x-C < 0 \text{ และ} \\ \dfrac{f(k)-f(x-C)}{k-x+c} & \text{เมื่อ } x+C > k \end{cases} \quad \text{.....(48)}$$

และ

$$C = \begin{cases} 2 & \text{ถ้า } k < 25 \\ 3 & \text{ถ้า } 25 \leq k \leq 39 \\ 4 & \text{ถ้า } 40 \leq k \leq 59 \\ 5 & \text{ถ้า } k \geq 60 \end{cases} \qquad \text{.....(49)}$$

เมื่อ k เป็นจำนวนข้อคำถามในแบบทดสอบ ที่มีการให้คะแนนแบบ 0 - 1 และมีค่าน้ำหนักคะแนนเท่ากัน

วิธีทวินาม (Binomial Procedure)

แอลเลนและเยน (Allen & Yen. 1979. : 242 - 243) ได้นำเสนอโมเดลความคลาดเคลื่อนแบบทวินามโดยนิยามและกำหนดข้อตกลงเบื้องต้นของโมเดล ดังนี้

โมเดลความคลาดเคลื่อนแบบทวินามนิยามว่าคะแนนที่สังเกตได้คือคะแนนที่ตอบข้อคำถามถูกของแบบทดสอบที่เป็นอิสระจากกันและข้อคำถามมีความยากเท่ากัน ความคลาดเคลื่อนมาตรฐานในการวัดแปรเปลี่ยนไปตามระดับของคะแนนจริง และไม่ขึ้นอยู่กับการแจกแจงคะแนนของกลุ่มตัวอย่างที่ทดสอบ

ข้อตกลงเบื้องต้นของโมเดลแบบทวินาม

1. คะแนนที่สังเกตได้จากการสอบของผู้สอบแต่ละคนเป็นผลรวมของคะแนนสำหรับข้อคำถามที่ตอบถูกเมื่อมีการให้คะแนน 1 ในการตอบถูก และ 0 ในการตอบผิด

2. โอกาสที่ตอบข้อคำถามถูก (เท่ากับหนึ่ง) คือ ζ_j สำหรับข้อคำถามแต่ละข้อ และทุกข้อสำหรับผู้เข้าสอบคนที่ j นั่นคือ ζ_j เป็นคะแนนจริงของผู้เข้าสอบคนที่ j ถ้าข้อคำถามมีความยากแตกต่างกัน หรือผู้เข้าสอบทำแบบทดสอบไม่เสร็จข้อตกลงนี้จะถูกละเมิด

3. $X_j = n\zeta_j + E_j$ เมื่อ E_j เป็นความคลาดเคลื่อนของการวัดสำหรับผู้เข้าสอบคนที่ j

4. ค่าคาดหวังของคะแนนที่สังเกตได้สำหรับผู้สอบคนหนึ่งที่มีคะแนนจริง ζ_j คือ $n\zeta_j : \varepsilon(X_j \mid n\zeta_j) = n\zeta_j$ ข้อตกลงนี้หมายถึง $\varepsilon(E_j \mid N\zeta_j) = 0$

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยใช้โมเดลแบบทวินามที่จะเสนอต่อไปนี้เป็นแนวคิดของ โคเลนและคนอื่น ๆ และวิธีของลอร์ด (Kolen et al. 1992 ; Lord. 1955, 1957 citing Brennan & Lee. 1999 : 10 - 11) มีดังนี้

ภายใต้โมเดลแบบทวินาม โอกาสที่จะตอบข้อคำถามถูก y ข้อ จากข้อคำถามทั้งหมด k ข้อ เป็นดังนี้

$$\Pr(y \mid \pi, k) = \binom{k}{y} \pi^y (1-\pi)^{k-y} \qquad \text{.....(50)}$$

เมื่อ π เป็นสัดส่วนของคะแนนจริงที่ตอบข้อคำถามถูก ซึ่งจากโมเดลนี้การประมาณค่า SEM ของผู้สอบแต่ละคนที่จะตอบข้อคำถามถูก x ข้อ ตามวิธีของลอร์ด คือ

$$\hat{\sigma}_{E(x)} = \sqrt{\frac{x(k-x)}{k-1}} = c_k\sqrt{\frac{x(k-x)}{k}} \qquad \text{.....(51)}$$

เมื่อ $\quad c_k = \sqrt{k/(k-1)}$ .....(52)

c_k เป็นองค์ประกอบการปรับแก้ที่ไม่ลำเอียง ค่า π ในสมการ (50) สามารถประมาณค่าได้จาก $\bar{x} = x/k$ นั่นคือ c_k ใช้เป็นค่าปรับแก้ $\hat{\sigma}^2_{E(x)}$ ให้เป็นค่าประมาณที่ไม่ลำเอียงของ $\sigma^2_{E(x)}$ และค่า $\bar{x}$ เป็นสัดส่วนของคะแนนที่สังเกตได้ของผู้เข้าสอบที่ตอบข้อคำถามถูกเมื่อข้อคำถามมีการให้คะแนนแบบทวิ (Dichotomous)

ดังนั้นในสมการ (51) กระจายได้เป็น

$$\hat{\sigma}_{E(x)} = c_k\sqrt{\sum_{y=0}^{k} y^2 \Pr(y \mid \bar{x},k) - \left[\sum_{y=0}^{k} y \Pr(y \mid \bar{x},k)\right]^2} \qquad \text{.....(53)}$$

สมการนี้เป็นผลมาจากที่ให้ Z เป็นตัวแปรสุ่มใด ๆ แล้ว $\sigma^2(Z) = \varepsilon Z^2 - (\varepsilon Z)^2$ เมื่อ ε เป็นค่าคาดหวัง สังเกต x และ y ในสมการ (53) จะมีบทบาทต่างกันแม้ว่าทั้ง x และ y หมายถึงจำนวนข้อที่ตอบถูกก็ตาม ค่าของ x เป็นคะแนนที่จะใช้ประมาณหาค่าความคลาดเคลื่อนมาตรฐาน แต่ y เป็นตัวแปรที่ชี้ถึงคะแนนใด ๆ หรือเป็นโอกาสที่คะแนนนั้นจะเกิดขึ้น (นั่นคือ ในสมการ (51) แสดงในเทอมของ x แต่สมการ (50) แสดงในเทอมของ y) จากนิยามความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละบุคคล ก็คือความคลาดเคลื่อนบนคะแนนจริง แต่ค่า SEMs ที่ถูกประมาณค่านี้ เกี่ยวกับการแทนค่าคะแนนที่สังเกตได้ด้วยคะแนนจริง ดังนั้นเพื่อที่จะให้ได้ค่า CSEM ของคะแนนแปลงรูป ตามวิธีของลอร์ด ก็จะแทนค่า y ในสมการ (53) ด้วย f(y) ดังนี้

$$\hat{\sigma}_{E(S|x)} = c_k\sqrt{\sum_{y=0}^{k} [f(y)]^2 \Pr(y \mid \bar{x},k) - \left[\sum_{y=0}^{k} f(y) \Pr(y \mid \bar{x},k)\right]^2} \qquad \text{.....(54)}$$

ซึ่งสมการ (54) ก็คือความคลาดเคลื่อนมาตรฐานในการวัดในระดับบุคคลที่ถูกประมาณค่าเป็นความคลาดเคลื่อนแบบ Δ - type ของคะแนนแปลงรูป สมการ (54) นี้แตกต่างอย่างมากจากวิธีของเฟลด์-ควอลส์ และการคำนวณก็ค่อนข้างยุ่งยากแต่ก็สามารถใช้โปรแกรมคอมพิวเตอร์ช่วยในการคำนวณได้

จากสมการ (54) เป็นการประมาณค่า SEM ของคะแนนแปลงรูป จะให้ค่าที่ไม่กำกวมก็ต่อเมื่อการแปลงคะแนนดิบไปเป็นคะแนนมาตรฐานต้องเป็นการแปลงแบบหนึ่งต่อหนึ่ง (one-to-one)

สมมติว่าคะแนนดิบ r ถูกแปลงเป็นคะแนนแปลงรูป ; s* ฉะนั้น ความคลาดเคลื่อนมาตรฐานในการวัดของ s* จะได้เป็น

$$\hat{\sigma}_{E(s*)} = \sqrt{\text{ave}[\sigma^2_{E(S|x)} \text{ for the r values of x such that } f(x) = s*} \qquad \text{.....(55)}$$

เบรนนันและลี (Brennan & Lee. 1999) ไม่ได้แสดงถึงการแปลงคะแนนดิบหลายคะแนนเป็นคะแนนแปลงรูปค่าเดียว (many - to - one) ในสมการ (55) สามารถเขียนในสมการอีกรูปหนึ่งได้ดังนี้

$$\hat{\sigma}_{E(S|x)} = c_k \hat{\sigma}_{E^*(S|x)} \quad \text{.....(56)}$$

เมื่อ $\hat{\sigma}_{E^*(S|x)} = \sqrt{\sum_{y=0}^{k}[f(y)]^2 \Pr(y \mid \bar{x},k) - \left[\sum_{y=0}^{k} f(y)\Pr(y \mid \bar{x},k)\right]^2}$ .....(57)

วิธีทวินามประกอบ (Compound Binomial Procedure)

นิยามโมเดลแบบทวินามประกอบ

เป็นโมเดลทั่วไปของโมเดลแบบทวินาม ซึ่งข้อคำถามแต่ละข้อมีความยากไม่เท่ากัน ความคลาดเคลื่อนมาตรฐานจะแปรเปลี่ยนไปตามระดับของคะแนนจริงและไม่ขึ้นอยู่กับการแจกแจงคะแนนของกลุ่มตัวอย่างที่ทดสอบ

เฟลต์ (Feldt. 1984) ได้สรุปอ้างอิงในการหา SEM ตามวิธีของลอร์ด ในแบบทดสอบที่ประกอบด้วยชุดของกลุ่มเนื้อหาที่ได้กำหนดไว้หรือเป็นแบบชั้น นั่นคือ ในวิธีการประมาณค่า SEM ตามวิธีของเฟลต์ ค่า SEM จะถูกมองว่าเป็นความคลาดเคลื่อนที่มีการแจกแจงแบบทวินามภายใต้กลุ่มเนื้อหาแต่ละกลุ่มหรือแต่ละชั้น และความคลาดเคลื่อนนี้ไม่มีความสัมพันธ์กันในระหว่างกลุ่มเนื้อหา ซึ่งโคเลน และคณะ ได้นำมาใช้เป็นกรอบแนวคิดสำหรับการหาค่า SEM ของคะแนนแปลงรูปที่เป็นความคลาดเคลื่อนแบบสัมบูรณ์ (Abosolute Error) ของคะแนนในแต่ละบุคคล โมเดลที่ได้นำมาพิจารณาในที่นี้มีความสัมพันธ์กันมากกับตารางโมเดลแบบเฉพาะเจาะจงของจาร์จัวราและเบรนนัน (Jarjoura & Brennan. 1982 ; citing Brennan & Lee. 1999 : 10 - 13)

สมมติว่าแบบทดสอบฉบับหนึ่งได้กำหนดกลุ่มเนื้อหาเป็น H กลุ่มหรือจำนวนชั้น และให้

$k = \{k_1,\ldots,k_h,\ldots,k_H\}$ = จำนวนข้อในแต่ละชั้น

$x = \{x_1,\ldots,x_h,\ldots,x_H\}$ = จำนวนข้อที่ตอบถูกในแต่ละชั้นของผู้สอบคนหนึ่ง

$\bar{x} = \{\bar{x}_1,\ldots,\bar{x}_h,\ldots,\bar{x}_H\}$ = สัดส่วนที่สังเกตได้ของข้อที่ตอบถูกในแต่ละชั้นของผู้สอบคนหนึ่ง

$\pi = \{\pi_1,\ldots,\pi_h,\ldots,\pi_H\}$ = สัดส่วนจริงของข้อที่ตอบถูกในแต่ละชั้นของผู้สอบคนหนึ่ง

$w = \{w_1,\ldots,w_h,\ldots,w_H\}$ = ค่าน้ำหนักของแต่ละชั้น

$x_t = \sum_{h=1}^{H} w_h x_h$ = คะแนนทั้งหมดที่ถูกถ่วงน้ำหนักในทุกชั้นของผู้เข้าสอบ .....(58)

y_h = จำนวนข้อที่ตอบถูกในชั้นที่ h และ

y_t = คะแนนทั้งหมดที่ถูกถ่วงน้ำหนักในทุกชั้น

โดยทั่วไปแบบทดสอบที่สร้างจะสอดคล้องกับตารางกำหนดลักษณะเฉพาะ นั่นคือค่าน้ำหนัก w จะเท่ากับ 1 และ x_t เป็นจำนวนข้อที่ตอบถูกในทุกชั้นทั้งหมดของผู้สอบ อย่างไรก็ตามทฤษฎีที่

พัฒนาขึ้นมานี้สามารถกำหนดน้ำหนักคะแนนในแต่ละชุดได้ และจะมีประโยชน์ถ้าข้อคำถามในแต่ละกลุ่มมีการให้น้ำหนักคะแนนเป็นสองหรือสามเท่ากับข้อคำถามในกลุ่มอื่น

ถ้าเราแปลงคะแนนดิบที่ถูกถ่วงน้ำหนักทั้งหมดเป็นคะแนนแปลงรูป $f(x_t)$ เมื่อเปรียบเทียบกับสมการ (56) และ (57) ภายใต้โมเดลทวินามประกอบ เมื่อกำหนดจำนวนชั้นเป็น H ดังนั้นการประมาณค่า SEM ของ คะแนนแปลงรูปในแต่ละคน ณ คะแนน x ในชั้น คือ

$$\hat{\sigma}_{E(S|x)} = c_k \hat{\sigma}_{E^*(S|x)} \qquad \text{.....(59)}$$

เมื่อ

$$\hat{\sigma}_{E^*(S|x)} = \sqrt{\sum_{y_t=0}^{\max(y_t)} [f(y_t)]^2 \Pr(y_t \mid \bar{x},k) - \left[\sum_{y_t=0}^{\max(y_t)} f(y_t)\Pr(y_t \mid \bar{x},k)\right]^2} \qquad \text{.....(60)}$$

การประมาณค่า SEM จากสมการ (59) และ (60) สำหรับบุคคลคนหนึ่งที่มีรูปแบบการตอบเป็นจำนวนข้อที่ตอบถูก x (หรือ $\bar{x}$) ในคะแนนแต่ละชั้นที่สังเกตได้ โดยทั่วไปค่า SEM นี้ไม่ใช่ค่าเฉลี่ยของ SEM ของผู้เข้าสอบทั้งหมดที่มีคะแนน X_t เดียวกัน

สมการ (59) และ (60) เป็นการนำเสนอการประมาณค่า 2 ค่าที่ต้องกำหนด c_k และคำนวณหา $\Pr(y_t \mid \bar{x},k)$ ก่อนที่อยู่ภายในแต่ละชั้นของการแจกแจงจำนวนข้อที่ตอบถูกที่มีการแจกแจงแบบทวินามนั่นเอง ในการคำนวณใช้ $\bar{x}_h$ แทนการประมาณค่าของ π_t ดังนี้

$$\Pr(y_h \mid \bar{x}_h,k_h) = \binom{k_h}{y_h} \bar{x}_h^{y_h}(1-\bar{x}_h)^{k_h-y_h}$$

ภายใต้ข้อตกลงที่ว่าความคลาดเคลื่อนระหว่างชั้นของเนื้อหาไม่มีความสัมพันธ์กัน จะได้

$$\Pr(y_1,\ldots y_H \mid \bar{x},k_h) = \prod_{h=1}^{H} \Pr(y_h \mid \bar{x}_h,k_h) \qquad \text{.....(61)}$$

เพราะฉะนั้น

$$\Pr(y_t \mid \bar{x},k) = \sum \Pr(y_1,\ldots,y_H \mid \bar{x},k) \qquad \text{.....(62)}$$

เมื่อหาผลรวมของ $y_1,\ldots,y_H$ ทั้งหมดจะทำให้ได้

$$\sum_{h=1}^{H} w_h y_h = y_t$$

ตัวอย่าง สมมติให้ k = {2,3,4} , w = {1,1,1} และคะแนนที่สังเกตในแต่ละชั้นของผู้เข้าสอบคนหนึ่งเป็น x = {1,1,0} ดังนั้นคะแนนเฉลี่ยแต่ละชั้นของคะแนนที่สังเกตได้เป็น $\bar{x}$ = {1/2, 1/3, 0} และ $y_t = 2$ จะทำให้มีรูปแบบของคะแนน y_h เมื่อ $y_t = 2$ ทั้งหมด 6 รูปแบบ ดังนี้ {2,0,0}, {0,2,0}, {0,0,2}, {1,1,0}, {1,0,1} และ {0,1,1} ฉะนั้นเมื่อจะหาโอกาสที่จะเกิดในแต่ละรูปแบบนี้ใช้สมการ (61) และถ้ารวมผลรวมของโอกาสที่จะเกิดทั้ง 6 รูปแบบก็จะมีค่าเท่ากับผลที่ได้ในสมการ (62)

เพื่อที่จะหา c_k ในสมการ (59) ให้นึกถึงวิธีการของเฟลด์ ที่ได้เสนอไว้ภายใต้โมเดลแบบทวินามประกอบ ที่มีค่าน้ำหนักเป็น 1 ฉะนั้นค่า SEM สำหรับคะแนนดิบของผู้เข้าสอบแต่ละคนก็คือ

$$\sqrt{\sum_{h=1}^{H} \frac{x_h(k_h - x_h)}{k_h - 1}}$$

ในการหาผลรวมค่าน้ำหนัก w ค่า SEM ของผู้สอบแต่ละคนที่รูปแบบหนึ่งของคะแนนที่สังเกตได้ x ตามวิธีของเฟลด์ จะได้เป็น

$$\hat{\sigma}_{E(x)} = \sqrt{\sum_{h=1}^{H} w_h^2 \frac{x_h(k_h - x_h)}{k_h - 1}} \qquad \text{.....(63)}$$

เพราะฉะนั้น c_k จะได้เป็น

$$c_k = \hat{\sigma}_{E(x)} / \hat{\sigma}_{E^*(x)} \qquad \text{.....(64)}$$

เมื่อ $\hat{\sigma}_{E^*(x)}$ ได้จากการแทนค่า $f(y_t)$ ในสมการ (60) ด้วย y_t สมการ (59) และ (63) จะให้ค่าเหมือนกัน เมื่อ $f(X_t) = X_t$ นั่นคือ เมื่อคะแนนแปลงรูปเป็นคะแนนดิบที่ถูกถ่วงน้ำหนักทั้งหมด

สรุป การประมาณค่า SEM ของคะแนนแปลงรูปแบบทวินามประกอบ ในผู้เข้าสอบแต่ละคน ณ ที่คะแนนที่สังเกต x ใช้สมการ (59), (60) และ (64) ซึ่งเป็นการคำนวณที่ละเอียดมากจำเป็นที่จะต้องมีโปรแกรมคอมพิวเตอร์ช่วยในการคำนวณ

ในการคำนวณ กำหนดให้ผู้เข้าสอบทั้งหมดมีคะแนนดิบทั้งหมดที่ถูกถ่วงน้ำหนัก (x_t) เหมือนกัน มีคะแนนแปลงเดียวกัน ดังนั้นภายใต้เงื่อนไขนี้ ตัวประมาณค่าที่ดีของ SEM ในคะแนนแปลงรูป ณ ที่ระดับคะแนน x_t เฉพาะ จะเป็น

$$\hat{\sigma}_{E(S|x_t)} = \sqrt{\text{ave}[\hat{\sigma}^2_{E(S|x)} \text{ such that } \sum_{h=1}^{H} w_h x_h = x_t]} \qquad \text{.....(65)}$$

สมการนี้ก็คือค่าเฉลี่ย SEM ของคะแนนแปลงรูปในคนทั้งหมดที่มีคะแนนดิบที่ถูกถ่วงน้ำหนักเดียวกัน ในสมการ (65) และสมการ (55) จะต่างกัน ในสมการ (65) ให้ค่าเฉลี่ย SEM ของผู้เข้าสอบที่มีคะแนนดิบ x_t ทั้งหมดเดียวกัน ส่วนสมการ (55) ให้ค่าเฉลี่ย SEM ณ ที่ระดับคะแนนแปลงรูปเฉพาะค่าหนึ่ง ซึ่งอาจมาจากคะแนนดิบหลายคะแนนก็ได้

จากวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปที่กล่าวไปแล้วนั้น จะเห็นได้ว่าแต่ละวิธีต่างก็มีข้อตกลงเบื้องต้นในการใช้ต่างกัน นั่นคือ วิธีที่ใช้ทฤษฎีการตอบข้อสอบ (วิธี IRT) เป็นวิธีที่เน้นเกี่ยวกับเรื่องการเป็นมิติเดียวของแบบทดสอบ วิธีทวินามและวิธีทวินามประกอบเป็นวิธีที่ใช้การแจกแจงแบบทวินามและทวินามประกอบ ส่วนวิธีของเฟลด์-ควอลส์ เป็นวิธีที่ง่ายไม่ขึ้นอยู่กับโมเดลทางจิตภาค ในการที่จะเลือกใช้วิธีใดเพื่อประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป นั้นนอกจากจะต้องคำนึงถึงข้อตกลงเบื้องต้นของการประมาณค่าแต่ละวิธีแล้วยังต้องคำนึงถึงประเภทของคะแนที่แปลง อีกทั้งคุณลักษณะของแบบทดสอบที่ใช้อีกด้วย

ตัวประมาณที่ดี

ในการวิจัยครั้งนี้นอกจากต้องการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีประมาณค่าต่างกันแล้ว ยังต้องการศึกษาอีกว่าถ้าสุ่มตัวอย่างจำลองจากประชากรเทียมที่ผู้วิจัยได้กำหนดไว้ มาขนาดละ 300 500 และ 800 คน แต่ละขนาดสุ่มซ้ำ 30 ครั้ง แล้ว ค่าความคลาดเคลื่อนมาตรฐานในการวัดจะมีค่าแตกต่างไปอย่างไรจากประชากร ในทฤษฎีการสุ่มตัวอย่างได้กล่าวถึงตัวประมาณที่ดีควรมีคุณสมบัติ ดังนี้ (นิยม ปุราคำ. 2517: 59 - 70 ; สุรินทร์ นิยมางกูร. 2542 : 42)

1. ไม่มีความลำเอียง (Unbiasedness) หมายความว่า ถ้าหาก $\hat{\theta}$ เป็นตัวประมาณของพารามิเตอร์ θ แล้ว $\hat{\theta}$ จะถูกเรียกว่าเป็นตัวประมาณที่ไม่มีความลำเอียง ในการประมาณ θ ถ้าหากว่า $E(\hat{\theta}) = \theta$ เมื่อ $E(\hat{\theta})$ คือค่าคาดหมายของ $\hat{\theta}$

2. มีความเป็น Consistency หมายความว่า ถ้า $\hat{\theta}_n$ เป็นตัวประมาณของพารามิเตอร์ θ ที่ขึ้นอยู่กับขนาดของตัวอย่าง (n) แล้ว $\hat{\theta}_n$ จะถูกเรียกว่าเป็น Consistent Estimator ในการประมาณ θ ถ้าหากว่าเมื่อ n มีค่ามาก แล้ว $\lim_{n \to \infty} \hat{\theta}_n = \theta$ จะมีค่าใกล้เคียงกับ $\hat{\theta}_n$ นั่นคือ

3. มีประสิทธิภาพสูง (High Efficiency) หมายความว่า ถ้า $\hat{\theta}_1$ และ $\hat{\theta}_2$ เป็นตัวประมาณที่ไม่ลำเอียงของพารามิเตอร์ θ ด้วยกันทั้งคู่ และ $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ แล้ว $\hat{\theta}_1$ จะถูกเรียกว่าเป็นตัวประมาณที่มีประสิทธิภาพสูงกว่า $\hat{\theta}_2$ และถ้า $V(\hat{\theta}_1)$ มีค่าต่ำที่สุด $\hat{\theta}_1$ จะถูกเรียกว่าเป็นตัวประมาณที่มีประสิทธิภาพสูงสุด บางครั้งตัวประมาณที่ไม่ลำเอียงและมีความแปรปรวนต่ำจะเรียกว่าเป็นตัวประมาณที่ดีที่สุด (Best Estimator)

4. คุณสมบัติอื่น เช่น สามารถคำนวณได้ง่าย สามารถจะหาค่าประมาณของความแปรปรวนได้ง่าย เป็นต้น

ในการวิจัยครั้งนี้ ผู้วิจัยได้ใช้คุณสมบัติของตัวประมาณ ได้แก่ ความไม่ลำเอียง และความแปรปรวนคลาดเคลื่อน เพื่อแสดงถึงความแม่นยำในการประมาณค่า

งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับความคลาดเคลื่อนมาตรฐานในการวัด

มีผู้ศึกษาเกี่ยวกับความคลาดเคลื่อนมาตรฐานในการวัด ดังนี้

เรวดี อินทสะระ (2530) ได้เปรียบเทียบความคลาดเคลื่อนและระดับความเพียงพอของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ โดยใช้กลุ่มทานผล ผลการศึกษาพบว่า ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราของรูปแบบอิงทฤษฎีการตอบข้อสอบมีค่าน้อยกว่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรารูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ ซึ่งมีค่าเท่ากับ 0.0896 และ 0.1505 ตามลำดับ เมื่อทดสอบค่าเฉลี่ยของคะแนนความคลาดเคลื่อนของการเทียบมาตราทั้งสองรูปแบบพบว่าคะแนน

ความคลาดเคลื่อนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นอกจากนี้แล้วยังมีงานวิจัยของสุมนา โสตถิผลอนันต์ (2539) ที่ได้เปรียบเทียบความคลาดเคลื่อนมาตรฐานที่เกิดจากการใช้วิธีการเทียบคะแนนแบบเส้นตรงตามแบบคะแนนจริงสัมพันธ์สี่วิธี ให้ผลแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

วลีมาศ แซ่อึ้ง (2537) ได้ศึกษาความคลาดเคลื่อนมาตรฐานและค่าประสิทธิภาพสัมพัทธ์ของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ ระหว่างมาตราวัดแบบลิเคอร์ทกับมาตราวัดแบบออสกูด ผลการวิจัยพบว่า ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ระหว่างมาตราวัดแบบลิเคอร์ทกับมาตราวัดแบบออสกูด มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ไอลดา คล้ายสำริด (2538) ได้เปรียบเทียบความคลาดเคลื่อนในการประมาณค่าความสามารถของผู้สอบในการทำแบบทดสอบวิชาคณิตศาสตร์ที่ใช้วิธีต่างกัน 4 วิธี ได้แก่ วิธีประมาณค่าตามค่าความยากของข้อที่ยากที่สุดที่ทำถูก วิธีประมาณค่าตามค่าเฉลี่ยความยากของข้อที่ทำถูก วิธีประมาณค่าโดยใช้วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข และวิธีประมาณค่าตามวิธีพร็อกซ์ ผลการวิจัยพบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถระหว่างวิธีประมาณค่าตามค่าความยากของข้อที่ยากที่สุดที่ทำยากกับวิธีประมาณค่าตามค่าเฉลี่ยความยากของข้อที่ทำถูก วิธีประมาณค่าโดยใช้วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข วิธีประมาณค่าตามวิธีพร็อกซ์ และระหว่างวิธีประมาณค่าตามค่าเฉลี่ยความยากของข้อที่ทำถูก กับวิธีประมาณค่าโดยใช้วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไข วิธีประมาณค่าตามวิธีพร็อกซ์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนวิธีประมาณค่าโดยใช้วิธีความเป็นไปได้สูงสุดแบบไม่มีเงื่อนไขกับวิธีประมาณค่าตามวิธีพร็อกซ์ มีค่าไม่แตกต่างกัน

รุ่งฤดี ทรัพย์นิธิ (2539) ได้เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติกแบบหนึ่ง สอง สามพารามิเตอร์ เมื่อใช้ขนาดกลุ่มตัวอย่างเป็น 250 คน 500 คน 750 คน และ 1,000 คน โดยใช้แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ จำนวน 20 ข้อ ผลการวิจัยพบว่า

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วยโมเดลหนึ่งพารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดลสองพารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกมีค่าแตกต่างกันในบางกลุ่มตัวอย่าง เมื่อประมาณค่าด้วยโมเดลสามพารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและการเดา มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วยโมเดลหนึ่งพารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดลสองและสามพารามิเตอร์ ปรากฏผลเช่นเดียวกันคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบมีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดลเมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันทุกขนาดกลุ่มตัวอย่าง ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากมีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดลเมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน มีค่าแตกต่างกัน จากทุกขนาดกลุ่มตัวอย่าง โดยโมเดลสามพารามิเตอร์ แตกต่างจากโมเดลหนึ่งและสองพารามิเตอร์

ผลจากการศึกษาพบว่าจำนวนกลุ่มตัวอย่างที่ให้ค่าประมาณพารามิเตอร์ที่ดีที่สุด คือ กลุ่มตัวอย่างจำนวน 750 คน

นอกจากนี้แล้วยังมีงานวิจัยที่เกี่ยวกับการประมาณค่าความเชื่อมั่นของมาตรวัดและศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัด มีดังนี้

ฉวีวรรณ ครองยุทธ (2539) ได้ประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่าที่มีขนาดกลุ่มตัวอย่างขนาด 40 80 160 250 และ 500 คน ที่คำนวณด้วยสูตรต่างกัน พบว่า สูตร α_λ และสูตร Ω เหมาะสมกับกลุ่มตัวอย่างขนาด 500 คน และสูตร Ω_w เหมาะสมกับกลุ่มตัวอย่างขนาด 80 คน โดยพิจารณาจากค่าความคลาดเคลื่อนมาตรฐานและค่าความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นที่ต่ำเมื่อเทียบกับกลุ่มอื่น

เกศเกล้า หมู่มิ่ง (2539) ได้ประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่าด้านมโนภาพแห่งตน โดยใช้สูตร α , r_{L1} , r_{L2} ,α_λ , Ω และ Ω_w พบว่า ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นด้วยสูตร α มีค่าต่ำสุด และสูตร r_{L1} มีค่าสูงสุด และความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นที่คำนวณด้วยสูตร r_{L1} มีค่าต่ำสุดและ สูตร α มีค่าสูงสุด

พนิดา ฟุ้งเกียรตินำสุข (2540) ได้หาค่าความคลาเคลื่อนมาตรฐานในการประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่า 3 รูปแบบได้แก่ แบบตัวเลข แบบกราฟิก และแบบบรรยาย ที่คำนวณด้วยสูตร r_{L2} และ Ω_w มีค่าตั้งแต่ .0456 ถึง .0617 และค่าความลำเอียงทางสถิติในการประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่าทั้ง 3 รูปแบบ ที่คำนวณด้วยสูตร r_{L2} และ Ω_w มีค่าตั้งแต่ - .0023 ถึง .0426

งานวิจัยที่เกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ

มีผู้ศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบ ไว้ดังนี้

นิภาพร ยอดเมือง (2539) ได้ศึกษาความคลาดเคลื่อนและความลำเอียงทางสถิติในการเทียบคะแนนเชิงเส้นตรงแบบคะแนนจริงสัมพันธ์ในกลุ่มตัวอย่างขนาด 30 60 120 และ 250 คน กับนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2538 จำนวน 1,826 คน โดยใช้แบบทดสอบการรับรู้ 3 ฉบับ เป็นแบบทดสอบเทียบคะแนน 2 ฉบับ และแบบทดสอบเชื่อมโยง 1 ฉบับ ผลการวิจัยพบว่า คะแนนเทียบเฉลี่ยของกลุ่มตัวอย่างทั้ง 4 ขนาด ไม่แตกต่างกัน ส่วนค่าความลำเอียงของสถิติในการ

เทียบคะแนนแตกต่างกันน้อยมาก และมีค่าเท่ากับ 0 เมื่อใช้กลุ่มตัวอย่างขนาด 60 และ 250 คน เมื่อคะแนนสังเกตเข้าใกล้คะแนนเฉลี่ย สำหรับความคลาดเคลื่อนมาตรฐานของการเทียบคะแนนในกลุ่มตัวอย่างทั้ง 4 ขนาด ก็แตกต่างกันเล็กน้อยเมื่อคะแนนสังเกตใกล้ค่าเฉลี่ย (38.8204)

อำนวยพร คำพันธ์ (2542) ได้ศึกษาผลของการจัดเรียงข้อสอบที่มีต่อค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะ โดยใช้วิธีการประมาณค่าสามวิธี ผลการวิจัยพบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะของแบบทดสอบที่จัดเรียงจากง่ายไปหายากมีค่าต่ำกว่าแบบทดสอบที่จัดเรียงแบบสุ่ม และค่าความคลาดเคลื่อนมาตรฐานในการวัดที่คำนวณโดยวิธีสามวิธีมีค่าแตกต่างกัน วิธีของลอร์ดและคีตส์ได้ค่าที่ใกล้เคียงสอดคล้องกันเป็นรูปพาราโบลาคว่ำ วิธีของธอร์นไดค์ ค่าที่ได้จะมีลักษณะสูงต่ำสลับกันในบางระดับคะแนน แต่ยังมีแนวโน้มเป็นรูปพาราโบลาคว่ำเช่นกัน และวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะตามวิธีของคีตส์จะได้ค่าที่ต่ำที่สุดอยู่ในช่วงปลายของระดับคะแนนทั้งสองฉบับ และค่าที่สูงที่สุดคือวิธีของธอร์นไดค์จากฉบับที่เรียงจากง่ายไปหายาก และวิธีของลอร์ดจากฉบับที่เรียงแบบสุ่มซึ่งอยู่ในช่วงกลางของระดับคะแนน

วิชุดา กิจธรรธรรม. (2544) ได้ศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนน ด้วยวิธีการสามวิธี คือ วิธีการวิเคราะห์ความแปรปรวน วิธีการตามทฤษฎีการตอบข้อสอบ และวิธีการของคีตส์ และเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนที่คำนวณด้วยวิธีการทั้งสามวิธีโดยการวิเคราะห์เส้นภาพ โดยศึกษาจากแบบทดสอบวัดความถนัดทางวิชาชีพพยาบาลของฟลานาแกนจำนวนห้าฉบับ ผลการวิจัยสรุปได้ว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนที่คำนวณด้วยวิธีการทั้งสามวิธี มีค่าไม่คงที่โดยเฉพาะอย่างยิ่งในช่วงปลายของคะแนน คือมีค่าต่ำในช่วงปลายคะแนนต่ำและช่วงปลายคะแนนสูง และสูงสุดในช่วงคะแนนตรงกลาง เมื่อเปรียบเทียบด้วยการวิเคราะห์เส้นภาพ พบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนที่คำนวณด้วยวิธีการของคีตส์มีค่าค่อนข้างคงที่และเมื่อเปรียบเทียบการคำนวณทั้งสามวิธีในแบบทดสอบแต่ละฉบับ พบว่าแบบทดสอบฉบับที่ห้า มีค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนต่ำสุด และเส้นภาพมีความราบเรียบ (Smooth) มากกว่าการคำนวณด้วยวิธีการอื่น ๆ

ลอร์ด (Lord. 1984 : 239 - 242) ได้ศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับความสามารถแตกต่างกัน จากการดำเนินการสอบด้วยแบบทดสอบฉบับเดียว และคำนวณค่าความคลาดเคลื่อนมาตรฐานในการวัดจากคะแนนที่เป็นจำนวนข้อในการตอบถูกโดยใช้วิธีการ 4 วิธี ได้แก่ วิธีที่ 1 (วิธีของลอร์ด) วิธีที่ 2 (ราสช์โมเดล) วิธีที่ 3 (ลอร์ดและโนวิค) และวิธีที่ 4 (วิธีของลอร์ดแบบทวินามประกอบ) จากแบบทดสอบ SCAT แบบอุปมาอุปไมยภาษา (School and College Ability Test) จำนวน 50 ข้อแบบ 4 ตัวเลือกที่ให้คะแนนตามจำนวนข้อที่ตอบถูก และเพื่อเปรียบเทียบวิธีการประมาณค่าทั้ง 4 ชุดของข้อมูลจริงที่ได้จากการตอบของกลุ่มตัวอย่างที่เป็นนักเรียนระดับชั้นมัธยมศึกษา จำนวน 3,000 คน ผลการศึกษาพบว่า วิธีที่ 3 จะให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงกว่าวิธีที่ 1 และวิธีที่ 2 สำหรับวิธีที่ 4 ให้ผลการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดสอดคล้องกับวิธีที่ 1 แต่เมื่อคะแนนที่ตอบถูกไม่เกิน 20 คะแนนแล้วพบว่าความคลาดเคลื่อนมาตรฐานที่ถูกประมาณค่าโดยวิธี 1 จะมีค่าน้อยกว่าที่ได้จากการประมาณค่าวิธีที่ 2 และวิธีที่ 4 ซึ่งอาจมี

สาเหตุมาจากผู้สอบที่ได้คะแนนต่ำ ไม่สามารถตอบข้อสอบได้หมดทุกข้ออาจจะด้วยเหตุผลที่ว่าไม่อาจเดาได้หรือไม่มีเวลาพอที่จะตอบได้ทัน ผู้สอบที่ตอบข้อคำถามได้น้อยกว่าจำนวนข้อทั้งหมดจะมีความคลาดเคลื่อนมาตรฐานในการวัดต่ำกว่าผู้สอบที่ตอบข้อคำถามได้ครบทุกข้อ ในวิธีที่ 2 จะให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนจากจำนวนข้อที่ตอบถูกที่ผู้สอบตอบข้อคำถามทั้งหมด วิธีที่ 3 และวิธีที่ 4 ก็เหมือนกันยกเว้นความสามารถจริงของผู้สอบถูกประมาณว่าต่ำกว่าความเป็นจริงถ้าผู้สอบไม่สามารถตอบข้อสอบทั้งหมดได้ วิธีที่ 1 และวิธีที่ 4 สามารถใช้กับแบบทดสอบที่ไม่เป็นมิติเดียว (Unidimension) ได้

เฟลด์ สเตฟเฟน และกุปตา (Feldt, Steffen & Gupta. 1985 : 351 - 361) ได้ศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด 5 วิธี ได้แก่ วิธี IRT วิธีทวินาม วิธีของคีตส์ วิธีทวินามประกอบ และวิธี ANOVA โดยใช้แบบทดสอบ 3 ฉบับ คือด้านคำศัพท์ (V) ด้านการอ่านวรรณคดี (L) และการใช้แหล่งสารสนเทศ (S_I) ทดสอบกับนักเรียนเกรด 9 และเกรด 11 แบบทดสอบที่ใช้สอบกับนักเรียนทั้งสองระดับนี้ใช้ชุดข้อสอบไม่เหมือนกัน ผลการวิเคราะห์ปรากฏว่าค่าความคลาดเคลื่อนมาตรฐานที่ได้จากการประมาณค่าทั้ง 5 วิธี มีค่าสูง ณ คะแนนที่อยู่ตรงกลาง มีค่าน้อยลงตรงคะแนนที่อยู่สุดโต่ง

บลิซ์ต และชามา (Blixt and Shama. 1986 : 545 - 550) ได้เปรียบเทียบวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานที่ระดับความสามารถแตกต่างกัน โดยใช้วิธีประมาณค่า 4 วิธีที่เสนอโดยลอร์ด ศึกษาจากกลุ่มตัวอย่างที่เป็นนักศึกษาปี 1 จำนวน 1,824 คน โดยการทำแบบทดสอบวิชาพีชคณิตจำนวน 35 ข้อ เป็นแบบปรนัย 4 ตัวเลือก มีค่าเฉลี่ยเท่ากับ 16.41 ค่าเบี่ยงเบนมาตรฐาน 7.45 และค่าความเชื่อมั่นของแบบทดสอบหาโดยวิธี KR. 20 ได้เท่ากับ .88 และค่าความคลาดเคลื่อนมาตรฐาน 2.54 ผลจากการศึกษาพบว่า วิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ ระดับคะแนนความสามารถหนึ่งมีค่าไม่แตกต่างกัน

ควอลส์ - เพนน์ (Qualls - Payne. 1992 : 213 - 225) ได้เปรียบเทียบวิธีการประมาณค่าความคลาดเคลื่อมาตรฐานในการวัดของระดับคะแนน โดยใช้วิธีประมาณค่า 6 วิธี คือ วิธีของ เฟลด์ คีตส์ จาร์รัวรา สามพารามิเตอร์ โมเลนคอปฟ์ และ ธอร์นไดค์ จากแบบทดสอบย่อยสองฉบับคือแบบวัดความถนัดด้านภาษา และคณิตศาสตร์ ของ ITBS (Iowa Tests of Basic Skills) ฟอร์ม G และฟอร์ม H ทำการสอบกับนักเรียนเกรด 3 5 และ 8 จำนวน 2,138 คน ผู้เข้าสอบแต่ละคนถูกกำหนดให้ทำแบบทดสอบซ้ำหรือแบบทดสอบคู่ขนาน ผลจากการศึกษาปรากฏว่า ความสัมพันธ์ระหว่างความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะและ คะแนนแปลงรูป ในทั้ง 6 วิธีมีค่าคล้ายคลึงกัน ค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบแบบคู่ขนานมีค่าสูงกว่าแบบทดสอบชนิดซ้ำ แต่ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดที่ได้จากการสอบครั้งเดียวในทุกวิธีมีค่าน้อยกว่าค่าที่ประมาณได้จากการสอบคู่ขนาน และวิธีของเฟลด์ ให้ค่าสูงกว่าวิธีอื่น และค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ได้จากการประมาณค่าทั้ง 6 วิธี ได้ค่าใกล้เคียงกัน แต่วิธีของเฟลด์ จะได้ค่าสูงกว่าวิธีอื่น

เฟลด์ และควอลส์ (Feldt & Qualls.1996 : 141 - 156) ได้ปรับปรุงวิธีการประมาณค่าความแปรปรวนคลาดเคลื่อนในการวัด ณ ระดับคะแนนเฉพาะหรือ ณ ที่แต่ละคะแนนในการทำแบบ

ทดสอบ จากวิธีของธอร์นไดค์ โดยใช้กับการทดสอบแบบคู่ขนาน คะแนนที่ได้จากการทดสอบมีทั้งแบบที่ใช้การจำลองสถานการณ์ (Simulated Test) และแบบทดสอบจริง (Actual Test) แบบทดสอบที่จำลองขึ้นมี 36 ข้อ มีความยากตั้งแต่ .40 ถึง .60 มีค่าความเชื่อมั่น .785 ให้คะแนนแบบ 0 - 1 คะแนนที่ได้แต่ละข้อและคะแนนทั้งหมดได้จากการจำลองสถานการณ์การตอบของผู้สอบ 1,000 คน คะแนนแบบทดสอบนี้นำมาจัดกลุ่มเป็น 3 แบบคือแบบที่มี 6 ส่วน 3 ส่วน และ 2 ส่วน ภายในแต่ละส่วนคู่ขนานกัน สำหรับแบบทดสอบฉบับจริง ข้อมูลได้จากการทำแบบทดสอบย่อยด้านคำศัพท์และความคิดรวบยอดทางคณิตศาสตร์ ของนักเรียนเกรด 4 และเกรด 7 แบบทดสอบแต่ละฉบับแบ่งออกเป็นส่วน ๆ แต่ละส่วนเป็นแบบแบบทดสอบคู่ขนานกัน

ผลของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบที่จำลองขึ้นมีค่าเฉลี่ยของความแปรปรวนคลาดเคลื่อนของแบบทดสอบที่แบ่งเป็น 6 ส่วน 3 ส่วน และ 2 ส่วน คือ 8.42, 8.16 และ 7.92 ตามลำดับ แบบทดสอบ V - 4 ที่ได้จากการสอบจริงมีค่าเฉลี่ยของความแปรปรวนคลาดเคลื่อนของแบบที่แบ่งเป็น 2 ส่วน 4 ส่วน และ 6 ส่วน คือ 6.03, 6.12 และ 6.08 ตามลำดับ สำหรับแบบทดสอบฉบับอื่น ก็ให้ค่าความแปรปรวนคลาดเคลื่อนต่ำเหมือนกัน และเห็นได้ชัดว่า ณ ที่ระดับคะแนนตรงช่วงกลาง ความแปรปรวนคลาดเคลื่อนมีค่าสูง

นอกจากงานวิจัยที่ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบที่ได้จากแบบทดสอบที่ประกอบด้วยข้อคำถามโดดแล้วยังมีงานวิจัยของ ลี (Lee. 2000 : 91 - 112 ; 161 - 180) ที่ใช้แบบทดสอบที่ประกอบด้วยข้อคำถามเป็นชุด (Testlets) แต่ละชุดคำถามเป็นสถานการณ์ ได้แก่ บทความ โคลง กลอน หรือแผนที่ที่กำหนดมาให้และนำไปใช้ตอบคำถามได้หลายข้อ ซึ่งมีลักษณะเป็นการฝ่าฝืนข้อตกลงในเรื่องของความเป็นมิติเดียว (Unidimensionality) และความเป็นอิสระของข้อคำถาม (Local Independence) เพื่อประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยใช้วิธีการประมาณค่า 5 วิธี ได้แก่ วิธี GIRT (Graded Response model) NIRT (Nominal Model) DIRT (Dichotomous Three - Parameter Logistic Model) และวิธีที่ใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) 2 รูปแบบคือ p x I หรือ p x (I : H) ข้อมูลที่ใช้เป็นข้อมูลจริงที่ได้จากการสอบ ITBS (Iowa Tests of Basic Skills) ฟอร์ม M ถึงฟอร์ม K ของนักเรียนเกรด 4 เกรด 7 และเกรด 8 ผลการศึกษาพบว่า ในแบบทดสอบเรื่องการอ่าน วิธี DIRT ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข (SEM) ต่ำสุดเมื่อเปรียบเทียบกับวิธีอื่น วิธี p x (I : H) ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสูงสุดตรงพิสัยคะแนนกลาง แต่วิธี IRT แบบพหุคะแนนให้ค่าที่ประมาณได้สูงกว่าตรงพิสัยของคะแนนที่สูงกว่า วิธี p x (I : H) ให้รูปร่างของค่าที่ประมาณเป็นโค้งที่ไม่สม่ำเสมอ แต่วิธีอื่นให้รูปร่างของโค้งที่ราบเรียบกว่า โดยทั่วไปวิธีที่ใช้ข้อคำถามโดดให้ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขต่ำกว่าวิธีที่ใช้ข้อสอบเป็นชุด แบบทดสอบเกี่ยวกับแผนที่ ผลการประมาณค่าที่ได้จะคล้ายกับที่พบในแบบทดสอบเกี่ยวกับการอ่าน แต่แบบทดสอบนี้ที่ใช้กับนักเรียนเกรด 4 วิธี p x I ให้ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขไม่แตกต่างจากวิธี GIRT และ วิธี NIRT ตรงพิสัยคะแนนช่วงกลาง และเมื่อใช้วิธีที่เน้นข้อคำถามโดดกับที่ใช้กับข้อคำถามเป็นชุด มีค่าความคลาดเคลื่อนมาตรฐานในการวัดแตกต่างกันน้อยมาก สำหรับแบบทดสอบเกี่ยวกับคำศัพท์ และข้อมูลการตอบที่ได้จากการจำลอง

ข้อมูล วิธี p x I ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสูงสุดตรงพิสัยของคะแนนตรงกลาง และวิธี NIRT และ GIRT ให้ค่าประมาณได้สูงกว่าตรงพิสัยของคะแนนที่ต่ำกว่าและสูงกว่า

นอกจากนี้แล้ว ลี (Lee) ยังได้ทำวิจัยเพื่อเปรียบเทียบวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบที่ประกอบด้วยข้อคำถามเป็นชุดโดยใช้วิธีการจำลองข้อมูลที่เป็นคะแนนการตอบแบบทดสอบความเข้าใจในการอ่านบทความ ผลการศึกษาพบว่า วิธีการประมาณค่าที่ของแบบทดสอบที่เน้นคำถามโดด วิธี p x I และ 3 PL ให้ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขต่ำกว่าของแบบทดสอบที่ประกอบด้วยข้อคำถามเป็นชุด วิธี p x (I : H) ให้ค่าลำเอียง (Bias) น้อยที่สุดแต่มีความคลาดเคลื่อนสุ่มของการประมาณมากที่สุด วิธี GRADED และ NOMINAL ให้การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขใกล้เคียงกันและให้ค่าที่ใกล้เคียงกับค่าที่แท้จริงตรงพิสัยคะแนนตรงกลาง แต่ให้ค่าที่ประมาณสูงกว่าความเป็นจริงตรงพิสัยคะแนนที่ต่ำกว่าและสูงกว่า ความลำเอียงของวิธีการประมาณค่าที่ใช้กับแบบทดสอบที่มีข้อคำถามโดดมีค่ามากขึ้นเมื่อระดับของความเกี่ยวข้องกันอย่างมีเงื่อนไขมีค่าสูงขึ้น

งานวิจัยที่เกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป

การศึกษาเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป มีผู้ศึกษาไว้ ดังนี้

โคเลน (Kolen. 1988 : 97 - 110) ได้ศึกษาการแปลงคะแนนดิบเป็นคะแนนมาตรฐานทั้งที่เป็นการแปลงแบบเส้นตรง (Linear Methods) และแบบไม่เป็นเส้นตรง (Nonlinear Methods) ซึ่งการประยุกต์ใช้วิธีแบบไม่เป็นเชิงเส้นตรงโดยใช้การแปลงแบบอาร์คซายน์ (Arcsine) เพื่อทำให้ความแปรปรวนคลาดเคลื่อนมีค่าเกือบเท่ากันในทุกคะแนนแปลงรูป ผลการศึกษาจากการแปลงคะแนนดิบของแบบทดสอบที่มี 100 ข้อ ไปเป็นคะแนนแปลงรูปที่ไม่ปัดเป็นจำนวนเต็ม โดยวิธี Constant Error ที่หาความเชื่อมั่นโดย KR-20 และ KR-21 และวิธี Variable Error แบบทวินามและแบบทวินามประกอบ ปรากฏว่าคะแนนแปลงรูปที่ได้มีค่าใกล้เคียงกันในช่วงคะแนนตรงกลาง วิธี Constant และวิธี Variable error ค่าคะแนนแปลงรูปต่างกันมากตรงคะแนนที่สูง และคะแนนที่ต่ำ วิธี Constant ที่หาความเชื่อมั่นโดย KR-20 ให้พิสัยของ คะแนนแปลงรูป กว้างกว่าวิธี Constant ที่หาความเชื่อมั่นโดย KR-21 และวิธี Variable แบบทวินามประกอบก็ให้พิสัยของคะแนนแปลงรูปมากกว่าแบบทวินามสำหรับแบบทดสอบที่มี 25 ข้อ มีลักษณะเป็นแบบทดสอบที่ยากและแบบทดสอบที่ง่าย พบว่าการแปลงโดยวิธี Binomial-Based Variable Error ของแบบทดสอบทั้งสองลักษณะนั้นให้ค่าเหมือนกัน ส่วนวิธีทวินามประกอบให้ค่าของคะแนนแปลงรูป ในแบบทดสอบทั้งสองลักษณะนั้นเหมือนกัน ยกเว้นคะแนนคะแนนแปลงรูปที่ตรงกับคะแนนดิบ 2 ในการหาความคลาดเคลื่อนในการวัดของการแปลงคะแนนดิบเป็นคะแนนสแตนไนน์ ซึ่งใช้การปัดเศษ พบว่าในแบบทดสอบที่มีความเชื่อมั่น .95 ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนในการวัดจะเพิ่มขึ้นประมาณ 42 % สัดส่วนของค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนในการวัดมีสูงขึ้นเล็กน้อยเมื่อแบบทดสอบมีความเชื่อมั่นลดลงและเมื่อความเชื่อ

มั่นของแบบทดสอบเป็น .70 ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนเพิ่มมากกว่า 5 % ดังนั้นจึงสรุปได้ว่า การปัดเศษสำหรับคะแนนสแตนไนน์ มีผลทำให้ความแม่นยำในการวัดลดลง

โคเลน แฮนสัน และเบรนนัน (Kolen, Hanson and Brennan. 1992 : 285 – 307) ได้ศึกษาผลของการแปลงคะแนน จำนวนค่าของคะแนนแปลง และการเทียบคะแนนที่มีต่อการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ใช้โมเดลของคะแนนจริงแบบเข้มงวด และความเชื่อมั่นที่หาโดยสูตร KR – 20 โดยใช้การแปลงคะแนน 4 แบบ ได้แก่ Normalization, ACT Scale, Stannine และแบบ Log-Odds ข้อมูลที่นำมาวิเคราะห์นี้ได้จากการตอบแบบทดสอบชุดใน ACT Assessment ของปี ค.ศ. 1988 ที่เป็นการสอบคัดเลือกเข้าวิทยาลัย วิชาภาษาอังกฤษ คณิตศาสตร์ การอ่าน และเหตุผลทางวิทยาศาสตร์ ซึ่งเป็นแบบทดสอบชนิดเลือกตอบ ผลการศึกษา พบว่า

ความเชื่อมั่นของคะแนนแปลงรูปแบบ Normalization มีค่าน้อยกว่าของคะแนนที่แปลงแบบ ACT Scale โดยเฉพาะวิชาคณิตศาสตร์และเหตุผลทางวิทยาศาสตร์ และความเชื่อมั่นของคะแนนที่แปลงแบบ Stannine มีค่าความเชื่อมั่นต่ำกว่าของคะแนนแปลงรูปแบบ Normalization

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป มีค่าเกือบเท่ากันตลอดทุกค่าของคะแนนแปลงรูป ยกเว้นตรงคะแนนที่อยู่สุดโต่ง (หรือคะแนนที่สูงมาก) ซึ่งจะมีค่าต่ำลง เมื่อใช้การแปลงคะแนนทั้งสี่แบบจะทำให้มีผลต่อการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข (CSEM) เพื่อให้สามารถเปรียบเทียบค่า CSEM กันได้ จึงหาค่า CSEM ด้วยค่าเบี่ยงเบนมาตรฐานซึ่งพบว่า การแปลงคะแนนแบบ ACT Scale ทำให้ CSEM มีค่าค่อนข้างคงที่ในเกือบทุกค่าคะแนน การแปลงแบบ Normalization และ Stannine ให้ค่า CSEM คล้ายกัน มีรูปร่างแบบ U-Shaped หัวกลับ ค่าของคะแนนที่อยู่ตรงกลางของคะแนนแปลงแบบ Normalization มีแนวโน้มว่าค่า CSEM จะมีค่าต่ำกว่าในคะแนนดิบ หรือคะแนนที่แปลงแบบ ACT Scale และการแปลงแบบ Log – Odds ให้ค่า CSEM มีรูปร่างเป็นแบบ U-Shaped หัวกลับเช่นเดียวกัน

ผลของจำนวนค่าของคะแนนแปลง ทำการศึกษาจากคะแนนแปลงรูปแบบ ACT Scale และกำหนดจำนวนค่าของสเกลให้มีจำนวนต่างกันเป็น 6, 12 และ 36 ค่า พบว่าเมื่อ คะแนนแปลงรูปที่มีจำนวน 6 ค่า CSEM จะมีค่าแกว่งไปมาขึ้นลง ที่เป็นเช่นนี้เพราะผลจากการแปลงคะแนนดิบหลายค่าให้เป็นคะแนนแปลงรูปเพียงค่าเดียว (Many-to-One) และค่าที่ต่ำสุดของ CSEM ตรงคะแนนจริงที่เป็นสัดส่วนในการตอบถูกจะตรงกับคะแนนจริงที่เป็นคะแนนเต็ม และค่าสูงสุดของ CSEM ของสัดส่วนในการตอบถูกจะตรงกับคะแนนจริงที่เป็นจำนวนเต็ม 2 ค่า และพบว่าค่า CSEM ของ คะแนนแปลงรูป ที่มี 6 ค่า สูงกว่าของคะแนนแปลงรูป ที่มี 12 และ 36 ค่า

ผลจากการเทียบคะแนนของแบบทดสอบทั้ง 4 วิชามีทั้งหมด 5 ฟอร์ม คือ ฟอร์ม A B C D และ E ซึ่งเป็นการเทียบคะแนนในวิชาเดียวกันกับแต่ละฟอร์มโดยนำคะแนนที่ได้จากการสอบที่ใช้รูปแบบกลุ่มอย่างสุ่มมาเทียบคะแนนแบบอิควิเปอร์เซ็นไทล์ (Equipercentile) ปรากฏว่า แบบทดสอบแต่ละวิชาในแต่ละฟอร์มมีค่า CSEM ใกล้เคียงกันมาก ยกเว้นบางฟอร์มของแบบทดสอบบางวิชา ได้แก่ วิชาภาษาอังกฤษค่า CSEM มีค่าสูงขึ้นตรงคะแนนที่อยู่สุดโต่งจำนวน 3 ฟอร์ม ซึ่งเป็นเพราะแบบทดสอบวิชาภาษาอังกฤษ 3 ฟอร์มนี้มีความยากมากกว่าอีก 2 ฟอร์ม จึงทำให้เกิดช่องว่าง

ในการแปลงตรงคะแนนที่สูง (ไม่มีใครได้คะแนนมากกว่า 35 คะแนน) นอกจากนี้แบบทดสอบการอ่านฟอร์ม B ถึงฟอร์ม E มีค่า CSEM สูงกว่าฟอร์ม A ซึ่งตรงนี้อาจเป็นเพราะฟอร์ม B ถึงฟอร์ม E มีค่าความเชื่อมั่นต่ำกว่าฟอร์ม A

เมย์และไนซ์วันเดอร์ (May & Nicewander. 1994) ได้ศึกษาความเชื่อมั่นและฟังก์ชันสารสนเทศของคะแนนตำแหน่งเปอร์เซ็นไทล์ (Percentile Ranks : PR) กับคะแนนที่ได้จากการตอบถูก (Number-Right Score : NR) โดยใช้ทฤษฎีการตอบข้อคำถาม ผลที่ได้ปรากฏว่าคะแนนตำแหน่งเปอร์เซ็นไทล์ให้สารสนเทศและความเชื่อมั่นน้อยกว่าคะแนนที่เป็นจำนวนข้อที่ตอบถูก และคะแนนตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบที่มีข้อคำถามง่ายหรือยากแต่มีค่าอำนาจจำแนกสูง ให้ค่าความเชื่อมั่นต่ำมาก และฟังก์ชั่นสารสนเทศของคะแนนตำแหน่งเปอร์เซ็นไทล์ให้ค่าสารสนเทศน้อยมากเกี่ยวกับคุณลักษณะแฝงที่วัด

อินส์โก (Insko. 1995) ได้ทำการเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขโดยใช้วิธีการประมาณค่าทั้งสี่วิธี ได้แก่ วิธี Polynomial, The Corrected Binomial Error Model, IRT และวิธี ANOVA โดยใช้รูปแบบ Monte Carlo สร้างข้อมูลเทียมขึ้นมา 120 ชุด (ในแต่ละเซลทำซ้ำ 10) รูปแบบการศึกษามีสามองค์ประกอบคือ 2 × 2 × 3 นั่นคือองค์ประกอบที่เกี่ยวกับความเชื่อมั่นของแบบทดสอบมี 2 ค่า คือ .75 และ .90 จำนวนข้อของแบบทดสอบมี 2 ระดับ คือ 20 และ 50 ข้อ และขนาดของกลุ่มตัวอย่างมี 3 ขนาดคือ 250, 500 และ 1000 คน ผลการวิจัยพบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดจากทั้งสี่วิธีที่ได้จากข้อมูลเทียมนำไปเปรียบเทียบกับเกณฑ์ พบว่าวิธี IRT เป็นวิธีที่แสดงให้เห็นว่ามีความสัมพันธ์อย่างสอดคล้องที่สุดกับเกณฑ์ของ CSEM รองลงมาได้แก่วิธี The Corrected Binomial Error Model, Polynomial และ ANOVA ตามลำดับลงมา

นอกจากการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ใช้โมเดลของคะแนนจริงแบบเข้มงวดแล้วโคเลน เซง และแฮนสัน (Kolen, Zeng & Hanson. 1996) ยังได้ศึกษาการประมาณค่า CSEM โดยใช้วิธี IRT และวิธีคะแนนจริงแบบเข้มงวดซึ่งมี 2 วิธี คือวิธีของเฟลด์ (Feldt. 1984) โดยประมาณค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของจำนวนข้อที่ตอบถูก และวิธีสัมประสิทธิ์แอลฟ่า (Alpha) โดยศึกษาจากผลการตอบแบบทดสอบ ACT ที่ให้คะแนนแบบ 0 –1 วิชาภาษาอังกฤษและคณิตศาสตร์ จำนวน 75 และ 60 ข้อ ตามลำดับ เป็นแบบเลือกตอบ แต่ละวิชามี 2 ฟอร์ม ซึ่งนำไปสอบกับกลุ่มตัวอย่างที่มีความเท่าเทียมกันทั้ง 2 ฟอร์ม และใช้โมเดลโลจีสติกแบบสามพารามิเตอร์ รวมทั้งการแจกแจงของ θ ที่ได้ภายหลัง โดยใช้โปรแกรม BILOG3 ประมาณค่าทั้ง 2 ฟอร์มของแต่ละแบบทดสอบแต่ละฟอร์ม ผลการศึกษาพบว่า ค่าความเชื่อมั่นที่คำนวณจากคะแนนดิบที่ใช้ทั้ง 3 วิธี มีค่าตั้งแต่ .8964 ขึ้นไป วิธีของเฟลด์ที่ใช้การประมาณความคลาดเคลื่อนมาตรฐานเฉลี่ยมีค่าความเชื่อมั่นต่ำสุด รองลงมาได้แก่วิธีสัมประสิทธิ์แอลฟ่า และวิธี IRT นอกจากนี้พบว่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขวิธีของเฟลด์ มีแนวโน้มมีค่ามากกว่าวิธีแอลฟ่าและ วิธี IRT

ต่อมา เฟลด์และควอลส์ (Feldt & Qualls. 1998. 169 – 174) ได้ทำการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบไม่เป็นเส้นตรงโดยใช้คะแนนเทียบชั้นเรียน (Grade Equivalent) จากการสอบแบบทดสอบความคิดรวบยอดทางคณิตศาสตร์

และการอ่าน ฟอร์ม G ของ ITBS (Iowa Tests of Basic Skills) ของนักเรียนเกรด 6 ปี ค.ศ. 1992 – 1993 จำนวน 1,000 คน แบบทดสอบความคิดรวบยอดทางคณิตศาสตร์ ฉบับสั้นมีจำนวน 39 ข้อ และแบบทดสอบการอ่าน ฉบับยาว มีจำนวน 56 ข้อ คะแนนเทียบชั้นของแบบทดสอบความคิดรวบยอดทางคณิตศาสตร์มีพิสัยตั้งแต่ 7 (เดือนที่ 7 ของชั้นอนุบาล) ถึง 106 (เดือนที่ 6 ของเกรด 10) และคะแนนเทียบชั้นของแบบทดสอบการอ่านมีพิสัยตั้งแต่ 4 ถึง 117

การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข (CSEM) ใช้สองวิธี คือ

วิธีแรก หาโพลิโนเมียลที่มีความสอดคล้องกับคู่ของคะแนนดิบและคะแนนเทียบชั้น ซึ่งปรากฏว่าได้โพลิโนเมียลอันดับ 10 ที่สอดคล้องกับข้อมูล และหาอนุพันธ์อันดับที่หนึ่งของฟังก์ชันที่แต่ละคะแนนดิบ ซึ่งเป็นความชันและนำค่าความชันเหล่านี้ไปคูณกับค่าความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขในคะแนนดิบที่ตรงกับคะแนนเทียบชั้น ซึ่งการหาค่าความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขในคะแนนดิบใช้วิธีทวินามประกอบ

สูตรที่ใช้คือ

$$\tilde{\sigma}_{E(GE)} = [f'(X_0)]\tilde{\sigma}_{E(X)} \qquad \text{ณ ที่ } X = X_0$$

วิธีที่สอง คำนวณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ณ แต่ละค่าของคะแนน ช่วงของคะแนนเทียบชั้น GE_U และ GE_L ของคะแนนดิบ X_0 ในวิชาความคิดรวบยอดทางคณิตศาสตร์ คือ $X_0 \pm 3$ และในวิชาการอ่านคือ $X_0 \pm 4$ ดังนั้นค่าความชันก็คือ $(GE_U - GE_L)/6$ และ $(GE_U - GE_L)/8$ ของแต่ละวิชาตามลำดับ และนำค่าเหล่านี้ไปคูณกับค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในคะแนนดิบตามวิธีของลอร์ดและคีตส์

สูตรที่ใช้คือ

$$\tilde{\sigma}_{E(GE)} = \left(\frac{GE_U - GE_L}{2C}\right)\tilde{\sigma}_{E(X)}$$

หรือ $$\tilde{\sigma}_{E(GE)} = \left(\frac{GE_U - GE_L}{k - X_0 + C}\right)\tilde{\sigma}_{E(X)}$$

หรือ $$\tilde{\sigma}_{E(GE)} = \left(\frac{GE_U - GE_L}{X_0 + C}\right)\tilde{\sigma}_{E(X)}$$

ผลการประมาณค่าความชันและความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข พบว่าค่าความชันจากวิธีการหาอนุพันธ์อันดับหนึ่ง $f'(X)$ ในวิชาความคิดรวบยอดทางคณิตศาสตร์และการอ่านมีค่าตั้งแต่ 1.575 ถึง 4.867 และ 1.500 ถึง 4.833 ตามลำดับ และวิธี $(GE_U - GE_L)/2C$ ในวิชาความคิดรวบยอดทางคณิตศาสตร์และการอ่านมีค่าตั้งแต่ 1.226 ถึง 4.574 และ 1.125 ถึง 4.250 ตามลำดับ สำหรับค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบวิธีทวินามประกอบ วิชาความคิดรวบยอดทางคณิตศาสตร์และการอ่านมีค่าตั้งแต่ 1.686 ถึง 3.129 และ 1.932 ถึง 3.724 ตามลำดับ วิธีของลอร์ด และคีตส์ วิชาความคิดรวบยอดทางคณิตศาสตร์และการอ่านมีค่าตั้งแต่ 1.605 ถึง 3.010 และ 1.808 ถึง 3.510 ตามลำดับ

การประมาณค่า SEM ของคะแนนดิบมีความคงที่ ซึ่งไม่เหมือนกับค่า SEM ที่ได้ในคะแนนเทียบชั้น GE ที่มีตัวแปรมากกว่าและคาดคะเนได้น้อยกว่าในแต่ละเนื้อหา ความสัมพันธ์ระหว่างคะแนนคะแนนเทียบชั้นและความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในวิชาคณิตศาสตร์และการอ่านมีค่าเปลี่ยนแปลงอย่างมากจากคะแนนเทียบชั้นค่าหนึ่งไปยังอีกค่าหนึ่ง ซึ่งปรากฏว่ามีค่ามากเมื่อคะแนนเทียบชั้นมีค่าสูง และมีแนวโน้มที่ค่าที่ได้จะต่ำตรงคะแนนแปลงรูปตรงกลาง ค่าเฉลี่ยแบบถ่วงน้ำหนักของการประมาณค่าตรงการแจกแจงคะแนนในครึ่งที่ต่ำกว่ามีแนวโน้มว่าจะมีค่ามากกว่าตรงการแจกแจงของคะแนนในครึ่งที่มากกว่า

ต่อมา วัง โคเลน และ แฮริส (Wang, Kolen & Harris. 2000) ได้ขยายขอบเขตงานวิจัยที่ Kolen Hanson และ Brennan (1992) ได้เคยศึกษาเกี่ยวกับ CSEM ของคะแนนที่แปลงรูป โดยใช้โมเดลของคะแนนจริงแบบเข้มงวด (Strong True-Score Model) และงานวิจัยของ Kolen, Zeng และ Hanson (1996) ที่ศึกษาทำนองเดียวกันนี้แต่ใช้วิธีแบบ IRT งานวิจัยทั้งสองเรื่องนี้เป็นการศึกษาในกรณีที่มีการตรวจให้คะแนนเป็นแบบทวิ (Dichotomous) ดังนั้น วัง โคเลน และ แฮริส จึงได้ขยายแนวคิดเกี่ยวกับการประมาณค่า CSEM ที่มีข้อคำถามเป็นแบบพหุวิภาค (Polytomous Items) โดยใช้วิธี IRT แบบพหุวิภาค (Polytomous) เพื่อประเมินความคงที่ในการตัดสินใจเกี่ยวกับการจำแนกระดับความสามารถที่ขึ้นอยู่กับแต่ละฟอร์มของแบบทดสอบเลือกตอบ วิธีทั่วไปของการประเมินค่า CSEM และค่าความเชื่อมั่นก็คือเพื่อให้ได้การแจกแจงความน่าจะเป็นของคะแนนแต่ละระดับเมื่อกำหนดค่าความสามารถ (θ) มาให้ และคำนวณค่าเฉลี่ยอย่างมีเงื่อนไข ค่าเบี่ยงเบนมาตรฐาน (หรือความแปรปรวน) อย่างมีเงื่อนไขของคะแนนแปลงรูปหรือคะแนนแต่ละระดับ ซึ่งข้อมูลได้มาจากการทดสอบ The American College Testing Program's Work Keys ที่มีกลุ่มตัวอย่างจำนวน 7,097 , 2,035 และ 1,793 คน เมื่อได้ข้อมูลมาแล้วทำการตรวจสอบความเหมาะสมของข้อมูล (Model Fit) ความคงที่ในการจำแนกชั้น (Classification Consistency) และความเชื่อมั่น ซึ่งพบว่าการตรวจสอบเหล่านี้ยอมรับได้ ผลจากการศึกษาเกี่ยวกับ CSEM พบว่า คะแนนที่ได้ในแต่ละระดับแบบใหม่มีค่าความเชื่อมั่นสูงกว่าแบบเก่า และความคงที่ในการจำแนกชั้นก็ดีกว่าแบบเก่า ซึ่งชี้ให้เห็นถึงประโยชน์ที่ได้จากวิธี IRT แบบพหุ (Polytomous)

นอกจากงานวิจัยที่เกี่ยวกับการศึกษาถึงรูปร่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีประมาณค่าหลายวิธี แล้วยังมีงานวิจัยของ ลี เบรนนัน และโคเลน (Lee, Brennan & Kolen. 2000) ที่ศึกษาถึงตัวประมาณค่าของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของ คะแนนแปลงรูป ที่ได้จากการแปลงคะแนนดิบแบบไม่เป็นเส้นตรง คือ แบบ DSS (Developmental Standard Scores) แบบ GE (Grade Equivalents) และ PR (Percentile Ranks) โดยใช้วิธีประมาณค่า 4 วิธี ได้แก่ วิธี IRT วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์ - ควอลส์ โดยพิจารณาถึงความแม่นยำสัมพัทธ์ ซึ่งศึกษากับข้อมูลจำลองเป็นชุดข้อมูล 2 มิติที่ฝ่าฝืนข้อตกลงของการเป็นมิติเดียวในวิธี IRT ในครั้งแรกจะใช้ข้อมูลจริงชุดหนึ่งเป็นจุดเริ่มในการกำหนดคุณลักษณะของการจำลอง เป็นข้อมูลที่ได้จากการสุ่มนักเรียนเกรด 8 จำนวน 3,000 คน ที่สอบวิชาความคิดรวบยอดทางคณิตศาสตร์และการกะประมาณในระดับ 14 ฟอร์ม K จำนวน 32 และ 24 ข้อตามลำดับ และนำมาปรับให้คะแนนเป็นแบบ 0 - 1 และใช้โมเดลโลจิสติกแบบสามพารามิเตอร์เพื่อประมาณค่า

คะแนนจริงหรือค่าประชากร และสุ่มค่า θ (θ_1,θ_2) ของแบบทดสอบย่อยทั้งสองมา 1,000 คู่ ให้ได้ค่าสหสัมพันธ์ของ θ_1 และ θ_2 เป็น .85 และ .50 ซึ่งเป็นค่าที่แทนระดับที่แตกต่างกันในการฝ่าฝืนข้อตกลงของการเป็นมิติเดียว ผลการศึกษาพบว่าวิธี IRT ให้ค่าความลำเอียง (Bias) เป็นลบ และความแปรปรวนคลาดเคลื่อนเฉลี่ย (Aev) มีค่าต่ำสุดของคะแนนแปลงทุกวิธีและของกรณีที่ค่าสหสัมพันธ์ระหว่าง θ_1 และ θ_2 เป็น .85 และ .50 วิธีเฟลด์ - ควอลส์ ให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยมากที่สุด ในสามวิธีที่ไม่ใช่วิธี IRT วิธีทวินามประกอบให้ค่าความแปรปรวนคลาดเคลื่อนมากกว่าวิธีทวินามและวิธีเฟลด์ - ควอลส์ในทุกรูปแบบของการแปลงคะแนน

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องเกี่ยวกับความคลาดเคลื่อนมาตรฐานในการวัดที่กล่าวมาทั้งหมดนี้เป็นการศึกษากับข้อมูลจริงและข้อมูลที่จำลองขึ้นมา ซึ่งยังมีการศึกษาในเรื่องเหล่านี้น้อยมาก ดังนั้นผู้วิจัยจึงสนใจเลือกศึกษาการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนที่ได้จากการแปลงรูปวิธีตำแหน่งเปอร์เซ็นไทล์ แบบคะแนนปกติ โดยใช้วิธีการประมาณค่าวิธีทฤษฎีการตอบข้อสอบ วิธีแบบทวินาม วิธีแบบทวินามประกอบ และวิธีเฟลด์ - ควอลส์ เพื่อศึกษาว่าเมื่อใช้วิธีการประมาณค่าที่ต่างกันและการแปลงรูปคะแนนที่แตกต่างกันแล้วรูปร่างของค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจะมีรูปร่างเป็นอย่างไร และตัวประมาณค่าที่นำมาศึกษานั้นเป็นตัวประมาณค่าที่ดีหรือไม่โดยพิจารณาจากความลำเอียงและความแปรปรวนของค่าที่ประมาณ ซึ่งศึกษาจากข้อมูลจำลองที่มีขนาด 300 500 และ 800 คน ทำซ้ำ 30

สมมติฐานการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ผู้วิจัยจึงได้กำหนดสมมติฐานในการวิจัย ดังนี้

1. ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ และคะแนนปกติ เมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีแบบทวินาม วิธีแบบทวินามประกอบ และวิธีเฟลด์-ควอลส์เมื่อนำมาเขียนกราฟ จะมีรูปร่างต่างกัน โดยที่เส้นกราฟจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์น่าจะให้ค่าที่ราบเรียบกว่าวิธีแบบคะแนนปกติ

2. ค่าความลำเอียง และความแปรปรวนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอยางมีเงื่อนไขคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ และคะแนนปกติ เมื่อใช้วิธีการประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์-ควอลส์ ที่ได้จากกลุ่มตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อนำมาเขียนกราฟ จะมีรูปร่างต่างกัน โดยที่ขนาดตัวอย่าง 800 น่าจะให้ค่าความลำเอียงเฉลี่ยและความแปรปรวนคลาดเคลื่อนเฉลี่ยน้อยกว่าเมื่อมีขนาดตัวอย่าง 300 และ 500 คน

บทที่ 3
วิธีดำเนินการวิจัย

ในการวิจัยเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ แบบคะแนนปกติ เมื่อใช้การประมาณค่าตามวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์ - ควอลส์ ผู้วิจัยได้ดำเนินการวิจัยโดยกำหนดขอบเขตของการวิจัย ดังนี้

1. กำหนดประชากรและกลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการวิจัย
3. การเก็บรวบรวมข้อมูล
4. สถิติและการวิเคราะห์ข้อมูล

กำหนดประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้ ได้แก่ นักเรียนชั้นมัธยมศึกษาตอนปลาย สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม ปีการศึกษา 2544 จำนวน 55 โรงเรียน นักเรียนจำนวน 21,534 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้แบ่งเป็น 2 ชุด คือ

ชุดที่ 1 กลุ่มตัวอย่างขนาดใหญ่

กลุ่มตัวอย่างชุดที่ 1 เป็นนักเรียนชั้นมัธยมศึกษาตอนปลาย สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม ปีการศึกษา 2544 ซึ่งเลือกมาจากประชากรโดยการสุ่ม จำนวน 1,500 คน และกลุ่มตัวอย่างที่เลือกมาอย่างเฉพาะเจาะจงจากนักเรียนชั้นมัธยมศึกษาตอนปลายของโรงเรียนขนาดใหญ่/ใหญ่พิเศษ จังหวัดร้อยเอ็ด ขอนแก่น นครพนม และยโสธร จำนวน 500 คน เพื่อให้ข้อมูลที่ได้มีการแจกแจงเข้าสู่การแจกแจงแบบปกติ ทั้งนี้กลุ่มตัวอย่างที่เลือกมาอย่างเจาะจงมีสภาพแวดล้อมการจัดการเรียนการสอนใกล้เคียงกัน และเป็นการสอบวัดเกี่ยวกับความถนัดทางการเรียน

สำหรับกลุ่มตัวอย่างที่สุ่มมาจากประชากรเป้าหมายได้มาโดยการสุ่มแบบสองขั้น (Two - Stage Random Sampling) มีขนาดโรงเรียนเป็นชั้น (Strata) และมีห้องเรียนเป็นหน่วยการสุ่ม มีขั้นตอนการสุ่ม ดังนี้

1. กำหนดขนาดโรงเรียนตามเกณฑ์มาตรฐานโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ออกเป็นขนาดใหญ่/ใหญ่พิเศษ ขนาดกลาง และขนาดเล็ก ปรากฏว่ามีโรงเรียนขนาดใหญ่/ใหญ่พิเศษ 9 โรง ขนาดกลาง 28 โรง และขนาดเล็ก 18 โรง

2. สุ่มโรงเรียนแต่ละขนาดมาประมาณ 15 % โดยวิธีการสุ่มอย่างง่ายได้โรงเรียนขนาดใหญ่/ใหญ่พิเศษ 2 โรง ขนาดกลาง 5 โรงและขนาดเล็ก 2 โรง

3. สุ่มนักเรียนชั้นมัธยมศึกษาปีที่ 4, 5 และ 6 จากโรงเรียนขนาดใหญ่โรงละ 6 ถึง 7 ห้อง ขนาดกลางโรงละ 5 ห้อง และขนาดเล็กโรงละ 3 ห้อง โดยวิธีการสุ่มแบ่งชั้น และใช้นักเรียนในห้องที่สุ่มได้เป็นกลุ่มตัวอย่าง ได้กลุ่มตัวอย่างทั้งหมด 1,603 คน จากนั้นสุ่มนักเรียนบางส่วนในแต่ละระดับชั้นออกเพื่อให้ได้กลุ่มตัวอย่างในแต่ละระดับชั้นละ 500 คน ผลที่ได้ปรากฏดังตาราง 1

ตาราง 1 กลุ่มตัวอย่างที่ใช้ในการวิจัย

ขนาด	โรงเรียน	อำเภอ	จำนวนนักเรียน			
			ม. 4	ม. 5	ม. 6	รวม
ใหญ่/	สารคามพิทยาคม	เมือง	81	104	71	256
ใหญ่พิเศษ	บรบือวิทยาคาร	บรบือ	75	76	39	190
กลาง	กันทรวิชัย	กันทรวิชัย	72	50	56	178
	เขวาไร่ศึกษา	โกสุมพิสัย	61	35	88	184
	มิตรภาพ	แกดำ	32	81	53	166
	มหาวิชานุกูล	เมือง	73	56	41	170
	ท่าขอนยางพิทยาคม	กันทรวิชัย	63	36	78	177
เล็ก	เขวาใหญ่พิทยาสรรค์	กันทรวิชัย	12	32	29	73
	เหล่ายาววิทยาคาร	บรบือ	31	30	45	106
	รวม		500	500	500	1,500
ใหญ่/	สตรีร้อยเอ็ด	เมือง		36	45	81
ใหญ่พิเศษ	ขามแก่นนคร	เมือง	30	42	44	116
	ปิยะมหาราชาลัย	เมือง	35	-	47	82
	ยโสธรพิทยาคม	เมือง	41	47	44	132
	คำเขื่อนแก้วชนูปถัมภ์	คำเขื่อนแก้ว	44	45	-	89
	รวมทั้งหมด		650	670	680	2,000

ชุดที่ 2 กลุ่มตัวอย่างจำลอง

กลุ่มตัวอย่างชุดนี้ได้จากการสุ่มคะแนนที่เป็นผลการตอบแบบทดสอบของนักเรียนชั้นมัธยมศึกษาตอนปลายจากกลุ่มตัวอย่างขนาดใหญ่จำนวน 2000 คน โดยใช้โปรแกรมคอมพิวเตอร์ SPSS for Windows ทำการสุ่มขนาดละ 300 500 และ 800 คน แต่ละขนาดทำการสุ่มซ้ำ 30 ครั้ง นำข้อมูลที่สุ่มได้ไปศึกษาถึงค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง เมื่อใช้วิธีประมาณค่าต่างกัน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบทดสอบวัดความถนัดทางการเรียนที่สร้างตามแนวของเทอร์สโตน (Thurstone) มี 3 ด้าน ที่คัดเลือกมาจากแบบทดสอบของโครงการวิจัยเกี่ยวกับการสร้างและพัฒนาแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม ประจำปีงบประมาณ 2539 - 2541 (สุทธิวรรณ พีรศักดิ์โสภณ. 2540 และ 2542) เป็นแบบปรนัยเลือกตอบห้าตัวเลือก ที่มีกระบวนการสร้างและการหาคุณภาพตามหลักของการสร้างเครื่องมือที่ดี ลักษณะของแบบทดสอบมีดังนี้

แบบทดสอบวัดความถนัดทางด้านตัวเลข (Numerical Ability) เป็นการวัดความเข้าใจทางด้านปริมาณ จำนวน ในเชิงความสัมพันธ์ของปริมาณตัวเลข ความคิดรวบยอด และหลักการทางคณิตศาสตร์เบื้องต้น ทักษะการคิดคำนวณ ความสามารถด้านเหตุผลเชิงปริมาณ

สำหรับแบบทดสอบวัดความถนัดทางด้านตัวเลขที่ใช้ในการวิจัยครั้งนี้ประกอบด้วยเรื่อง ทักษะการคำนวณ คณิตศาสตร์เหตุผล และอนุกรมธรรมดา

ทักษะการคิดคำนวณ เป็นการวัดความคล่องแคล่วในการบวก ลบ คูณ หาร ตัวเลข

ตัวอย่างข้อสอบ

คำถามข้อ 0 $(3.25 - 1.25) \times 10 = ?$

ก. 2

ข. 9

ค. 10

ง. 15.75

จ. 20

คณิตศาสตร์เหตุผล เป็นการวัดความสามารถในการคิดหาเหตุผล

ตัวอย่างข้อสอบ

คำถามข้อ 0) ก + ข มีค่าอยู่ระหว่าง 6 กับ 14 ค่าของ ก และ ข ควรเป็นเท่าใด

ก. ก มีค่า 1 , ข มีค่า 5

ข. ก มีค่า 3 , ข มีค่า 11

ค. ก มีค่า 5 , ข มีค่า 8

ง. ก มีค่า 7 , ข มีค่า 10

จ. ก มีค่า 8 , ข มีค่า 7

อนุกรมธรรมดา ลักษณะข้อสอบจะมีตัวเลขเรียงกันอยู่อย่างมีกฎเกณฑ์ อาจเป็นกฎเกณฑ์ใดกฎเกณฑ์หนึ่ง ให้พิจารณาว่าตัวเลขที่เรียงกันนั้นมีการลด การเพิ่มอย่างไร หรือมีแนวโน้มเช่นไร แล้วให้หาคำตอบในช่องที่เว้นว่างไว้ให้

ตัวอย่างข้อสอบ

คำถามข้อ 0) 1 3 5 7 9 ...

ก. 10

ข. 11

ค. 12

ง. 13

จ. 14

แบบทดสอบวัดความถนัดทางด้านภาษา (Verbal Ability) ประกอบด้วยเรื่องคำตรงข้ามและความเข้าใจภาษา ลักษณะของข้อคำถามเป็นดังนี้

คำตรงข้าม เป็นการวัดความสามารถในการหาความหมายของคำ หรือข้อความ ที่มีความหมายตรงข้ามกับคำหรือข้อความที่กำหนดให้

ตัวอย่างข้อสอบ

คำถามข้อ 0) หยิบ มีความหมายตรงข้ามกับคำในข้อใด

ก. วาง

ข. ทิ้ง

ค. ถือ

ง. จับ

จ. ปล่อย

ความเข้าใจภาษา เป็นการวัดความสามารถในการอ่านจับใจความ แปลความ และวิเคราะห์เรื่องราวจากสิ่งที่อ่าน อาจได้แก่ ข้อความ โคลงกลอน บทสนทนา

ตัวอย่าง อ่านข้อความต่อไปนี้แล้วนำไปตอบคำถามข้อ 0) – ข้อ 00)

"คนเป็นอันมากที่เกิดมาแล้วย่อมแก่และเจ็บตายไปเปล่า หาได้ทำให้ชีวิตของตนเป็นประโยชน์สักเท่าใดไม่ ยิ่งผู้ที่บริบูรณ์ด้วยทรัพย์สมบัติ มักมัวเมาในหมู่สตรี บำรุงบำเรอ อันเข้าใจว่าเป็นความสุข ยิ่งทำให้ชีวิตเปล่าจากประโยชน์"

ข้อ 0) ข้อความนี้กล่าวในทำนองใด

ก. ต่อว่า

ข. ตักเตือน

ค. แนะนำ

ง. ตำหนิ

จ. สั่งสอน

ข้อ 00) ข้อความนี้ต้องการเตือนสติใครเป็นสำคัญ

ก. คนร่ำรวย

ข. คนขี้เกียจ

ค. คนมัวเมา

ง. คนแก่

จ. คนโง่

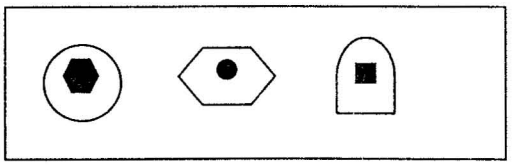
แบบทดสอบวัดความถนัดทางด้านเหตุผล (Reasoning Ability) เป็นการวัดความสามารถในการคิดอย่างมีวิจารณญาณ หาเหตุผลค้นคว้าหาความสำคัญ ความสัมพันธ์ และหลักการทั้งหลายที่สร้างกฎ หรือทฤษฎี หลักวินิจฉัย ลงสรุปได้อย่างถูกต้อง

สำหรับแบบทดสอบวัดความถนัดทางด้านเหตุผลที่ใช้ในการวิจัยครั้งนี้เป็นแบบภาพ ประกอบด้วยเรื่องการจำแนกประเภท อุปมาอุปไมย และอนุกรมภาพ

การจำแนกประเภท ลักษณะข้อสอบเป็นแบบภาพทรงเรขา เป็นการวัดความสามารถในการจัดจำแนกแยกประเภทของสิ่งต่าง ๆ โดยอาศัยโครงสร้าง มโนภาพ หน้าที่ เป็นต้น เปรียบเทียบกับกลุ่มนั้น ลักษณะของข้อสอบจะกำหนดภาพมาให้ 5 ภาพ ให้หาว่าภาพใดที่ต่างจากภาพอื่น ๆ จากตัวเลือก ก ถึง จ

ตัวอย่างข้อสอบ

คำถาม 0) ข้อใดจัดอยู่ในกลุ่มภาพที่กำหนดให้

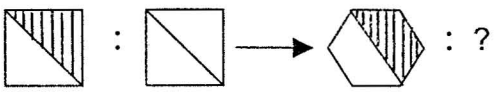




ก ข ค ง จ

แบบอุปมาอุปไมย เป็นการวัดความสามารถในการวิเคราะห์ความสัมพันธ์ระหว่างสิ่งของและเรื่องราวต่าง ๆ แล้วให้ขยายหลักการนั้นออกไปสู่สิ่งอื่นหรือสถานการณ์อื่นที่มีความสัมพันธ์เป็นทำนองเดียวกันหรือลักษณะเดียวกันกับของเดิม ในการวิจัยครั้งนี้เป็นแบบอุปมาอุปไมยภาพ รูปแบบของคำถาม มีดังนี้

ตัวอย่างข้อสอบ

คำถาม ข้อ 0) 



ก ข ค ง จ

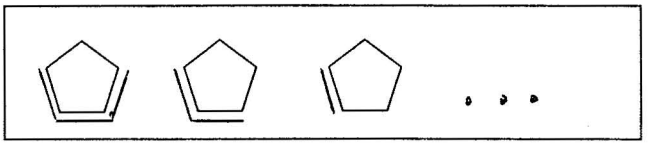
แบบอนุกรมภาพ

อนุกรมภาพหนึ่งมิติ ลักษณะข้อสอบจะเป็นภาพที่มีการเรียงลำดับอย่างใดอย่างหนึ่ง ซึ่งมีแนวโน้มไปในทิศทางเดียวกัน ให้พิจารณาว่าภาพที่กำหนดมาให้ 3 ภาพ หรือ 4 ภาพ นั้นมีแนวโน้มไปในทิศทางใดแล้วพิจารณาเลือกคำตอบจากข้อ ก ถึง จ ที่มีแนวโน้มเช่นเดียวกับภาพชุดแรกที่กำหนดให้

อนุกรมภาพหลายมิติ ข้อสอบจะกำหนดรูปภาพมาให้ชุดหนึ่ง ภาพนั้นจะมีการเรียงลำดับอย่างใดอย่างหนึ่ง ให้หารูปที่ขาดหายไป

ตัวอย่างข้อสอบ

ข้อคำถาม 0)





ก ข ค ง จ

การคัดเลือกข้อสอบ

จากแบบทดสอบของโครงการวิจัยเกี่ยวกับการสร้างและพัฒนาแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม ประจำปีงบประมาณ 2539 - 2541 ผู้วิจัยได้คัดเลือกข้อสอบที่มีค่าความยากตั้งแต่ .20 - .80 และค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป เพื่อใช้ในการวิจัยครั้งนี้ ดังรายละเอียดต่อไปนี้

แบบทดสอบวัดความถนัดด้านตัวแลขมี 30 ข้อ ประกอบด้วยเรื่องทักษะการคำนวณ 10 ข้อ คณิตศาสตร์เหตุผล 10 ข้อ และอนุกรมธรรมดา 10 ข้อ มีค่าความยากตั้งแต่ .27 - .73 และค่าอำนาจจำแนกตั้งแต่ .38 - .60

แบบทดสอบวัดความถนัดด้านภาษามี 40 ข้อ ประกอบด้วยเรื่องคำตรงข้าม จำนวน 20 ข้อ และความเข้าใจภาษา จำนวน 20 ข้อ มีค่าความยากตั้งแต่ .22 - .77 และค่าอำนาจจำแนกตั้งแต่ .20 - .54

แบบทดสอบวัดความถนัดด้านเหตุผลแบบภาพ มี 30 ข้อ ประกอบด้วยเรื่องการจำแนกประเภทจำนวน 10 ข้อ อุปมาอุปไมยจำนวน 10 ข้อ และอนุกรมภาพจำนวน 10 ข้อ มีค่าความยากตั้งแต่ .22 - .78 และค่าอำนาจจำแนกตั้งแต่ .21 - .67

การทดลองเครื่องมือ

เมื่อคัดเลือกข้อสอบวัดความถนัดแต่ละด้านได้แล้ว ก่อนที่จะนำไปใช้เพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ผู้วิจัยได้นำเครื่องมือทั้ง 3 ฉบับ ไปทดลองกับกลุ่มตัวอย่างที่ไม่ใช่กลุ่มตัวอย่างจริงที่เป็นนักเรียนชั้นมัธยมศึกษาตอนปลายของโรงเรียนประชาพัฒนา โรงเรียนผดุงนารี และโรงเรียนสาธิตมหาวิทยาลัยมหาสารคาม รวมทั้งหมด 244 คน เพื่อหาเวลาที่

เหมาะสมของการทำแบบทดสอบแต่ละฉบับและภาษาที่ใช้ในคำชี้แจง จากการทดลองครั้งนี้ทำให้ต้องกำหนดจำนวนเวลาที่ใช้ในการทำแบบทดสอบแต่ละฉบับ ดังนี้

แบบทดสอบวัดความถนัดด้านตัวเลข จำนวน 30 ข้อ ให้เวลาทำ 30 นาที

แบบทดสอบวัดความถนัดด้านเหตุผล จำนวน 30 ข้อ ให้เวลาทำ 30 นาที

แบบทดสอบวัดความถนัดด้านภาษา จำนวน 40 ข้อ ให้เวลาทำ 25 นาที

นอกจากนี้ยังได้หาค่าความยาก ค่าอำนาจจำแนกและค่าความเชื่อมั่นแบบทดสอบแต่ละด้าน ปรากฏผลดังตาราง 2

ตาราง 2 ค่าพิสัยความยาก (p) อำนาจจำแนก (r) และค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน เมื่อใช้สูตร KR - 20, KR - 21 และ สูตร r_B

แบบทดสอบ	พิสัยของ p	พิสัยของ r	KR - 20	KR - 21	r_B
ความถนัดด้านตัวเลข	.21 - .78	.17 - .81	.8361	.8136	.8386
ความถนัดด้านเหตุผล	.17 - .61	.20 - .69	.7730	.7612	.7793
ความถนัดด้านภาษา	.15 - .81	.00 - .69	.7371	.7088	.7408

วิธีการเก็บรวบรวมข้อมูล

การดำเนินการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลตั้งแต่วันที่ 25 สิงหาคม - 30 ตุลาคม 2544 ตามขั้นตอนต่อไปนี้

1. ทำหนังสือขอความร่วมมือจากผู้อำนวยการโรงเรียนมัธยมศึกษาที่เป็นกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล เพื่อกำหนดวัน เวลา และสถานที่สอบ

2. ดำเนินการสอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ภาษา และเหตุผลกับกลุ่มตัวอย่างจำนวน 1500 คน ก่อนการสอบได้ชี้แจงให้นักเรียนกลุ่มตัวอย่างทราบถึงวัตถุประสงค์ของการสอบและโน้มน้าวจิตใจให้นักเรียนมีความร่วมมือในการสอบเพื่อให้ได้ผลการสอบที่เชื่อถือได้

3. นำผลการตอบแบบทดสอบแต่ละฉบับมาตรวจสอบว่ามีการแจกแจงแบบปกติหรือไม่ ผลปรากฏว่าข้อมูลที่ได้จากการตอบแบบทดสอบทั้งสามฉบับมีการแจกแจงแบบไม่ปกติ ดังนั้น ผู้วิจัยจึงได้เก็บข้อมูลเพิ่มเติมจากนักเรียนโรงเรียนสตรีร้อยเอ็ด ขามแก่นนคร ปิยะมหาราชาลัย ยโสธรพิทยาคม และคำเขื่อนแก้วชนูปถัมภ์ และนำผลมาตรวจสอบการแจกแจงแบบปกติ ปรากฏว่าข้อมูลที่ได้มีการแจกแจงเข้าสู่การแจกแจงแบบปกติ (Approximately Normality)

4. กลุ่มตัวอย่างที่ตอบแบบทดสอบทั้ง 3 ฉบับนี้ ผู้วิจัยกำหนดให้เป็นประชากรเทียม (Pseudo - population) จากนั้นจึงสุ่มคะแนนของนักเรียนในประชากรเทียมออกเป็นกลุ่มตัวอย่างขนาดเล็ก 300 500 และ 800 คน แต่ละขนาดสุ่มซ้ำ 30 ครั้ง

การจัดกระทำข้อมูล

เมื่อเก็บรวบรวมข้อมูลได้แล้ว ผู้วิจัยได้จัดกระทำข้อมูลดังนี้

สำหรับกลุ่มตัวอย่างขนาดใหญ่

การจัดกระทำข้อมูลกับกลุ่มตัวอย่างขนาดใหญ่ ดำเนินการดังนี้

1. นำข้อมูลการตอบของนักเรียนแต่ละคนและบันทึกลงในไฟล์แต่ละไฟล์

2. วิเคราะห์ข้อมูลเบื้องต้น ได้แก่ ค่าสถิติพื้นฐานคือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ ความโด่ง ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้โปรแกรม SPSS for WINDOWS และโปรแกรม r_B

3. ตรวจสอบการแจกแจงของคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ ปรากฏว่าคะแนนจากแบบทดสอบแต่ละฉบับมีการแจกแจงเข้าสู่การแจกแจงแบบปกติ (Approximately Normality)

4. ตรวจสอบความเป็นมิติเดียวของคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ ปรากฏว่าคะแนนจากแบบทดสอบวัดทุกฉบับมีความเป็นมิติเดียว (ดูรายละเอียดผลการวิเคราะห์ในภาคผนวก ข)

5. แปลงคะแนนดิบของนักเรียนแต่ละคนจากแบบทดสอบแต่ละฉบับเป็นคะแนนตำแหน่งเปอร์เซ็นไทล์ และคะแนนแบบปกติ โดยใช้โปรแกรม T- score และ โปรแกรมสำเร็จรูป Exel

6. ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนจากแบบทดสอบแต่ละฉบับตามวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์-ควอลส์ โดยใช้โปรแกรม CBINSS และโปรแกรมเฟลด์-ควอลส์ ที่ผู้วิจัยได้พัฒนาขึ้น

สำหรับตัวอย่างจำลอง

เมื่อได้ข้อมูลจากกลุ่มตัวอย่างขนาดใหญ่แล้วกำหนดให้เป็นประชากรเทียม และดำเนินการสุ่มให้เป็นกลุ่มตัวอย่างจำลอง มีวิธีการดังนี้

1. จากข้อมูลขนาดใหญ่ที่บันทึกลงในแต่ละไฟล์ ทำการสุ่มโดยใช้โปรแกรม SPSS for WINDOWS ให้ได้ขนาด 300 500 และ 800 คน แต่ละขนาดสุ่มซ้ำ 30 ครั้ง ดังนั้นจากคะแนนแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านได้ข้อมูลเป็นตัวอย่างจำลองจำนวน 90 ตัวอย่าง รวมเป็นข้อมูลที่เป็นตัวอย่างจำลองทั้งหมด 270 ตัวอย่าง

2. จากข้อมูลที่สุ่มได้แต่ละตัวอย่างจำลอง นำมาหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และความเชื่อมั่น และนำคะแนนมาแปลงเป็นตำแหน่งเปอร์เซ็นไทล์ และคะแนนแบบปกติ

3. ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านในแต่ละตัวอย่างจำลอง ตามวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์-ควอลส์ โดยใช้โปรแกรม CBINSS และโปรแกรมเฟลด์ - ควอลส์

4. ศึกษาถึงความแม่นยำในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแต่ละแบบที่มีวิธีการประมาณค่าต่างกัน โดยพิจารณาจากค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย

สถิติและการวิเคราะห์ข้อมูล

เมื่อเก็บรวบรวมข้อมูลมาได้แล้ว ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป SPSS for Windows, โปรแกรม CBINSS โปรแกรม r_B และโปรแกรมเฟลด์-ควอลส์ ดังนี้

การวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างขนาดใหญ่

สำหรับการวิเคราะห์ข้อมูลจากกลุ่มตัวอย่างขนาดใหญ่ ดำเนินการดังนี้

1. ตรวจสอบการแจกแจงแบบปกติของข้อมูล
2. ตรวจสอบความเป็นมิติเดียวของข้อมูล
3. วิเคราะห์ข้อมูลเบื้องต้น เพื่อหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ ความโด่ง และค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ
3. แปลงคะแนนดิบให้เป็นคะแนนตำแหน่งเปอร์เซ็นไทล์ และคะแนนแบบปกติ
4. ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงรูปต่าง ๆ โดยใช้โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้น จากสูตรตามวิธีการดังต่อไปนี้

4.1 วิธีทฤษฎีการตอบข้อสอบ (Kolen, Zeng & Hanson. 1996. 130 -131)

$$\sigma^2[s(X) \mid \theta] = \sum_{i=0}^{K}[s(i) - \xi(\theta)]^2 \Pr(X = i \mid \theta)$$

เมื่อ

$\sigma^2[s(X) \mid \theta]$ แทน ความแปรปรวนคลาดเคลื่อนอย่างมีเงื่อนไขของ Scale score ณ ระดับความสามารถ θ

$\xi(\theta)$ แทน ค่าเฉลี่ยการแจกแจงอย่างมีเงื่อนไข

เมื่อ $\xi(\theta) = \sum_{i=0}^{K}[s(i)\Pr(X = i \mid \theta)$

$s(i)$ แทน คะแนนแปลงรูป

$\Pr(X = i|\theta)$ แทน การแจกแจงอย่างมีเงื่อนไขของคะแนนดิบที่เป็นจำนวนข้อที่ตอบถูกเมื่อกำหนดความสามารถ θ มาให้

$i = 0, 1, 2,...,K$; K เป็นจำนวนข้อ ณ ระดับความสามารถ θ

4.2 วิธีทวินาม (Brennan & Lee. 1999 : 9 - 10)

$$\hat{\sigma}_{E(S|x)} = c_k\hat{\sigma}_{E^*(S|x)}$$

เมื่อ $\hat{\sigma}_{E^*(S|x)} = \sqrt{\sum_{y=0}^{k}[f(y)]^2 \Pr(y \mid \bar{x},k) - \left[\sum_{y=0}^{k} f(y)\Pr(y \mid \bar{x},k)\right]^2}$

$\hat{\sigma}_{E^*(S|x)}$ แทน ความคลาดเคลื่อนอย่างมีเงื่อนไขของ Scale score

$\Pr(y \mid \bar{x},k)$ แทน โอกาสที่จะตอบข้อคำถามถูก y ข้อจากข้อคำถามทั้งหมด k ข้อ

เมื่อ $\bar{x} = \frac{x}{k}$; X เป็นจำนวนข้อที่ตอบถูก และ k เป็นจำนวนข้อทั้งหมด

และ $c_k = \sqrt{\frac{k}{k-1}}$

4.3 วิธีทวินามประกอบ (Brennan & Lee. 1999 : 10 - 12)

$$\hat{\sigma}_{E(S|x)} = c_k \hat{\sigma}_{E^*(S|x)}$$

เมื่อ $$c_k = \frac{\hat{\sigma}_{E(x)}}{\hat{\sigma}_{E^*(x)}}$$

$$\hat{\sigma}_{E^*(S|x)} = \sqrt{\sum_{y_t=0}^{\max(y_t)}[f(y_t)]^2 \Pr(y_t \mid \bar{x},k) - \left[\sum_{y_t=0}^{\max(y_t)} f(y)\Pr(y_t \mid \bar{x},k)\right]^2}$$

4.3.1 คำนวณค่า $\hat{\sigma}_{E(x)}$ และ $\hat{\sigma}_{E^*(x)}$ จากสูตรต่อไปนี้ ((Lee, Brennan & Kolen. 2000 : 4 - 5)

$$\hat{\sigma}_{E(x)} = \sqrt{\sum_{h=1}^{H} \frac{x_h(k_h - x_h)}{k_h - 1}}$$

เมื่อ X แทน คะแนนของผู้สอบคนหนึ่ง ๆ

k แทน จำนวนข้อคำถามทั้งหมด

และ

$$\hat{\sigma}_{E^*(x)} = \sqrt{\sum_{h=1}^{H} \frac{x_h(k_h - x_h)}{k_h}}$$

เมื่อ x_h แทน คะแนนของผู้สอบคนหนึ่ง ๆ ในแต่ละกลุ่ม/ชั้น h

k_h แทน จำนวนข้อคำถามในกลุ่ม/ชั้น h

4.4 วิธีเฟลด์ - ควอลส์ (Feldt & Qualls. 1998 : 162 - 164)

$$\hat{\sigma}_{E(S|x)} = f'(X)\hat{\sigma}_{E(X|x)}$$

เมื่อ

$$f'(X) = \begin{cases} \dfrac{f(x+C)-f(x-C)}{2C} & x-C \geq 0,\ x+C \leq k \\ \dfrac{f(x+C)-f(0)}{x+C} & x-C < 0 \\ \dfrac{f(k)-f(x-C)}{k-x+C} & x+C > k \end{cases}$$

k เป็นจำนวนข้อคำถามในแบบทดสอบ และ

$$C = \begin{cases} 2 \text{ ถ้า } k < 25 \\ 3 \text{ ถ้า } 25 \leq k \leq 39 \\ 4 \text{ ถ้า } 40 \leq k \leq 59 \\ 5 \text{ ถ้า } k \geq 60 \end{cases}$$

$\hat{\sigma}_{E(s|x)}$ แทน ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของ Scale score : s ที่เท่ากับ x

$f'(x)$ แทน อนุพันธ์ของ f(x) หรือความชันของ f(x)

X แทน คะแนนที่สังเกตได้

$\hat{\sigma}_{E(X|x)}$ แทน ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแต่ละบุคคลที่ X = x

4.4.1 จากวิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงตามวิธีของเฟลด์ - ควอลส์ ต้องประมาณค่า $\hat{\sigma}_{E(X|x)}$ ของคะแนนดิบก่อน ในการวิจัยครั้งนี้ประมาณค่า $\hat{\sigma}_{E(X|x)}$ จากคะแนนดิบโดยใช้สูตรต่อไปนี้

4.4.1.1 ใช้สูตรที่คีตส์ ได้ปรับมาจากลอร์ด ดังนี้ (Lord. 1955 ; Keats. 1957 citing Feldt & Qualls. 1998 : 168)

$$_{KL}\hat{\sigma}_{E(x)} = \left[\left(\frac{1-r_{xx'}}{1-21r_{xx'}}\right)\left(\frac{X(k-X}{k-1}\right)\right]^{1/2}$$

เมื่อ X แทน คะแนนของผู้สอบคนหนึ่ง ๆ

k แทน จำนวนข้อคำถามทั้งหมด

$_{21}r_{xx'}$ แทน ความเชื่อมั่นที่หาโดยสูตรของคูเดอร์ริชาร์ดสัน (Kuder-Richardson) สูตรที่ 21

และ $r_{xx'}$ แทน ความเชื่อมั่นที่หาโดยสูตรของคูเดอร์ริชาร์ดสัน 20 และสูตรสัมประสิทธิ์ความเชื่อมั่น r_B

สูตรสัมประสิทธิ์ความเชื่อมั่น r_B หาได้จากสูตร (บุญเชิด ภิญโญอนันตพงษ์. 2542 : 59 - 72)

$$r_B = \frac{1}{1-\sum\lambda_i^2}\left[1-\frac{\sum P_i(1-P_i)}{S_x^2}\right]$$

เมื่อ P_i แทน ความยากของข้อสอบแต่ละข้อ

S_x^2 แทน ความแปรปรวนของคะแนนทั้งฉบับ

$\lambda_i = [\sum B_{ji} - P_i(\bar{X}-1)]/S_x^2$

$\bar{X} = \sum P_i$

B_{ji} แทน สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใด ๆ ถูก

4.4.1.2 ใช้สูตรแบบทวินามประกอบ (Feldt. 1984 : 888) ดังนี้

$$CB\,\hat{\sigma}^2_{E(x)} = \sum_{h=1}^{H}\frac{x_h(k_h - x_h)}{k_h - 1}$$

เมื่อ x_h แทน คะแนนของผู้สอบคนหนึ่ง ๆ ในแต่ละกลุ่ม/ชั้น h

k_h แทน จำนวนข้อคำถามในกลุ่ม/ชั้น h

การวิเคราะห์ข้อมูลจากตัวอย่างจำลอง

เมื่อสุ่มผลการตอบที่ได้จากตัวอย่างขนาดใหญ่มาขนาดละ 300 500 และ 800 คน แต่ละขนาดสุ่มซ้ำ 30 ครั้ง แต่ละครั้งดำเนินการวิเคราะห์เช่นเดียวกับกรณีของกลุ่มตัวอย่างขนาดใหญ่ และศึกษาถึงความแม่นยำในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง โดยใช้สูตรต่อไปนี้ (Lee, Brennan & Kolen. 2000 : 9)

1. ความลำเอียง (Bias) หาได้จากสูตร

$$\text{abias} = \frac{1}{N}\sum_{g=1}^{N}\left[\left(\frac{1}{R}\sum_{r=1}^{R}\hat{e}_{gr}\right)-e_g\right]$$

2. ความแปรปรวนคลาดเคลื่อน (Error Variance) หาได้จากสูตร

$$\text{aev} = \frac{1}{NR}\sum_{g=1}^{N}\sum_{r=1}^{R}\left(\hat{e}_{gr}-\frac{1}{R}\sum_{r=1}^{R}\hat{e}_{gr}\right)^2$$

เมื่อ N แทน จำนวนคนในตัวอย่างแต่ละขนาด

R แทน จำนวนการทำซ้ำ

e_g แทน ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่แท้จริงของคนที่ g

$\hat{e}_{gr}$ แทน ความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ประมาณค่าได้ของคนที่ g จากการทำซ้ำครั้งที่ r

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การนำเสนอผลการวิจัยเกี่ยวกับการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปครั้งนี้ เพื่อให้เกิดความเข้าใจตรงกันผู้วิจัยขอกำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล ดังต่อไปนี้

สัญลักษณ์และอักษรย่อที่ใช้ในการนำเสนอผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้แทนความหมายในการวิเคราะห์ข้อมูล มีดังนี้

X	แทน	คะแนนที่ได้จากการทดสอบ
K	แทน	จำนวนข้อของแบบทดสอบ
$\overline{X}$	แทน	คะแนนเฉลี่ย
PR	แทน	คะแนนแปลงรูปตำแหน่งเปอร์เซ็นไทล์
Z	แทน	คะแนนแปลงรูปแบบคะแนนปกติ
S	แทน	ความเบี่ยงเบนมาตรฐานของคะแนน
S^2	แทน	ความแปรปรวนของคะแนน
C.V.	แทน	สัมประสิทธิ์การกระจายของคะแนน
KR-20	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรคูเดอร์ - ริชาร์ดสัน ที่ 20
KR-21	แทน	ค่าความเชื่อมั่นที่คำนวณจากสูตรคูเดอร์ - ริชาร์ดสัน ที่ 21
r_B	แทน	ค่าความเชื่อมั่นที่ได้จากสูตรสัมประสิทธิ์ r_B
f'(X)	แทน	อนุพันธ์ของ f(X) หรือ ความชันของ f(X)
$F\text{-}Q_{R20}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงที่คำนวณจากวิธีของเฟลด์ – ควอลส์เมื่อคำนวณความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของคะแนนดิบที่ใช้ความเชื่อมั่นของตัวเศษโดยสูตร KR – 20
$F\text{-}Q_{RB}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงที่คำนวณจากวิธีของเฟลด์ – ควอลส์เมื่อคำนวณค่าความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของคะแนนดิบที่ใช้ความเชื่อมั่นของตัวเศษโดยสูตร R_B

$F\text{-}Q_{CB}$	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงที่คำนวณจากวิธีของเฟลต์ – ควอลส์เมื่อคำนวณความคลาดเคลื่อนมาตรฐานอย่างมีเงื่อนไขของคะแนนดิบที่ใช้สูตรทวินามประกอบ
BINSS	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงเมื่อใช้วิธีทวินาม
CBINSS	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงเมื่อใช้วิธีทวินามประกอบ
IRTSS	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ
θ	แทน	ระดับความสามารถที่ได้จากทฤษฎีการตอบข้อสอบ
CSEMs	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง
abias	แทน	ค่าความลำเอียงเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัด
aev	แทน	ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัด

ผลการวิเคราะห์ข้อมูล

ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านที่ได้จากตัวอย่างขนาดใหญ่และจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ตอนที่ 2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่

ตอนที่ 3 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านในกลุ่มตัวอย่างจำลองขนาด 300 500 และ 800 คน

ตอนที่ 4 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ผลการวิเคราะห์ข้อมูลนำเสนอตามลำดับ ดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน

1.1 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่

ผู้วิจัยได้นำผลการตอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผลที่เก็บรวบรวมได้จากตัวอย่างขนาดใหญ่มาตรวจให้คะแนนและหาค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบแต่ละด้าน ได้ผลดังตาราง 3 และ 4

ตาราง 3 ค่าสถิติพื้นฐานจากตัวอย่างขนาดใหญ่

แบบทดสอบ	K	พิสัยของคะแนน	$\overline{X}$	S.D	C.V.	ความโด่ง	ความเบ้
ด้านตัวเลข	30	1 - 29	15.2165	6.8001	.4469	-1.040	.1010
ด้านภาษา	40	3 - 36	19.7480	6.2342	.3157	-.6910	-.1670
ด้านเหตุผล	30	1 - 29	14.1970	6.5202	.4593	-.8720	.1480

ผลการวิเคราะห์ตามตาราง 3 พบว่า คะแนนเฉลี่ยของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขมีค่าเกินครึ่งหนึ่งของคะแนนเต็ม แต่แบบทดสอบวัดความถนัดทางการเรียนด้านภาษาและด้านเหตุผลมีค่าเฉลี่ยน้อยกว่าครึ่งหนึ่งของคะแนนเต็ม เมื่อพิจารณาค่าสัมประสิทธิ์การกระจายพบว่าคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลมีการกระจายมากที่สุด รองลงมาได้แก่คะแนนจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านภาษา ตามลำดับ สำหรับค่าความโด่งของคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านมีค่าเป็นลบ นั่นคือการแจกแจงของคะแนนเป็นเส้นโค้งมียอดสูง และค่าความเบ้จากคะแนนแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผลมีค่าเป็นบวก แต่คะแนนจากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าความเบ้เป็นลบ

ตาราง 4 ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างขนาดใหญ่

แบบทดสอบ	k	KR-20	KR-21	r_B
ด้านตัวเลข	30	.8819	.8667	.8836
ด้านภาษา	40	.7819	.7618	.7834
ด้านเหตุผล	30	.8609	.8525	.8616

ผลการวิเคราะห์ตามตาราง 4 พบว่า ค่าความเชื่อมั่นเมื่อประมาณค่าโดยใช้สูตร KR - 20 สูตร KR - 21 และสูตร r_B ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขให้ค่าสูงสุด รองลงมาได้แก่ด้านเหตุผลและด้านภาษา ตามลำดับ และเมื่อพิจารณาจากแบบทดสอบแต่ละด้านและแต่ละวิธีที่ใช้ในการประมาณค่าความเชื่อมั่น พบว่า สูตร r_B ให้ค่าความเชื่อมั่นสูงกว่าสูตร KR-20 และ KR-21

1.2 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ผู้วิจัยทำการสุ่มผลการตอบแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่มาขนาดละ 300 500 และ 800 คน สุ่มมาขนาดละ 30 ครั้ง แล้วหาค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบ ได้ผลดังตาราง 5 - 7

ตาราง 5 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ค่าสถิติ	ขนาดตัวอย่างจำลอง		
	300 คน	500 คน	800 คน
ค่าเฉลี่ยสูงสุด	16.0100	15.7960	15.5913
ค่าเฉลี่ยต่ำสุด	14.6467	14.5740	14.9000
ค่าเฉลี่ยทั้งหมด	15.2533	15.2412	15.2352
S สูงสุด	7.1402	7.0373	6.9732
S ต่ำสุด	6.5069	6.5606	6.6401
S เฉลี่ย	6.8259	6.7868	6.8130
C.V. สูงสุด	0.4707	0.4640	0.4623
C.V. ต่ำสุด	0.4191	0.4223	0.4328
C.V. เฉลี่ย	0.4476	0.4454	0.4473
ความโด่งสูงสุด	-0.9040	-0.9530	-0.9570
ความโด่งต่ำสุด	-1.1980	-1.1690	-1.1370
ความโด่งเฉลี่ย	-1.0490	-1.0500	-1.0470
ความเบ้สูงสุด	-0.2260	0.1860	0.1710
ความเบ้ต่ำสุด	-0.0600	-0.1300	-0.0250
ความเบ้เฉลี่ย	0.1080	0.0980	0.1000
KR -20 สูงสุด	0.8943	0.8909	0.8891
KR -20 ต่ำสุด	0.8693	0.8721	0.8759
KR -20 เฉลี่ย	0.8831	0.8812	0.8825

ตาราง 5 (ต่อ)

ค่าสถิติ	ขนาดตัวอย่าง		
	300 คน	500 คน	800 คน
KR -21 สูงสุด	0.8806	0.8769	0.8750
KR -21 ต่ำสุด	0.8536	0.8543	0.8586
KR -21 เฉลี่ย	0.8677	0.8660	0.8673
r_B สูงสุด	0.8958	0.8924	0.8906
r_B ต่ำสุด	0.8710	0.8737	0.8776
r_B เฉลี่ย	0.8848	0.8829	0.8841

ผลการวิเคราะห์ตามตาราง 5 พบว่า คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 300 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 16.010 ต่ำสุดเท่ากับ 14.6467 ค่าสัมประสิทธิ์การกระจายที่สูงสุดคือ .4707 ต่ำสุดคือ .4191 ค่าความเชื่อมั่นมีค่าสูงกว่า .8000 ทุกตัวอย่างที่สุ่มมา โดยที่ค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างจำลองขนาด 500 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 15.7960 ต่ำสุดเท่ากับ 14.5740 ค่าสัมประสิทธิ์การกระจายที่สูงสุด คือ .4640 ต่ำสุดคือ .4223 ค่าความเชื่อมั่นมีค่าสูงกว่า .8000 ทุกตัวอย่างที่สุ่มมา โดยที่ค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างขนาด 800 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 15.5913 ต่ำสุดเท่ากับ 14.9000 ค่าสัมประสิทธิ์การกระจายที่สูงสุดคือ .4623 ต่ำสุดคือ .4328 ค่าความเชื่อมั่นมีค่าสูงกว่า .8000 ทุกตัวอย่างที่สุ่มมา โดยค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

ตาราง 6 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ค่าสถิติ	ขนาดตัวอย่างจำลอง		
	300 คน	500 คน	800 คน
ค่าเฉลี่ยสูงสุด	20.4100	20.2920	19.9975
ค่าเฉลี่ยต่ำสุด	19.1567	19.3040	19.4313
ค่าเฉลี่ยทั้งหมด	19.6953	19.7647	19.7156
S สูงสุด	6.5756	6.5043	6.4369
S ต่ำสุด	5.7926	5.8750	6.0863
S เฉลี่ย	6.1874	6.2484	6.2660
C.V. สูงสุด	0.3381	0.3337	0.3252
C.V. ต่ำสุด	0.2900	0.2895	0.3070
C.V. เฉลี่ย	0.3143	0.3162	0.3178
ความโด่งสูงสุด	-0.3150	-0.5130	-0.5780
ความโด่งต่ำสุด	-0.9210	-0.8110	-08110
ความโด่งเฉลี่ย	-0.6470	-0.6760	-0.7080
ความเบ้สูงสุด	0.0360	-0.0350	-0.800
ความเบ้ต่ำสุด	-02710	-0.2680	-0.2690
ความเบ้เฉลี่ย	-0.1610	-0.1550	-0.1690
KR -20 สูงสุด	0.8049	0.8009	0.7971
KR -20 ต่ำสุด	0.7441	0.7527	0.7698
KR -20 เฉลี่ย	0.7785	0.7830	0.7842
KR -21 สูงสุด	0.7886	0.7834	0.7781
KR -21 ต่ำสุด	0.7200	0.7286	0.7488
KR -21 เฉลี่ย	0.7572	0.7625	0.7643
r_B สูงสุด	0.8066	0.8025	0.7985
r_B ต่ำสุด	0.7456	0.7546	0.7714
r_B เฉลี่ย	0.7802	0.7846	0.7858

ผลการวิเคราะห์ตามตาราง 6 พบว่า คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านภาษาจากตัวอย่างขนาด 300 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 20.4100 ต่ำสุดเท่ากับ 19.1567 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .3381 ต่ำสุดคือ .2900 ค่าความเชื่อมั่นส่วนใหญ่มีค่าต่ำกว่า .8000 แต่มากกว่า .7200 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองขนาด 500 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 20.2920 ต่ำสุดเท่ากับ19.3040 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .3337 ต่ำสุดคือ .2895 ค่าความเชื่อมั่นส่วนใหญ่มีค่าต่ำกว่า .8000 แต่สูงกว่า .7500 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองขนาด 800 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าเฉลี่ยใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 19.9975 ต่ำสุดเท่ากับ 19.4313 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .3252 ต่ำสุดคือ .3070 ค่าความเชื่อมั่นส่วนใหญ่มีค่าต่ำกว่า .8000 แต่สูงกว่า .7500 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

ตาราง 7 ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ค่าสถิติ	ขนาดตัวอย่างจำลอง		
	300 คน	500 คน	800 คน
ค่าเฉลี่ยสูงสุด	15.0100	14.9800	14.5125
ค่าเฉลี่ยต่ำสุด	1.7833	13.7800	13.8025
ค่าเฉลี่ยทั้งหมด	14.3239	14.2260	14.2322
S สูงสุด	6.7728	6.7826	6.7300
S ต่ำสุด	6.1006	6.2628	6.3350
S เฉลี่ย	6.5008	6.5419	6.6975
C.V. สูงสุด	0.4794	0.4764	0.4801
C.V. ต่ำสุด	0.4258	0.4366	0.4421
C.V. เฉลี่ย	0.4539	0.4600	0.4707
ความโด่งสูงสุด	-0.6970	-0.7560	-0.7740
ความโด่งต่ำสุด	-1.0160	-1.0020	-0.9700
ความโด่งเฉลี่ย	-0.8676	-0.8751	-0.8788
ความเบ้สูงสุด	0.2580	0.2450	0.2550
ความเบ้ต่ำสุด	-0.0590	-0.1400	0.0760
ความเบ้เฉลี่ย	0.1249	0.1425	0.1401
KR -20 สูงสุด	0.8738	0.8740	0.8714
KR -20 ต่ำสุด	0.8368	0.8465	0.8511
KR -20 เฉลี่ย	0.8598	0.8624	0.8606
KR -21 สูงสุด	0.8659	0.8662	0.8639
KR -21 ต่ำสุด	0.8264	0.8372	0.8419
KR -21 เฉลี่ย	0.8511	0.8536	0.8564
r_B สูงสุด	0.8745	0.8746	0.8720
r_B ต่ำสุด	0.8381	0.8474	0.8519
r_B เฉลี่ย	0.8608	0.8340	0.8614

ผลการวิเคราะห์ตามตาราง 7 พบว่า คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างขนาด 300 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 15.0100 ต่ำสุดเท่ากับ 13.7833 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .4794 ต่ำสุดคือ .4258 ค่าความเชื่อมั่นส่วนใหญ่มีค่าสูงกว่า .8000 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างขนาด 500 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง แต่ละครั้งมีค่าใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 14.9800 ต่ำสุดเท่ากับ 13.7800 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .4764 ต่ำสุดคือ .4366 ค่าความเชื่อมั่นส่วนใหญ่มีค่าสูงกว่า .8000 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

คะแนนเฉลี่ยของแบบทดสอบความถนัดทางการเรียนด้านเหตุผลจากตัวอย่างจำลองขนาด 800 คน เมื่อทำการสุ่มซ้ำ30 ครั้ง แต่ละครั้งมีค่าใกล้เคียงกัน ค่าเฉลี่ยสูงสุดเท่ากับ 14.5125 ต่ำสุดเท่ากับ 13.8025 สัมประสิทธิ์การกระจายค่าที่สูงสุดคือ .4801 ต่ำสุดคือ .4421 ค่าความเชื่อมั่นส่วนใหญ่มีค่าสูงกว่า .8000 ซึ่งค่าความเชื่อมั่นที่คำนวณจากสูตร r_B ให้ค่าสูงกว่าสูตร KR -20 และ KR-21

ตอนที่ 2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน

ผู้วิจัยได้นำคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล จากตัวอย่างขนาดใหญ่มาแปลงเป็นคะแนนแบบตำแหน่งเปอร์เซ็นไทล์ และแบบคะแนนปกติ แล้วประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงนี้ โดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์- ควอลส์ ได้ผลดังตาราง 8 – 13

ตาราง 8 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

X	PR	BINSS	CBINSS	เฟลด์-ควอลส์			
				f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
30	99.43	0	0	0.8933	0	0	0
29	99.43	1.0656	0	1.3325	1.2548	1.2468	1.3325
28	98.30	2.4779	2.3992	1.7420	2.2794	2.2649	2.4251
27	96.75	3.9686	3.9275	1.9800	3.1161	3.0961	3.3048
26	94.10	5.3714	5.3106	2.5500	4.5472	4.5183	4.7883
25	90.72	6.6339	6.5800	3.0450	5.9530	5.9152	6.2830
24	87.55	7.7539	7.7860	3.4917	7.3266	7.2801	7.8165
23	84.13	8.7365	8.5974	3.7000	8.2092	8.1570	8.5716
22	80.03	9.5861	9.4142	3.8283	8.8810	8.8243	9.2441
21	75.80	10.3081	10.0791	4.0550	9.7482	9.6858	10.0842
20	71.90	10.9091	10.6930	4.2433	10.4933	10.4267	10.8467
19	67.75	11.3949	11.3454	4.2883	10.8405	10.7719	11.4346
18	63.22	11.7713	11.6269	4.2867	11.0163	10.9464	11.5249
17	58.67	12.0463	11.8646	4.3867	11.4031	11.3308	11.9035
16	54.30	12.2308	12.0148	4.4533	11.6548	11.5804	12.1298
15	50.08	12.3361	12.1486	4.4700	11.7244	11.6497	12.2346
14	45.58	12.3714	12.1295	4.4367	11.6112	11.5371	12.0685
13	41.03	12.3424	12.2344	4.4217	11.4941	11.4212	12.0856
12	36.40	12.2494	11.9898	4.4967	11.5560	11.4827	11.9839
11	32.05	12.0835	11.9782	4.5467	11.4935	11.4208	12.0772
10	27.77	11.8181	11.7044	4.5800	11.3259	11.2540	11.8742
9	23.10	11.4022	11.3428	4.5700	10.9863	10.9159	11.5551
8	18.30	10.763	10.7483	4.4967	10.4314	10.3648	11.0068
7	13.55	9.8232	9.8460	4.2283	9.3814	9.3218	9.9380
6	8.98	8.5333	8.5520	3.6833	7.7287	7.6798	8.2045
5	5.07	6.9079	6.8028	2.9917	5.8487	5.8116	6.1399
4	2.40	5.0556	4.9570	2.2500	4.0122	3.9868	4.2392
3	1.00	3.1855	3.1103	1.4883	2.3423	2.3273	2.4980
2	0.35	1.5748	1.4990	1.0040	1.3137	1.3054	1.4036
1	0.05	0.4809	0.4207	0.5875	0.5532	0.5497	0.5875
0	0.05	0	0	0.3167	0	0	0

ผลการวิเคราะห์ตามตาราง 8 พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ จากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ให้ค่าสูงตรงกลางของคะแนนแปลง และให้ค่าลดต่ำลงตรงด้านปลายของคะแนนแปลง เมื่อพิจารณาแต่ละวิธีที่ใช้ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข พบว่า วิธีทวินามและวิธี

ทวินามประกอบให้ค่าสูงสุดตรงคะแนนแปลง PR ที่ 45.58 และ 41.03 ตามลำดับ แต่วิธีเฟลต์ – ควอลส์ คำนวณจากสูตร $F\text{-}Q_{R20}$, $F\text{-}Q_{RB}$ และ $F\text{-}Q_{CB}$ ให้ค่าสูงสุดตรงคะแนนแปลง PR ที่ 50.08

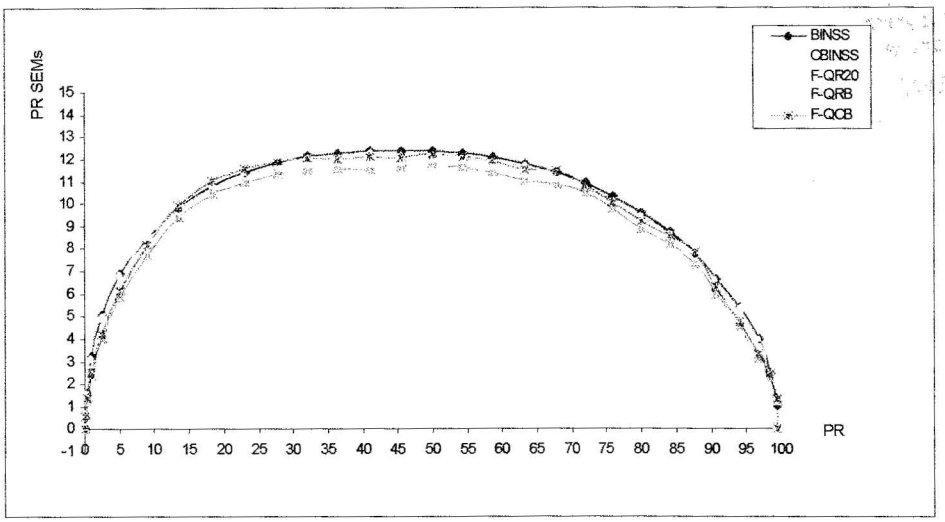
ตาราง 9 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติ (Z) จากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลต์ – ควอลส์ และวิธีทฤษฎีการตอบข้อสอบ

				เฟลต์-ควอลส์					
X	Z	BINSS	CBINSS	f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	θ	IRTSS
30	2.53	0	0	0.2267	0	0	0	2.37	0
29	2.53	0.3380	0.3309	0.2425	0.2284	0.2269	0.2425	2.37	0.2242
28	2.12	0.5174	0.5254	0.2420	0.3167	0.3146	0.3369	2.05	0.3563
27	1.85	0.5759	0.5827	0.2300	0.3620	0.3597	0.3839	1.67	0.3888
26	1.56	0.5702	0.5722	0.2550	0.4547	0.4518	0.4788	1.43	0.3714
25	1.32	0.5380	0.5389	0.2133	0.4171	0.4144	0.4401	1.33	0.3527
24	1.15	0.5001	0.4953	0.1917	0.4022	0.3996	0.4291	1.28	0.3549
23	1.00	0.4661	0.4633	0.1633	0.3624	0.3601	0.3783	1.16	0.3697
22	0.84	0.4391	0.4297	0.1433	0.3325	0.3304	0.3460	0.96	0.3792
21	0.70	0.4188	0.4127	0.1350	0.3245	0.3225	0.3357	0.66	0.3736
20	0.58	0.4034	0.3944	0.1300	0.3215	0.3194	0.3323	0.51	0.3555
19	0.46	0.3912	0.3859	0.1217	0.3076	0.3056	0.3245	0.42	0.3358
18	0.34	0.3812	0.3738	0.1167	0.2998	0.2979	0.3138	0.39	0.3268
17	0.22	0.3726	0.3623	0.1150	0.2989	0.2970	0.3121	0.34	0.3345
16	0.11	0.3654	0.3561	0.1150	0.3010	0.2990	0.3132	0.27	0.3548
15	0.00	0.3598	0.3516	0.1150	0.3016	0.2997	0.3148	0.08	0.3782
14	-0.11	0.3563	0.3450	0.1150	0.3010	0.2990	0.3128	-0.11	0.3964
13	-0.23	0.3556	0.3442	0.1167	0.3033	0.3014	0.3190	-0.31	0.4058
12	-0.35	0.3585	0.3550	0.1233	0.3170	0.3149	0.3286	-0.43	0.4075
11	-0.47	0.3654	0.3586	0.1317	0.3328	0.3307	0.3498	-0.56	0.4053
10	-0.59	0.3762	0.3691	0.1450	0.3586	0.3563	0.3759	-0.70	0.4025
9	-0.74	0.3900	0.3772	0.1650	0.3967	0.3941	0.4172	-0.87	0.3999
8	-0.90	0.4055	0.3983	0.1950	0.4524	0.4495	0.4773	-0.98	0.3947
7	-1.10	0.4220	0.4135	0.2317	0.5140	0.5107	0.5446	-1.23	0.3822
6	-1.34	0.4401	0.4397	0.2650	0.5560	0.5525	0.5903	-1.38	0.3584
5	-1.64	0.4609	0.4580	0.3000	0.5865	0.5828	0.6157	-1.55	0.3234
4	-1.98	0.4837	0.4791	0.3650	0.6509	0.6467	0.6877	-1.76	0.2829
3	-2.33	0.4973	0.5034	0.3250	0.5115	0.5082	0.5455	-1.74	0.2451
2	-2.70	0.4674	0.4703	0.3300	0.4318	0.4291	0.4614	-1.97	0.2065
1	-3.29	0.3256	0.3289	0.3275	0.3084	0.3064	0.3275	-2.18	0.1350
0	-3.29	0	0	0.3200	0	0	0	-2.18	0

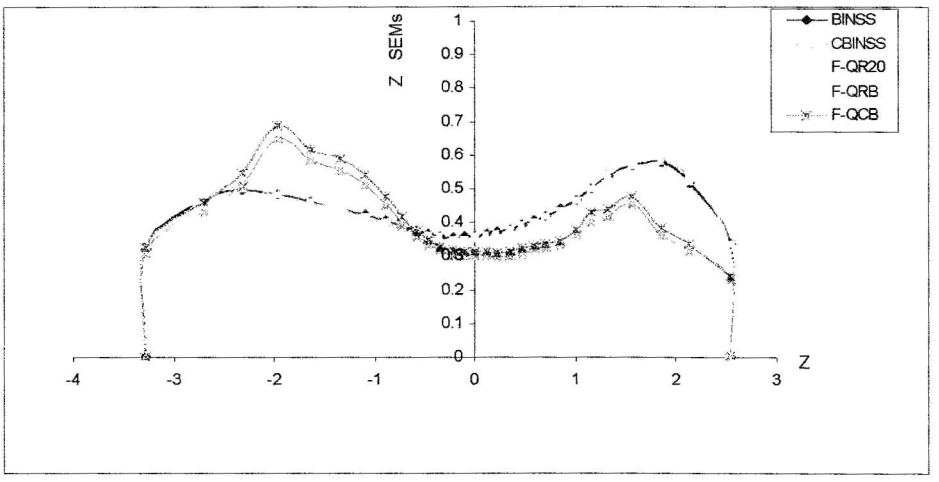
ผลการวิเคราะห์ตามตาราง 9 พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ เมื่อใช้วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงสุดที่ Z เท่ากับ 1.85 และให้ค่าต่ำสุดตรงคะแนน Z เท่ากับ -3.29 สำหรับวิธีของเฟลด์ – ควอลส์เมื่อใช้ สูตร $F\text{-}Q_{R20}$, $F\text{-}Q_{RB}$ และ $F\text{-}Q_{CB}$ ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสูงเมื่อ Z เท่ากับ -1.98 และมีค่าต่ำสุดเมื่อ Z เท่ากับ 2.53

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้วิธีทฤษฎีการตอบข้อสอบให้ค่าสูงสุดตรง θ เท่ากับ –0.43 และมีค่าต่ำลงเมื่อ θ ห่างไปจาก –0.86 แต่มีบางระดับของ θ ที่มีค่ามากขึ้นค่าความคลาดเคลื่อนมาตรฐานในการวัดไม่แน่นอนและมีค่าต่ำลงเมื่อ θ ที่มีค่ามากสุดและน้อยสุด

เพื่อให้เห็นความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่ ผู้วิจัยจึงได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ มาเขียนเป็นเส้นกราฟ ดังภาพประกอบ 2 และ 3



ภาพประกอบ 2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์



ภาพประกอบ 3 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และ วิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้วิธีการประมาณค่าโดยวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ที่ใช้สูตรต่างกัน รูปร่างของกราฟจะคล้ายกัน นั่นคือมีรูปร่างคล้ายตัวยูหัวกลับ ซึ่งค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าค่อนข้างคงที่ตรงช่วงกลางของคะแนน PR และจะมีค่าลดต่ำลงตรงด้านปลายของ PR ทั้งทางด้านมากและด้านน้อย

ผลการวิเคราะห์ตามภาพประกอบ 3 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้วิธีการประมาณค่าโดยวิธีทวินาม วิธีทวินามประกอบ จะให้เส้นกราฟคล้ายคลึงกัน และวิธีของเฟลด์ – ควอลส์เมื่อใช้สูตรต่างกันรูปร่างของกราฟเหมือนกัน ลักษณะของเส้นกราฟในแต่ละวิธีที่ประมาณค่ามีลักษณะหยักขึ้นลง มีลักษณะคล้ายโค้งสองยอด (Bimodal) โดยเฉพาะตรงที่คะแนน Z ใกล้ศูนย์จะมีค่าต่ำ และค่าความคลาดเคลื่อนมาตรฐานในการวัดจะสูงและลดลงอีกเมื่อคะแนน Z เริ่มมีค่ามากขึ้นหรือลดลง เมื่อพิจารณาในแต่ละวิธีที่ใช้ในการประมาณค่าพบว่าวิธีเฟลด์ - ควอลส์จากสูตร $F\text{-}Q_{CB}$ จะให้ค่าสูงกว่าวิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ $F\text{-}Q_{RB}$

เมื่อวิเคราะห์ตามภาพประกอบ 2 และ 3 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้วิธีประมาณค่าตามวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ พบว่า รูปร่างของกราฟจะคล้ายกันเมื่อมีคะแนนแปลงเหมือนกัน แต่รูปร่างกราฟจะต่างกันเมื่อใช้คะแนนแปลงต่างกัน

ตาราง 10 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

X	PR	BINSS	CBINSS	เฟลด์-ควอลส์			
				f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
40	99.97	0	0	0	0	0	0
39	99.97	0	0	0.0180	0.0172	0.0172	0
38	99.97	0.0498	0	0.0367	0.0490	0.0488	0
37	99.97	0.1677	0	0.0743	0.1200	0.1196	0
36	99.97	0.4225	0	0.1175	0.2161	0.2154	0.2287
35	99.88	0.8818	0.8872	0.2213	0.4487	0.4471	0.4735
34	99.75	1.6022	1.5775	0.4150	0.9086	0.9055	0.9521
33	99.45	2.6150	2.5710	0.7238	1.6863	1.6804	1.7618
32	99.03	3.9157	3.9417	1.1150	2.7348	2.7254	2.8843
31	98.20	5.4626	5.4434	1.5913	4.0744	4.0604	4.2679
30	96.65	7.1829	7.1689	2.1650	5.7483	5.7286	6.0171
29	94.18	8.9828	8.9225	2.7663	7.5737	7.5477	7.8820
28	91.05	10.7597	10.7486	3.4188	9.6063	9.5732	10.0295
27	87.15	12.4143	12.3578	4.0400	11.6029	11.5629	12.0477
26	82.43	13.8613	13.8671	4.5625	13.3439	13.2979	13.9321
25	77.32	15.0381	15.0486	4.9475	14.6867	14.6362	15.3155
24	71.68	15.9104	15.8683	5.2375	15.7329	15.6790	16.3414
23	65.88	16.4737	16.4260	5.3663	16.2662	16.2104	16.8922
22	60.15	16.7493	16.6930	5.3663	16.3697	16.3134	16.9992
21	54.6	16.7760	16.6533	5.3775	16.4659	16.4094	17.0345
20	49.15	16.5972	16.4976	5.3188	16.3068	16.2504	16.8932
19	44.22	16.2493	16.0776	5.2100	15.9530	15.8983	16.4558
18	39.50	15.7528	15.6777	5.0213	15.3173	15.2646	15.9013
17	34.30	15.1125	14.9802	4.7838	14.5005	14.4508	14.9813
16	29.13	14.3253	14.2938	4.5275	13.6002	13.5535	14.1600
15	24.20	13.3909	13.3582	4.2713	12.6792	12.6356	13.2018
14	19.98	12.3221	12.2009	4.0038	11.7098	11.6693	12.1002
13	16.33	11.1466	11.0378	3.6850	10.5833	10.5468	10.9538
12	12.93	9.9022	9.8119	3.3013	9.2762	9.2442	9.6165
11	10.05	8.6258	8.5480	2.8250	7.7346	7.7080	8.0240

ตาราง 10 (ต่อ)

X	PR	BINSS	CBINSS	เฟลต์-ควอลส์			
				f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
10	7.47	7.3458	7.2177	2.3850	6.3324	6.3107	6.5261
9	4.82	6.0798	6.0169	1.9913	5.0986	5.0811	5.2919
8	2.72	4.8440	4.7869	1.6000	3.9243	3.9109	4.0762
7	1.60	3.6650	3.5955	1.2525	2.9182	2.9081	3.0218
6	0.90	2.5869	2.4931	0.9300	2.0362	2.0292	2.0966
5	0.40	1.6655	1.6537	0.5988	1.2142	1.2100	1.2739
4	0.13	0.9488	0.8960	0.3363	0.6185	0.6164	0.6391
3	0.03	0.4556	0	0.2243	0.3622	0.3610	0.3816
2	0.03	0.1656	0	0.1450	0.1938	0.1931	0
1	0.03	0.0316	0	0.0740	0.0708	0.0706	0
0	0.03	0	0	0.0250	0	0	0

ผลการวิเคราะห์ตามตาราง 10 พบว่า ค่าความคลาดเคลื่อนมาตรฐาน.ในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา ให้ค่าสูงตรงกลางของคะแนนแปลง และให้ค่าลดต่ำลงตรงด้านปลายของคะแนนแปลง เมื่อพิจารณาแต่ละวิธี พบว่า วิธีทวินามและวิธีเฟลต์ – ควอลส์ สูตร $F\text{-}Q_{R20}$, $F\text{-}Q_{RB}$ และ $F\text{-}Q_{CB}$ ให้ค่าสูงสุดตรงกับคะแนน PR เท่ากับ 54.60 โดยที่วิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{CB}$ ให้ค่าสูงสุด และวิธีทวินามประกอบให้ค่าสูงสุดตรงกับคะแนน PR เท่ากับ 60.15

ตาราง 11 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติ (Z) จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์ และวิธีทฤษฎีการตอบข้อสอบ

X	Z	BINSS	CBINSS	เฟลด์-ควอลส์				θ	IRTSS
				f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$		
40	3.43	0	0	0	0	0	0	2.78	0
39	3.43	0.0248	0	0.0780	0.0747	0.0744	0	2.78	0.0194
38	3.43	0.1098	0	0.1033	0.1381	0.1376	0	2.78	0.0903
37	3.43	0.2312	0	0.1271	0.2053	0.2046	0	2.78	0.1893
36	3.43	0.3538	0	0.1363	0.2506	0.2498	0.2653	2.78	0.2808
35	3.04	0.4548	0.4597	0.1663	0.3371	0.3360	0.3559	2.52	0.3462
34	2.81	0.5289	0.5314	0.2000	0.4379	0.4364	0.4588	2.11	0.3866
33	2.54	0.5805	0.5815	0.2325	0.5417	0.5398	0.5659	2.17	0.4123
32	2.34	0.6155	0.6217	0.2613	0.6408	0.6386	0.6759	2.03	0.4315
31	2.10	0.6375	0.6400	0.2388	0.6113	0.6092	0.6405	1.79	0.4472
30	1.83	0.6479	0.6500	0.2350	0.6239	0.6218	0.6531	1.60	0.4593
29	1.57	0.6473	0.6461	0.2238	0.6126	0.6105	0.6377	1.44	0.4672
28	1.34	0.6370	0.6375	0.2213	0.6217	0.6195	0.6492	1.29	0.4716
27	1.13	0.6191	0.6165	0.2113	0.6067	0.6046	0.6301	1.12	0.4737
26	0.93	0.5961	0.5956	0.1963	0.5740	0.5720	0.5994	0.95	0.4751
25	0.75	0.5703	0.5694	0.1813	0.5380	0.5362	0.5612	0.82	0.4767
24	0.57	0.5439	0.5405	0.1700	0.5107	0.5089	0.5304	0.68	0.4791
23	0.41	0.5187	0.5151	0.1600	0.4850	0.4833	0.5037	0.50	0.4824
22	0.26	0.4959	0.4923	0.1500	0.4576	0.4560	0.4752	0.36	0.4862
21	0.12	0.4770	0.4714	0.1438	0.4402	0.4387	0.4555	0.22	0.4900
20	-0.02	0.4625	0.4579	0.1400	0.4292	0.4277	0.4447	0.05	0.4931
19	-0.15	0.4528	0.4462	0.1388	0.4249	0.4234	0.4384	-0.11	0.4952
18	-0.27	0.4477	0.4443	0.1375	0.4194	0.4180	0.4354	-0.24	0.4958
17	-0.40	0.4466	0.4414	0.1375	0.4168	0.4154	0.4306	-0.45	0.4950
16	-0.55	0.4491	0.4471	0.1388	0.4168	0.4154	0.4341	-0.57	0.4928
15	-0.70	0.4548	0.4527	0.1413	0.4193	0.4179	0.4367	-0.76	0.4892
14	-0.84	0.4641	0.4586	0.1463	0.4277	0.4263	0.4421	-0.85	0.4841
13	-0.98	0.4778	0.4725	0.1575	0.4523	0.4508	0.4682	-1.07	0.4770
12	-1.13	0.4962	0.4918	0.1713	0.4812	0.4795	0.4990	-1.25	0.4672
11	-1.28	0.519	0.5153	0.1800	0.4928	0.4911	0.5113	-1.37	0.4536

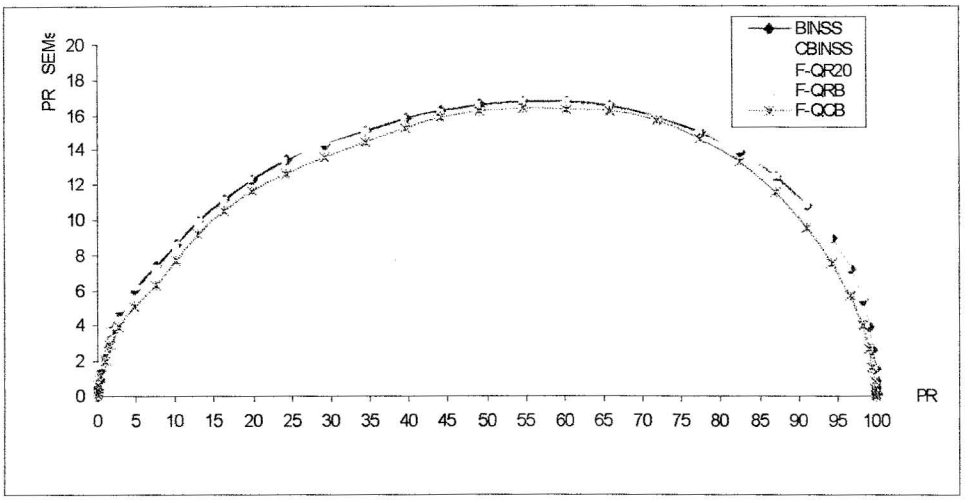
ตาราง 11 (ต่อ)

X	Z	BINSS	CBINSS	เฟลต์-ควอลส์				θ	IRTSS
				f'(X)	F-Q_{R20}	F-Q_{RB}	F-Q_{CB}		
11	-1.28	0.519	0.5153	0.1800	0.4928	0.4911	0.5113	-1.37	0.4536
10	-1.44	0.5441	0.5372	0.1913	0.5078	0.506	0.5235	-1.53	0.4348
9	-1.66	0.5686	0.5658	0.2088	0.5345	0.5327	0.5549	-1.73	0.4095
8	-1.92	0.5889	0.5866	0.2350	0.5764	0.5744	0.5987	-1.86	0.3777
7	-2.14	0.6006	0.5974	0.2688	0.6262	0.624	0.6485	-2.07	0.3420
6	-2.37	0.5975	0.5925	0.2488	0.5446	0.5427	0.5609	-2.19	0.3095
5	-2.65	0.5690	0.5730	0.2213	0.4487	0.4471	0.4708	-2.30	0.2884
4	-3.01	0.5009	0.4992	0.1888	0.3472	0.3460	0.3588	-2.07	0.2760
3	-3.43	0.3823	0	0.1843	0.2976	0.2966	0.3136	-2.65	0.2502
2	-3.43	0.2224	0	0.1767	0.2361	0.2353	0	-2.65	0.1818
1	-3.43	0.0688	0	0.156	0.1493	0.1488	0	-2.65	0.0719
0	-3.43	0	0	0.105	0	0	0	-2.65	0

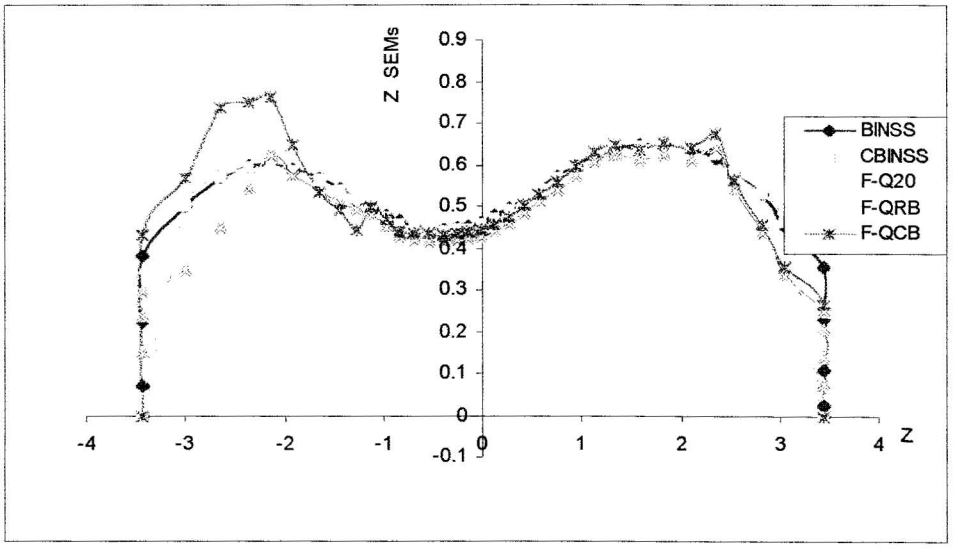
ผลการวิเคราะห์ตามตาราง 11 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ จากแบบทดสอบวัดความถนัดทางการเรียนด้าน เมื่อใช้วิธีทวินามและวิธีทวินามประกอบ วิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้ค่าใกล้เคียง โดยที่วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงสุดที่ Z เท่ากับ 1.83 วิธีเฟลต์ – ควอลส์ทุกสูตร ให้ค่าสูงสุดที่ Z เท่ากับ 2.34 และให้ค่าต่ำทางด้านปลายของ Z

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีทฤษฎีการตอบข้อสอบให้ค่าสูงสุดตรง θ เท่ากับ –0.24 และมีค่าลดต่ำลงเมื่อ θ ห่างไปจาก –0.24 ไปทางด้านปลายของค่า θ ที่มีค่ามากและค่าน้อย

เพื่อให้เห็นความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทย์และแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่ ผู้วิจัยจึงได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงเมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์ มาเขียนเป็นเส้นกราฟดังภาพประกอบ 4 และ 5



ภาพประกอบ 4 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์



ภาพประกอบ 5 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 4 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าโดยวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้รูปร่างของกราฟคล้ายกัน นั่นคือมีรูปร่างคล้ายตัวยูหัวกลับ ซึ่งค่าสูงสุดจะอยู่ตรงกับคะแนน PR ที่ 60.15 และจะมีค่า ลดต่ำลงตรงด้านปลายของคะแนน PR ที่ค่ามากและค่าน้อย

ผลการวิเคราะห์ตามภาพประกอบ 5 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงแบบคะแนนปกติ จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีการประมาณค่าโดยวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์สูตร F-Q_{R20}, F-Q_{RB} และ F-Q_{CB} รูปร่างของกราฟไม่ราบเรียบแต่จะคล้ายกันโดยเฉพาะตรงคะแนน Z อยู่ในช่วง -1 ถึง +1 มีความแตกต่างกันตรงคะแนน Z ที่มีค่าสูงและต่ำทางด้านปลาย โดยเฉพาะตรงที่คะแนน Z มีค่าเป็นลบ ทางด้านปลายกราฟจะแตกต่างกันมาก ซึ่งวิธีเฟลด์ - ควอลส์สูตร F-Q_{CB} เส้นกราฟจะสูงเมื่อ Z มีค่าต่ำมากทางด้านปลายแต่จะมีค่าสูง เมื่อคะแนน Z มีค่าบวกหรือมีค่ามากทางด้านปลาย

เมื่อพิจารณาตามภาพประกอบ 4 และ 5 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทย์และคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ พบว่า รูปร่างของกราฟจะคล้ายกันเมื่อมีคะแนนแปลงเหมือนกัน แต่รูปร่างกราฟจะต่างกันเมื่อใช้คะแนนแปลงต่างกัน

ตาราง 12 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

X	PR	BINSS	CBINSS	เฟลด์-ควอลส์			
				f'(X)	$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
30	99.85	0	0	0.6267	0	0	0
29	99.85	0.6786	0.6442	0.8000	0.7769	0.7750	0.8000
28	99.18	1.5690	1.4726	1.0340	1.3954	1.3919	1.4105
27	97.97	2.6322	2.5934	1.2917	2.0964	2.0911	2.1681
26	96.65	3.8425	3.7188	1.7750	3.2642	3.2561	3.3156
25	94.68	5.1253	5.0476	2.2883	4.6137	4.6021	4.7217
24	92.10	6.3914	6.3703	2.7200	5.8861	5.8711	6.0582
23	89.20	7.5715	7.4108	3.1033	7.1007	7.0830	7.1610
22	85.45	8.6320	8.4831	3.4633	8.2857	8.2646	8.3770
21	81.65	9.5723	9.2798	3.7617	9.3255	9.3022	9.2968
20	78.03	10.4114	10.2439	3.9700	10.1247	10.0989	10.2544
19	73.90	11.1715	11.0026	4.1083	10.7104	10.6833	10.8592
18	69.53	11.8651	11.7944	4.3117	11.4272	11.3983	11.6911
17	65.38	12.4889	12.1622	4.5467	12.1887	12.1578	12.1945
16	60.80	13.0238	12.6896	4.7250	12.7523	12.7202	12.7426
15	55.78	13.4406	13.2115	4.8633	13.1553	13.1218	13.2572
14	50.75	13.7071	13.444	5.0383	13.598	13.5637	13.6587
13	45.55	13.7955	13.5081	5.1367	13.7704	13.7354	13.8015
12	40.35	13.6851	13.5041	5.1383	13.6181	13.5837	13.7684
11	35.15	13.3639	13.2588	5.1000	13.2957	13.2620	13.5282
10	29.98	12.8275	12.6245	4.9583	12.6452	12.6130	12.7569
9	24.95	12.0788	11.9696	4.7750	11.8377	11.8081	12.0440
8	20.15	11.1274	10.9539	4.5417	10.8655	10.8378	10.9894
7	15.80	9.9880	9.9565	4.2100	9.6329	9.6089	9.8858
6	11.70	8.6766	8.5901	3.7833	8.1871	8.1663	8.3574
5	7.90	7.2096	7.1814	3.2300	6.5123	6.4959	6.6996
4	4.72	5.6106	5.5490	2.6083	4.7967	4.7847	4.9226
3	2.25	3.9317	3.8752	1.9250	3.1243	3.1164	3.2201
2	0.77	2.2849	2.2202	1.5500	2.0917	2.0865	2.1590
1	0.15	0.8703	0.8012	1.1425	1.1095	1.1067	1.1425
0	0.15	0.0000	0.0000	0.7000	0	0	0

ผลการวิเคราะห์ตามตาราง 12 พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีประมาณค่าแต่ละวิธีให้ค่าสูงสุดตรงกับ PR เดียวกันคือ 45.55 แต่วิธีทวินามให้ค่าสูงสุด ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจะมีค่าลดลงตรงคะแนน PR สูงและต่ำทางด้านปลาย

ตาราง 13 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติ (Z) จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์ และวิธีทฤษฎีการตอบข้อสอบ

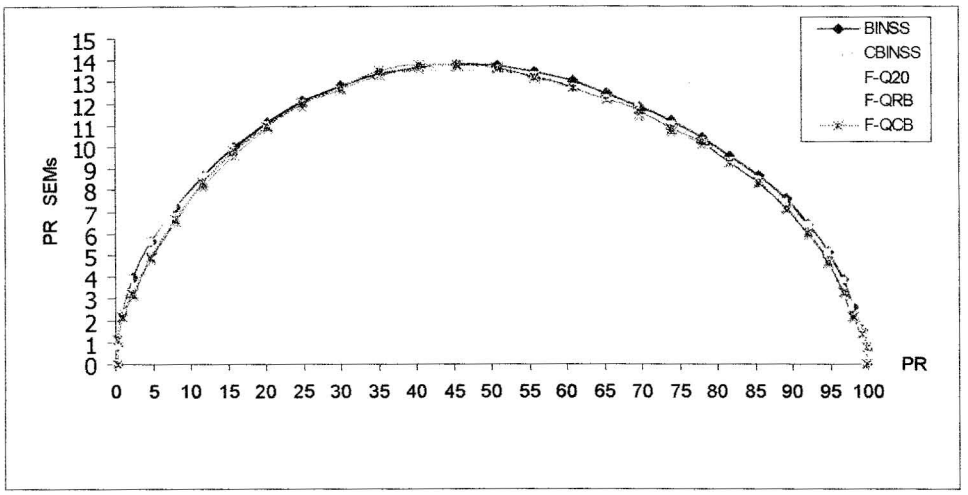
X	Z	BINSS	CBINSS	เฟลด์-ควอลส์				θ	IRTSS
				f'(X)	F-Q_{R20}	F-Q_{RB}	F-Q_{CB}		
30	2.97	0	0	0.3067	0	0	0	2.50	0
29	2.97	0.3256	0.3289	0.2850	0.2768	0.2761	0.2850	2.50	0.1659
28	2.40	0.4674	0.4703	0.2720	0.3671	0.3661	0.3710	2.29	0.2955
27	2.05	0.4973	0.5034	0.2600	0.4220	0.4209	0.4364	2.00	0.3592
26	1.83	0.4837	0.4791	0.2883	0.5302	0.5289	0.5385	1.73	0.3738
25	1.61	0.4609	0.4580	0.2233	0.4503	0.4491	0.4608	1.54	0.3680
24	1.41	0.4401	0.4397	0.1917	0.4148	0.4137	0.4270	1.43	0.3645
23	1.24	0.4220	0.4135	0.1767	0.4042	0.4032	0.4077	1.33	0.3708
22	1.06	0.4055	0.3983	0.1617	0.3868	0.3858	0.3911	1.18	0.3810
21	0.90	0.3900	0.3772	0.1500	0.3719	0.3709	0.3707	0.99	0.3861
20	0.77	0.3762	0.3691	0.1400	0.3570	0.3561	0.3616	0.76	0.3821
19	0.64	0.3654	0.3586	0.1317	0.3433	0.3424	0.3481	0.62	0.3719
18	0.51	0.3585	0.3550	0.1250	0.3313	0.3305	0.3389	0.49	0.3629
17	0.40	0.3556	0.3442	0.1250	0.3351	0.3343	0.3353	0.43	0.3623
16	0.27	0.3563	0.345	0.1250	0.3374	0.3365	0.3371	0.36	0.3721
15	0.15	0.3598	0.3516	0.1250	0.3381	0.3373	0.3407	0.25	0.3882
14	0.02	0.3654	0.3561	0.1300	0.3509	0.3500	0.3524	0.06	0.4034
13	-0.11	0.3726	0.3623	0.1317	0.3530	0.3521	0.3539	-0.17	0.4123
12	-0.24	0.3812	0.3738	0.1383	0.3666	0.3657	0.3706	-0.31	0.4139
11	-0.38	0.3912	0.3859	0.1433	0.3737	0.3727	0.3801	-0.48	0.4116
10	-0.52	0.4034	0.3944	0.1483	0.3783	0.3773	0.3816	-0.59	0.4103
9	-0.68	0.4188	0.4127	0.1583	0.3925	0.3915	0.3993	-0.69	0.4132
8	-0.84	0.4391	0.4297	0.1717	0.4107	0.4096	0.4155	-0.86	0.4177
7	-1.00	0.4661	0.4633	0.1917	0.4386	0.4375	0.4501	-1.08	0.4171
6	-1.19	0.5001	0.4953	0.2200	0.4761	0.4749	0.4860	-1.30	0.4042
5	-1.41	0.538	0.5389	0.2633	0.5309	0.5296	0.5461	-1.50	0.3749
4	-1.67	0.5702	0.5722	0.3283	0.6038	0.6023	0.6196	-1.65	0.3297
3	-2.00	0.5759	0.5827	0.2967	0.4815	0.4803	0.4963	-1.82	0.2697
2	-2.42	0.5174	0.5254	0.3120	0.4210	0.4200	0.4346	-2.05	0.1909
1	-2.97	0.3380	0.3390	0.3250	0.3156	0.3148	0.3250	-2.10	0.0896
0	-2.97	0	0	0.3233	0	0	0	-2.10	0

ผลการวิเคราะห์ตามตาราง 13 พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีประมาณค่าแต่ละวิธีค่าสูงสุดตรงกับค่า Z ที่เป็นลบ นั่นคือ วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงสุด

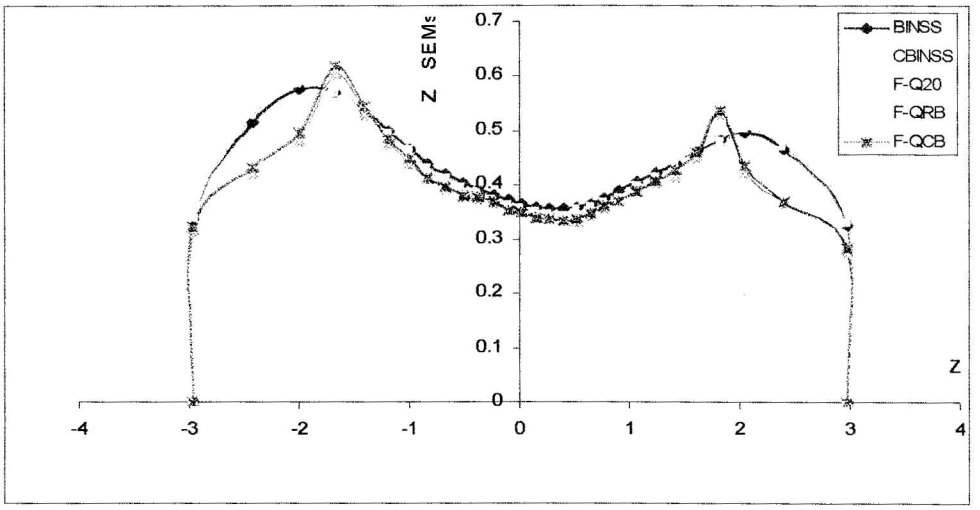
ตรงค่า Z เท่ากับ –2.00 และวิธีเฟลต์ - ควอลส์ให้ค่าสูงสุดตรงค่า Z เท่ากับ –1.67 ค่าที่ต่ำสุดในแต่ละวิธีจะตรงกับค่า Z ทางด้านปลาย

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลเมื่อใช้วิธีทฤษฎีการตอบข้อสอบให้ค่าสูงสุดตรง θ เท่ากับ –0.86 และมีค่าลดต่ำลงเมื่อ θ ห่างไปจาก –0.86 แต่มีบางระดับของ θ ที่มีค่าความคลาดเคลื่อนสูงขึ้น และลดต่ำลงอีกจนถึง θ ที่มีค่ามากและค่าน้อยทางด้านปลาย

เพื่อให้เห็นความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ และแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่ ผู้วิจัยจึงได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงเมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์ มาเขียนเป็นกราฟดังภาพประกอบ 6 และ 7



ภาพประกอบ 6 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์



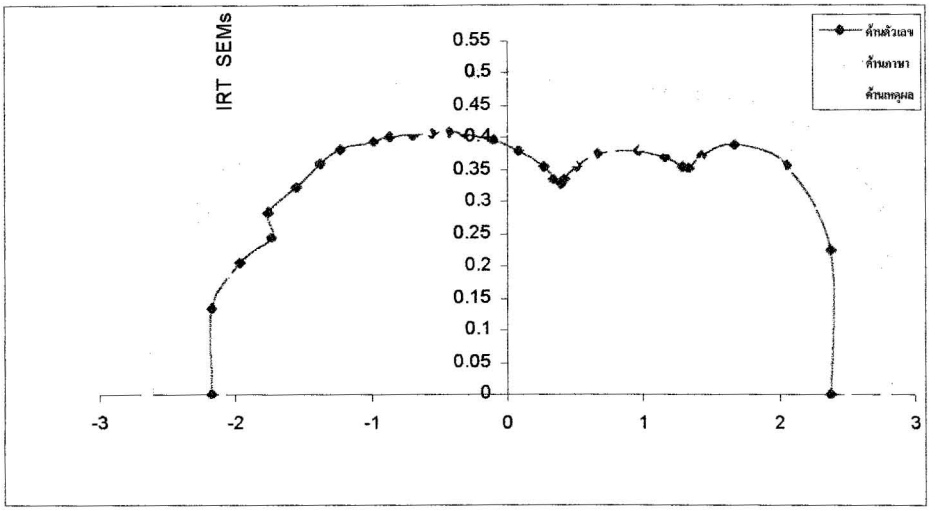
ภาพประกอบ 7 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 6 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีประมาณค่าโดยวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์สูตร F-Q_{R20} , F-Q_{RB} และ F-Q_{CB} จะให้รูปร่างของกราฟเหมือนกัน รูปร่างคล้ายตัวยูหัวกลับ แต่เบ้ขวา ค่าสูงสุดจะตรงกับคะแนน PR ที่ต่ำกว่า 50 และจะมีค่าลดต่ำลงตรงด้านปลายของคะแนน PR ที่ค่ามากและค่าน้อย

ผลการวิเคราะห์ตามภาพประกอบ 7 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ Z จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีการประมาณค่าแต่ละวิธีจะเส้นกราฟใกล้เคียงกันตรงคะแนน Z ระหว่าง -1 ถึง +1 แต่เส้นกราฟจะไม่ราบเรียบเมื่อค่า Z มีค่ามากขึ้นและน้อยลง วิธีทวินามและวิธีทวินามประกอบจะให้เส้นกราฟคล้ายกัน และวิธีของเฟลด์ – ควอลส์แต่ละสูตรก็จะให้เส้นกราฟคล้ายกันและให้ค่าสูงสุด แต่วิธีทวินามและวิธีทวินามประกอบเส้นกราฟจะโด่งขึ้นตรงค่า Z ที่มากและน้อย

เมื่อพิจารณาตามภาพประกอบ 6 และ 7 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลเมื่อใช้วิธีประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ พบว่า รูปร่างของกราฟจะคล้ายกันเมื่อคะแนนแปลงเหมือนกัน แต่รูปร่างกราฟจะต่างกันเมื่อคะแนนแปลงต่างกัน

เพื่อให้เห็นความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทฤษฎีการตอบข้อสอบ ผู้วิจัยจึงได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขมาเขียนเป็นเส้นกราฟ ดังภาพประกอบ 8



ภาพประกอบ 8 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล จากตัวอย่างขนาดใหญ่ เมื่อใช้วิธีทฤษฎีการตอบข้อสอบ

ผลการวิเคราะห์ตามภาพประกอบ 8 ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ พบว่า เส้นกราฟมีลักษณะเหมือนตัวยูหัวกลับและมีรอยหยักบางช่วงไม่สม่ำเสมอ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงกว่าด้านตัวเลขและด้านเหตุผลในทุกช่วงระดับของค่า θ และมีค่าลดลงอย่างต่อเนื่องเมื่อ θ มีค่าไปทางด้านปลายทั้งสอง สำหรับแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผลเส้นกราฟมีสูงหรือต่ำในบางระดับของ θ และลาดลดต่ำลงเมื่อ θ มีค่าไปทางด้านปลายทั้งสองเช่นเดียวกัน

ตอนที่ 3 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน ในกลุ่มตัวอย่างจำลองขนาด 300 500 และ 800 คน

3.1 พิสัยของค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากตัวอย่างจำลองขนาด 300 500 และ 800 คน จากการสุ่มซ้ำ 30 ครั้ง

ผู้วิจัยได้ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอยางมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน โดยสุ่มตัวอย่างขนาด 300 500 และ 800 คน มาขนาดละ 30 ครั้ง แต่ละครั้งที่ได้นำคะแนนมาแปลงเป็นคะแนนแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ และประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยวิธีทวินาม ทวินามประกอบ วิธีของเฟลด์ - ควอลส์ และวิธีทฤษฎีการตอบข้อสอบ (ดูรายละเอียดจากภาคผนวก จ) ได้ผลสรุปดังตาราง 14 - 16

ตาราง 14 พิสัยค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

แบบทดสอบ	ขนาดตัวอย่าง	BINSS	CBINSS	เฟลด์ - ควอลส์			วิธีที่ให้ค่าสูง	วิธีที่ให้ค่าต่ำ
				$F\text{-}Q_{r20}$	$F\text{-}Q_{rB}$	$F\text{-}Q_{CB}$		
ตัวเลข	300 คน	9.5650 ถึง 10.5673	9.4677 ถึง 10.4543	8.9475 ถึง 9.9212	8.8820 ถึง 9.8447	9.2718 ถึง 10.3718	BINSS	$F\text{-}Q_{rB}$
	500 คน	9.6156 ถึง 11.9049	9.4846 ถึง 10.2508	9.0011 ถึง 9.7951	8.9386 ถึง 9.7111	9.4243 ถึง 10.2336	BINSS	$F\text{-}Q_{rB}$
	800 คน	9.7313 ถึง 10.2488	9.6417 ถึง 10.1073	9.1070 ถึง 9.8904	9.0367 ถึง 9.8252	9.5912 ถึง 10.0663	BINSS	$F\text{-}Q_{rB}$
ภาษา	300 คน	12.4274 ถึง 14.4749	12.3892 ถึง 14.4133	12.1338 ถึง 13.1766	12.0922 ถึง 13.1378	12.5895 ถึง 13.7071	BINSS	$F\text{-}Q_{rB}$
	500 คน	12.5623 ถึง 3.5702	12.4954 ถึง 13.9345	11.9559 ถึง 13.3132	11.9124 ถึง 13.2629	12.4204 ถึง 13.8723	BINSS	$F\text{-}Q_{rB}$
	800 คน	12.6283 ถึง 13.4172	12.6939 ถึง 13.3253	12.1003 ถึง 12.7976	12.0584 ถึง 12.7531	12.5700 ถึง 13.2659	BINSS	$F\text{-}Q_{rB}$
เหตุผล	300 คน	9.5048 ถึง 11.4509	9.3708 ถึง 11.2428	9.0641 ถึง 11.0916	9.0407 ถึง 11.0497	9.1781 ถึง 11.2588	BINSS	$F\text{-}Q_{rB}$
	500 คน	10.0341 ถึง 11.0382	9.8818 ถึง 10.8482	9.6943 ถึง 10.6957	9.6704 ถึง 10.6639	9.8267 ถึง 10.8079	BINSS	$F\text{-}Q_{rB}$
	800 คน	10.1036 ถึง 12.0173	9.9369 ถึง 11.8481	9.8381 ถึง 11.0125	9.8137 ถึง 10.9758	9.9980 ถึง 11.1448	BINSS	$F\text{-}Q_{rB}$

ผลการวิเคราะห์ตามตาราง 14 พบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนตำแหน่งเปอร์เซ็นไทล์ที่ได้จากตัวอย่างแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง เมื่อใช้วิธีการประมาณค่าแต่ละวิธีมีค่าใกล้เคียงกัน และมีค่าพิสัยแคบ เมื่อพิจารณาจากแต่ละวิธีพบว่าวิธีทวินามให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุด แต่วิธีเฟลด์ - ควอลส์สูตร F-Q_{rB} ให้ค่าต่ำสุดในแบบทดสอบแต่ละด้านและแต่ละขนาดตัวอย่าง

ตาราง 15 พิสัยค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน

แบบทดสอบ	ขนาดตัวอย่าง	BINSS	CBINSS	เฟลด์ - ควอลส์			วิธีที่ให้ค่าสูง	วิธีที่ให้ค่าต่ำ
				F-Q_{r20}	F-Q_{rB}	F-Q_{CB}		
ตัวเลข	300 คน	.3811 ถึง .4383	.3779 ถึง .4339	.3463 ถึง .4020	.3436 ถึง .3993	.3603 ถึง .4215	BINSS	F-Q_{rB}
	500 คน	.1770 ถึง .4459	.1752 ถึง .4550	.1582 ถึง .4019	.1572 ถึง .3993	.1664 ถึง .4244	BINSS	F-Q_{rB}
	800 คน	.3942 ถึง .4214	.3897 ถึง .4162	.3617 ถึง .3848	.3603 ถึง .3820	.3796 ถึง .4052	BINSS	F-Q_{rB}
ภาษา	300 คน	.4695 ถึง .5713	.4661 ถึง .5691	.4442 ถึง .5209	.4427 ถึง .5187	.4618 ถึง .5401	BINSS	F-Q_{rB}
	500 คน	.4881 ถึง .5679	.4884 ถึง .5647	.4600 ถึง .5316	.4427 ถึง .5187	.4786 ถึง .5541	BINSS	F-Q_{rB}
	800 คน	.4963 ถึง .6303	.4926 ถึง .5321	.4669 ถึง .5049	.4563 ถึง .5027	.4854 ถึง .5250	BINSS	F-Q_{rB}

ตาราง 15 (ต่อ)

แบบทดสอบ	ขนาดตัวอย่าง	BINSS	CBINSS	เฟลด์ – ควอลส์			วิธีที่ให้ค่าสูง	วิธีที่ให้ค่าต่ำ
				$F\text{-}Q_{r20}$	$F\text{-}Q_{rB}$	$F\text{-}Q_{CB}$		
เหตุผล	300 คน	.3717 ถึง .4432	.3670 ถึง .4353	.3411 ถึง .4154	.3403 ถึง .4139	.3464 ถึง .4221	BINSS	$F\text{-}Q_{rB}$
	500 คน	.3936 ถึง .4282	.3876 ถึง .4222	.3688 ถึง .4050	.3676 ถึง .4041	.3735 ถึง .4109	BINSS	$F\text{-}Q_{rB}$
	800 คน	.4006 ถึง .4773	.3950 ถึง .4715	.3775 ถึง .4464	.3766 ถึง .4449	.3833 ถึง .4535	BINSS	$F\text{-}Q_{rB}$

ผลการวิเคราะห์ตามตาราง 15 พบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนปกติที่ได้จากตัวอย่างจำลองแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง เมื่อใช้วิธีการประมาณค่าแต่ละวิธีให้ค่าใกล้เคียงกัน และมีค่าพิสัยแคบ เมื่อพิจารณาจากแต่ละวิธีพบว่าวิธีทวินามให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุด แต่วิธีเฟลด์ - ควอลส์สูตร $F\text{-}Q_{rB}$ ให้ค่าต่ำสุดในแบบทดสอบแต่ละด้านและแต่ละขนาดตัวอย่าง

ตาราง 16 พิสัยค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ

แบบทดสอบ	ขนาดตัวอย่าง		
	300 คน	500 คน	800 คน
ด้านตัวเลข	.3482 - .3883	.3279 - .3789	.3942 - .4212
ด้านภาษา	.4439 - .5064	.4462 - .5064	.4555 - .4903
ด้านเหตุผล	.3675 - .4101	.3662 - .3994	.3717 - .3979

ผลการวิเคราะห์ตามตาราง 16 พบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างจำลองแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง มีค่าใกล้เคียงกัน แบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีพิสัยค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุด รองลงมาได้แก่ด้านเหตุผล และด้านตัวเลขตามลำดับ ในแบบทดสอบด้านตัวเลขและด้านเหตุผล เมื่อขนาดตัวอย่างจำลอง 500 คน มีค่าความคลาดเคลื่อนมาตรฐานในการวัดน้อยกว่าเมื่อขนาดตัวอย่างจำลองเป็น 300 และ 800 คน

3.2 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงที่ได้จากตัวอย่างจำลองขนาด 300 500 และ 800 คน และจากตัวอย่างขนาดใหญ่

ผู้วิจัยได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อทำการสุ่มซ้ำ 30 ครั้ง มาหาค่าเฉลี่ย ได้ผลดังตาราง 17 และ 18

ตาราง 17 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองแต่ละขนาดที่ใช้คะแนนตำแหน่งเปอร์เซ็นไทล์ (PR) และคะแนนปกติ (Z) เมื่อใช้วิธีประมาณค่าต่างกัน

แบบทดสอบ	ขนาดตัวอย่าง	คะแนนแปลง	ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดวิธี				
			BINSS	CBINSS	เฟลด์ - ควอลส์		
					$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
ด้านตัวเลข	300 คน	PR	10.0499	9.9361	9.3897	9.3208	9.7893
		Z	0.4113	0.4074	0.3737	0.3709	0.3911
	500 คน	PR	10.0939	9.9152	9.4171	9.3509	9.8587
		Z	0.4047	0.4014	0.3701	0.3675	0.3889
	800 คน	PR	9.9850	9.8573	9.3618	9.2994	9.8039
		Z	0.4090	0.4046	0.3749	0.3724	0.3935
ด้านภาษา	300 คน	PR	13.3248	13.2556	12.6092	12.5606	13.1172
		Z	0.5201	0.5171	0.4858	0.4839	0.5057
	500 คน	PR	13.1172	13.0515	12.4861	12.4390	12.9819
		Z	0.5182	0.5149	0.4870	0.4852	0.5066
	800 คน	PR	13.0528	12.9863	12.4489	12.4022	12.9280
		Z	0.5222	0.5156	0.4884	0.4866	0.5076
ด้านเหตุผล	300 คน	PR	12.5472	10.3676	10.1706	10.1401	10.2886
		Z	0.4141	0.4076	0.3892	0.3881	0.3950
	500 คน	PR	10.4625	10.2874	10.1398	10.1140	10.2619
		Z	0.4146	0.4082	0.3910	0.3900	0.3969
	800 คน	PR	10.5389	10.3611	10.1626	10.1362	10.2699
		Z	0.4199	0.4132	0.3962	0.3952	0.4024

ผลการวิเคราะห์ตามตาราง 17 พบว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแต่ละแบบในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาทุกขนาดตัวอย่างให้ค่าสูงสุด เมื่อพิจารณาจากแต่ละวิธีที่ใช้ในการประมาณค่า พบว่า วิธีที่ให้ค่าเฉลี่ยความคลาดเคลื่อนสูงสุดคือวิธีทวินาม และวิธีที่ให้ค่าเฉลี่ยน้อยสุดคือวิธีเฟลด์ - ควอลส์สูตร F-Q_{RB}

ตาราง 18 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างแต่ละขนาดเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ

แบบทดสอบ	ขนาดตัวอย่าง			ตัวอย่างขนาดใหญ่
	300 คน	500 คน	800 คน	2000 คน
ด้านตัวเลข	.3681	.3646	.4090	.3645
ด้านภาษา	.4754	.4728	.4699	.4730
ด้านเหตุผล	.3858	.3811	.3851	.3828

ผลการวิเคราะห์ตามตาราง 18 พบว่า ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีทฤษฎีการตอบข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาให้ค่าสูงสุด รองลงมาได้แก่ ด้านเหตุผลและด้านตัวเลข ตามลำดับ ในทุกขนาดตัวอย่างจำลองและในตัวอย่างขนาดใหญ่

3.3 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากตัวอย่างขนาดใหญ่เมื่อใช้วิธีประมาณค่าแต่ละวิธี

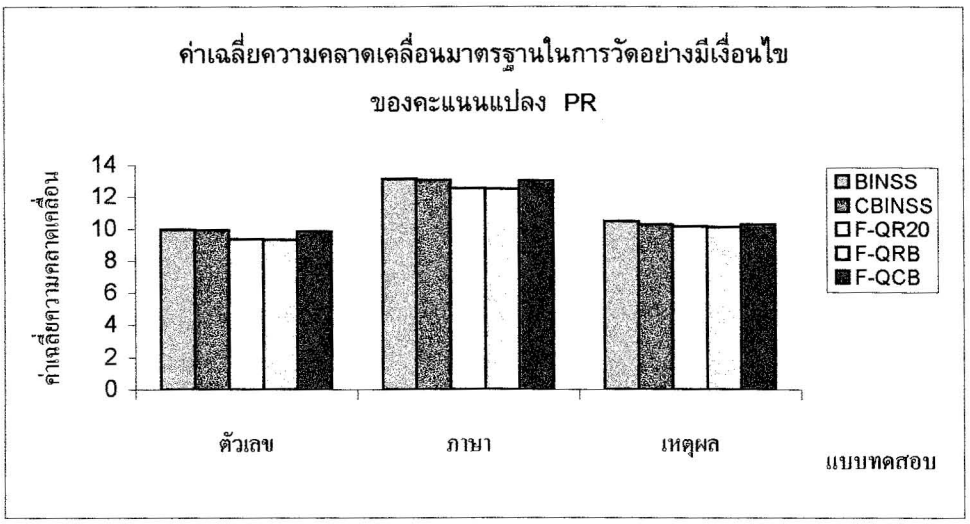
ผู้วิจัยได้นำค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน เมื่อใช้วิธีประมาณค่าแต่ละวิธีจากตัวอย่างขนาดใหญ่ 2000 คน มาหาค่าเฉลี่ย ได้ผลดังตาราง 19

ตาราง 19 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากกลุ่มตัวอย่างขนาดใหญ่

แบบทดสอบ	ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดวิธี				
	BINSS	CBINSS	เฟลด์ - ควอลส์		
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
คะแนนตำแหน่งเปอร์เซ็นไทล์ (PR)					
ด้านตัวเลข	9.9507	9.8915	9.3574	9.2898	9.8077
ด้านภาษา	13.1077	13.0348	12.5130	12.4699	12.9939
ด้านเหตุผล	10.4581	10.2808	10.1397	10.1140	10.2576
คะแนนปกติ (Z)					
ด้านตัวเลข	0.4074	0.4051	0.3759	0.3732	0.3950
ด้านภาษา	0.5200	0.5166	0.4897	0.4879	0.5088
ด้านเหตุผล	0.4152	0.4085	0.3933	0.3923	0.3991

ผลการวิเคราะห์ตามตาราง 19 พบว่า ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีประมาณค่าแต่ละวิธี พบว่าทั้งคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และคะแนนปกติ วิธีทวินามให้ค่าสูงสุด และวิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{RB}$ ให้ค่าต่ำสุด และพบว่าในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าความคลาดเคลื่อนมาตรฐานสูงสุด รองลงมาได้แก่แบบด้านเหตุผลและด้านตัวเลข ตามลำดับ

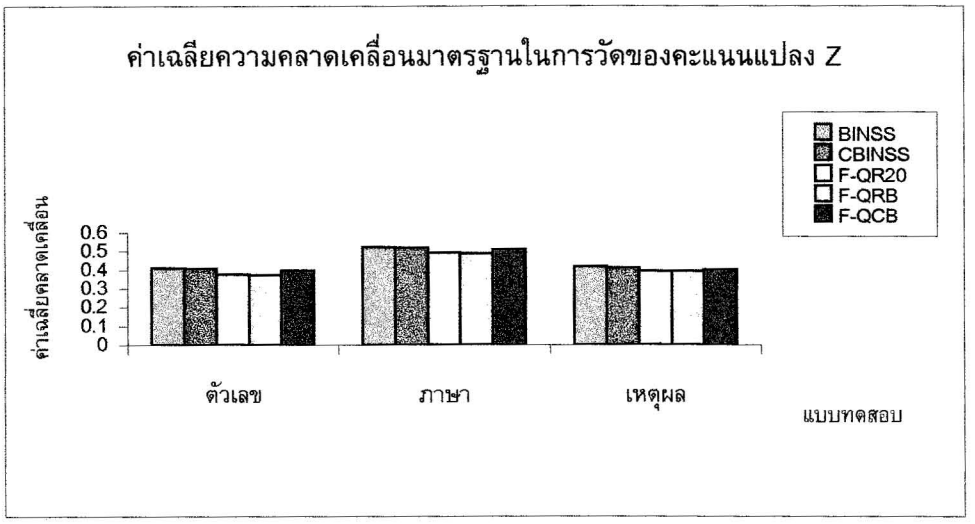
เพื่อให้เห็นความแตกต่างของค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติจากแบบทดสอบวัดความถนัดแต่ละด้าน เมื่อใช้วิธีประมาณค่าต่างกัน ผู้วิจัยจึงนำเสนอเป็นกราฟ ดังภาพประกอบ 9 และ 10



ภาพประกอบ 9 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านเหตุผล และด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 9 เมื่อใช้วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์แต่ละวิธีในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน พบว่า ให้เส้นกราฟใกล้เคียงกัน โดยค่าฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าสูงสุด รองลงมา ได้แก่ แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล และด้านตัวเลข ตามลำดับ

เมื่อพิจารณาในแต่ละวิธีแล้วพบว่า วิธีทวินามให้ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดสูงสุด วิธีเฟลต์ – ควอลส์สูตร F-Q_{RB} ให้ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดต่ำสุด



ภาพประกอบ 10 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านเหตุผล และด้านภาษา จากตัวอย่างขนาดใหญ่ เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 10 เมื่อใช้วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติแต่ละวิธีในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน พบว่า ให้เส้นกราฟใกล้เคียงกัน โดยค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าสูงสุด รองลงมาได้แก่ แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล และด้านตัวเลข ตามลำดับ

และเมื่อพิจารณาจากแต่ละวิธีแล้วพบว่า วิธีทวินามให้ค่าสูงสุด และวิธีของเฟลด์ – ควอลส์ สูตร F-Q_{RB} ให้ค่าต่ำสุด

ตอนที่ 4 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

จากตัวอย่างจำลองแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง นำคะแนนมาแปลงเป็นคะแนนตำแหน่งเปอร์เซ็นไทล์และคะแนนปกติ และประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยวิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์สูตร F-Q_{R20} , F-Q_{RB} และ F-Q_{CB} และวิธีทฤษฎีการตอบข้อสอบ และหาค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงในแต่ละวิธีและแต่ละขนาดตัวอย่าง แล้วนำมาประมาณค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ได้ผลดังตาราง 20 - 22

ตาราง 20 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข

ขนาดตัวอย่าง	ค่าสถิติ	BINSS	CBINSS	เฟลด์ - ควอลส์		
				$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
300 คน						
PR	abias	.09918	.04458	.03226	.03099	-.06936
	aev	.11285	.10060	.08036	.07985	.10089
Z	abias	.00391	.00233	-.00229	-.00232	-.02059
	aev	.00023	.00021	.00024	.00024	.00073
500 คน						
PR	abias	.01432	.02364	.05971	.06100	.00006
	aev	.17046	.03309	.04231	.04139	.04462
Z	abias	-.00276	-.00366	-.00582	-.00572	-.02285
	aev	.00194	.00201	.00017	.00171	.00242
800 คน						
PR	abias	.03428	-.03419	.00440	.00958	-.05472
	aev	.02085	.01794	.02606	.02695	.02103
Z	abias	.00155	-.00047	-.00103	-.00083	-.01818
	aev	.00005	.00005	.00004	.00004	.00038

ผลการวิเคราะห์ตามตาราง 20 ค่าความลำเอียงเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข พบว่าวิธีทวินามให้ค่าสูงสุด วิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{CB}$ ให้ค่าความลำเอียงเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 300 และ 800 คน และวิธีทวินามประกอบให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 800 คน แต่เมื่อขนาดตัวอย่างเท่ากับ 500 คนวิธีเฟลด์ - ควอลส์สูตร $F\text{-}Q_{RB}$ ให้ค่าสูงสุด

เมื่อพิจารณาจากแต่ละวิธี ค่าความลำเอียงเฉลี่ยในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข พบว่า วิธีทวินามให้ค่าความลำเอียงเฉลี่ยสูงสุดเมื่อตัวอย่างมีขนาด 300 คน แต่จากตัวอย่างขนาด 500 คนให้ค่าความลำเอียงเฉลี่ยต่ำกว่าจากค่าในตัวอย่างขนาด 800 คน วิธีทวินามประกอบให้ค่าความลำเอียงเฉลี่ยสูงสุดเมื่อขนาดตัวอย่างเท่ากับ 300 คน แต่จากตัวอย่างขนาด 800 คนให้ค่าความลำเอียงเฉลี่ยติดลบ วิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ

สูตร F-Q_{RB} เมื่อตัวอย่างมีขนาด 500 คนให้ค่าความลำเอียงเฉลี่ยสูงกว่าเมื่อขนาดตัวอย่างเท่ากับ 300 คน และให้ค่าน้อยสุดเมื่อตัวอย่างมีขนาด 800 คน

ค่าความลำเอียงเฉลี่ยในคะแนนแปลงแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข วิธีเฟลด์ – ควอลส์ทุกสูตรให้ค่าเป็นลบในทุกขนาดตัวอย่าง แต่วิธีทวินามให้ค่าเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 500 คน และวิธีทวินามประกอบให้ค่าเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 500 และ 800 คน วิธีทวินามให้ค่าความลำเอียงเฉลี่ยสูงสุดเมื่อขนาดตัวอย่างเท่ากับ 300 และ 800 คน

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขเมื่อขนาดตัวอย่างเท่ากับ 300 และ 500 คน วิธีทวินามให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยสูงสุด แต่เมื่อตัวอย่างมีขนาด 800 คน วิธีเฟลด์ – ควอลส์สูตร F-Q_{RB} ให้ค่าสูงสุด

เมื่อพิจารณาจากวิธีการประมาณค่าแต่ละวิธีในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข พบว่า วิธีทวินามประกอบ และวิธีเฟลด์ – ควอลส์ทุกสูตรเมื่อตัวอย่างมีขนาดใหญ่ขึ้นค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจะลดลง แต่ในวิธีทวินามพบว่าค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อตัวอย่างมีขนาด 500 คนให้ค่าสูงกว่าเมื่อตัวอย่างมีขนาด 300 คน และมีค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 800 คน

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในคะแนนปกติของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข มีค่าต่ำในแต่ละวิธีที่ใช้ประมาณค่า และให้ค่าใกล้เคียงกันเมื่อตัวอย่างมีขนาดเท่ากัน

ตาราง 21 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา

ขนาดตัวอย่าง	ค่าสถิติ	BINSS	CBINSS	เฟลด์ - ควอลส์		
				$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
300 คน						
PR	abias	.21709	.22086	.10181	.09634	.12329
	aev	.24561	.24771	.10132	.10006	.11867
Z	abias	.00011	.00048	-.00407	-.00425	-.00311
	aev	.00046	.00047	.00031	.00031	.00032
500 คน						
PR	abias	.00947	.01674	-.02129	-.02528	-.01206
	aev	.12191	.12259	.11706	.11786	.12755
Z	abias	-.00201	.00021	-.00286	-.00301	-..00217
	aev	.00021	.00021	.00019	.00019	.00020
800 คน						
PR	abias	-.05497	-.04841	-.05849	-.06208	-.06590
	aev	.03370	.02857	.03183	.03235	.03259
Z	abias	.00223	-.00103	-.00148	-.00154	-.00119
	aev	.00050	.00010	.00009	.00009	.00009

ผลการวิเคราะห์ตามตาราง 21 ค่าความลำเอียงเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าแต่ละวิธีในตัวอย่างแต่ละขนาด พบว่า วิธีเฟลด์ - ควอลส์ทุกสูตรให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 500 คน และทุกวิธีให้ค่าความลำเอียงเฉลี่ยเป็นลบเมื่อตัวอย่างมีขนาดเท่ากับ 800 คน นอกนั้นให้ค่าเป็นบวก

ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าแต่ละวิธีในขนาดตัวอย่างแต่ละขนาด พบว่า วิธีเฟลด์ - ควอลส์ทุกสูตรให้ค่าเป็นลบในทุกขนาดตัวอย่าง วิธีทวินามให้ค่าเป็นลบเมื่อตัวอย่างมีขนาดเท่ากับ 500 คน และวิธีทวินามประกอบให้ค่าเป็นลบเมื่อตัวอย่างมีขนาดเท่ากับ 800 คน เมื่อพิจารณาจากแต่ละวิธีพบว่าวิธี

ทวินาม วิธีทวินามประกอบ และวิธีเฟลต์ - ควอลส์ทุกสูตร มีค่าความลำเอียงเฉลี่ยน้อยลงเมื่อขนาดตัวอย่างใหญ่ขึ้น

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าแต่ละวิธีในตัวอย่างแต่ละขนาด พบว่า วิธีทวินามประกอบให้ค่าสูงสุดในทุกขนาดตัวอย่าง และวิธีเฟลต์ – ควอลส์สูตร F-Q_{RB} ให้ค่าต่ำสุดเมื่อตัวอย่างมีขนาด 300 คน นอกจากนี้พบว่า วิธีทวินามและวิธีทวินามประกอบให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยต่ำลงเมื่อตัวอย่างมีขนาดใหญ่ขึ้น แต่วิธีเฟลต์ - ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้ค่าสูงสุดเมื่อตัวอย่างมีขนาด 500 คน และวิธีเฟลต์ - ควอลส์สูตร F-Q_{CB} ให้ค่าสูงสุดเมื่อตัวอย่างมีขนาด 300 คน และมีค่าน้อยลงเมื่อตัวอย่างมีขนาดใหญ่ขึ้

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในคะแนนแปลงแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีประมาณค่าแต่ละวิธีในตัวอย่างขนาด 300 คน พบว่าวิธีทวินามประกอบให้ค่าสูงสุดและวิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้ค่าต่ำใกล้เคียงกัน สำหรับตัวอย่างขนาด 500 คนค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยมีค่าเกือบไม่แตกต่างกัน และในตัวอย่างขนาด 800 คน วิธีเฟลต์ - ควอลส์สูตร F-Q_{CB} ให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยสูงสุด

ตาราง 22 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล

ขนาดตัวอย่าง	ค่าสถิติ	BINSS	CBINSS	เฟลด์ – ควอลส์		
				$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$
300 คน						
PR	abias	.08910	.08684	.03087	.02606	.04271
	aev	.11788	.11292	.15465	.15247	.15840
Z	abias	-.00107	-.00095	-.00411	-.00428	-.00376
	aev	.00017	.00016	.00028	.00028	.00027
500 คน						
PR	abias	.00448	.00663	.00014	-.00007	.00430
	aev	.05139	.04906	.04773	.04768	.04702
Z	abias	-.00053	-.00030	-.00238	-.00238	-.00219
	aev	.00008	.00008	.00008	.00008	.00008
800 คน						
PR	abias	.08078	.08030	.02288	.02219	.02482
	aev	.12141	.12108	.04531	.04461	.04606
Z	abias	.00468	.00229	.00286	.00284	.00304
	aev	.00019	.00019	.00015	.00015	.00015

ผลการวิเคราะห์ตามตาราง 22 ค่าความลำเอียงเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีประมาณค่าแต่ละวิธีในตัวอย่างแต่ละขนาด พบว่า วิธีเฟลด์ - ควอลส์สูตร $F\text{-}Q_{RB}$ ให้ค่าเป็นลบเมื่อตัวอย่างมีขนาดเท่ากับ 500 คน นอกนั้นให้ค่าความลำเอียงเฉลี่ยเป็นบวก วิธีทวินามให้ค่าสูงสุดและวิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{RB}$ ให้ค่าต่ำสุดในทุกขนาดตัวอย่าง นอกจากนี้พบว่า ทุกวิธีที่ใช้ในการประมาณค่าให้ค่าความลำเอียงเฉลี่ยสูงสุดเมื่อขนาดตัวอย่างเท่ากับ 300 คน และให้ค่าต่ำเมื่อขนาดตัวอย่างเท่ากับ 500 คน

ค่าความลำเอียงเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล ทุกวิธีที่ใช้ประมาณค่าให้ค่าความลำเอียงเฉลี่ยเป็นลบเมื่อตัวอย่างมีขนาดเท่ากับ 300 และ 500 คน นอกนั้นให้ค่าเป็นลบ

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์จากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล พบว่า วิธีเฟลด์ – ควอลส์สูตร F-Q_{CB} ให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยสูงสุดและวิธีทวินามประกอบให้ค่าต่ำสุดเมื่อตัวอย่างมีขนาด 300 คน วิธีทวินามให้ค่าสูงสุดและวิธีเฟลด์ – ควอลส์สูตร F-Q_{CB} ให้ค่าต่ำสุดเมื่อตัวอย่างมีขนาด 500 คน วิธีทวินามให้ค่าสูงสุดและวิธีเฟลด์ – ควอลส์สูตร F-Q_{R20} ให้ค่าต่ำสุดเมื่อตัวอย่างมีขนาด 800 คน นอกจากนี้พบว่า วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงเมื่อตัวอย่างมีขนาด 800 คน และให้ค่าต่ำตัวอย่างมีขนาด 500 คน แต่วิธีเฟลด์ – ควอลส์ทุกให้ค่าลดลงเมื่อตัวอย่างมีขนาดใหญ่ขึ้น

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในคะแนนแปลงแบบคะแนนปกติจากแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล พบว่าแต่ละวิธีที่ใช้ประมาณค่าเมื่อขนาดตัวอย่างต่างกันให้ค่าต่ำใกล้เคียงกันเมื่อตัวอย่างมีขนาด 300 และ 800 คน และเมื่อตัวอย่างมีขนาด 500 คน ทุกวิธีที่ใช้ให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยต่ำกว่าเมื่อตัวอย่างมีขนาด 300 และ 800 คน

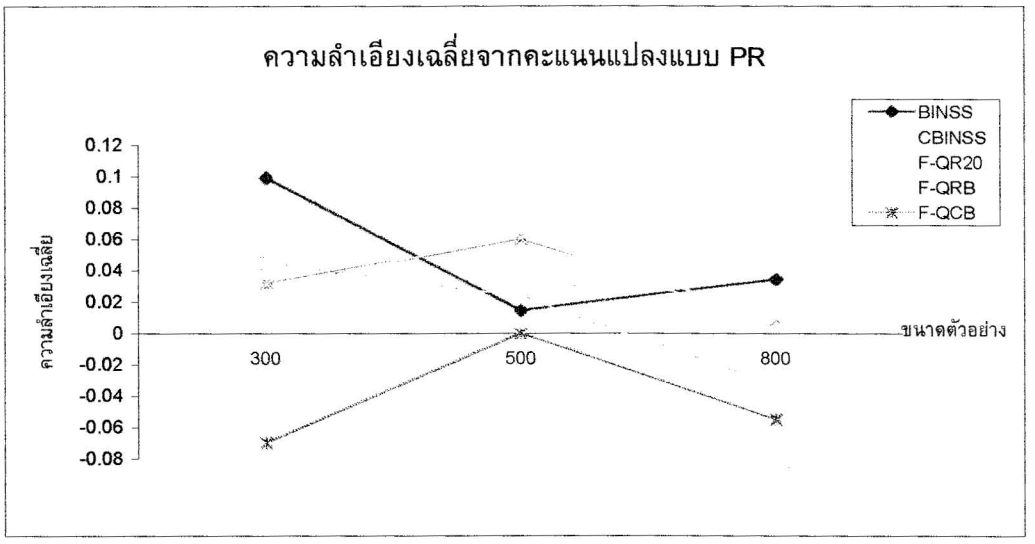
ตาราง 23 ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ

แบบทดสอบ	ค่าสถิติ	ขนาดตัวอย่าง		
		300 คน	500 คน	800 คน
ด้านตัวเลข	abias	.00358	-.00045	-.00017
	aev	.00013	.00017	.00005
ด้านภาษา	abias	-.04457	-.04717	-.05009
	aev	.00219	.00242	.00256
ด้านเหตุผล	abias	.00298	-.00179	.00229
	aev	.00012	.00007	.00005

ผลการวิเคราะห์ตามตาราง 23 พบว่า ค่าความลำเอียงเฉลี่ยเมื่อใช้วิธีประมาณค่าโดยทฤษฎีการตอบข้อสอบ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 500 และ 800 คน แบบทดสอบวัดความถนัดทางการเรียนด้านภาษาให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 300 500 และ 800 คน ส่วนแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลมีค่าเป็นลบเมื่อตัวอย่างมีขนาด 500 คน

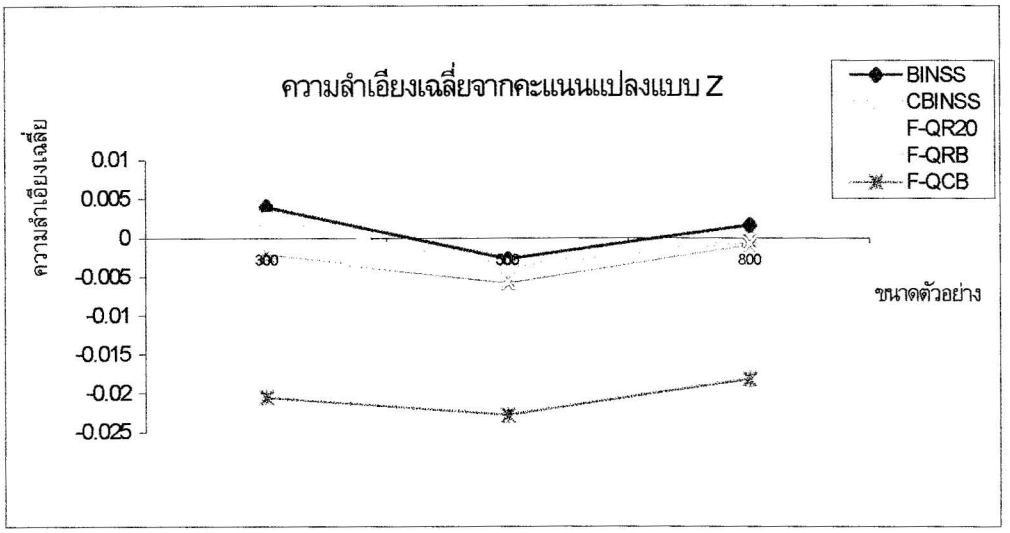
ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในแบบทดสอบฉบับเดียวกันเมื่อขนาดตัวอย่างต่างกันให้ค่าใกล้เคียงกัน และพบว่าในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยสูงสุด

เพื่อให้เห็นความแตกต่างของค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงในแบบทดสอบแต่ละฉบับเมื่อใช้วิธีการประมาณค่าต่างกัน ผู้วิจัยจึงนำเสนอเป็นกราฟดังภาพประกอบ 11 ถึง 24



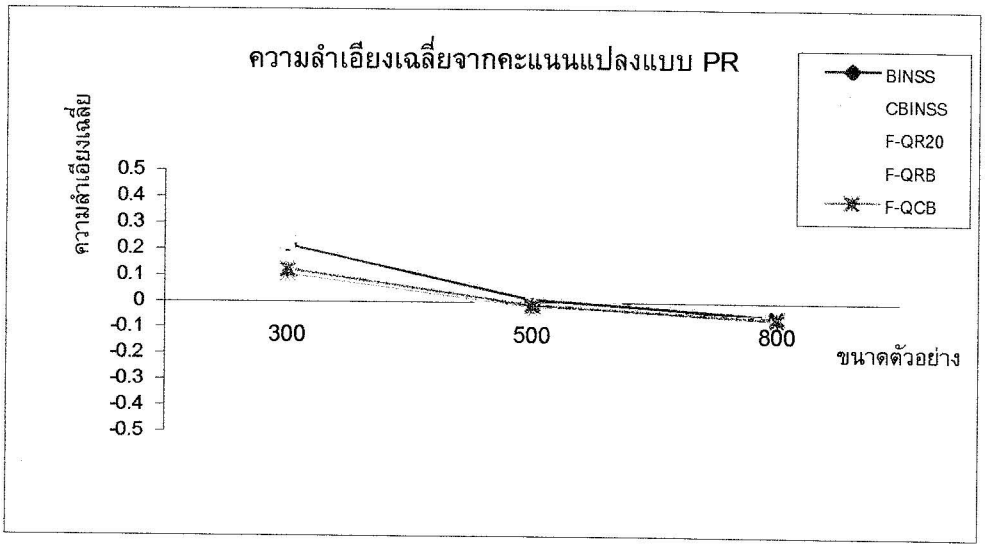
ภาพประกอบ 11 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 11 เส้นกราฟที่แสดงความลำเอียงเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้การประมาณค่าวิธีของเฟลต์ – ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้เส้นกราฟเกือบเป็นเส้นเดียวกัน แต่วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลต์ – ควอลส์สูตร F-Q_{CB} ให้ค่าความลำเอียงเฉลี่ยแตกต่างกันไปตามขนาดของตัวอย่าง เส้นกราฟขึ้นลงไม่เป็นไปในทางเดียวกัน



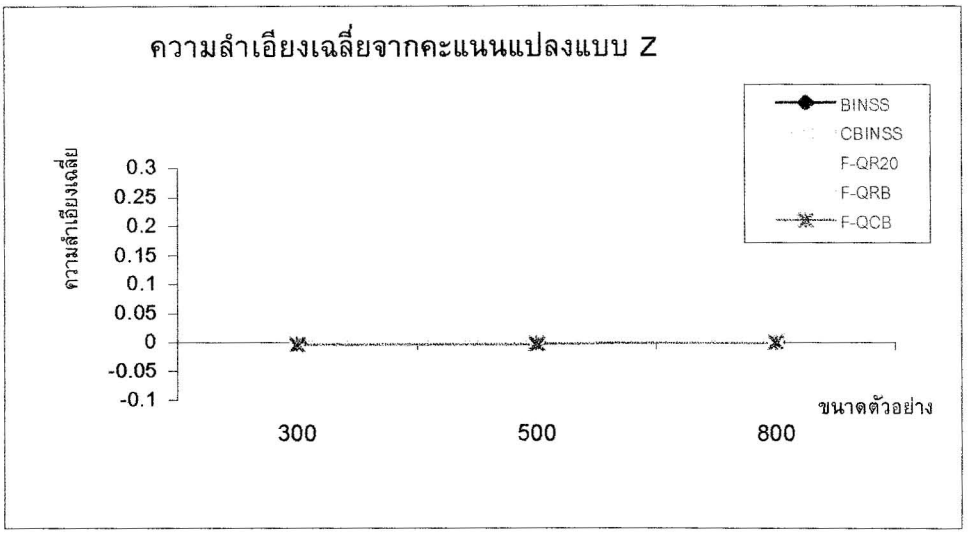
ภาพประกอบ 12 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 12 เส้นกราฟที่แสดงความลำเอียงเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อใช้วิธีประมาณค่าแต่ละวิธีให้เส้นกราฟที่ค่อนข้างเป็นไปในแนวทางเดียวกัน โดยที่วิธีทวินามให้ค่าความลำเอียงเฉลี่ยสูงสุด วิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{CB}$ ให้ค่าต่ำสุด วิธีทวินามและวิธีทวินามประกอบให้เส้นกราฟเกือบเป็นเส้นเดียวกัน แต่ละวิธีค่าความลำเอียงเฉลี่ยที่ต่ำสุดจะเป็นลบตรงเมื่อขนาดตัวอย่าง 500 คน วิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$, $F\text{-}Q_{RB}$ และ $F\text{-}Q_{CB}$ ให้ค่าความลำเอียงเฉลี่ยต่ำสุดเมื่อตัวอย่างจำลองมีขนาด 500 เช่นเดียวกับวิธีทวินามและวิธีทวินามประกอบแต่ที่ขนาดตัวอย่าง 800 คนค่าความลำเอียงเฉลี่ยสูงขึ้น



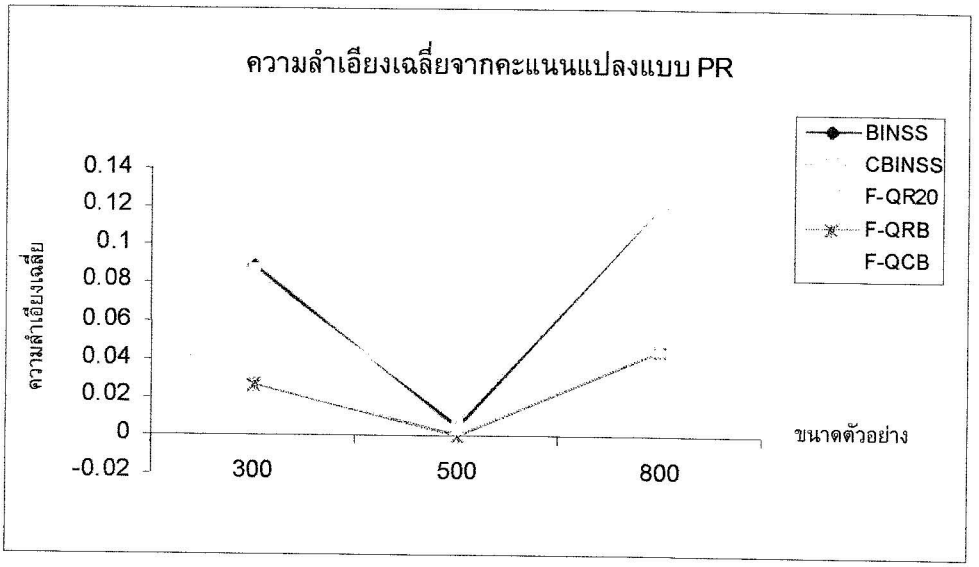
ภาพประกอบ 13 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 13 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา พบว่า วิธีวิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} , F-Q_{RB} ให้เส้นกราฟเกือบเป็นเส้นเดียวกัน วิธีเฟลต์ – ควอลส์สูตร F-Q_{CB} จากตัวอย่างแต่ละขนาดให้ค่าสูงแตกต่างจากวิธีวิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} , F-Q_{RB} เส้นกราฟจะต่ำสุดเมื่อขนาดตัวอย่าง 800 คน และเส้นกราฟสูงเมื่อตัวอย่างมีขนาด 300 คน



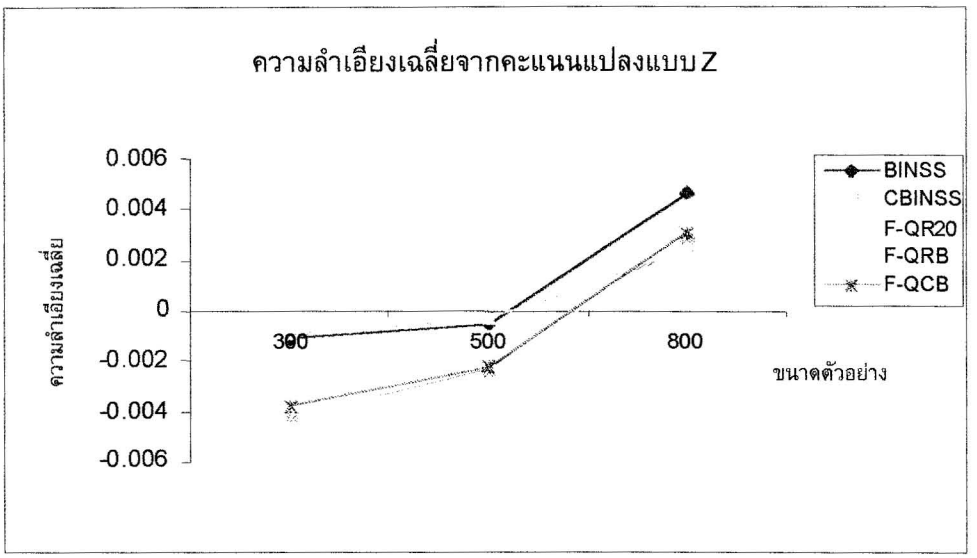
ภาพประกอบ 14 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ - ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 14 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา พบว่า วิธีเฟลต์ – ควอลส์สูตร F-Q_{CB} จากตัวอย่างแต่ละขนาดให้ค่าต่ำแตกต่างจากวิธีวิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} และสูตร F-Q_{RB} ซึ่งให้เส้นกราฟเกือบเป็นเส้นเดียวกัน แต่ละวิธีเส้นกราฟจะเกือบอยู่ตรงแกนที่มีค่าเป็นศูนย์



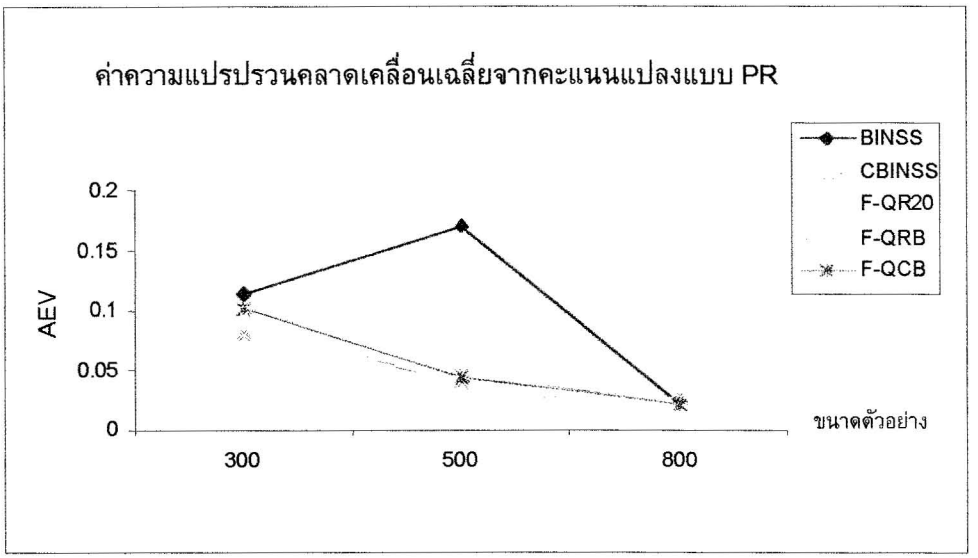
ภาพประกอบ 15 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 15 เส้นกราฟที่แสดงความลำเอียงเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลค่อนข้างเหมือนกัน นั่นคือให้ค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 500 คน วิธีเฟลด์ – ควอลส์สูตร F-Q_{CB} จะให้ค่าสูงต่างไปจากวิธีทวินาม วิธีทวินามประกอบ วิธีเฟลด์ – ควอลส์ สูตร F-Q_{R20} และสูตร F-Q_{RB}



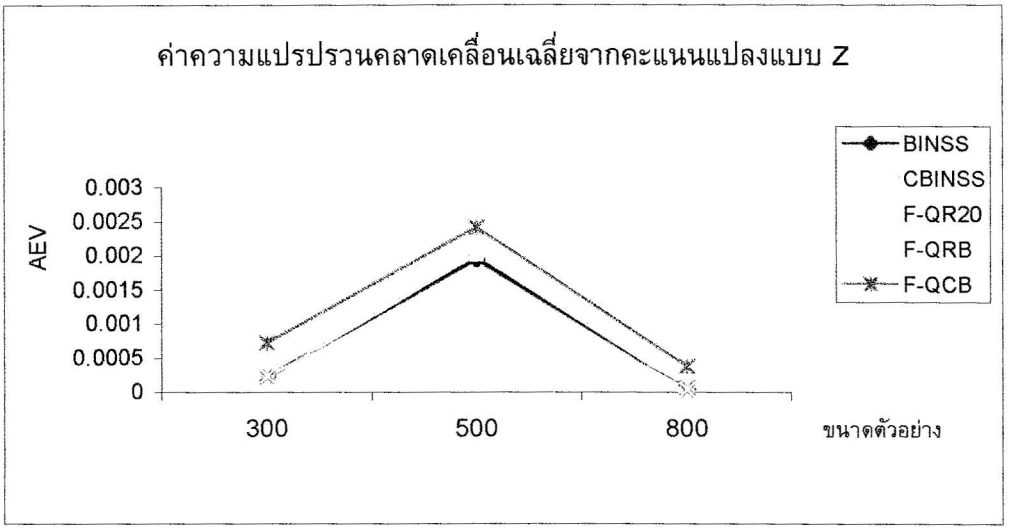
ภาพประกอบ 16 ค่าความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 16 เส้นกราฟแสดงความลำเอียงเฉลี่ยจากคะแนนแปลงรูปแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล วิธีเฟลด์ – ควอลส์สูตร $F\text{-}Q_{CB}$ ให้ค่าต่ำสุดในทุกขนาดตัวอย่าง วิธีทวินามให้ค่าสูงสุดเมื่อขนาดตัวอย่าง 800 คน และค่าความลำเอียงเฉลี่ยในแต่ละวิธีให้ค่าต่ำสุดเมื่อขนาดตัวอย่าง 300 คน



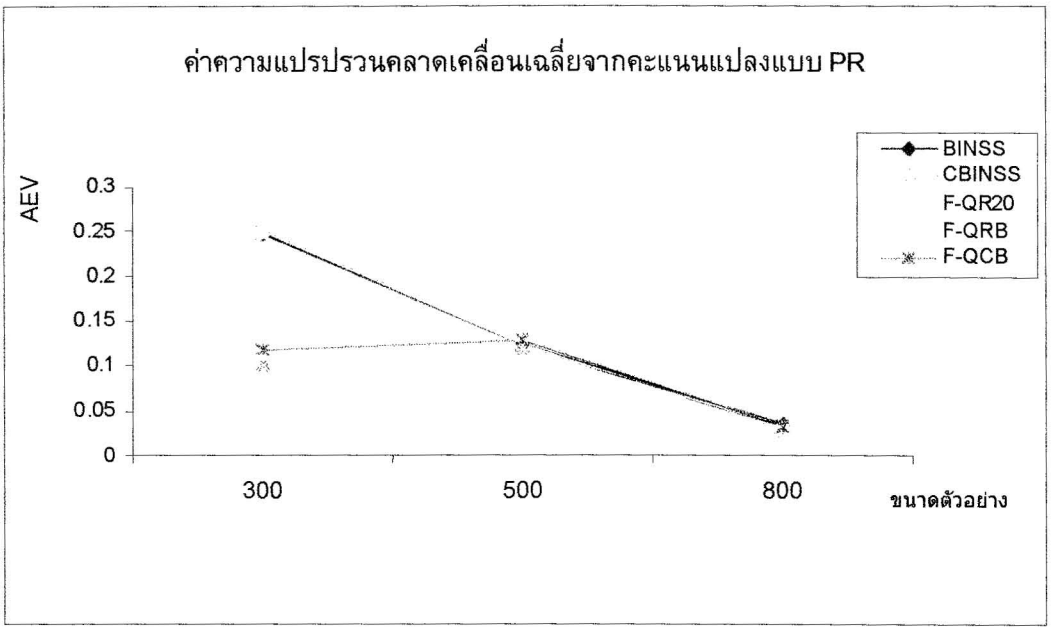
ภาพประกอบ 17 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 17 กราฟแสดงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข วิธีทวินามจะให้ค่าสูงสุดเมื่อขนาดตัวอย่างเท่ากับ 300 และ 500 คน วิธีเฟลต์ – ควอลส์ สูตร F-Q_{RB} ให้ค่าสูงสุดเมื่อขนาดตัวอย่างเท่ากับ 800 คน วิธีทวินามประกอบให้ค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 800 คน



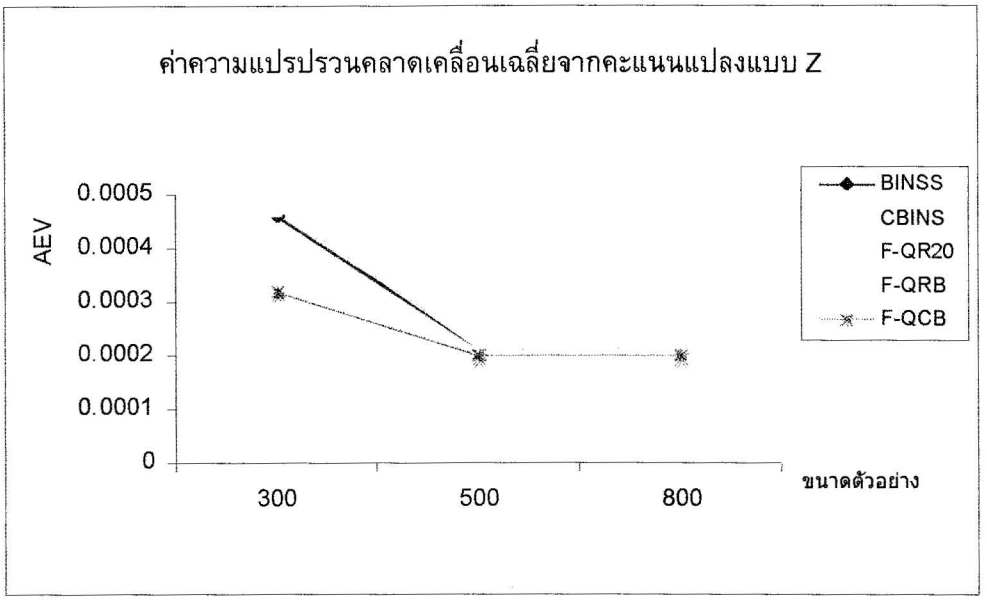
ภาพประกอบ 18 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 18 กราฟแสดงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข วิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{CB}$ ให้ค่าสูงสุดในทุกขนาดตัวอย่าง วิธีทวินาม วิธีทวินามและวิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ $F\text{-}Q_{RB}$ ให้ค่าใกล้เคียงกันมากเมื่อขนาดตัวอย่างเท่ากับ 300 และ 800 คน และวิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$ ให้ค่าต่ำสุดในทุกขนาดตัวอย่าง



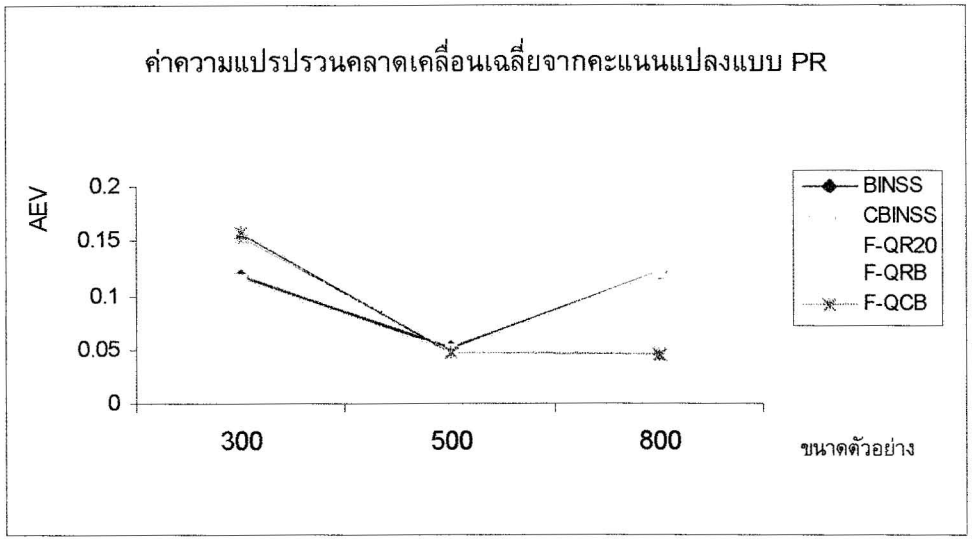
ภาพประกอบ 19 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 19 กราฟแสดงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา วิธีเฟลต์ – ควอลส์ สูตร $F\text{-}Q_{CB}$ เส้นกราฟสูงในทุกขนาดตัวอย่าง วิธีทวินาม วิธีทวินามและวิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ $F\text{-}Q_{RB}$ ให้ค่าใกล้เคียงกันมากในทุกตัวอย่างเช่นเดียวกัน



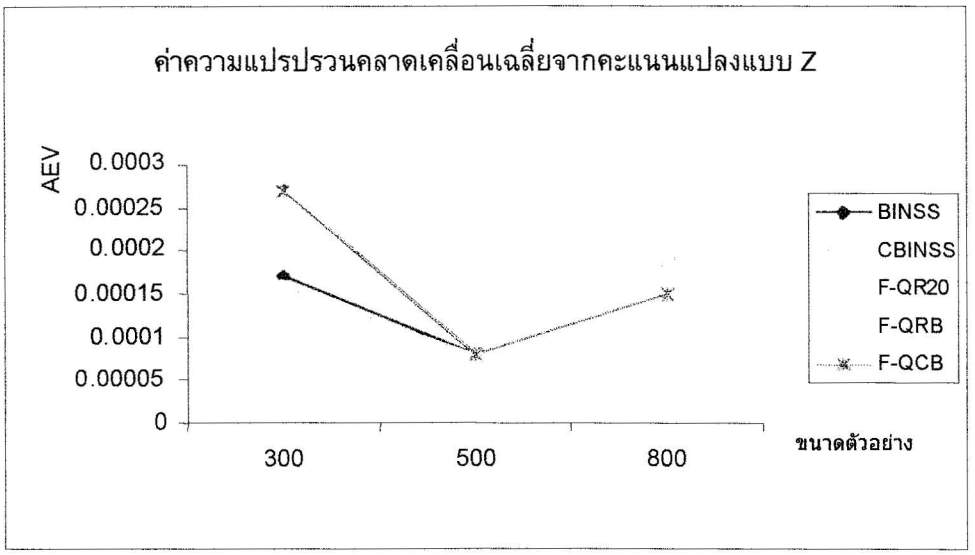
ภาพประกอบ 20 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 20 กราฟแสดงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา แต่ละวิธีเมื่อขนาดตัวอย่างมากขึ้นค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยมีแนวโน้มลดลง ยกเว้นวิธีทวินามเมื่อขนาดตัวอย่างเท่ากับ 800 คน ค่ากลับสูงขึ้นเล็กน้อย สำหรับวิธีเฟลต์ – ควอลส์ สูตร $F\text{-}Q_{CB}$ เส้นกราฟสูงในทุกขนาดตัวอย่าง และวิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ $F\text{-}Q_{RB}$ ให้ค่าใกล้เคียงกันมากในทุกตัวอย่าง



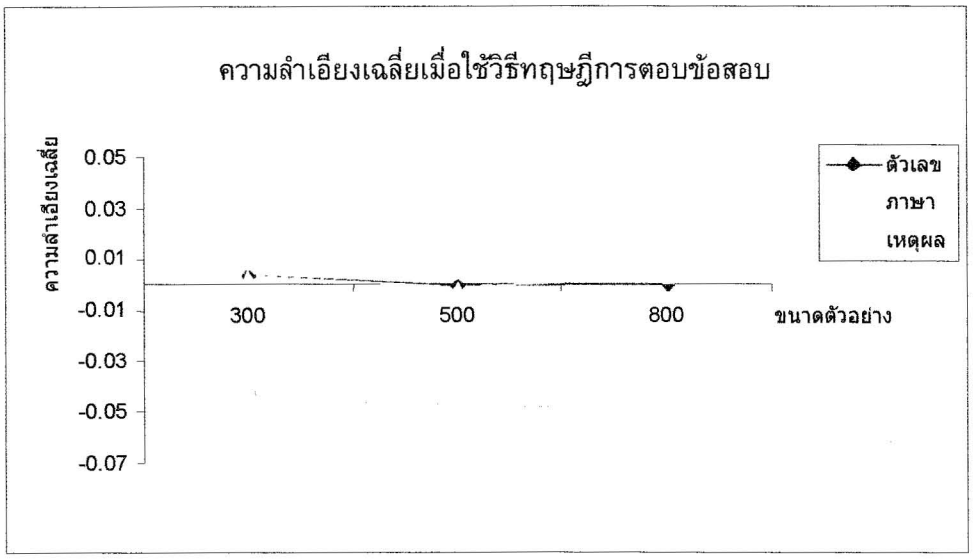
ภาพประกอบ 21 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างจำลองขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 21 กราฟแสดงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล วิธีเฟลต์ – ควอลส์ สูตร $F\text{-}Q_{CB}$ เส้นกราฟสูงในทุกขนาดตัวอย่าง วิธีทวินามและวิธีทวินามประกอบให้ค่าใกล้เคียงกันในทุกขนาดตัวอย่างเช่นเดียวกับวิธีเฟลต์ – ควอลส์สูตร $F\text{-}Q_{R20}$ และ $F\text{-}Q_{RB}$ ให้ค่าใกล้เคียงกันในทุกขนาดตัวอย่าง



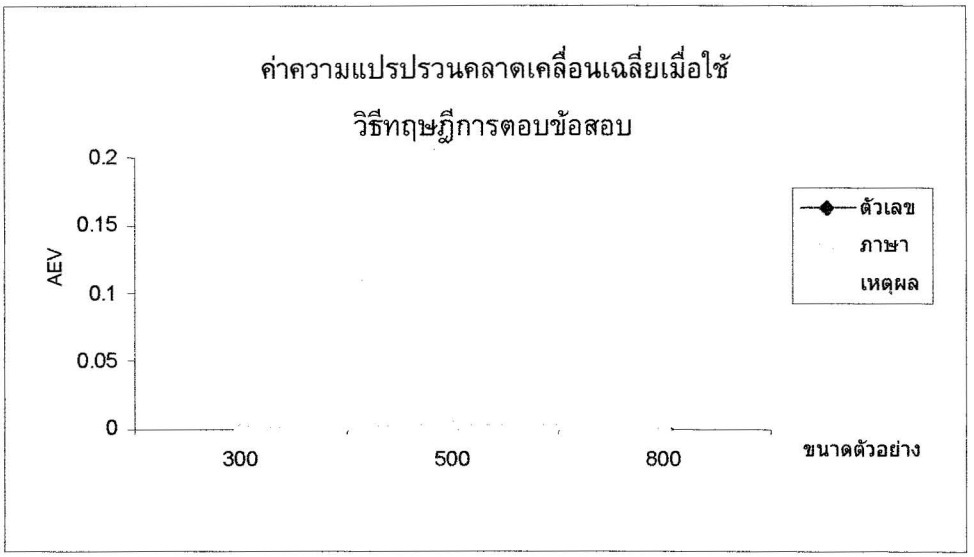
ภาพประกอบ 22 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของตัวอย่างขนาด 300 500 และ 800 คน เมื่อใช้การประมาณค่าวิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลต์ – ควอลส์

ผลการวิเคราะห์ตามภาพประกอบ 22 กราฟแสดงค่าความแปรปรวนเฉลี่ยจากคะแนนแปลงแบบคะแนนปกติ ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล วิธีทวินามและวิธีทวินามประกอบ วิธีเฟลต์ – ควอลส์สูตร F-Q_{R20} และ F-Q_{RB} ให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเท่ากัน เส้นกราฟเป็นเส้นเดียวกัน วิธีเฟลต์ – ควอลส์สูตร F-Q_{CB} ให้ค่าสูงเมื่อกลุ่มตัวอย่างมีขนาด 300 คนและ 500 คน แต่วิธีทวินามประกอบให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 800 คน และก็ให้ค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 500 คน แต่วิธีเฟลต์ – ควอลส์สูตร F-Q_{CB} ให้ค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 800 คน



ภาพประกอบ 23 ความลำเอียงเฉลี่ยเมื่อใช้วิธีทฤษฎีการตอบข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ผลการวิเคราะห์ตามภาพประกอบ 23 ความลำเอียงเฉลี่ยในแบบทดสอบวัดความถนัดทางการเรียนเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ พบว่า ด้านภาษาในทุกขนาดตัวอย่างให้เส้นกราฟต่ำสุด แต่ด้านตัวเลขและด้านเหตุผลให้ค่าเกือบเป็นเส้นกราฟเดียวกัน ด้านตัวเลขมีค่าเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 500 คน และ 800 คน แต่ด้านเหตุผลเส้นกราฟอยู่ต่ำกว่าแกนกลางเมื่อขนาดตัวอย่างเท่ากับ 500 คน



ภาพประกอบ 24 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อใช้วิธีทฤษฎีการตอบข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ผลการวิเคราะห์ตามภาพประกอบ 24 ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในแบบทดสอบวัดความถนัดทางการเรียนเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ พบว่า แบบทดสอบแต่ละฉบับที่ได้จากตัวอย่างแต่ละขนาดปรากฏว่าเส้นกราฟนั้นเกือบอยู่บนแกนกลางของกราฟ มีค่าต่ำเกือบเป็นศูนย์

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ เมื่อใช้วิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์ โดยใช้คะแนนที่ได้จากการตอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล ของนักเรียนชั้นมัธยมศึกษาตอนปลาย ว่ามีความคงที่หรือไม่ และศึกษาคุณสมบัติของตัวประมาณค่าว่ามีความแม่นยำหรือไม่ โดยพิจารณาจากค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย ซึ่งสรุปสาระสำคัญได้ดังนี้

จุดมุ่งหมายของการวิจัย

1. เพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ จากคะแนนแบบทดสอบวัดความถนัดทางการเรียน
2. เพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้การประมาณค่าด้วยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ - ควอลส์ จากคะแนนแบบทดสอบวัดความถนัดทางการเรียน
3. เพื่อศึกษาความแม่นยำในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัด ได้แก่ ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย จากคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ โดยใช้วิธีประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขวิธีทฤษฎีการตอบข้อสอบ วิธีแบบทวินาม วิธีแบบทวินามประกอบ และวิธีของเฟลด์ - ควอลส์

วิธีดำเนินการ

ข้อมูลที่นำมาศึกษา

ข้อมูลที่นำมาศึกษาในการวิจัยครั้งนี้เป็นคะแนนที่ได้จากการตอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา ด้านเหตุผล ของนักเรียนชั้นมัธยมศึกษาตอนปลาย สังกัดกรมสามัญศึกษา จังหวัดมหาสารคาม ขอนแก่น ยโสธร และนครพนม ปีการศึกษา 2544 จำนวน 2000 คน ให้เป็นประชากรเทียม (Pseudo - population) และทำการสุ่มประชากรเทียมนี้โดยใช้โปรแกรมคอมพิวเตอร์ SPSS for Windows ให้เป็นตัวอย่างจำลองขนาด 300 500 และ 800 คน แต่ละขนาดทำการสุ่มซ้ำ 30 ครั้ง ได้ข้อมูลที่จำลองมาจากตัวอย่างขนาดใหญ่ 270 ตัวอย่าง นำข้อมูลที่สุ่มมานี้ไปศึกษาถึงคุณสมบัติของตัวประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง เมื่อใช้วิธีการประมาณค่าต่างกัน

การเก็บรวบรวมข้อมูลและจัดกระทำกับข้อมูล

ผู้วิจัยได้เก็บรวบรวมข้อมูลตั้งแต่วันที่ 25 สิงหาคม - 30 ตุลาคม 2544 ตามขั้นตอนดังนี้

1. ทำหนังสือขอความร่วมมือจากผู้อำนวยการโรงเรียนมัธยมศึกษาที่เป็นกลุ่มตัวอย่างในการเก็บรวบรวมข้อมูล เพื่อกำหนดวัน เวลา และสถานที่สอบ

2. ดำเนินการสอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ภาษา และเหตุผล กับกลุ่มตัวอย่าง ก่อนการสอบได้ชี้แจงให้นักเรียนกลุ่มตัวอย่างทราบถึงวัตถุประสงค์ของการสอบและโน้มน้าวจิตใจให้นักเรียนมีความร่วมมือในการสอบเพื่อให้ได้ผลการสอบที่เชื่อถือได้

3. นำผลการตอบแบบทดสอบแต่ละฉบับมาตรวจสอบว่ามีการแจกแจงแบบปกติหรือไม่ ผลปรากฏว่าข้อมูลที่ได้จากการตอบแบบทดสอบทั้งสามฉบับมีการแจกแจงแบบไม่ปกติ ดังนั้น ผู้วิจัยจึงได้เก็บข้อมูลเพิ่มเติมจากนักเรียนโรงเรียนสตรีร้อยเอ็ด ขอนแก่นวิทยายน ปิยะมหาราชาลัย ยโสธรพิทยาคม และคำเขื่อนแก้วชนูปถัมภ์ และนำผลมาตรวจสอบการแจกแจงแบบปกติ ผลปรากฏว่าข้อมูลที่ได้มีการแจกแจงเข้าสู่การแจกแจงแบบปกติ (Approximately Normality)

4. กลุ่มตัวอย่างที่ตอบแบบทดสอบทั้ง 3 ฉบับนี้ ผู้วิจัยกำหนดให้เป็นประชากรเทียม (Pseudo - population) จากนั้นจึงสุ่มคะแนนของนักเรียนในประชากรเทียมออกเป็นกลุ่มตัวอย่างจำลองขนาด 300 500 และ 800 คน แต่ละขนาดสุ่มซ้ำ 30 ครั้ง

เมื่อเก็บรวบรวมข้อมูลได้แล้ว ผู้วิจัยได้จัดกระทำข้อมูลดังนี้

สำหรับกลุ่มตัวอย่างขนาดใหญ่

การจัดกระทำข้อมูลกับกลุ่มตัวอย่างขนาดใหญ่ ดำเนินการดังนี้

1. นำข้อมูลการตอบของนักเรียนแต่ละคนและบันทึกลงในไฟล์แต่ละไฟล์

2. วิเคราะห์ข้อมูลเบื้องต้น ได้แก่ ค่าสถิติพื้นฐานคือ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ ความโด่ง ค่าความเชื่อมั่นของแบบทดสอบแต่ละฉบับ โดยใช้โปรแกรม SPSS for WINDOWS และโปรแกรม r_B

3. ตรวจสอบการแจกแจงของคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ ปรากฏว่าคะแนนจากแบบทดสอบแต่ละฉบับมีการแจกแจงเข้าสู่การแจกแจงแบบปกติ (Approximately Normality)

4. ตรวจสอบความเป็นมิติเดียวของคะแนนที่ได้จากแบบทดสอบแต่ละฉบับ ปรากฏว่าคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนทุกฉบับมีความเป็นมิติเดียว

5. แปลงคะแนนดิบของนักเรียนแต่ละคนจากแบบทดสอบแต่ละฉบับเป็นคะแนนตำแหน่งเปอร์เซ็นไทล์ และคะแนนแบบปกติ โดยใช้โปรแกรม T- score และ โปรแกรมสำเร็จรูป Exel

6. ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านตามวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ

และวิธีเฟลด์-ควอลส์ โดยใช้โปรแกรม CBINSS และโปรแกรมเฟลด์-ควอลส์ ที่ผู้วิจัยได้พัฒนาขึ้น

สำหรับตัวอย่างจำลอง

เมื่อได้ข้อมูลจากกลุ่มตัวอย่างขนาดใหญ่แล้วกำหนดให้เป็นประชากรเทียม และดำเนินการสุ่มให้เป็นกลุ่มตัวอย่างจำลอง มีวิธีการดังนี้

1. จากข้อมูลขนาดใหญ่ที่บันทึกลงในแต่ละไฟล์ ทำการสุ่มโดยใช้โปรแกรม SPSS for WINDOWS ให้ได้ขนาด 300 500 และ 800 คน แต่ละขนาดของตัวอย่างสุ่มซ้ำ 30 ครั้ง ดังนั้นจากคะแนนแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านได้ข้อมูลเป็นตัวอย่างจำลอง 90 ตัวอย่าง รวมเป็นข้อมูลที่เป็นตัวอย่างจำลองทั้งหมด 270 ตัวอย่าง

2. จากข้อมูลที่สุ่มได้แต่ละตัวอย่างของแต่ละขนาด นำมาหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และความเชื่อมั่น และนำคะแนนมาแปลงเป็นตำแหน่งเปอร์เซ็นไทล์ และคะแนนแบบปกติ

3. ประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงจากแบบทดสอบแต่ละฉบับตามวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์-ควอลส์ โดยใช้โปรแกรม CBINSS และโปรแกรมเฟลด์ - ควอลส์

4. ศึกษาคุณสมบัติของตัวประมาณค่าโดยพิจารณาจากค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแต่ละแบบที่มีวิธีการประมาณค่าต่างกัน

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบทดสอบวัดความถนัดทางการเรียนที่สร้างตามแนวของเทอร์สโตน (Thurstone) มี 3 ด้าน ซึ่งคัดเลือกมาจากแบบทดสอบของโครงการวิจัยเกี่ยวกับการสร้างและพัฒนาแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย ประจำปีงบประมาณ 2539 - 2541 (สุทธิวรรณ พีรศักดิ์โสภณ. 2540 และ 2542) เป็นแบบปรนัยเลือกตอบห้าตัวเลือก แต่ละข้อที่คัดเลือกมาใช้ในการวิจัยมีค่าความยากตั้งแต่ .20 - .80 และค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป ก่อนนำไปใช้จริงผู้วิจัยได้ทดลองอีกครั้งกับกลุ่มตัวอย่างจำนวน 244 คน ได้ผลการวิเคราะห์ดังนี้

แบบทดสอบวัดความถนัดทางด้านตัวเลข (Numerical Ability) ประกอบด้วยเรื่องทักษะการคำนวณ คณิตศาสตร์เหตุผล และอนุกรมธรรมดา จำนวน 30 ข้อ คุณภาพด้านความเชื่อมั่นเมื่อใช้สูตร KR -20 , KR -21 และสูตร r_B มีค่า .8361, .8136 และ .8386 ตามลำดับ มีค่าความยากตั้งแต่ .21 - .78 ค่าอำนาจจำแนกตั้งแต่ .17 -.81

แบบทดสอบวัดความถนัดทางด้านภาษา (Verbal Ability) ประกอบด้วยเรื่องคำตรงข้ามและความเข้าใจภาษา จำนวน 40 ข้อ คุณภาพด้านความเชื่อมั่นเมื่อใช้สูตร KR -20 , KR -21 และ

สูตร r_B มีค่า .7371, .7088 และ .7408 ตามลำดับ มีค่าความยากตั้งแต่ .15 - .81 และค่าอำนาจจำแนกตั้งแต่ .11 - .57

แบบทดสอบวัดความถนัดทางด้านเหตุผล (Reasoning Ability) ประกอบด้วยเรื่องการจำแนกประเภท อุปมาอุปไมย และอนุกรมภาพ จำนวน 30 ข้อ คุณภาพด้านความเชื่อมั่นเมื่อใช้สูตร KR -20 , KR -21 และสูตร r_B มีค่า .7730, .7612 และ .7793 ตามลำดับ มีค่าความยากตั้งแต่ .17 - .61 ค่าอำนาจจำแนกตั้งแต่ .20 - .69

สรุปผลการวิจัย

ผู้วิจัยสรุปผลการวิจัย ได้ดังนี้

1. ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านที่ได้จากตัวอย่างขนาดใหญ่ และจากตัวอย่างจำลองขนาด 300 500 และ 800 คน ดังสรุปในตาราง 24 และ 25

ตาราง 24 สรุปค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่

ค่าสถิติ	แบบทดสอบวัดความถนัดทางการเรียน		
	ด้านตัวเลข	ด้านภาษา	ด้านเหตุผล
จำนวนข้อ	30	40	30
คะแนนต่ำสุด	1	3	1
คะแนนสูงสุด	29	36	29
ค่าเฉลี่ย	15.2165	19.7480	14.1970
ค่าเบี่ยงเบนมาตรฐาน	6.8001	6.2342	6.5202
สัมประสิทธิ์การกระจาย	.4469	.3157	.4593
ความโด่ง	-1.040	-.6910	-.8720
ความเบ้	.1010	-.1670	.1480
KR - 20	.8819	.7819	.8609
KR - 21	.8667	.7618	.8525
r_B	.8836	.7834	.8616

สรุป แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขมีค่าเฉลี่ยเกินครึ่งหนึ่งของคะแนนเต็ม แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลมีค่าสัมประสิทธิ์การกระจายสูงสุด แบบทดสอบทั้งสามด้านมีการแจกจงโด่งต่ำ แบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีการแจกแจงเบ้ทางลบ แบบ

ทดสอบวัดความถนัดทางการเรียนด้านตัวเลขมีค่าความเชื่อมั่นสูงสุด และวิธีสัมประสิทธิ์ความเชื่อมั่น r_B ให้ค่าสูงสุดในทุกแบบทดสอบ

ตาราง 25 สรุปค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างจำลองขนาด 300 500 และ 800 คน

ค่าสถิติ	แบบทดสอบวัดความถนัดทางการเรียน								
	ด้านตัวเลข			ด้านภาษา			ด้านเหตุผล		
	300 คน	500 คน	800 คน	300 คน	500 คน	800 คน	300 คน	500 คน	800 คน
ค่าเฉลี่ยทั้งหมด	15.2533	15.2412	15.2352	19.6953	19.7647	19.7156	14.3239	14.2260	14.2322
S เฉลี่ย	6.8259	6.7868	6.8130	6.1874	6.2484	6.2660	6.5008	6.5419	6.6975
C.V. เฉลี่ย	.4476	.4454	.4473	.3143	.3162	.3178	.4539	.4600	.4707
ความโด่งเฉลี่ย	-1.0490	-1.0500	-1.0470	-.6470	-.6760	-.7080	-.8676	-.8751	-.8788
ความเบ้เฉลี่ย	.1080	.0980	.1000	-.1610	-.1550	-.1690	.1249	.1425	.1401
KR - 20 เฉลี่ย	.8831	.8812	.8825	.7785	.7830	.7842	.8598	.8624	.8606
KR - 21 เฉลี่ย	.8677	.8660	.8673	.7572	.7625	.7643	.8511	.8536	.8564
r_B เฉลี่ย	.8848	.8829	.8841	.7802	.7846	.7858	.8608	.8340	.8614

สรุป ค่าสถิติพื้นฐานและความเชื่อมั่นจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อขนาดตัวอย่างเป็น 300 500 และ 800 คน มีดังนี้

1. ค่าเฉลี่ยของคะแนนแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขจากตัวอย่างแต่ละขนาดส่วนใหญ่มีค่าเกินครึ่งหนึ่งของคะแนนเต็ม แต่ในแบบทดสอบด้านภาษาและเหตุผลส่วนใหญ่มีค่าไม่เกินครึ่งหนึ่งของคะแนนเต็ม

2. ค่าสัมประสิทธิ์การกระจายเฉลี่ยของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลมีค่าสูงกว่าแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านภาษาในทุกขนาดตัวอย่าง

3. รูปร่างการแจกแจงคะแนนของแบบทดสอบวัดความถนัดทางการเรียนทุกฉบับมีความโด่งเป็นลบในทุกขนาดตัวอย่าง นั่นคือข้อมูลมีการแจกแจงที่มียอดสูง แต่ความเบ้ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผลส่วนใหญ่เบ้ขวาในทุกขนาดตัวอย่าง และแบบทดสอบด้านภาษาในทุกขนาดตัวอย่างมีลักษณะเบ้ซ้าย

4. ค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อคำนวณจากสูตร KR -20, KR -21 และ r_B ทุกขนาดตัวอย่างสูตร r_B ให้ค่าสูงสุด ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข เมื่อขนาดตัวอย่าง 500 คน ค่าความเชื่อมั่นที่ได้ในแต่ละวิธีมีค่าต่ำกว่าเมื่อขนาดตัวอย่างเท่ากับ 300 และ 800 คน

2. ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างขนาดใหญ่ เมื่อใช้การประมาณค่าโดยวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์ - ควอลส์ ได้ผลสรุปดังตาราง 27

ตาราง 26 สรุปค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลง (CSEM) ในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน จากตัวอย่างขนาดใหญ่

วิธี	ค่า CSEM ของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์		
	ด้านตัวเลข	ด้านภาษา	ด้านเหตุผล
BINSS	.4809 - 12.3714	.0316 - 16.7760	.6786 - 13.7955
	ค่าสูงสุดตรงกับ PR 45.58	ค่าสูงสุดตรงกับ PR 54.60	ค่าสูงสุดตรงกับ PR 45.55
CBINSS	.4207 -12.2344	.8872 - 16.6930	.6442 - 13.5081
	ค่าสูงสุดตรงกับ PR 50.08	ค่าสูงสุดตรงกับ PR 60.15	ค่าสูงสุดตรงกับ PR 45.55
F-Q_{r20}	.5532 - 11.7244	.0172 - 16.4659	.7769 - 13.7704
	ค่าสูงสุดตรงกับ PR 50.08	ค่าสูงสุดตรงกับ PR 54.60	ค่าสูงสุดตรงกับ PR 45.55
F-Q_{rB}	.5497 - 11.6497	.0172 - 16.4094	.7750 - 13.7354
	ค่าสูงสุดตรงกับ PR 50.08	ค่าสูงสุดตรงกับ PR 54.60	ค่าสูงสุดตรงกับ PR 45.55
F-Q_{CB}	.5875 - 12.2346	.2287 - 17.0345	.8000 - 13.8015
	ค่าสูงสุดตรงกับ PR 50.08	ค่าสูงสุดตรงกับ PR 54.60	ค่าสูงสุดตรงกับ PR 45.55

หมายเหตุ ค่าต่ำสุดเป็นค่าที่ไม่นับค่าศูนย์

ตาราง 26 (ต่อ)

วิธี	ค่า CSEM ของคะแนนแปลงแบบคะแนนปกติ		
	ด้านตัวเลข	ด้านภาษา	ด้านเหตุผล
BINSS	.3259 - .5759	.0248 - .6479	.3256 - .5759
	ค่าสูงสุดตรงกับ Z 1.85	ค่าสูงสุดตรงกับ Z 1.83	ค่าสูงสุดตรงกับ Z -2
CBINSS	.3289 - .5827	.4597 - .6500	.3289 - .5827
	ค่าสูงสุดตรงกับ Z 1.85	ค่าสูงสุดตรงกับ Z 1.83	ค่าสูงสุดตรงกับ Z -2
F-Q_{r20}	.2284 - .6509	.0747 - .6408	.2768 - .6038
	ค่าสูงสุดตรงกับ Z -1.98	ค่าสูงสุดตรงกับ Z 2.34	ค่าสูงสุดตรงกับ Z -1.67
F-Q_{rB}	.2269 - .6467	.0744 - .6386	.2761 - .6023
	ค่าสูงสุดตรงกับ Z-1.98	ค่าสูงสุดตรงกับ Z 2.34	ค่าสูงสุดตรงกับ Z -1.67
F-Q_{CB}	.2425 - 6877	.2653 - .6759	.2850 - .6196
	ค่าสูงสุดตรงกับ Z -1.98	ค่าสูงสุดตรงกับ Z 2.34	ค่าสูงสุดตรงกับ Z -1.67
IRTSS	.1350 - .4075	.0194 - .4958	.0896 - .4177
	ค่าสูงสุดตรงกับ θ -.43	ค่าสูงสุดตรงกับ θ -.27	ค่าสูงสุดตรงกับ θ -.84

หมายเหตุ ค่าต่ำสุดเป็นค่าที่ไม่นับค่าศูนย์

สรุป ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ (PR) จากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านให้ค่าสูงตรงช่วงกลางของตำแหน่งเปอร์เซ็นไทล์ และให้ค่าลดลงต่ำลงตรงด้านปลายของคะแนนแปลง รูปร่างของเส้นกราฟคล้ายรูปพาราโบลาคว่ำหรือรูปตัวยูหัวกลับ และคะแนนตำแหน่งเปอร์เซ็นไทล์ที่ตรงกับคะแนนดิบเป็นศูนย์หรือคะแนนเต็มค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าเป็นศูนย์ทุกวิธีที่ใช้ในการประมาณค่า

ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข วิธีทวินามให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงตรงกับ PR ที่ 45.58 วิธีทวินามประกอบและวิธีเฟลด์ – ควอลส์ที่คำนวณจากสูตร F-Q_{R20}, F-Q_{RB} และ F-Q_{CB} ให้ค่าสูงสุดตรงคะแนนแปลง PR ที่ 50 และค่าที่สูงสุดเป็นค่าที่ได้จากวิธีทวินาม สำหรับแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา วิธีทวินาม วิธีเฟลด์ – ควอลส์ทุกสูตรให้ค่าความคลาดเคลื่อนสูงสุดตรงคะแนน PR ที่ 54.60 แต่วิธีทวินามประกอบให้ค่าสูงสุดตรงคะแนน PR ที่ 60.15 และค่าที่สูงสุดเป็นค่าที่ได้จากวิธีเฟลด์ – ควอลส์สูตร F-Q_{CB} แต่ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลทุกวิธีที่ใช้ในการประมาณค่าให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุดตรงคะแนน PR ที่ 45.55 และค่าสูงสุดเป็นค่าที่ได้จากวิธีเฟลด์ – ควอลส์สูตร F-Q_{CB}

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติ (Z) จากแบบทดสอบแต่ละฉบับมีค่าไม่คงที่ ลักษณะกราฟคล้ายโค้งสองยอด (Bimodal) ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวแลข วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงตรงกับคะแนน Z เท่ากับ 1.85 และวิธีของเฟลด์ – ควอลส์ทุกสูตรให้ค่าสูงตรงกับคะแนน Z เท่ากับ -1.98 ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงตรงกับคะแนน Z เท่ากับ 1.83 และวิธีของเฟลด์ – ควอลส์ทุกสูตรให้ค่าสูงตรงกับคะแนนเท่ากับ Z เท่ากับ 2.34 และในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงตรงกับ Z เท่ากับ -2 และวิธีของเฟลด์ – ควอลส์ทุกสูตรให้ค่าสูงตรงกับคะแนน Z เท่ากับ -1.98

สำหรับค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีทฤษฎีการตอบข้อสอบในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา มีค่าสูงกว่าในแบบทดสอบด้านตัวเลขและด้านเหตุผล

3. ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านจากตัวอย่างจำลองขนาด 300 500 และ 800 คน ที่สุ่มซ้ำขนาดละ 30 ครั้ง เมื่อใช้วิธีประมาณค่าโดยทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีเฟลด์ - ควอลส์ ได้ผลสรุปดังนี้

3.1 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนตำแหน่งเปอร์เซ็นไทล์ที่ได้จากตัวอย่างจำลองแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง เมื่อใช้วิธีการประมาณค่าแต่ละวิธีให้ค่าใกล้เคียงกัน และใกล้เคียงกับค่าที่ได้จากตัวอย่างขนาดใหญ่ เมื่อพิจารณาโดยรวมวิธีทวินามให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุดและวิธีเฟลด์ - ควอลส์สูตร F-Q_{rB} ให้ค่าต่ำสุดในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านและแต่ละขนาดตัวอย่าง

3.2 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนปกติที่ได้จากตัวอย่างจำลองแต่ละขนาดที่สุ่มซ้ำ 30 ครั้ง เมื่อใช้วิธีการประมาณค่าแต่ละวิธีมีค่าพิสัยแคบ และมีค่าใกล้เคียงกับค่าที่ได้จากตัวอย่างขนาดใหญ่ เมื่อพิจารณาโดยรวมวิธีทวินามจะให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุดและวิธีเฟลด์ - ควอลส์สูตร F-Q_{rB} ให้ค่าต่ำสุดในทุกแบบทดสอบและทุกขนาดตัวอย่าง

4. ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปจากแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านในกลุ่มตัวอย่างจำลองที่มีขนาด 300 500 และ 800 คน สรุปผลได้ดังนี้

ค่าเฉลี่ยความลำเอียงและค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติเมื่อใช้วิธีการประมาณค่าต่างกันจากตัวอย่างขนาด 300 500 และ 800 คน มีค่าต่ำและไม่แตกต่างกันมากนัก รูปร่างของเส้นกราฟจะต่างกัน

แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข วิธีทวินามและวิธีทวินามประกอบให้ค่าความลำเอียงเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน แต่วิธีเฟลด์ - ควอลส์ให้

ค่าสูงเมื่อตัวอย่างมีขนาด 500 คน และมีค่าเป็นลบสำหรับการแปลงแบบคะแนนปกติในทุกขนาดตัวอย่าง ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์วิธีทวินามให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 500 คน แต่วิธีทวินามประกอบและวิธีเฟลด์ - ควอลส์ให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน ในการแปลงแบบคะแนนปกติให้ค่าใกล้เคียงกันในทุกขนาดตัวอย่างและทุกวิธี

แบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้คะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีทวินาม วิธีทวินามประกอบและวิธีเฟลด์ - ควอลส์ให้ค่าความลำเอียงเฉลี่ยสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน วิธีทวินามและวิธีทวินามประกอบให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 800 คน แต่วิธีเฟลด์ - ควอลส์ให้ค่าเป็นบวกเมื่อตัวอย่างมีขนาด 500 และ 800 คน ในการแปลงแบบคะแนนปกติ วิธีทวินามและวิธีทวินามประกอบให้ค่าพอกันในทุกขนาดตัวอย่างโดยที่วิธีทวินามให้ค่าเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 500 คน วิธีทวินามประกอบให้ค่าเป็นลบเมื่อตัวอย่างมีขนาด 800 คน แต่วิธีเฟลด์ - ควอลส์ให้ค่าติดลบในทุกขนาดตัวอย่าง ค่าความแปรปรวนความคลาดเคลื่อนเฉลี่ยในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน แต่วิธีเฟลด์ - ควอลส์ให้ค่าสูงเมื่อตัวอย่างมีขนาด 500 คน และการแปลงแบบคะแนนปกติค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยใกล้เคียงกันในทุกขนาดตัวอย่างและทุกวิธีที่ใช้ในการประมาณค่า

แบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีทวินามและวิธีทวินามประกอบให้ค่าความลำเอียงเฉลี่ยสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน แต่วิธีเฟลด์ - ควอลส์ให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 500 คน และคะแนนแปลงแบบคะแนนปกติทุกวิธีที่ใช้ให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 800 คน แต่มีค่าเป็นลบเมื่อขนาดตัวอย่างเท่ากับ 300 และ 500 คน ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีทวินามและวิธีทวินามประกอบให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 800 คน แต่วิธีเฟลด์ - ควอลส์ให้ค่าสูงเมื่อขนาดตัวอย่างเท่ากับ 300 คน สำหรับคะแนนแปลงแบบปกติค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยให้ค่าใกล้เคียงกันในทุกวิธีและทุกขนาดตัวอย่าง

ค่าความลำเอียงเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีทฤษฎีการตอบข้อสอบในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาให้ติดลบในทุกขนาดตัวอย่าง ด้านตัวเลขและด้านเหตุผลให้ค่าใกล้เคียงกัน ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยสำหรับแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านในทุกขนาดตัวอย่างให้ค่าเกือบเป็นค่าเดียวกัน

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยในแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้วิธีทฤษฎีการตอบข้อสอบ เมื่อขนาดตัวอย่างต่างกันให้ค่าใกล้เคียงกัน

อภิปรายผล

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ความแม่นยำในการประมาณค่าโดยตรวจสอบจากค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย จากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ และคะแนนปกติ เมื่อใช้การประมาณค่าวิธีทฤษฎีการตอบข้อสอบ วิธีทวินาม วิธีทวินามประกอบ และวิธีของเฟลด์ – ควอลส์ที่คำนวณจากผลคูณของฟังก์ชั่นของคะแนนดิบ (ในที่นี้คือ $f'(X)$) กับค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนดิบตามสูตรของ คีตส์ - ลอร์ด และสูตรคำนวณโดยวิธีทวินามประกอบของคะแนนดิบ สำหรับสูตรของคีตส์ - ลอดร์ดใช้ค่าความเชื่อมั่นจากสูตร KR-20 และ r_B โดยใช้คะแนนจากการตอบแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ภาษา และเหตุผล สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย

จากการศึกษาพบว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแต่ละแบบมีค่าไม่คงที่และรูปร่างแตกต่างกันเมื่อคะแนนแปลงต่างกัน ซึ่งสอดคล้องกับงานวิจัยของโคเลน แฮนสันและเบรนนัน (Kolen, Hanson & Brennan. 1992) ที่ใช้คะแนนแปลงต่างกัน ได้แก่ ACT scale คะแนนปกติ คะแนนสแตนไนน์ และล็อก-ออด (Log – Odds) ในแบบทดสอบภาษาอังกฤษ คณิตศาสตร์ การอ่านและวิทยาศาสตร์เหตุผล ปรากฏว่าได้ค่าต่างกันมีรูปร่างของกราฟไม่เหมือนกัน

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษาและด้านเหตุผล จากวิธีประมาณค่าแบบทวินาม ทวินามประกอบ วิธีของเฟลด์ – ควอลส์ สูตร $F - Q_{R20}$,$F - Q_{RB}$ และ $F - Q_{CB}$ ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ แต่ละระดับคะแนน PR มีค่าค่อนข้างคงที่ในช่วง PR ตรงกลาง และค่าจะลดต่ำลงตรงด้านปลาย PR และค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าเป็นศูนย์เมื่อคะแนนแปลงนั้นตรงกับคะแนนดิบศูนย์หรือคะแนนเต็ม หรือตรงคะแนนแปลงที่มีค่าเท่ากัน ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาคะแนนต่ำสุดเป็น 3 และคะแนนสูงสุดเป็น 36 ดังนั้นเมื่อแปลงคะแนนแล้วตรงคะแนนส่วนทางด้านปลายจึงมีค่าเท่ากัน ทำนองเดียวกันในแบบทดสอบด้านตัวเลขและด้านเหตุผลคะแนนต่ำสุดเป็น 1 และคะแนนสูงสุดเป็น 29 ดังนั้น ค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงทางด้านปลายตรงนี้เมื่อใช้แต่ละวิธีในการประมาณค่าจึงเป็นศูนย์ ซึ่งเป็นข้อจำกัดของการศึกษาครั้งนี้ ดังนั้นถ้าจะแก้ปัญหาของคะแนนแปลงที่ตรงกับคะแนนศูนย์หรือคะแนนเต็มจึงควรใช้การปรับแก้ (Lee, Brennan & Kolen. 2000 : 10) เมื่อพิจารณาจากเส้นกราฟค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าสูงตรงกลางของคะแนนแปลงและค่อยลดต่ำลงตรงปลายของกราฟ รูปร่างกราฟคล้ายตัวยูหัวกลับหรือพาราโบลาคว่ำ อย่างไรก็ตามลักษณะของกราฟนี้สอดคล้องกับผลการวิจัยของเบรนนันและลี (Brennan & Lee. 1996) นอกจากนี้แล้วรูปร่างของกราฟคล้ายกับกราฟค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ แต่ละคะแนนดิบที่ลอร์ด (Lord. 1984) บลิกซ์และชามา (Blixt & Shama. 1986) เฟลด์ สเตฟเฟนและกุปตา (Feldt, Steffen & Gupta. 1985) ได้ศึกษาไว้ ทั้งนี้เพราะตามข้อตกลงของทฤษฎีคะแนนจริงแบบมาตรฐานเดิมค่าคาดหวังของคะแนนที่สอบได้หรือค่าเฉลี่ยของคะแนนก็คือคะแนน

จริงนั่นเอง และคะแนนจริงนี้จะได้จากค่าเฉลี่ยของการวัดซ้ำอย่างเป็นอิสระจากกัน (Allen & Yen. 1979 : 57 - 58) ดังนั้นในการสอบวัด ถ้าการแจกแจงเข้าสู่การแจกแจงแบบปกติค่าเฉลี่ยของคะแนนที่สอบได้หรือคะแนนจริงจะอยู่ตรงกึ่งกลางของโค้ง และ ณ ที่คะแนนตรงกลางตรงนี้จะเป็นค่าที่ทำให้เกิดการกระจายหรือค่าเบี่ยงเบนไปจากคะแนนที่สอบได้มากที่สุด จึงทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัด ณ ที่คะแนนตรงกลางมีค่าสูง และอีกประการหนึ่งก็คือในการแปลงคะแนนดิบเป็นคะแนน PR รูปร่างของกราฟค่าความคลาดเคลื่อนมาตรฐานในการวัดเป็นรูปตัวยูหัวกลับคล้ายของคะแนนดิบ เพราะการแปลงคะแนนแบบนี้เป็นการแปลงเกือบเป็นแนวเส้นตรง นั่นคือคะแนนที่สอบได้ยิ่งมาก ตำแหน่งเปอร์เซ็นไทล์ก็ยิ่งสูง และที่ตรงตำแหน่งเปอร์เซ็นไทล์จะเป็นตำแหน่งที่ทำให้เกิดการกระจายมากกว่าที่ตำแหน่งอื่นเช่นเดียวกัน

เมื่อพิจารณาลักษณะของกราฟค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในแบบทดสอบความถนัดทางการเรียนด้านภาษาจะค่อนข้างเบ้ซ้ายเล็กน้อย ที่เป็นดังนี้อาจเป็นเพราะค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและเหตุผลให้ค่าสูงสุดตรง PR ที่ต่ำกว่า 50 แต่ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจะให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงกว่าในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผล โดยที่วิธีทวินามและวิธีเฟลต์ – ควอลส์ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงสุดตรงที่ PR เท่ากับ 54.60 วิธีทวินามประกอบให้ค่าสูงสุดตรงกับ PR เท่ากับ 60.15 และอาจเกิดจากค่าสัมประสิทธิ์การกระจายของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าต่ำกว่าด้านตัวเลขและเหตุผล อีกทั้งแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีจำนวนข้อ 40 ข้อและแบบทดสอบแบ่งเป็นสองส่วนเท่ากันในขณะที่แบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผลมีจำนวน 30 ข้อและแบ่งเป็นสามส่วนเท่ากัน มีจำนวนข้อเท่ากันและแต่ละฉบับประกอบด้วยข้อคำถามสามส่วนแต่ละส่วนมี 10 ข้อ เมื่อจำนวนข้อไม่เท่ากันในการประมาณค่าโดยวิธีเฟลต์ - ควอลส์ ซึ่งใช้ฟังก์ชั่นของคะแนนดิบก็จะต่างกันไปตามจำนวนข้อของแบบทดสอบด้วย จึงอาจทำให้รูปร่างของกราฟค่อนข้างเบ้ไปทางซ้ายเล็กน้อย

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบคะแนนปกติในแบบทดสอบทั้งสามฉบับเมื่อใช้วิธีการประมาณค่าแบบทวินาม ทวินามประกอบ วิธีของเฟลต์ – ควอลส์ สูตร $F - Q_{R20}$,$F - Q_{RB}$ และ $F - Q_{CB}$ ปรากฏว่าให้ค่าไม่คงที่ ลักษณะของเส้นกราฟหยักขึ้นลงคล้ายโค้งสองยอดซึ่งจะแตกต่างจากเมื่อใช้คะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ที่เป็นดังนี้เพราะการแปลงแบบตำแหน่งเปอร์เซ็นไทล์เป็นการแปลงเกือบเป็นเส้นตรง คะแนนดิบยิ่งมากตำแหน่งเปอร์เซ็นไทล์ก็จะยิ่งสูงนั่นคือตำแหน่งเปอร์เซ็นไทล์จะเป็นสเกลที่บ่งบอกให้ทราบว่ามีข้อมูลอยู่ร้อยละเท่าไรที่ได้คะแนนต่ำกว่าหรือเท่ากับคะแนน ณ ตำแหน่งเปอร์เซ็นไทล์นั้น แต่คะแนนปกติ (Z) เป็นคะแนนที่บอกให้ทราบว่าคะแนนนั้นเบี่ยงเบนออกไปจากค่ากลางคิดเป็นกี่หน่วยเบี่ยงเบนมาตรฐาน ถ้าคะแนน Z เป็นบวกแสดงว่าคะแนนนั้นมีค่ามากกว่าคะแนนเฉลี่ย แต่ถ้าคะแนน Z เป็นลบแสดงว่าคะแนนนั้นมีค่าน้อยกว่าคะแนนเฉลี่ย

ด้วยเหตุนี้จึงทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขมีค่าสูงต่ำไปตามค่าของคะแนนปกติ

ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาและด้านเหตุผลทุกวิธีที่ใช้ประมาณค่าให้รูปร่างของเส้นกราฟใกล้เคียงกันในช่วง Z มีค่าระหว่าง -1 ถึง +1 แต่จะแตกต่างกันตรงด้านปลายของค่า Z ทั้งค่าบวกและลบ ซึ่งอาจเป็นเพราะมีผู้ที่ได้คะแนนสูงและต่ำเป็นจำนวนน้อยมาก แต่ในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขให้รูปร่างของเส้นกราฟค่อนข้างแตกต่างจากแบบทดสอบวัดความถนัดทางด้านภาษาและด้านตัวเลขในระหว่างค่า Z เป็น - 1 ถึง +1 ซึ่งในระหว่างค่า Z นี้ให้ค่าต่างกันเล็กน้อยในแต่ละวิธีที่ใช้ประมาณค่า ที่เป็นดังนี้อาจมาจากคะแนนเฉลี่ยของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขมีค่าสูงกว่าครึ่งหนึ่งของคะแนนเต็มแต่คะแนนเฉลี่ยแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาและด้านเหตุผลมีค่าต่ำกว่าครึ่งหนึ่งของคะแนนเต็ม จึงเป็นผลทำให้การแปลงเป็นคะแนนปกติแตกต่างกันในระดับคะแนนดิบเดียวกัน และจากผลการแปลงคะแนนจะพบว่าตรงคะแนน 15 ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขาตรงกับคะแนน Z เป็นศูนย์ แต่ในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผลให้ค่า Z ที่มากกว่าศูนย์เล็กน้อยและในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาตรงคะแนน 20 ให้ค่า Z ติดลบเล็กน้อยจะไม่เท่ากับศูนย์พอดีเหมือนกับในแบบทดสอบด้านตัวเลข จึงอาจทำให้รูปร่างของเส้นกราฟจะแตกต่างในช่วงระหว่างค่า Z เท่ากับ -1 ถึง +1 และเมื่อพิจารณาในแบบทดสอบฉบับเดียวกันวิธีทวินามและวิธีทวินามประกอบให้ค่าใกล้เคียงกัน และวิธีเฟลด์ - ควอลส์ทุกสูตรก็ให้ค่าใกล้เคียงกัน อย่างไรก็ตามค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบปกติมีค่าค่อนข้างแกว่งมาก ณ ที่ค่า Z ทางด้านปลายทั้งทางบวกและทางลบ ซึ่งนี้ก็สอดคล้องกับงานวิจัยของโคเลน แฮนสัน และเบรนนัน (Kolen, Hanson & Brennan. 1992 : 298) ที่เมื่อใช้คะแนนแปลงแบบปกติแล้วลักษณะของกราฟหยักขึ้นลงคล้ายโค้งสองยอดเช่นเดียวกัน

สำหรับวิธีทฤษฎีการตอบข้อสอบซึ่งมีพื้นฐานมาจากโมเดลทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างพฤติกรรมการตอบข้อสอบกับระดับความสามารถของผู้ตอบ (θ) การแจกแจงความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในทฤษฎีนี้เป็นแบบทวินามประกอบและจะขึ้นอยู่กับพารามิเตอร์ของข้อสอบ ((Kolen, Hanson & Brennan. 1992 : 304) ในการวิจัยครั้งนี้ผู้วิจัยประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดโดยใช้แบบสามพารามิเตอร์ ปรากฏว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้านเมื่อใช้คะแนนแปลงเป็นค่า θ ให้ค่าใกล้เคียงกับค่าที่ได้จากแต่ละวิธีในคะแนนแปลงแบบปกติ (Z) ทั้งนี้เพราะแท้จริงแล้วระดับความสามารถ θก็คือคะแนนจริงของผู้สอบที่ประมาณได้จากคะแนนรวมจากการตอบแบบทดสอบนั่นเองโดยปรับให้เป็นคะแนนที่มีค่าเฉลี่ยเป็นศูนย์และส่วนเบี่ยงเบนมาตฐานเป็นหนึ่ง เช่นเดียวกับคะแนนปกติ (Hambleton & Swaminathan. 1985 : 54 - 55) จึงทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าใกล้เคียงกัน

ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาค่าความคลาดเคลื่อนมาตรฐานในการวัดค่อนข้างคงที่ในช่วงระดับ θ ตรงกลาง นอกจากนี้พบว่าในแต่ละระดับค่า θ ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงกว่าในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผล ที่เป็นเช่นนี้อาจเป็นเพราะฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาให้ค่าสูงสุดเป็น 4.0700 ณ ที่ระดับความสามารถ 0.5 ส่วนฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขให้ค่าสูงสุดเป็น 11.4397 ณ ที่ระดับความสามารถ 0.3 และด้านเหตุผลให้ฟังก์ชันสารสนเทศของแบบทดสอบสูงสุดเป็น 7.4700 ณ ที่ระดับความสามาร 0.7 จะเห็นได้ว่าฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามีค่าต่ำกว่าในแบบทดสอบด้านตัวเลขและด้านเหตุผล ทั้งที่จำนวนข้อมากกว่าแบบทดสอบทั้งสองด้านนี้ จึงเป็นเหตุทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขสูงกว่าในแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผล แฮมเบิลตัน สวามินาทานและโรเจอร์ (Hambleton, Swaminathan & Rogers. 1991 : 94 - 95) ได้กล่าวว่าขนาดของความคลาดเคลื่อนมาตรฐานในการวัดขึ้นอยู่กับองค์ประกอบ 1) จำนวนข้อคำถาม ถ้าข้อคำถามมากความคลาดเคลื่อนก็จะน้อยลง 2) คุณภาพของข้อคำถาม ถ้าข้อคำถามมีอำนาจจำแนกสูงค่าความคลาดเคลื่อนจะน้อยลงและตอบคำถามถูกโดยปราศจากการเดา และ 3) ความสัมพันธ์ระหว่างความยากของข้อคำถามกับความสามารถของผู้ตอบ ค่าความคลาดเคลื่อนจะน้อยลงถ้าค่าพารามิเตอร์ความยากมีค่าใกล้กับค่าพารามิเตอร์ความสามารถของผู้ตอบ จากองค์ประกอบที่ส่งผลต่อความคลาดเคลื่อนมาตรฐานดังกล่าวจึงอาจเป็นสาเหตุให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาสูงกว่าด้านตัวเลขและด้านเหตุผล ประกอบกับธรรมชาติของการทดสอบแบบทดสอบวัดความถนัดทางการเรียนด้านภาษามักมีความเป็นปรนัยในคำถามน้อยกว่า ดังนั้นจึงทำคุณภาพของข้อคำถามไม่ดีเท่าที่ควร และจากข้อค้นพบอีกประการก็คือค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขที่ได้จากแบบทดสอบทั้งสามฉบับนี้มีค่าไม่คงที่ตลอดช่วงระดับความสามารถ θ ก็เนื่องจากว่าฟังก์ชันสารสนเทศของแบบทดสอบทั้งสามฉบับนี้มีค่าต่ำ เพราะ แฮมเบิลตัน สวามินาทานและโรเจอร์ (Hambleton, Swaminathan & Rogers. 1991 : 95) กล่าวว่า ขนาดของความคลาดเคลื่อนมาตรฐานในการวัดจะมีความคงที่ถ้าฟังก์ชันสารสนเทศของแบบทดสอบมีค่าสูงกว่า 25

สำหรับผลที่ได้จากการศึกษาคุณสมบัติของตัวประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข ในงานวิจัยครั้งนี้ศึกษาโดยใช้ค่าสถิติที่แสดงถึงความแม่นยำในการประมาณค่าได้แก่ ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย ปรากฏว่า ค่าความลำเอียงเฉลี่ยที่ได้จากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขที่ใช้คะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ เมื่อขนาดตัวอย่างจำลอง 300 คน วิธีประมาณค่าของเฟลด์ – ควอลส์สูตร $F - Q_{CB}$ ให้ค่าความลำเอียงเฉลี่ยเป็นลบ แต่วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์สูตร $F - Q_{CB}$ ให้ค่าความลำเอียงเฉลี่ยเป็นลบเมื่อตัวอย่างจำลองมีขนาด 800 คน นอกนั้นให้ค่าความลำเอียงเฉลี่ยเป็นบวก และสำหรับคะแนนแปลงแบบคะแนนปกติ วิธีของเฟลด์ – ควอลส์ทุกสูตรให้ค่าเป็นลบเมื่อขนาดตัวอย่างจำลองเป็น 300 คน แต่เมื่อ

ตัวอย่างจำลองขนาด 800 คน วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์ทุกสูตรให้ค่าความลำเอียงเฉลี่ยเป็นลบ นอกนั้นให้ค่าเป็นบวก ค่าความลำเอียงเฉลี่ยเป็นลบเป็นค่าที่บอกให้ทราบว่าในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขเมื่อใช้วิธีการแปลงคะแนนต่างกันจากวิธีการประมาณค่าต่างกันนั้นค่าที่ประมาณได้มีค่าต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขในตัวอย่างขนาดใหญ่หรือจากประชากรเทียม (2000 คน) ที่ผู้วิจัยได้กำหนดไว้ ที่เป็นดังนี้อาจเนื่องมาจากตัวอย่างจำลองที่สุ่มมาได้แต่ละขนาดที่ได้จากการสุ่มซ้ำขนาดละ 30 ครั้ง เพราะเมื่อพิจารณาจากค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ปรากฏว่า ค่าความเชื่อมั่นของแบบทดสอบเมื่อขนาดตัวอย่างจำลอง 300 และ 800 คน ส่วนใหญ่มีค่าความเชื่อมั่นสูงกว่าค่าความเชื่อมั่นของแบบทดสอบที่ได้จากตัวอย่างขนาดใหญ่ แต่เมื่อขนาดตัวอย่างจำลอง 500 คน ส่วนใหญ่มีค่าความเชื่อมั่นต่ำกว่าค่าที่ได้จากตัวอย่างขนาดใหญ่และต่ำกว่าเมื่อขนาดตัวอย่างจำลอง 300 และ 800 คน ที่เป็นดังนี้อาจเป็นผลมาจากกระบวนการสุ่มตัวอย่างจำลองแต่ละขนาดโดยใช้โปรแกรมสำเร็จรูป SPSS for WINDOWS และเมื่อพิจารณาถึงค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยเฉลี่ยที่ได้จากตัวอย่างจำลองแต่ละขนาดและแต่ละวิธีก็ปรากฏว่าวิธีที่ให้ค่าความลำเอียงเฉลี่ยเป็นลบจะเป็นวิธีที่ค่าประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากตัวอย่างจำลองมีค่าต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างขนาดใหญ่นั่นเอง ในทฤษฎีการทดสอบแบบมาตรฐานเดิมถ้าแบบทดสอบที่มีค่าความคลาดเคลื่อนมาตรฐานในการวัดต่ำ ค่าความเชื่อมั่นก็จะสูง

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ย ซึ่งเป็นค่าที่แสดงถึงปริมาณความแตกต่างที่ค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างจำลองแต่ละขนาดเบี่ยงเบนไปจากค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างขนาดใหญ่ จากผลการวิจัยครั้งนี้เมื่อใช้คะแนนจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขของคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีทวินามประกอบและวิธีของเฟลด์ – ควอลส์ทุกสูตรเมื่อตัวอย่างมีขนาดใหญ่ขึ้นค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยจะลดลง แต่วิธีทวินามค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อตัวอย่างมีขนาด 500 คนมีค่าสูงกว่าเมื่อตัวอย่างมีขนาด 300 คน และให้ค่าต่ำสุดเมื่อตัวอย่างมีขนาด 800 คน ที่เป็นดังนี้ก็เป็นผลมาจากที่กล่าวไปแล้วว่าค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขเมื่อตัวอย่างจำลองขนาด 500 คน มีค่าต่ำกว่าเมื่อตัวอย่างจำลองมีขนาด 300 และ 800 คน จึงเป็นผลทำให้ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อตัวอย่างจำลองมีขนาด 500 คนสูงกว่าเมื่อตัวอย่างจำลองมีขนาด 300 และ 800 คน

ค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา เมื่อใช้วิธีการประมาณค่าต่างกัน ผลการวิจัยปรากฏว่า ในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีเฟลด์ – ควอลส์ทุกสูตรให้ค่าความลำเอียงเป็นลบเมื่อตัวอย่างจำลองขนาด 500 และ 800 คน แต่วิธีทวินามและวิธีทวินามประกอบให้ค่าเป็นลบเมื่อตัวอย่างจำลองขนาด 800 คน และในคะแนนแปลงแบบคะแนนปกติ วิธีเฟลด์ – ควอลส์ทุกสูตรให้ค่าเป็นลบในทุกขนาดตัวอย่างจำลอง แต่วิธี

ทวินามให้ค่าความลำเอียงเฉลี่ยเป็นลบเมื่อตัวอย่างมีขนาด 500 คน วิธีทวินามประกอบให้ค่าความลำเอียงเฉลี่ยเป็นลบเมื่อตัวอย่างมีขนาด 800 คน นั่นคือ ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดที่ได้จากวิธีเหล่านี้จากขนาดตัวอย่างจำลองดังกล่าวให้ค่าประมาณที่ต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างขนาดใหญ่ จึงทำให้ค่าความลำเอียงเป็นลบ และเมื่อพิจารณาจากค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาจากตัวอย่างจำลองแต่ละขนาด ที่ได้จากการสุ่มซ้ำ 30 ครั้ง ก็พบว่า ค่าความเชื่อมั่นโดยส่วนใหญ่ในแต่ละขนาดตัวอย่างมีค่าสูงกว่าค่าความเชื่อมั่นที่ได้จากตัวอย่างขนาดใหญ่ เมื่อค่าความเชื่อมั่นสูงค่าความคลาดเคลื่อนมาตรฐานในการวัดก็จะต่ำ ดังนั้นค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างจำลองโดยส่วนใหญ่จึงมีค่าต่ำกว่าค่าความคลาดเคลื่อนมาตรฐานในการวัดจากตัวอย่างขนาดใหญ่จึงทำให้ค่าความลำเอียงเป็นลบ

ค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยเมื่อใช้คะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา แต่ละวิธีที่ใช้ในการประมาณค่าให้ค่าความแปรปรวนต่ำใกล้เคียงกันมาก โดยที่วิธีทวินามและวิธีทวินามประกอบให้ค่าน้อยลงเมื่อตัวอย่างจำลองมีขนาดใหญ่ขึ้น วิธีเฟลด์ – ควอลส์ทุกสูตรให้ค่าสูงเมื่อตัวอย่างจำลองมีขนาด 500 คน และกลับให้ค่าต่ำลงเมื่อตัวอย่างจำลองมีขนาด 800 คน ที่เป็นดังนี้เป็นผลมาจากค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาในตัวอย่างจำลองขนาด 300 คน ให้ค่าความเชื่อมั่นสูงสุด รองลงมาได้แก่ ค่าความเชื่อมั่นจากตัวอย่างจำลองขนาด 500 และ 800 คน ตามลำดับ

ค่าความลำเอียงเฉลี่ยและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์และคะแนนปกติในแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล เมื่อใช้วิธีการประมาณค่าต่างกัน ผลการวิจัยปรากฏว่าในคะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีเฟลด์ – ควอลส์สูตร F – Q_{RB} ให้ค่าความลำเอียงเฉลี่ยเป็นลบ เมื่อตัวอย่างจำลองมีขนาด 500 คน และในคะแนนแปลงแบบคะแนนปกติ ทุกวิธีที่ใช้ประมาณค่าให้ค่าความลำเอียงเฉลี่ยเป็นลบในตัวอย่างจำลองขนาด 300 และ 500 คน แต่ในตัวอย่างจำลองขนาด 800 คนให้ค่าเป็นบวก นั่นคือ ค่าความคลาดเคลื่อนมาตรฐานในการวัดทีประมาณค่าได้จากตัวอย่างแต่ละขนาดและแต่ละวิธีที่ใช้ให้ค่าต่ำกว่าค่าที่ได้จากตัวอย่างขนาดใหญ่จึงทำให้ค่าความลำเอียงเฉลี่ยเป็นลบ และเป็นที่น่าสังเกตว่าเมื่อใช้คะแนนแปลงแบบตำแหน่งเปอร์เซ็นไทล์ วิธีประมาณค่าของเฟลด์ – ควอลส์สูตร F – Q_{RB} ให้ค่าความลำเอียงเฉลี่ยเป็นลบ นอกนั้นให้ค่าเป็นบวก ที่เป็นดังนี้เพราะเมื่อพิจารณาจากค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาของตัวอย่างจำลองขนาด 500 คนที่ใช้สูตรความเชื่อมั่น r_B ให้ค่าสูงสุด เมื่อค่าความเชื่อมั่นสูงความคลาดเคลื่อนมาตรฐานในการวัดก็ต่ำ ดังนั้นในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดโดยใช้วิธีเฟลด์ – ควอลส์สูตร F – Q_{RB} จึงให้ค่าต่ำกว่าค่าที่ได้จากตัวอย่างขนาดใหญ่ สำหรับคะแนนแปลงแบบคะแนนปกติค่าความลำเอียงเป็นลบในขนาดตัวอย่าง 300 และ 500 คน แต่ค่าที่ได้มีค่าใกล้เคียงกันมาก แต่เมื่อตัวอย่างจำลองมีขนาดใหญ่ขึ้น ค่าความลำเอียงเฉลี่ยกลับเป็นบวก ที่เป็นดังนี้อาจมาจากค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล

เมื่อตัวอย่างจำลองมีขนาด 800 คน ค่าความเชื่อมั่นที่ได้กลับมีค่าต่ำกว่าค่าที่ได้จากตัวอย่างขนาด 300 และ 500 คน อีกทั้งวิธีการแปลงคะแนนแบบปกติค่าของคะแนนแปลงมีทั้งค่าบวกและค่าลบ จึงทำให้ค่าความลำเอียงเฉลี่ยของตัวอย่างจำลองขนาด 800 คนมีค่าเป็นลบ

ในวิธีทฤษฎีการตอบข้อสอบค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการวัดจากแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษา และด้านเหตุผล ไม่แตกต่างกันนัก แต่ในแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาให้ค่าความลำเอียงเฉลี่ยเป็นลบในทุกขนาดตัวอย่างจำลอง ที่เป็นดังนี้อาจเนื่องมาจากฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษาในตัวอย่างจำลองแต่ละขนาดมีค่ามากกว่าในตัวอย่างขนาดใหญ่ที่มีค่าเพียง 4.0700 ซึ่งฟังก์ชันสารสนเทศนี้มีความสัมพันธ์กับค่าความคลาดเคลื่อนมาตรฐานในการวัด นั่นคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าเท่ากับส่วนกลับของรากที่สองของฟังก์ชั่นสารสนเทศของแบบทดสอบ (Hambleton, Swaminathan & Rogers. 1991) และจากผลการวิจัยได้ว่าเมื่อขนาดตัวอย่างจำลองใหญ่ขึ้น ค่าประมาณความคลาดเคลื่อนมาตรฐานในการวัดจะลดลง ทำให้ค่าความลำเอียงเฉลี่ยมีค่าน้อยลง และบางขนาดของตัวอย่างจำลองเป็นลบ ทั้งนี้เพราะในทฤษฎีการตอบข้อสอบซึ่งเป็นทฤษฎีที่แสดงความสัมพันธ์ระหว่างความสามารถที่แท้จริงกับพฤติกรรมการตอบข้อสอบ จำนวนพารามิเตอร์ที่ใช้ยิ่งมากขนาดตัวอย่างควรให้มีจำนวนมากพอ และในการวิจัยครั้งนี้ผู้วิจัยใช้โมเดลแบบสามพารามิเตอร์ ดังนั้นในตัวอย่างขนาด 300 คน จึงให้ค่าประมาณความคลาดเคลื่อนมาตรฐานสูงกว่าเมื่อขนาดตัวอย่างจำลองเป็น 500 และ 800 คน เมื่อพิจารณาถึงค่าความแปรปรวนคลาดเคลื่อนเฉลี่ยมีค่าต่ำมากและมีค่าใกล้เคียงกันมีค่าเกือบเป็นศูนย์ นั่นแสดงว่าในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดเมื่อใช้วิธีทฤษฎีการตอบข้อสอบให้ผลใกล้เคียงกับค่าที่ได้จากตัวอย่างขนาดใหญ่

ถึงแม้ว่าจะไม่สามารถสรุปออกมาได้อย่างแน่ชัดว่าคะแนนแปลงแบบใดเมื่อใช้วิธีการประมาณค่าต่างกันจากตัวอย่างจำลองแต่ละขนาด จะให้ค่าความแม่นยำในการประมาณค่ามากกว่ากันซึ่งอาจเป็นผลมาจากจำนวนครั้งที่สุ่มซ้ำอาจน้อยเกินไป เพราะการวิจัยครั้งนี้ผู้วิจัยใช้การสุ่มซ้ำเพียง 30 ครั้ง อย่างไรก็ตามผลที่ได้ก็สอดคล้องกับงานวิจัยของ ลี เบรนนัน และโคเลน (Lee, Brennan & Kolen. 2000 ; Feldt & Qualls. 1998) ที่ได้ศึกษาถึงตัวประมาณค่าของความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงแบบ DSS GE และ PR โดยพิจารณาจากค่าความลำเอียงและความแปรปรวนคลาดเคลื่อนเฉลี่ย จากข้อมูลจำลองที่กำหนดให้ความเชื่อมั่นในการวัดเป็น .85 และ .50 ผลการวิจัยพบว่าได้รูปร่างของตัวประมาณค่าต่างกัน และยังได้กล่าวอีกว่าในความแม่นยำของการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของแต่ละวิธีขึ้นอยู่กับองค์ประกอบหลายอย่างเช่น รูปแบบของคะแนนแปลง ลักษณะของแบบทดสอบและวิธีการที่ใช้ในการประมาณค่า

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. ในการรายงานผลการสอบของผู้สอบ เมื่อมีการแปลงคะแนนอาจใช้คะแนนในรูปของจำนวนเต็มเพื่อไม่ให้ดูยุ่งยากสำหรับการแปลความหมาย และประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่แต่ละคะแนนแปลงนั้น อีกทั้งควรประมาณค่าคะแนนจริงเป็นช่วงด้วยระดับความเชื่อมั่นที่ต้องการ

2. ขนาดตัวอย่างที่ใช้ในการประมาณค่าเมื่อศึกษากับตัวอย่างขนาด 300 500 และ 800 คน ให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ได้ไม่แตกต่างจากกรณีที่ศึกษากับตัวอย่างขนาดใหญ่ ดังนั้นเมื่อนำไปใช้จริงไม่จำเป็นต้องใช้กับขนาดตัวอย่างที่ใหญ่

3. จากวิธีที่ใช้ในการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขจากวิธีทวินาม วิธีทวินามประกอบ วิธีของเฟลด์ – ควอลส์สูตร $F - Q_{R20}$,$F - Q_{RB}$ และ $F - Q_{CB}$ และวิธีทฤษฎีการตอบข้อสอบ ให้ค่าไม่แตกต่างกันนัก แต่การประมวลผลวิธีทวินามและทวินามประกอบจะมีความยุ่งยากในการคำนวณมากกว่า ดังนั้นเพื่อสะดวกในทางปฏิบัติจึงควรใช้วิธีของเฟลด์ – ควอลส์ เว้นแต่ว่าแบบทดสอบนั้นประกอบด้วยเนื้อหาหลายตอนหรือหลายส่วน

4. การแปลงคะแนนดิบเป็นคะแนนปกติ เมื่อใช้วิธีการประมาณค่าจากการศึกษาครั้งนี้ ทำให้ค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขมีความเป็นไปได้ในการนำไปใช้มากกว่าการแปลงแบบตำแหน่งเปอร์เซ็นไทล์ เพราะค่าความคลาดเคลื่อนมาตรฐานในการวัดตรงกลางของคะแนนแปลงมีค่าไม่สูงไปกว่าคะแนนตรงด้านปลาย

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรทำการศึกษากับแบบทดสอบที่มีการแปลงคะแนนที่แตกต่างไปจากนี้กับแบบทดสอบที่มีลักษณะการตรวจให้คะแนนต่างกัน หรือที่มีรูปแบบคำถามหลายรูปแบบ และเนื้อหาที่หลากหลาย เพื่อศึกษาถึงค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไข โดยใช้วิธีการประมาณค่าที่แตกต่างไปจากนี้

2. ควรศึกษาวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขโดยคำนึงถึงแหล่งความแปรปรวนมากกว่าหนึ่งแหล่ง ที่อาจเกิดขึ้นจากการวัดแต่ละครั้ง โดยวิธีทฤษฎีสัมประสิทธิ์การสรุปอ้างอิง

3. ควรศึกษาวิธีการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขตามวิธีของเฟลด์ – ควอลส์โดยใช้การคำนวณค่าความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนดิบที่เกี่ยวข้องกับค่าความเชื่อมั่นที่ต่างไปจากวิธีที่ใช้ในการวิจัยครั้งนี้

4. ควรศึกษากับจำนวนกลุ่มตัวอย่างขนาดอื่น ๆ และกลุ่มตัวอย่างที่มีการแจกแจงคะแนนต่างรูปแบบกัน

5. ควรศึกษากับข้อมูลจำลองโดยกำหนดค่าความเชื่อมั่นที่แตกต่างกัน เพื่อประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดและความแม่นยำในการวัด

6. ควรศึกษาถึงการประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดของคะแนนแปลงที่มีจำนวนช่วงของคะแนนแปลงแตกต่างกัน เช่น คะแนนแปลงที่มี 10 15 20 30 ค่า เป็นต้น

7. ควรศึกษาค่าความคลาดเคลื่อนมาตฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงที่มีค่าสุดโต่งหรือคะแนนแปลงทางด้านปลาย อาจโดยใช้สูตรการปรับแก้ เพื่อไม่ให้ค่าที่ได้ออกมาเป็นศูนย์

บรรณานุกรม

บรรณานุกรม

เกศเกล้า หมู่มิ่ง. (2539). ***การประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่า.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

ฉวีวรรณ ครองยุทธ. (2539). ***การประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่าที่มีขนาดกลุ่มตัวอย่างกัน.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

พนิดา ฟุ้งเกียรตินำสุข (2540). ***การประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่าที่มีรูปแบบการตอบแตกต่างกัน.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

นิภาพร ยอดเมือง. (2539). ***ความคลาดเคลื่อนและความลำเอียงทางสถิติในการเทียบคะแนนเชิงเส้นตรงแบบคะแนนจริงสัมพันธ์ในกลุ่มตัวอย่างขนาดเล็ก.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

นิยม ปุราคำ. (2517). ***ทฤษฎีของการสำรวจสถิติจากตัวอย่างและการประยุกต์ : Sample Survey Theory and Applications.*** กรุงเทพ ฯ : ศ.ส. การพิมพ์.

บุญเชิด ภิญโญอนันตพงษ์. (2542, กันยายน - ธันวาคม). "งานวิจัย : สัมประสิทธิ์ r_B การประมาณค่าความเชื่อมั่นสำหรับแบบทดสอบผลสัมฤทธิ์ชนิดเลือกตอบที่ประกอบด้วยความยากรายข้อต่างกัน ," ***วิชาการศึกษาศาสตร์.*** 1(1) : 59 - 72.

เยาวดี วิบูลย์ศรี. (2540). ***การวัดผลและการสร้างแบบสอบผลสัมฤทธิ์ (Measurement and Achievement Test Construction).*** พิมพ์ครั้งที่ 2. กรุงเทพฯ : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย

รุ่งฤดี ทรัพย์นิธิ. (2539). ***การศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติกแบบหนึ่ง สอง สาม พารามิเตอร์ และขนาดกลุ่มตัวอย่างต่างกัน.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

เรวดี อินทสะระ. (2530). ***การเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตราระหว่างรูปแบบอิงทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

วลีมาศ แซ่อึ้ง. (2537). ***การศึกษาความคลาดเคลื่อนมาตรฐานและประสิทธิภาพสัมพัทธ์ของการประมาณค่าเจตคติต่อวิชาคณิตศาสตร์ ระหว่างมาตราวัดแบบลิเคอร์ทกับมาตราวัดแบบออสกูด.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

วิชุดา กิจธรธรรม. (2544). ***การประมาณค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนของแบบทดสอบวัดความถนัดทางวิชาชีพพยาบาลตามแนวคิดของฟลานาแกน ด้วยวิธีการวิเคราะห์ความแปรปรวน วิธีการตามทฤษฎีการตอบข้อสอบ และวิธีการของคีต.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

สุรินทร์ นิยมางกูร. (2542). ***เทคนิคการสุ่มตัวอย่าง : Sampling Techniques.*** พิมพ์ครั้งที่ 4 กรุงเทพ ฯ : มหาวิทยาลัยเกษตรศาสตร์.

สุทธิวรรณ พีรศักดิ์โสภณ. (2540, 2542). ***รายงานโครงการวิจัยเรื่องการสร้างและพัฒนาแบบทดสอบวัดความถนัดทางการเรียนสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย.*** คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม.

สุมนา โสตถิผลอนันต์. (2539). ***การศึกษาวิธีการเทียบคะแนนแบบเส้นตรงตามแบบจำลองคะแนนจริงสัมพันธ์.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

สำนักงานคณะกรรมการการศึกษาแห่งชาติ. (2542). *พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542. สำนักงานคณะกรรมการการศึกษาแห่งชาติ สำนักนายกรัฐมนตรี.* กรุงเทพฯ : สำนักงานฯ.

อนันต์ ศรีโสภา. (2525). **ทฤษฎีการวัดและการทดสอบ.** กรุงเทพมหานคร : ไทยวัฒนาพานิช.

อำนวยพร คำพันธุ์. (2542). ***ผลของการจัดเรียงข้อสอบที่มีต่อค่าความคลาดเคลื่อนมาตรฐานในการวัดที่ระดับคะแนนเฉพาะ.*** ปริญญานิพนธ์ กศ.ม.(การวัดผลการศึกษา) กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

ไอลดา คล้ายสำริด. (2538) ***การเปรียบเทียบความคลาดเคลื่อนในการประมาณค่า ประสิทธิภาพสัมพัทธ์ในการประมาณค่าและความสามารถของผู้สอบ ในแบบทดสอบวิชาคณิตศาสตร์ที่ใช้วิธีต่างกัน.*** ปริญญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพ ฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ.

Allen, Mary J. & Yen, Wendy M. (1979). ***Introduction to Measurement Theory.*** Belmont California : Brook/Cole Publishing.

Angoff, W.H. (1971). "Norms, and Equivalent Score," ***in Educational Measurement.*** Edited by L.R. p. 508 - 600. Washington, D.C. : American Council on Education.

Blixt, Sonya L. & Shama, Deborah D. (1986). "An Empirical Investigation of the Standard errors of Measurement at Different Ability Level," ***Educational and Psychological Measurement.*** 46 : 545 - 550.

Brennan, R.L. & Lee, Won-Chan. (1999, February). "Conditional Scale-Score Standard Errors of Measurement under Binomial and Compound Binomial Assumption," ***Educational and Psychological Measurement.*** 59(1) : 5 - 24.

Feldt, Leonard S. (1984). "Some Relationships between the Binomial Error Model and Classical Test Theory," ***Educational and Psychological Measurement.*** 44 : 883 - 591.

Feldt, Leonard S., Steffen, M. & Gupta, N.C. (1985, December). "A Comparison of Five Methods for Estimating Standard Error of Measurement a Specific Score Levels," ***Applied Psychological Measurement.*** 9(4) : 351 - 361.

Feldt, Leonard S.& Qualls, Audrey L. (1996, Summer). " Estimation of Measurement Error Variance at Specific Score Levels," ***Journal of Educational Measurement.*** 33(2) : 141 - 156.

---------. (1998). " Approximating Scale Score Standard Error of Measurement From the Raw Score Standard Error," ***Applied Measurement in Education.*** 11(2) : 159 – 177.

Feldt, Leonard S., Steffen Manfred & Gupta Naim C. (1985, December). "A Comparison of Five Methods for Estimating the Standard Error of Measurement at Specific Score Levels," ***Applied Psychological Measurement.*** 9(4) : 351 - 361.

Glas, Dees A.W. and Beguin, Anton A. (1996) ***Appropriateness of IRT Observed Score Equating Researh Report 96-04.*** (Online). Available http://ericae.net/ericdc/ED424238.htm.

Hambleton, Ronald K. & Swaminathan, Hariharan. (1985). ***Item Response Theory : Principles and Applications).*** Boston : Kluwer - Nijhoff Publishing.

Hambleton, Ronald K., Swaminathan, H & Rogers, H.Jane. (1991). ***Fundamentals of Item Response Theory.*** Newbury Park : Sage.

Insko, William Robert,Jr. (1995, April). ***Conditional Standard Errors of Measurement : A Monte Carlo Comparison of Four Estimation Methods.*** (CD-ROM). Available DAI-A (1995, April)

Keats, J.A. (1957). "Estimation of Error Variances of Test Scores," ***Psychometrika.*** 22 : 29-41.

Kolen, Michael J. (1988, Summer). "Defining Score Scales in Relation to Measurement Error," ***Journal of Educational Measurement.*** 25(2) : 97 - 110.

Kolen, Michael J. & Brennan, Robert L. (1995). ***Testing Equating : Methods and Practices.*** New York : Springer - Verlag.

Kolen, Michael J., Hanson, Bradley A. & Brennan, Robert L. (1992, Winter). "Conditional Standard Errors of Measurement for Scale Scores," ***Journal of Educational Measurement.*** 29(4) : 285 – 307.

Kolen, Michael J., Zeng, Lingjia & Hanson, Bradley A. (1996, Summer). "Conditional Standards Errors of Measurement for Scale Scores Using IRT," ***Journal of Educational Measurement.*** 33(2) : 129 – 140.

Lee, Guemin. (2000). "Estimating Conditional Standard Errors of Measurement for Tests Composed of Testlets," ***Applied Measurement in Education.*** 13(2) : 161 - 180.

Lee, Won-Chan., Brennan, R.L. & Kolen, M.J. (2000, Spring). "Estimators of Conditional Scale - Score Standard Errors of Measurement : A Simulation Study," ***Journal of Educational Measurement.*** 37(1) : 1 - 20.

Lin, Miao - Hsiang & Hsiung, Chao. (1992, December). "Four Bootstrap Confidence Intervals for the Binomial - Error Model," ***Psychometrika.*** 57(4) : 449 - 520.

Livingston, Samuel A. (1982, Summer). "Estimation of the conditional stadard error of measurement for stratified tests," ***Journal of Educational Measurement.*** 19(2) : 135-138

Lord, Federic M. (1955,). "Estimating Test Reliability'," ***Educational and Psychological Measurement.*** 15 : 325 – 336.

----------. (1957, Winter). "Do tests of the same length have the same standard errors of measurement?," ***Educational and Psychological Measurement.*** 17 : 510 – 521.

----------. (1959). "Tests of the same length do have the same standard error of measurement," ***Educational and Psychological Measurement.*** XIX(2) : 233 - 239.

----------. (1965). "A Strong True Score Theory, with Applications," ***Psychometrika.*** 30 : 239-270.

----------. (1969, September). "Estimating True - Score Distributions in Psychological Testing (An Empirical Bayes Estimation Problems)," ***Psychometrika.*** 34(3) : 259 - 299.

----------. (1984, Fall). ***"Standard errors of measurement at different ability levels," Journal of Educational Measurement.*** 21(3) : 239 – 243.

Lord, Federic M. & Novick, Melvin R. (1968). ***Statistical Theories of Mental Test Scores.*** CA : Addison - Wesley

Lord, Federic M. & Wingersky, M.S. (1984). "Comparison of IRT True Score and Equipercentile Observed - score "Equating"," ***Applied Psychological Measurement.*** 8 : 453 - 461.

May, Kim & Nicewander, Alan. (1994, Winter). "Reliability and Information Functions for Percentile Ranks," ***Journal of Educational Measurement.*** 31(4) : 313 - 325.

Mary, Hanson, Harris and Sconing. (1999, April 19-23). ***Issues in Creating and Reporting Concordance Results Based on Equipercentile Methods.*** (Online). Available http://ericae.net/ericdc/ED431033.htm.

Maryland School Performance Program : How is MSPAP scored? (2000). ***Scaling and Equating.*** (Online). Available http://www.mdk12.org/mspp/mspp/how- scored/95tech_report/scaling_equating.htm.

Mollenkopf, William G. (1949, September). "Variation ofthe Standard Error of Measurement," ***Psychometrika.*** 14(3) : 189 – 215.

Popham, W James, (1990). ***Modern Educational Measurement : A Practitioner's Perspective.*** Englewood Cliffs, N.J. : Prentice Hall, 1990.

Qualls - Payne, Audrey L. (1992, Fall). "A Comparison of Score Level Estimates of the Standard Error of Measurement," ***Journal of Educational Measurement.*** 29(3) : 213 - 225.

Stockburger, David W. (2000). "Linear Transformations," ***in Introductory Statistics : Concepts, Models, and Applications.*** (Online). Available http://www.psychstat.smsu.edu/introbook/sbk15m.htm

University of Minnesota, The College of Education & Human Development. (2000). ***The Equipercentile Method.*** Available http://edpsy.cole.umn.edu/oea/New_Site/Reports/Briefcase/HypertextBC1-1/equipct.htm

Wang, Tianyou., Kolen, Michael J. & Harris, Deborah J. (2000, Summer). "Psychometric Properties of Scale Scores and Performance Levels for Performance Assessments Using Polytomous IRT," ***Journal of Educational Measurement.*** 37(2) : 141 - 162.

Woodruff, David. (1990, Fall). " Conditional standard error of measurement in prediction," ***Journal of Educational Measurement.*** 27(3) : 191 – 208.

----------. (1991, Fall). " Stepping Up Test Score Conditional Variance," ***Journal of Educational Measurement.*** 28(3) : 191 – 196.

ภาคผนวก

ภาคผนวก ก

ค่าความยากและค่าอำนาจจำแนกรายข้อของแบบทดสอบวัดความถนัดแต่ละด้าน

ค่า p_i, $(1 - p_i)$, ΣB_{ji} และ λ_i รายข้อของแบบทดสอบวัดความถนัดแต่ละด้าน

จากการทดลองก่อนเก็บข้อมูลจริงกับตัวอย่าง 244 คน

ตาราง 27 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
1	ก	.05	.00	6	ก	.33	.08	11	ก*	.56	.63
	ข	.10	.20		ข	.09	.12		ข	.13	.23
	ค*	.78	.40		ค	.06	.09		ค	.14	.15
	ง	.02	.00		ง*	.37	.43		ง	.08	.12
	จ	.05	.00		จ	.13	.11		จ	.09	.12
2	ก	.10	.05	7	ก*	.48	.81	12	ก	.18	.34
	ข	.01	.00		ข	.06	.12		ข	.03	.00
	ค	.03	.00		ค	.23	.31		ค	.04	.00
	ง	.20	.34		ง	.09	.18		ง	.22	.25
	จ*	.65	.48		จ	.12	.18		จ*	.53	.69
3	ก*	.59	.69	8	ก	.19	.26	13	ก	.08	.09
	ข	.12	.14		ข*	.38	.41		ข*	.53	.69
	ค	.25	.48		ค	.08	-.03		ค	.18	.23
	ง	.03	.00		ง	.08	-.03		ง	.08	.17
	จ	.00	00		จ	.26	.21		จ	.13	.20
4	ก	.13	.11	9	ก	.17	.18	14	ก*	.46	.61
	ข	.15	.23		ข	.25	-.02		ข	.15	.17
	ค	.09	.06		ค	.17	.18		ค	.15	.11
	ง	.08	.03		ง*	.33	.32		ง	.07	.05
	จ*	.55	.44		จ	.08	-.03		จ	.18	.29
5	ก	.07	.08	10	ก	.18	.14	15	ก	.08	.05
	ข	.12	.08		ข	.13	.14		ข	.14	-.06
	ค*	.27	.26		ค*	.42	.49		ค	.12	.11
	ง	.15	.23		ง	.13	.11		ง*	.50	.26
	จ	.38	.67		จ	.12	.06		จ	.15	.15

ตาราง 27 (ต่อ)

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
16	ก	.08	.08	21	ก	.04	.00	26	ก*	.38	.52
	ข*	.42	.25		ข*	.61	.51		ข	.17	.00
	ค	.15	.08		ค	.13	.23		ค	.23	.31
	ง	.24	-.05		ง	.05	.00		ง	.11	.09
	จ	.12	.14		จ	.18	.17		จ	.10	.14
17	ก	.21	.17	22	ก	.02	.00	27	ก*	.42	.46
	ข	.12	.03		ข	.03	.00		ข	.18	-.12
	ค*	.34	.43		ค	.03	.00		ค	.15	.14
	ง	.12	.12		ง*	.62	.44		ง	.05	.11
	จ	.19	.11		จ	.29	.40		จ	.18	.32
18	ก	.08	.05	23	ก	.15	.14	28	ก	.06	.00
	ข	.15	.09		ข	.15	.08		ข	.63	.09
	ค*	.12	.17		ค	.21	.14		ค*	.22	.21
	ง	.24	.20		ง	.11	.03		ง	.07	.11
	จ	.40	.52		จ*	.38	.37		จ	.02	.00
19	ก*	.40	.55	24	ก	.05	.08	29	ก	.49	.09
	ข	.24	.14		ข	.06	.03		ข	.28	.02
	ค	.13	.20		ค	.22	.28		ค*	.09	.15
	ง	.12	.12		ง*	.58	.49		ง	.06	.03
	จ	.10	.08		จ	.06	.09		จ	.04	.00
20	ก	.21	.17	25	ก	.08	.05	30	ก	.21	.02
	ข*	.23	.31		ข	.12	.09		ข	.10	-.14
	ค	.14	.09		ค*	.39	.60		ค	.22	.14
	ง	.28	-.12		ง	.16	.20		ง*	.21	.35
	จ	.14	.15		จ	.22	.25		จ	.22	.32

ตาราง 28 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
1	ก*	.77	.00	6	ก	.18	-.05	11	ก	.38	.25
	ข	.05	.00		ข*	.50	.08		ข	.04	.00
	ค	.12	-.09		ค	.27	.02		ค	.07	.02
	ง	.02	.00		ง	.03	.00		ง	.12	.06
	จ	.05	.00		จ	.02	.00		จ*	.38	.34
2	ก	.02	.00	7	ก	.08	.03	12	ก	.21	-.02
	ข	.18	.35		ข	.07	.02		ข	.16	.05
	ค	.07	.11		ค	.31	.12		ค	.19	.17
	ง	.05	.00		ง*	.44	.23		ง*	.25	.32
	จ*	.68	.57		จ	.10	.08		จ	.18	.11
3	ก	.01	.00	8	ก	.08	.06	13	ก	.01	.00
	ข	.08	.02		ข	.07	.05		ข*	.52	.51
	ค	.18	.25		ค*	.50	.57		ค	.13	-.14
	ง	.10	.17		ง	.22	.32		ง	.11	.21
	จ*	.62	.40		จ	.12	.14		จ	.24	.41
4	ก	.16	.17	9	ก	.15	.12	14	ก*	.51	.25
	ข*	.52	.46		ข	.18	.08		ข	.27	-.05
	ค	.11	.06		ค	.11	.15		ค	.05	.02
	ง	.16	.14		ง*	.44	.35		ง	.04	.00
	จ	.05	.00		จ	.08	-.05		จ	.13	.26
5	ก*	.71	.46	10	ก	.21	.02	15	ก	.25	.11
	ข	.19	.35		ข	.22	-.03		ข	.05	.02
	ค	.07	.05		ค	.07	.05		ค*	.52	.26
	ง	.02	.00		ง	.19	.35		ง	.13	.08
	จ	.02	.00		จ*	.31	.40		จ	.05	.00

ตาราง 28 (ต่อ)

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
16	ก	.10	.17	21	ก*	.60	.37	26	ก	.12	.11
	ข	.16	.05		ข	.15	.14		ข*	.52	.43
	ค*	.55	.44		ค	.14	.03		ค	.10	.14
	ง	.02	.00		ง	.01	.00		ง	.11	.15
	จ	.18	.20		จ	.11	.18		จ	.15	.03
17	ก	.10	.14	22	ก*	.48	.17	27	ก*	.35	.26
	ข	.15	.08		ข	.03	.00		ข	.11	.12
	ค	.08	.17		ค	.13	-.02		ค	.35	.15
	ง*	.50	.63		ง	.10	-.05		ง	.04	.00
	จ	.16	.26		จ	.25	.17		จ	.15	-.05
18	ก	.08	.15	23	ก	.22	.11	28	ก	.09	.15
	ข	.09	.15		ข	.12	.14		ข	.15	.20
	ค*	.32	.37		ค	.07	.14		ค	.15	-.03
	ง	.36	-.08		ง	.28	.08		ง*	.48	.40
	จ	.15	.14		จ*	.30	.44		จ	.12	.09
19	ก	.48	.21	24	ก	.18	.14	29	ก*	.41	.29
	ข	.14	-.09		ข	.10	.05		ข	.28	-.06
	ค*	.32	.20		ค*	.62	.31		ค	.10	.08
	ง	.03	.00		ง	.04	.00		ง	.11	.12
	จ	.04	.00		จ	.07	.08		จ	.10	.17
20	ก	.18	.32	25	ก	.18	-.02	30	ก	.20	.06
	ข	.04	.00		ข*	.45	.44		ข*	.41	.44
	ค	.28	-.15		ค	.02	.00		ค	.07	.02
	ง*	.23	.28		ง	.15	.14		ง	.24	.26
	จ	.28	.06		จ	.20	.28		จ	.08	.11

ตาราง 28 (ต่อ)

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
31	ก	.17	.12	35	ก	.14	.21	39	ก*	.59	.44
	ข	.08	.06		ข	.08	.11		ข	.25	.28
	ค*	.28	.09		ค*	.68	.40		ค	.03	.00
	ง	.42	-.14		ง	.06	.03		ง	.06	.00
	จ	.05	.05		จ	.03	.00		จ	.19	.23
32	ก	.15	.12	36	ก	.08	-.02	40	ก	.48	.41
	ข*	.55	.57		ข	.08	.09		ข*	.15	.11
	ค	.10	.14		ค	.11	.06		ค	.07	.02
	ง	.08	.09		ง	.32	.09		ง	.07	.02
	จ	.12	.21		จ*	.40	.25		จ	.08	.02
33	ก	.12	.09	37	ก	.19	.05				
	ข	.07	.02		ข	.17	.12				
	ค	.08	.08		ค	.05	.05				
	ง*	.36	.51		ง	.19	.02				
	จ	.36	.32		จ*	.38	.26				
34	ก*	.71	.15	38	ก	.08	.08				
	ข	.08	.09		ข	.05	.02				
	ค	.08	.08		ค*	.81	.20				
	ง	.02	.00		ง	.03	.00				
	จ	.10	-.05		จ	.01	.00				

ตาราง 29 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
1	ก	.05	-.02	6	ก	.14	.15	11	ก	.27	.17
	ข	.23	.25		ข	.08	.05		ข	.07	.08
	ค	.08	.06		ค	.08	.12		ค*	.42	.46
	ง	.08	.09		ง	.16	.17		ง	.13	.14
	จ*	.55	.41		จ*	.54	.49		จ	.12	.08
2	ก*	.45	.25	7	ก	.11	.09	12	ก*	.52	.57
	ข	.11	.03		ข	.08	-.02		ข	.08	.11
	ค	.18	.09		ค*	.53	.38		ค	.19	.23
	ง	.13	.02		ง	.13	.14		ง	.14	.12
	จ	.12	.12		จ	.15	.17		จ	.07	.11
3	ก	.08	.02	8	ก	.05	.11	13	ก	.09	.15
	ข	.12	.17		ข	.16	.11		ข*	.34	.40
	ค*	.41	.35		ค	.22	.29		ค	.10	.08
	ง	.22	.03		ง*	.39	.54		ง	.40	.03
	จ	.17	.12		จ	.17	.03		จ	.07	.14
4	ก	.07	.05	9	ก	.22	.15	14	ก*	.45	.60
	ข	.05	.02		ข	.15	.21		ข	.09	.18
	ค*	.29	.25		ค*	.41	.51		ค	.12	.15
	ง	.29	.12		ง	.13	.08		ง	.20	.09
	จ	.27	.11		จ	.09	.06		จ	.13	.17
5	ก	.17	-.09	10	ก	.15	.14	15	ก	.15	.21
	ข*	.24	.20		ข	.14	.00		ข	.10	.11
	ค	.31	.18		ค	.12	.18		ค	.22	.26
	ง	.19	.08		ง	.27	.02		ง	.12	.11
	จ	.09	.03		จ*	.32	.32		จ*	.41	.69

ตาราง 29 (ต่อ)

ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r	ข้อที่	ตัวเลือก	p	r
16	ก	.19	.23	21	ก*	.61	.41	26	ก	.18	.11
	ข	.09	.00		ข	.09	.06		ข*	.39	.32
	ค	.12	.08		ค	.13	.20		ค	.12	.08
	ง*	.40	.46		ง	.07	.02		ง	.19	.08
	จ	.20	.15		จ	.09	.15		จ	.12	.08
17	ก	.12	.05	22	ก	.12	.21	27	ก*	.32	.52
	ข	.22	.05		ข	.25	.09		ข	.18	.28
	ค*	.35	.35		ค	.07	.08		ค	.25	.03
	ง	.14	.03		ง	.12	.02		ง	.18	.12
	จ	.17	.25		จ*	.44	.38		จ	.05	.11
18	ก	.28	.35	23	ก	.23	.12	28	ก	.15	.02
	ข	.18	.18		ข	.13	.23		ข*	.31	.37
	ค*	.39	.57		ค	.12	.08		ค	.21	.23
	ง	.07	.08		ง*	.45	.52		ง	.15	.02
	จ	.06	-.03		จ	.08	.09		จ	.18	.12
19	ก	.08	.06	24	ก	.11	.21	29	ก	.08	.06
	ข	.22	.17		ข	.15	.14		ข	.15	-.02
	ค	.18	.08		ค	.14	.18		ค	.25	.11
	ง	.28	.03		ง*	.46	.58		ง*	.31	.34
	จ*	.25	.34		จ	.12	.06		จ	.20	.18
20	ก	.16	.08	25	ก	.08	.05	30	ก	.11	.15
	ข	.21	.05		ข	.15	.28		ข	.32	-.02
	ค	.16	.02		ค	.15	.28		ค	.19	.08
	ง*	.28	.32		ง*	.38	.43		ง	.16	-.05
	จ	.18	.20		จ	.13	.08		จ*	.21	.17

ตาราง 30 ค่า p_i, $(1 - p_i)$, ΣB_{ji} และ λ_i รายข้อแบบทดสอบวัดความถนัดด้านตัวเลข

ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i	ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i
1	0.7910	0.2090	9.9262	0.0247	16	0.4180	0.5820	5.3811	0.0170
2	0.6557	0.3443	8.5123	0.0288	17	0.2910	0.7090	4.5164	0.0344
3	0.5410	0.4590	7.7910	0.0462	18	0.3770	0.6230	5.6434	0.0385
4	0.5287	0.4713	7.0943	0.0300	19	0.2664	0.7336	4.3975	0.0392
5	0.3197	0.6803	5.2008	0.0448	20	0.1557	0.8443	2.6803	0.0261
6	0.3770	0.6230	5.3402	0.0296	21	0.5820	0.4180	7.9672	0.0376
7	0.4631	0.5369	7.1639	0.0541	22	0.6516	0.3484	8.5697	0.0318
8	0.3525	0.6475	5.0574	0.0296	23	0.4590	0.5410	6.1844	0.0268
9	0.3074	0.6926	4.5410	0.0296	24	0.5861	0.4139	7.8279	0.0322
10	0.3115	0.6885	4.7951	0.0357	25	0.3525	0.6475	5.4385	0.0407
11	0.6148	0.3852	8.4098	0.0396	26	0.3975	0.6025	5.7910	0.0359
12	0.4672	0.5328	6.9344	0.0460	27	0.3811	0.6189	5.4221	0.0306
13	0.5164	0.4836	7.3852	0.0426	28	0.1475	0.8525	2.3238	0.0184
14	0.3689	0.6311	5.6803	0.0423	29	0.1189	0.8811	1.6639	0.0088
15	0.5000	0.5000	6.4057	0.0194	30	0.1844	0.8156	2.9713	0.0250

ตาราง 31 ค่า p_i, $(1 - p_i)$, ΣB_{ji} และ λ_i รายข้อของแบบทดสอบวัดความถนัดด้านเหตุผล

ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i	ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i
1	0.5779	0.4221	6.6230	0.0303	16	0.3730	0.6270	4.7910	0.0391
2	0.4467	0.5533	5.0902	0.0223	17	0.3156	0.6844	4.1230	0.0357
3	0.3934	0.6066	4.6393	0.0256	18	0.3033	0.6967	4.6762	0.0613
4	0.2705	0.7295	3.3648	0.0242	19	0.2295	0.7705	3.3975	0.0410
5	0.2213	0.7787	2.7008	0.0178	20	0.2541	0.7459	3.2254	0.0252
6	0.5492	0.4508	6.5041	0.0367	21	0.5656	0.4344	6.7787	0.0409
7	0.5451	0.4549	6.3279	0.0316	22	0.4098	0.5902	5.2377	0.0419
8	0.3648	0.6352	4.7008	0.0388	23	0.4344	0.5656	5.4180	0.0394
9	0.3525	0.6475	4.5369	0.0373	24	0.4221	0.5779	5.5287	0.0483
10	0.3115	0.6885	3.7541	0.0233	25	0.3852	0.6148	4.8852	0.0380
11	0.3934	0.6066	4.9713	0.0381	26	0.3689	0.6311	4.5123	0.0301
12	0.5287	0.4713	6.4672	0.0432	27	0.2951	0.7049	4.0205	0.0396
13	0.3074	0.6926	4.0287	0.0352	28	0.2582	0.7418	3.3238	0.0273
14	0.3934	0.6066	5.1393	0.0445	29	0.2582	0.7418	3.2377	0.0241
15	0.3566	0.6434	5.0041	0.0534	30	0.1885	0.8115	2.2951	0.0150

ตาราง 32 ค่า p_i, $(1 - p_i)$, ΣB_{ji} และ λ_i รายข้อของแบบทดสอบวัดความถนัดด้านภาษา

ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i	ข้อ	p_i	$(1-p_i)$	ΣB_{ji}	λ_i
1	0.8279	0.1721	15.7992	0.0039	21	0.6311	0.3689	12.7992	0.0263
2	0.7213	0.2787	14.8156	0.0359	22	0.5164	0.4836	10.4221	0.0200
3	0.5697	0.4303	11.5656	0.0241	23	0.2828	0.7172	6.4221	0.0330
4	0.5451	0.4549	11.3852	0.0330	24	0.6598	0.3402	13.2213	0.0226
5	0.7377	0.2623	15.0082	0.0322	25	0.5164	0.4836	10.8770	0.0340
6	0.5041	0.4959	9.8893	0.0107	26	0.5656	0.4344	11.8156	0.0343
7	0.4590	0.5410	9.1885	0.0154	27	0.3689	0.6311	7.6680	0.0212
8	0.4877	0.5123	10.5041	0.0393	28	0.4959	0.5041	10.2131	0.0255
9	0.3689	0.6311	7.7828	0.0247	29	0.4221	0.5779	8.5451	0.0171
10	0.3361	0.6639	7.2377	0.0271	30	0.3852	0.6148	8.1844	0.0275
11	0.3607	0.6393	7.5697	0.0229	31	0.2664	0.7336	5.2664	0.0069
12	0.2377	0.7623	5.2008	0.0217	32	0.5533	0.4467	11.7336	0.0389
13	0.6230	0.3770	12.9262	0.0350	33	0.3402	0.6598	7.5492	0.0343
14	0.5328	0.4672	10.6844	0.0185	34	0.7213	0.2787	14.1270	0.0146
15	0.5779	0.4221	11.5615	0.0192	35	0.6803	0.3197	13.9221	0.0322
16	0.6107	0.3893	12.6025	0.0322	36	0.3852	0.6148	7.9139	0.0192
17	0.5205	0.4795	11.2746	0.0439	37	0.4057	0.5943	8.3525	0.0208
18	0.3443	0.6557	7.2910	0.0239	38	0.7582	0.2418	14.7992	0.0138
19	0.3115	0.6885	6.4508	0.0171	39	0.6066	0.3934	12.4508	0.0299
20	0.2049	0.7951	4.5369	0.0203	40	0.4877	0.5123	10.1557	0.0285

ภาคผนวก ข

การตรวจสอบความเป็นมิติเดียวของแบบทดสอบวัดความถนัดทางการเรียนแต่ละด้าน

ตาราง 33 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	7.072	23.574	23.574
2	1.509	5.030	28.603
3	1.364	4.547	33.150
4	1.052	3.507	36.657
5	1.005	3.350	40.007
6	.989	3.296	43.303
7	.947	3.156	46.459
8	.914	3.047	49.507
9	.888	2.960	52.467
10	.835	2.782	55.249
11	.816	2.720	57.970
12	.813	2.711	60.681
13	.787	2.623	63.303
14	.784	2.615	65.918
15	.754	2.513	68.431
16	.737	2.457	70.888
17	.734	2.448	73.336
18	.713	2.375	75.711
19	.692	2.308	78.019
20	.681	2.271	80.290
21	.674	2.247	82.537
22	.661	2.202	84.738
23	.643	2.142	86.881
24	.622	2.075	88.956
25	.616	2.053	91.009
26	.579	1.929	92.937
27	.569	1.896	94.833
28	.553	1.842	96.675
29	.527	1.755	98.430
30	.471	1.570	100.000

Extraction Method: Principal Component Analysis.

ตาราง 34 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	4.584	11.459	11.459
2	1.821	4.554	16.013
3	1.296	3.241	19.254
4	1.247	3.116	22.370
5	1.220	3.051	25.421
6	1.189	2.973	28.394
7	1.119	2.798	31.192
8	1.090	2.724	33.916
9	1.078	2.694	36.610
10	1.062	2.655	39.265
11	1.022	2.554	41.819
12	1.018	2.546	44.364
13	.989	2.471	46.836
14	.980	2.450	49.286
15	.978	2.444	51.730
16	.951	2.377	54.107
17	.940	2.350	56.457
18	.925	2.314	58.770
19	.911	2.278	61.048
20	.900	2.251	63.299
21	.889	2.222	65.520
22	.857	2.144	67.664
23	.845	2.114	69.777
24	.838	2.096	71.873
25	.823	2.058	73.931
26	.803	2.007	75.938
27	.801	2.004	77.942
28	.796	1.990	79.932
29	.784	1.961	81.893
30	.753	1.882	83.775
31	.741	1.852	85.627
32	.732	1.829	87.455
33	.723	1.806	89.262
34	.699	1.747	91.009
35	.682	1.706	92.715
36	.661	1.653	94.368
37	.626	1.564	95.932
38	.609	1.523	97.455
39	.578	1.445	98.900
40	.440	1.100	100.000

Extraction Method: Principal Component Analysis.

ตาราง 35 ค่า Eigenvalues ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	6.093	20.308	20.308
2	1.419	4.731	25.040
3	1.287	4.291	29.331
4	1.158	3.861	33.192
5	1.071	3.570	36.762
6	1.043	3.477	40.239
7	.991	3.304	43.542
8	.955	3.183	46.726
9	.931	3.102	49.828
10	.875	2.918	52.746
11	.852	2.839	55.586
12	.840	2.800	58.386
13	.827	2.757	61.143
14	.813	2.710	63.853
15	.789	2.629	66.483
16	.778	2.593	69.075
17	.775	2.585	71.660
18	.761	2.538	74.198
19	.748	2.493	76.691
20	.718	2.392	79.083
21	.695	2.317	81.399
22	.678	2.261	83.660
23	.670	2.235	85.895
24	.656	2.187	88.082
25	.636	2.119	90.201
26	.619	2.063	92.265
27	.611	2.035	94.300
28	.596	1.985	96.285
29	.564	1.880	98.165
30	.551	1.835	100.000

Extraction Method: Principal Component Analysis.

ภาคผนวก ค

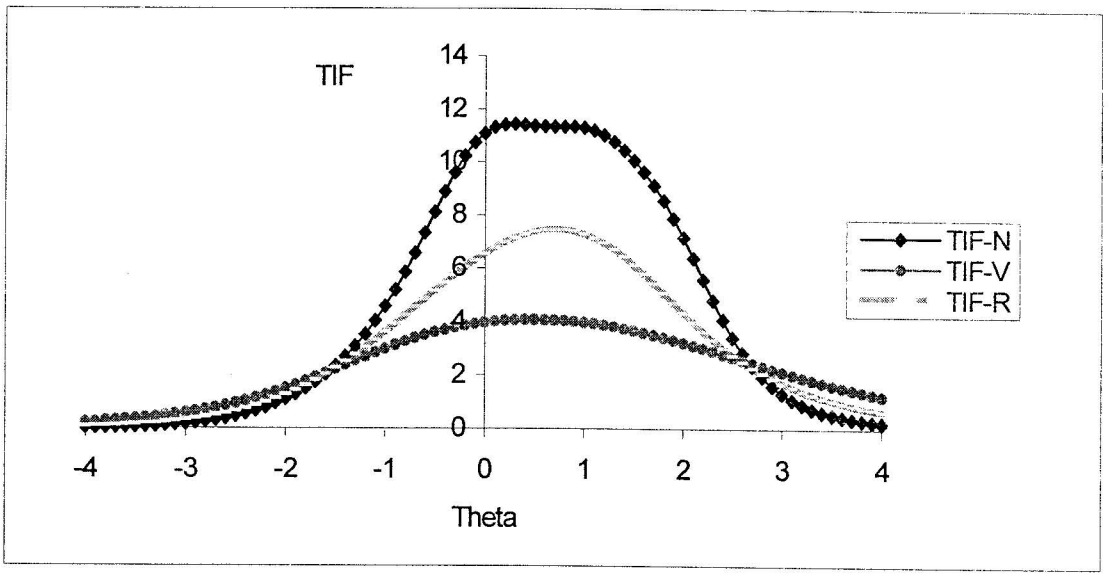
ค่าพารามิเตอร์ข้อสอบและฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข ด้านภาษาและด้านเหตุผล ที่ได้จากตัวอย่างขนาดใหญ่ 2000 คน

ตาราง 36 ค่าพารามิเตอร์ข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลขและด้านเหตุผล

ด้านตัวเลข				ด้านเหตุผล			
ข้อที่	a	b	c	ข้อที่	a	b	c
1	0.877	-1.376	0.127	1	0.587	-0.658	0.084
2	1.049	-0.766	0.143	2	0.358	0.106	0.102
3	1.108	-0.385	0.178	3	0.482	0.211	0.095
4	1.008	-0.288	0.111	4	0.571	0.893	0.110
5	1.339	0.233	0.090	5	0.493	1.989	0.037
6	1.006	0.494	0.101	6	0.784	-0.223	0.080
7	1.71	-0.082	0.134	7	0.456	-0.23	0.092
8	0.678	0.335	0.051	8	1.018	0.487	0.185
9	0.716	0.772	0.121	9	0.605	0.454	0.069
10	0.905	0.534	0.151	10	0.453	0.947	0.082
11	0.746	-0.602	0.091	11	0.785	-0.011	0.062
12	1.396	-0.195	0.068	12	0.575	-0.427	0.065
13	0.962	-0.225	0.101	13	0.601	0.623	0.040
14	1.159	0.142	0.128	14	0.731	-0.057	0.042
15	1.609	1.497	0.387	15	0.882	-0.007	0.041
16	0.323	0.929	0.074	16	0.523	0.089	0.078
17	1.051	0.66	0.141	17	0.807	0.324	0.071
18	0.988	0.643	0.147	18	0.618	0.703	0.079
19	1.325	0.945	0.144	19	1.061	0.778	0.106
20	1.784	1.103	0.121	20	0.619	0.816	0.102
21	0.808	-0.568	0.089	21	0.898	-0.667	0.054
22	0.775	-0.808	0.129	22	0.694	0.209	0.045
23	0.598	-0.15	0.119	23	0.635	0.093	0.101
24	0.726	-0.103	0.155	24	0.557	0.138	0.084
25	0.833	0.301	0.171	25	0.805	0.216	0.151
26	1.268	1.125	0.324	26	0.836	0.475	0.171
27	1.222	0.759	0.279	27	0.492	1.303	0.087
28	1.421	1.640	0.145	28	0.79	1.098	0.238
29	1.794	1.901	0.152	29	0.754	1.226	0.113
30	1.033	1.402	0.131	30	1.271	1.374	0.099
Mean	1.074	0.329	0.143	Mean	0.691	0.409	0.092
S	0.360	0.799	0.071	S	0.203	0.621	0.045

ตาราง 37 ค่าพารามิเตอร์ข้อสอบของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา

ข้อที่	a	b	c	ข้อที่	a	b	c
1	0.209	-4.102	0.129	21	0.406	-0.770	0.122
2	0.876	-0.698	0.068	22	0.275	0.604	0.114
3	0.654	-0.046	0.137	23	0.556	1.207	0.047
4	0.471	0.240	0.178	24	0.434	-0.484	0.125
5	0.749	-1.296	0.088	25	0.501	-0.045	0.112
6	0.249	0.940	0.119	26	0.501	-0.057	0.123
7	0.637	0.892	0.202	27	0.291	2.268	0.106
8	0.566	0.443	0.069	28	0.513	0.593	0.158
9	0.451	1.128	0.094	29	0.293	1.675	0.133
10	0.522	1.143	0.142	30	0.577	1.281	0.091
11	0.255	2.840	0.193	31	0.311	4.004	0.215
12	0.657	1.459	0.105	32	0.574	0.106	0.113
13	0.300	-0.803	0.107	33	0.701	2.070	0.251
14	0.296	-0.508	0.115	34	0.370	-0.686	0.102
15	0.497	-0.336	0.125	35	0.693	-0.211	0.122
16	0.290	-0.438	0.120	36	0.391	2.015	0.137
17	0.537	0.244	0.100	37	0.232	3.151	0.152
18	0.327	1.960	0.092	38	0.325	-0.522	0.142
19	0.216	3.342	0.151	39	0.423	-0.055	0.101
20	0.503	2.200	0.123	40	0.422	0.962	0.136
				Mean	0.451	0.643	0.126
				S	0.164	1.493	0.040



ภาพประกอบ 25 ฟังก์ชันสารสนเทศของแบบทดสอบวัดความถนัดทางการเรียนแต่ละฉบับ

ภาคผนวก ง

ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียน
จากตัวอย่างจำลองขนาด 300 500 และ 800 คน ที่ได้จากการสุ่มซ้ำ

ตาราง 38 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างจำลองขนาด 300 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	15.4667	6.7810	0.4384	-1.090	0.149	0.8818	0.8659	0.8832
2	14.6800	6.6883	0.4556	-1.040	0.196	0.8773	0.8611	0.8788
3	14.7267	6.5069	0.4418	-1.038	0.165	0.8672	0.8513	0.8692
4	15.0300	6.9371	0.4616	-1.005	0.110	0.8880	0.8733	0.8898
5	15.4000	6.8793	0.4467	-1.054	0.096	0.8850	0.8707	0.8866
6	15.1367	6.9281	0.4577	-1.028	0.031	0.8884	0.8729	0.8900
7	15.4833	6.8000	0.4392	-1.115	0.034	0.8831	0.8669	0.8848
8	15.2967	6.8394	0.4471	-1.031	0.150	0.8846	0.8687	0.8865
9	15.8700	7.1402	0.4499	-1.198	-0.053	0.8964	0.8828	0.8979
10	15.0667	6.9635	0.4622	-1.130	0.155	0.8883	0.8745	0.8904
11	15.0900	6.5675	0.4352	-1.020	0.136	0.8723	0.8546	0.8739
12	14.7067	6.6800	0.4542	-1.043	0.207	0.8757	0.8607	0.8773
13	15.0800	7.0189	0.4654	-1.074	0.176	0.8908	0.8770	0.8923
14	15.5867	6.9414	0.4453	-1.018	0.084	0.8883	0.8737	0.8899
15	16.0100	6.9679	0.4352	-1.131	-0.060	0.8919	0.8754	0.8935
16	14.9700	6.5691	0.4388	-0.904	0.226	0.8716	0.8547	0.8737
17	15.6133	6.5432	0.4191	-0.948	0.029	0.8693	0.8536	0.8710
18	15.0133	6.6589	0.4435	-0.947	0.158	0.8761	0.8595	0.8775
19	15.0300	6.7953	0.4521	-1.036	0.118	0.8817	0.8665	0.8837
20	15.4700	7.1198	0.4602	-1.104	0.075	0.8947	0.8816	0.8964
21	15.0867	7.1012	0.4707	-1.051	0.143	0.8943	0.8806	0.8958
22	15.4433	6.5877	0.4266	-1.024	0.145	0.8734	0.8559	0.8754
23	15.6033	6.9906	0.4480	-1.078	0.017	0.8908	0.8760	0.8921
24	15.4200	7.0732	0.4587	-1.181	-0.005	0.8929	0.8795	0.8949
25	15.5267	6.5803	0.4238	-0.928	0.140	0.8729	0.8555	0.8743
26	15.3567	6.6244	0.4314	-0.990	0.168	0.8744	0.8578	0.8763
27	14.6467	6.8041	0.4645	-1.042	0.177	0.8820	0.8670	0.8841
28	15.2133	7.1065	0.4671	-1.119	0.080	0.8953	0.8809	0.8967
29	15.2800	6.8492	0.4482	-1.077	0.088	0.8846	0.8692	0.8866
30	15.2967	6.7325	0.4401	-1.022	0.092	0.8792	0.8634	0.8810
รวมทั้งหมด	15.2533	6.8259	0.4476	-1.049	0.108	0.8831	0.8677	0.8848

ตาราง 39 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างจำลองขนาด 500 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	15.3960	7.0373	0.4571	-1.169	0.121	0.8903	0.8779	0.8918
2	15.7960	6.6714	0.4223	-1.028	0.031	0.8774	0.8607	0.8791
3	15.4080	6.6947	0.4345	-1.022	0.086	0.8769	0.8615	0.8788
4	15.3520	6.8388	0.4455	-1.020	0.083	0.8826	0.8687	0.8843
5	15.0440	6.9069	0.4591	-1.054	0.141	0.8869	0.8718	0.8885
6	15.6520	6.8446	0.4373	-1.115	-0.013	0.8849	0.8692	0.8863
7	15.3640	6.8359	0.4449	-1.123	0.121	0.8836	0.8685	0.8852
8	15.7040	6.6420	0.4229	-1.042	0.022	0.8756	0.8590	0.8774
9	15.0420	6.8934	0.4583	-1.093	0.110	0.8859	0.8712	0.8875
10	14.9680	6.6398	0.4436	-1.022	0.100	0.8743	0.8585	0.8759
11	14.5740	6.6586	0.4569	-1.068	0.181	0.8734	0.8596	0.8756
12	15.0800	6.7898	0.4503	-1.004	0.177	0.8802	0.8662	0.8819
13	15.1740	6.8141	0.4491	-1.065	0.085	0.8829	0.8674	0.8845
14	14.7800	6.6823	0.4521	-1.057	0.053	0.8766	0.8608	0.8784
15	15.2200	6.8957	0.4531	-1.045	0.161	0.8859	0.8714	0.8873
16	15.4400	6.7404	0.4366	-1.037	0.125	0.8797	0.8639	0.8810
17	15.3820	6.9769	0.4536	-1.117	0.076	0.8897	0.8752	0.8915
18	15.6240	6.7247	0.4304	-0.994	-0.001	0.8799	0.8632	0.8817
19	15.5440	7.0116	0.4511	-1.083	0.028	0.8909	0.8769	0.8924
20	14.8700	6.6246	0.4455	-0.953	0.186	0.8735	0.8577	0.8754
21	15.0360	6.6858	0.4447	-0.988	0.121	0.8764	0.8609	0.8781
22	15.1820	6.7510	0.4447	-1.017	0.081	0.8804	0.8643	0.8821
23	15.0780	6.9956	0.4640	-1.089	0.126	0.8893	0.8759	0.8906
24	15.0620	6.6909	0.4442	-1.107	0.115	0.8778	0.8612	0.8797
25	15.4280	6.8254	0.4424	-1.048	0.089	0.8831	0.8681	0.8847
26	15.2260	6.5606	0.4309	-1.029	0.111	0.8720	0.8543	0.8737
27	15.214	6.8283	0.4488	-1.130	0.075	0.8831	0.8681	0.8850
28	14.888	6.6880	0.4492	-0.943	0.157	0.8763	0.8610	0.8779
29	15.264	6.8314	0.4475	-1.017	0.080	0.8835	0.8683	0.8852
30	15.444	6.8235	0.4418	-1.033	0.107	0.8831	0.8680	0.8846
รวมทั้งหมด	15.2412	6.7868	0.4454	-1.050	0.098	0.8812	0.8660	0.8829

ตาราง 40 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างจำลองขนาด 800 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	15.2513	6.7083	0.4399	-1.046	0.085	0.8775	0.8621	0.8794
2	15.0555	6.6937	0.4446	-0.961	0.171	0.8765	0.8613	0.8782
3	15.4638	6.8139	0.4406	-1.072	0.054	0.8825	0.8675	0.8841
4	15.3463	6.8575	0.4469	-1.044	0.071	0.8843	0.8696	0.8860
5	15.0975	6.8289	0.4523	-1.025	0.119	0.8841	0.8681	0.8858
6	15.1938	6.8899	0.4535	-1.049	0.131	0.8851	0.8711	0.8867
7	14.9700	6.7812	0.4530	-1.000	0.146	0.8808	0.8658	0.8824
8	15.3325	6.7075	0.4375	-1.025	0.102	0.8786	0.8621	0.8786
9	15.1088	6.7792	0.4487	-0.984	0.164	0.8808	0.8657	0.8824
10	15.3425	6.6401	0.4328	-1.036	0.051	0.8760	0.8586	0.8778
11	15.4725	6.7666	0.4373	-1.081	0.046	0.8809	0.8652	0.8825
12	15.3875	6.8281	0.4437	-1.053	0.058	0.8839	0.8682	0.8855
13	15.5913	6.8433	0.4389	-1.079	-0.025	0.8841	0.8691	0.8860
14	15.0863	6.8381	0.4533	-1.058	0.128	0.8837	0.8686	0.8854
15	15.2225	6.9044	0.4536	-1.137	0.065	0.8860	0.8718	0.8879
16	15.1800	6.8000	0.4480	-1.047	0.128	0.8813	0.8667	0.8830
17	15.2125	6.9732	0.4584	-1.059	0.144	0.8891	0.8750	0.8906
18	15.3450	6.7684	0.4411	-1.076	0.043	0.8804	0.8652	0.8821
19	15.2825	6.7668	0.4428	-1.054	0.081	0.8801	0.8651	0.8816
20	15.2700	6.9302	0.4538	-1.044	0.129	0.8872	0.8730	0.8888
21	14.9000	6.8888	0.4623	-1.044	0.131	0.8856	0.8710	0.8872
22	15.4663	6.8699	0.4442	-1.083	0.021	0.8848	0.8702	0.8863
23	15.5600	6.8401	0.4396	-1.089	0.082	0.8845	0.8689	0.8860
24	15.0538	6.9440	0.4613	-1.056	0.136	0.8878	0.8736	0.8895
25	14.9038	6.7451	0.4526	-1.025	0.153	0.8793	0.8640	0.8812
26	14.9150	6.6825	0.4480	-0.957	0.168	0.8759	0.8607	0.8776
27	15.4375	6.9406	0.4496	-1.086	0.054	0.8882	0.8736	0.8899
28	15.3788	6.7524	0.4391	-1.007	0.115	0.8802	0.8644	0.8817
29	15.1313	6.7738	0.4477	-1.040	0.097	0.8812	0.8654	0.8830
30	15.0975	6.8331	0.4526	-1.101	0.139	0.8833	0.8683	0.8850
รวมทั้งหมด	15.2352	6.8130	0.4473	-1.047	0.100	0.8825	0.8673	0.8841

ตาราง 41 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างจำลองขนาด 300 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	19.1967	6.2263	0.3243	-0.841	-0.074	0.7812	0.7615	0.7828
2	19.9000	6.4962	0.3264	-0.518	-0.155	0.8015	0.7826	0.8028
3	19.6633	6.3628	0.3236	-0.754	-0.126	0.7917	0.7724	0.7934
4	19.6167	5.9768	0.3047	-0.584	-0.161	0.7624	0.7386	0.7641
5	20.1033	6.2168	0.3092	-0.553	-0.257	0.7838	0.7603	0.7856
6	19.3133	6.0407	0.3128	-0.591	-0.201	0.7671	0.7449	0.7690
7	20.3667	6.0016	0.2947	-0.608	-0.271	0.7657	0.7410	0.7676
8	20.0633	6.0322	0.3007	-0.708	-0.121	0.7658	0.7438	0.7676
9	19.7233	6.0984	0.3092	-0.515	-0.147	0.7729	0.7499	0.7747
10	19.4500	6.5756	0.3381	-0.682	-0.129	0.8049	0.7886	0.8066
11	19.7300	6.0231	0.3053	-0.606	-0.116	0.7659	0.7430	0.7677
12	19.4600	6.2838	0.3229	-0.596	-0.225	0.7865	0.7661	0.7879
13	19.4833	6.2095	0.3187	-0.612	-0.188	0.7800	0.7598	0.7817
14	19.8133	6.1382	0.3098	-0.572	-0.092	0.7750	0.7534	0.7763
15	19.5000	6.3243	0.3243	-0.769	-0.209	0.7901	0.7694	0.7921
16	19.8967	6.2500	0.3141	-0.628	-0.227	0.7834	0.7631	0.7853
17	19.6600	6.4110	0.3261	-0.921	-0.207	0.7942	0.7762	0.7961
18	19.1633	6.1815	0.3226	-0.617	-0.018	0.7784	0.7577	0.7797
19	19.2533	6.2439	0.3243	-0.619	0.036	0.7846	0.7629	0.7864
20	20.0900	6.1270	0.3050	-0.793	-0.202	0.7730	0.7524	0.7744
21	19.1567	6.2183	0.3246	-0.627	-0.207	0.7791	0.7609	0.7808
22	19.8633	6.1185	0.3080	-0.732	-0.249	0.7732	0.7517	0.7749
23	20.1033	5.9507	0.2960	-0.667	-0.246	0.7613	0.7360	0.7635
24	20.1967	6.2499	0.3095	-0.841	-0.214	0.7850	0.7631	0.7868
25	19.6200	6.1107	0.3115	-0.441	-0.200	0.7745	0.7511	0.7759
26	19.3100	6.4045	0.3317	-0.835	-0.065	0.7951	0.7759	0.7970
27	19.2433	6.3739	0.3312	-0.818	-0.102	0.7925	0.7735	0.7943
28	20.4100	5.9182	0.2900	-0.580	-0.120	0.7576	0.7329	0.7594
29	19.6167	6.2658	0.3194	-0.472	-0.115	0.7846	0.7645	0.7859
30	19.8933	5.7926	0.2912	-0.315	-0.219	0.7441	0.7200	0.7456
รวมทั้งหมด	19.6953	6.1874	0.3143	-0.647	-0.161	0.7785	0.7572	0.7802

ตาราง 42 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างจำลองขนาด 500 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	19.918	6.4550	0.3241	-0.662	-0.105	0.7969	0.7795	0.7986
2	20.0300	6.3392	0.3165	-0.756	-0.134	0.7902	0.7704	0.7919
3	19.6560	6.0705	0.3088	-0.517	-0.136	0.7686	0.7474	0.7700
4	20.1960	6.3351	0.3137	-0.698	-0.212	0.7888	0.7701	0.7910
5	19.5000	6.2360	0.3198	-0.771	-0.075	0.7840	0.7621	0.7854
6	19.6180	6.1939	0.3157	-0.680	-0.187	0.7798	0.7584	0.7815
7	19.6260	6.0841	0.3100	-0.604	-0.164	0.7719	0.7487	0.7735
8	19.4300	6.1960	0.3189	-0.761	-0.139	0.7792	0.7587	0.7808
9	20.1100	6.3355	0.3150	-0.609	-0.268	0.7901	0.7701	0.7917
10	19.5860	6.1595	0.3145	-0.588	-0.204	0.7763	0.7554	0.7778
11	19.9360	6.2381	0.3129	-0.708	-0.203	0.7822	0.7621	0.7837
12	19.9020	6.1089	0.3069	-0.646	-0.138	0.7728	0.7508	0.7745
13	19.8540	6.2731	0.3160	-0.811	-0.117	0.7848	0.7650	0.7865
14	19.7780	6.4653	0.3269	-0.600	-0.198	0.7986	0.7803	0.8004
15	19.9180	6.1634	0.3094	-0.569	-0.237	0.7775	0.7556	0.7788
16	19.3840	6.2532	0.3226	-0.805	-0.035	0.7839	0.7636	0.7853
17	19.9720	6.1952	0.3102	-0.581	-0.156	0.7806	0.7584	0.7822
18	19.8600	6.3716	0.3208	-0.721	-0.100	0.7919	0.7730	0.7936
19	19.6900	6.4219	0.3262	-0.665	-0.190	0.7977	0.7770	0.7994
20	20.2920	5.8750	0.2895	-0.685	-0.240	0.7527	0.7286	0.7546
21	19.9240	6.3837	0.3204	-0.730	-0.188	0.7930	0.7740	0.7947
22	19.8860	6.3081	0.3172	-0.811	-0.214	0.7871	0.7679	0.7888
23	19.7900	6.0335	0.3049	-0.597	-0.062	0.7664	0.7439	0.7681
24	20.0040	6.3832	0.3191	-0.513	-0.178	0.7927	0.7739	0.7943
25	19.4900	6.5043	0.3337	-0.714	-0.168	0.8009	0.7834	0.8025
26	19.4840	6.0245	0.3092	-0.665	-0.120	0.7654	0.7432	0.7671
27	19.7100	6.1143	0.3102	-0.655	-0.102	0.7723	0.7514	0.7738
28	19.7320	6.1017	0.3092	-0.566	-0.170	0.7711	0.7502	0.7729
29	19.3040	6.3700	0.3300	-0.793	-0.137	0.7930	0.7732	0.7944
30	19.3600	6.4577	0.3336	-0.793	-0.083	0.7989	0.7799	0.8004
รวมทั้งหมด	19.7647	6.2484	0.3162	-0.676	-0.155	0.7830	0.7625	0.7846

ตาราง 43 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างจำลองขนาด 800 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	19.5913	6.2770	0.3204	-0.794	-0.080	.7848	.7654	.7863
2	19.6600	6.3279	0.3219	-0.811	-0.181	.7890	.7696	.7905
3	19.8750	6.4322	0.3236	-0.690	-0.172	.7971	.7778	.7985
4	19.7800	6.3614	0.3216	-0.745	-0.136	.7919	.7722	.7937
5	19.7725	6.2307	0.3151	-0.669	-0.212	.7813	.7615	.7827
6	19.8250	6.0863	0.3070	-0.663	-0.114	.7698	.7488	.7714
7	19.9975	6.3919	0.3196	-0.707	-0.269	.7937	.7746	.7951
8	19.4313	6.1898	0.3185	-0.627	-0.185	.7786	.7582	.7802
9	19.7500	6.2473	0.3163	-0.680	-0.138	.7831	.7629	.7846
10	19.8380	6.1741	0.3112	-0.699	-0.138	.7774	.7566	.7790
11	19.6363	6.2243	0.3170	-0.755	-0.144	.7801	.7610	.7817
12	19.6950	6.2332	0.3165	-0.628	-0.139	.7816	.7617	.7830
13	19.6563	6.3264	0.3219	-0.638	-0.245	.7878	.7695	.7895
14	19.7313	6.2882	0.3187	-0.578	-0.175	.7861	.7663	.7877
15	19.8838	6.3291	0.3183	-0.649	-0.183	.7899	.7696	.7914
16	19.4738	6.1820	0.3175	-0.690	-0.098	.7771	.7575	.7788
17	19.6413	6.3864	0.3252	-0.735	-0.165	.7914	.7743	.7931
18	19.6450	6.1722	0.3142	-0.691	-0.231	.7782	.7565	.7797
19	19.5200	6.2513	0.3203	-0.778	-0.180	.7834	.7633	.7850
20	19.8913	6.4369	0.3236	-0.705	-0.213	.7968	.7781	.7984
21	19.6563	6.2355	0.3172	-0.685	-0.135	.7825	.7618	.7843
22	19.6075	6.2474	0.3186	-0.731	-0.151	.7825	.7630	.7841
23	19.7813	6.1456	0.3107	-0.784	-0.149	.7750	.7541	.7769
24	19.6363	6.1690	0.3142	-0.734	-0.132	.7764	.7562	.7780
25	19.7888	6.2000	0.3133	-0.719	-0.174	.7788	.7589	.7807
26	19.9438	6.3072	0.3162	-0.715	-0.265	.7877	.7678	.7894
27	19.7275	6.3149	0.3201	-0.672	-0.204	.7886	.7685	.7903
28	19.6238	6.2876	0.3204	-0.792	-0.127	.7859	.7663	.7876
29	19.7588	6.2584	0.3167	-0.754	-0.158	.7840	.7638	.7859
30	19.6488	6.2659	0.3189	-0.707	-0.172	.7853	.7645	.7871
รวมทั้งหมด	19.7156	6.2660	0.3178	-0.708	-0.169	0.7842	0.7643	0.7858

ตาราง 44 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 300 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	14.4200	6.5687	0.4555	-0.805	0.134	0.8636	0.8549	0.8643
2	14.3633	6.6928	0.4660	-0.839	0.065	0.8693	0.8616	0.8701
3	15.0100	6.7445	0.4493	-0.921	0.099	0.8738	0.8639	0.8745
4	14.3267	6.1006	0.4258	-0.739	0.204	0.8368	0.8264	0.8381
5	14.6467	6.4974	0.4436	-0.893	0.074	0.8596	0.8508	0.8604
6	14.5167	6.5681	0.4525	-0.894	0.198	0.8632	0.8548	0.8639
7	14.6800	6.6361	0.4521	-0.926	-0.059	0.8662	0.8584	0.8673
8	13.9067	6.4703	0.4653	-0.933	0.124	0.8581	0.8501	0.8593
9	13.9233	6.4021	0.4598	-0.820	0.258	0.8559	0.8462	0.8565
10	14.5000	6.5119	0.4491	-0.926	0.038	0.8606	0.8517	0.8613
11	14.6100	6.5407	0.4477	-0.875	0.106	0.8620	0.8532	0.8628
12	14.0067	6.4248	0.4587	-0.822	0.059	0.8573	0.8473	0.8581
13	14.3833	6.5232	0.4535	-0.799	0.140	0.8617	0.8525	0.8627
14	14.3533	6.4447	0.4490	-0.898	0.114	0.8578	0.8480	0.8586
15	14.1333	6.4468	0.4561	-0.728	0.129	0.8582	0.8484	0.8590
16	14.3533	6.4473	0.4492	-0.906	0.118	0.8565	0.8482	0.8571
17	14.3533	6.4067	0.4464	-0.864	0.116	0.8551	0.8458	0.8561
18	13.7833	6.4939	0.4711	-0.914	0.178	0.8600	0.8517	0.8613
19	14.0800	6.5235	0.4633	-0.851	0.166	0.8597	0.8529	0.8605
20	14.1600	6.3726	0.4500	-0.919	0.184	0.8530	0.8440	0.8538
21	14.6100	6.7178	0.4598	-0.865	0.098	0.8712	0.8627	0.8717
22	14.1267	6.7728	0.4794	-0.941	0.132	0.8735	0.8659	0.8741
23	13.9267	6.4415	0.4625	-0.856	0.146	0.8575	0.8485	0.8583
24	14.3100	6.4622	0.4516	-1.016	0.084	0.8587	0.8491	0.8599
25	14.1667	6.4460	0.4550	-0.842	0.110	0.8572	0.8483	0.8583
26	14.6300	6.5745	0.4494	-0.944	0.160	0.8603	0.8551	0.8637
27	14.3000	6.6012	0.4616	-0.897	0.128	0.8658	0.8568	0.8664
28	14.6100	6.4521	0.4416	-0.792	0.114	0.8573	0.8482	0.8579
29	14.5400	6.5374	0.4496	-0.905	0.120	0.8618	0.8531	0.8627
30	13.9867	6.2012	0.4434	-0.697	0.210	0.8436	0.8336	0.8448
รวมทั้งหมด	14.3239	6.5008	0.4539	-0.8676	0.1249	0.8598	0.8511	0.8608

ตาราง 45 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 500 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	14.1100	6.5062	0.4611	-0.925	0.065	0.8602	0.8518	0.8612
2	13.9540	6.5107	0.4666	-0.803	0.245	0.8604	0.8523	0.8612
3	14.0640	6.6715	0.4744	-0.965	0.190	0.8679	0.8608	0.8687
4	13.9260	6.4996	0.4667	-0.812	0.224	0.8601	0.8518	0.8608
5	14.1600	6.6825	0.4719	-0.895	0.120	0.8689	0.8613	0.8698
6	14.0840	6.7093	0.4764	-0.910	0.161	0.8701	0.8628	0.8709
7	14.1600	6.5700	0.4640	-0.826	0.099	0.8647	0.8553	0.8657
8	14.2500	6.5902	0.4625	-1.002	0.120	0.8642	0.8563	0.8649
9	14.1160	6.4035	0.4536	-0.809	0.171	0.8555	0.8459	0.8562
10	14.4720	6.5308	0.4513	-0.898	0.147	0.8637	0.8528	0.8645
11	14.2900	6.6116	0.4627	-0.858	0.217	0.8654	0.8574	0.8662
12	14.3280	6.4233	0.4483	-0.872	0.144	0.8558	0.8468	0.8567
13	14.1480	6.6262	0.4683	-0.839	0.143	0.8662	0.8583	0.8671
14	14.3700	6.5436	0.4554	-0.897	-0.014	0.8625	0.8536	0.8632
15	14.2220	6.2628	0.4404	-0.758	0.165	0.8465	0.8372	0.8474
16	13.9880	6.5103	0.4654	-0.756	0.240	0.8600	0.8523	0.8607
17	14.1660	6.6467	0.4692	-0.914	0.109	0.8675	0.8594	0.8682
18	14.0360	6.4257	0.4578	-0.816	0.173	0.8565	0.8474	0.8571
19	14.4360	6.4982	0.4501	-0.837	0.107	0.8600	0.8510	0.8606
20	14.5180	6.3389	0.4366	-0.891	0.120	0.8604	0.8416	0.8512
21	14.4980	6.5940	0.4548	-0.954	0.111	0.8648	0.8562	0.8654
22	14.9800	6.6266	0.4424	-0.947	-0.011	0.8660	0.8578	0.8670
23	14.4800	6.5589	0.4530	-0.863	0.150	0.8632	0.8543	0.8639
24	14.1580	6.3987	0.4519	-0.884	0.174	0.8538	0.8456	0.8544
25	14.0280	6.5496	0.4669	-0.868	0.141	0.8625	0.8544	0.8634
26	14.1340	6.6429	0.4700	-0.946	0.156	0.8671	0.8592	0.8679
27	13.7800	6.4398	0.4673	-0.842	0.199	0.8572	0.8486	0.8581
28	14.3680	6.5399	0.4552	-0.866	0.170	0.8631	0.8534	0.8636
29	14.2660	6.7826	0.4754	-0.949	0.141	0.8740	0.8662	0.8746
30	14.2900	6.5628	0.4593	-0.852	0.098	0.8630	0.8547	0.8641
รวมทั้งหมด	14.2260	6.5419	0.4600	-0.8751	0.1425	0.8624	0.8536	0.8340

ตาราง 46 ค่าสถิติพื้นฐานและความเชื่อมั่นของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากตัวอย่างจำลองขนาด 800 คน

ตัวอย่าง	ค่าสถิติพื้นฐาน					ค่าความเชื่อมั่น		
ขนาด	$\bar{X}$	S.D	C.V.	ความโด่ง	ความเบ้	KR-20	KR-21	r_B
1	14.1775	6.6824	0.4713	-0.900	0.173	0.8691	0.8613	0.8698
2	14.1975	6.3859	0.4498	-0.783	0.190	0.8530	0.8448	0.8536
3	14.3750	6.5269	0.4540	-0.889	0.119	0.8611	0.8527	0.8619
4	14.2038	6.6857	0.4707	-0.913	0.127	0.8693	0.8614	0.8700
5	14.2450	6.5926	0.4628	-0.919	0.112	0.8643	0.8564	0.8650
6	14.5088	6.4602	0.4453	-0.947	0.083	0.8571	0.8488	0.8577
7	14.2775	6.5599	0.4595	-0.885	0.147	0.8628	0.8546	0.8634
8	14.4463	6.5965	0.4566	-0.887	0.150	0.8646	0.8564	0.8652
9	14.2086	6.5403	0.4603	-0.873	0.167	0.8620	0.9803	0.8626
10	14.1013	6.5051	0.4613	-0.871	0.167	0.8604	0.8518	0.8611
11	14.3775	6.6664	0.4637	-0.873	0.077	0.8645	0.8602	0.8693
12	14.1900	6.4123	0.4519	-0.819	0.157	0.8551	0.8463	0.8558
13	14.3388	6.4892	0.4526	-0.902	0.076	0.8591	0.8506	0.8599
14	14.2338	6.4718	0.4547	-0.910	0.098	0.8583	0.8497	0.8588
15	14.1475	6.4856	0.4584	-0.845	0.157	0.8591	0.8506	0.8598
16	14.0638	6.3350	0.4504	-0.774	0.255	0.8511	0.8419	0.8519
17	14.4925	6.6639	0.4598	-0.859	0.090	0.8687	0.8600	0.8695
18	14.2575	6.4386	0.4516	-0.809	0.134	0.8572	0.8478	0.8579
19	14.2650	6.5231	0.4573	-0.887	0.172	0.8611	0.8526	0.8620
20	14.1550	6.6384	0.4690	-0.968	0.180	0.8667	0.8590	0.8674
21	13.9313	6.4677	0.4643	-0.799	0.197	0.8588	0.8500	0.8596
22	14.1688	6.3733	0.4498	-0.849	0.086	0.8539	0.8441	0.8549
23	14.1488	6.4669	0.4571	-0.864	0.120	0.8577	0.8496	0.8584
24	14.0188	6.7300	0.4801	-0.970	0.180	0.8714	0.8639	0.8720
25	13.8025	6.4448	0.4669	-0.819	0.178	0.8582	0.8489	0.8591
26	14.1763	6.5273	0.4604	-0.928	0.085	0.8606	0.8529	0.8614
27	14.4163	6.4459	0.4471	-0.948	0.114	0.8568	0.8480	0.8575
28	14.4425	6.5267	0.4519	-0.928	0.077	0.8612	0.8526	0.8620
29	14.0850	6.4575	0.4585	-0.854	0.200	0.8585	0.8491	0.8591
30	14.5125	6.4163	0.4421	-0.892	0.134	0.8558	0.8462	0.8566
รวมทั้งหมด	14.2322	6.6975	0.4707	-0.8788	0.1401	0.8606	0.8564	0.8614

ภาคผนวก จ

ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูป
แบบตำแหน่งเปอร์เซ็นไทล์และแบบคะแนนปกติ จากแบบทดสอบวัดความถนัด
ทางการเรียนแต่ละด้าน เมื่อขนาดตัวอย่างจำลอง 300 500 และ 800 คน

ตาราง 47 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาด 300 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	10.1092	10.0167	9.4747	9.4156	9.9686	0.4340	0.4297	0.3939	0.3915	0.4152	0.3746
2	10.3732	10.3034	9.6969	9.6366	10.2316	0.4282	0.4256	0.3918	0.3894	0.4140	0.3883
3	10.5673	10.4332	9.9212	9.8447	10.3537	0.4270	0.4217	0.3904	0.3874	0.4081	0.3770
4	9.8341	9.7668	9.1820	9.1070	9.6852	0.3894	0.3867	0.3567	0.3538	0.3768	0.3695
5	9.8734	9.7573	9.2416	9.1775	9.6761	0.4049	0.4005	0.3712	0.3686	0.3896	0.3572
6	9.9848	9.8601	9.2687	9.2007	9.7517	0.4214	0.4167	0.3834	0.3806	0.4040	0.3538
7	10.0533	9.9205	9.3253	9.2577	9.8087	0.4237	0.4187	0.3819	0.3791	0.4030	0.3723
8	10.0155	9.9050	9.3444	9.2676	9.8463	0.3970	0.3932	0.3627	0.3598	0.3832	0.3660
9	9.5937	9.4677	8.9475	8.8820	9.3754	0.4181	0.4133	0.3775	0.3747	0.3967	0.3532
10	9.9330	9.8356	9.3089	9.2205	9.7584	0.4114	0.4076	0.3763	0.3727	0.3951	0.3674
11	10.5372	10.4448	9.8244	9.7612	10.3718	0.4328	0.4288	0.3928	0.3902	0.4154	0.3813
12	10.4239	10.3085	9.8223	9.7585	10.2605	0.4383	0.4339	0.4020	0.3993	0.4215	0.3807
13	9.7697	9.6390	9.1349	9.0724	9.5577	0.4001	0.3953	0.3666	0.3641	0.3849	0.3597
14	9.7884	9.7160	9.1613	9.0934	9.6661	0.3924	0.3897	0.3593	0.3566	0.3794	0.3635
15	9.7613	9.6163	8.9935	8.9281	9.4983	0.4086	0.4029	0.3675	0.3648	0.3885	0.3540
16	10.5654	10.4543	9.6001	9.5218	10.0860	0.4149	0.4114	0.3789	0.3758	0.3991	0.3748
17	10.4089	10.2728	9.8233	9.7592	10.1219	0.4111	0.4068	0.3682	0.3658	0.3812	0.3785
18	10.2631	10.1363	9.7543	9.6998	10.0831	0.4089	0.4042	0.3796	0.3775	0.3949	0.3733
19	10.0924	9.9603	9.5773	9.4947	9.7660	0.4006	0.4050	0.3745	0.3712	0.3852	0.3720
20	9.5650	9.4771	9.2604	9.1831	9.5082	0.3811	0.3779	0.3692	0.3661	0.3810	0.3482
21	9.5923	9.5195	9.0311	8.9663	9.3035	0.3946	0.3919	0.3470	0.3446	0.3603	0.3538
22	10.4193	10.3069	9.2988	9.2257	9.6792	0.4259	0.4215	0.3463	0.3436	0.3611	0.3842

ตาราง 47 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	9.6560	9.5330	9.3759	9.3199	9.6630	0.3934	0.3892	0.3868	0.3845	0.4021	0.3582
24	9.6802	9.5583	9.0259	8.9631	9.2718	0.4041	0.3995	0.3588	0.3563	0.3720	0.3596
25	10.4453	10.2726	9.2077	9.1554	9.4770	0.4140	0.4078	0.3489	0.3469	0.3612	0.3796
26	10.3989	10.2392	9.7709	9.6944	10.1815	0.4295	0.4233	0.3829	0.3799	0.4006	0.3759
27	10.1264	10.0132	9.5659	9.4816	9.8335	0.4186	0.4143	0.3946	0.3911	0.4090	0.3699
28	9.6121	9.5312	9.2303	9.1680	9.4207	0.3917	0.3887	0.3732	0.3707	0.3848	0.3510
29	9.9535	9.8373	9.3661	9.2821	9.8556	0.4100	0.4056	0.3782	0.3748	0.3987	0.3741
30	10.1003	9.9802	9.1552	9.0867	9.6174	0.4142	0.4101	0.3487	0.3461	0.3672	0.3715
รวม	10.0499	9.9361	9.3897	9.3208	9.7893	0.4113	0.4074	0.3737	0.3709	0.3911	0.3681

ตาราง 48 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาด 500 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	9.7313	9.8555	9.2769	9.2129	9.6817	0.4459	0.4550	0.3768	0.3742	0.3941	0.3567
2	10.0341	9.9236	9.5189	9.4523	10.0271	0.4030	0.3986	0.3738	0.3712	0.3944	0.3708
3	10.1871	10.0801	9.5629	9.4861	10.0295	0.4102	0.4061	0.3768	0.3738	0.3958	0.3361
4	9.9036	9.7833	9.333	9.2672	9.7482	0.3929	0.3886	0.3619	0.3594	0.3787	0.3615
5	9.8619	9.7229	9.2223	9.1576	9.6716	0.4025	0.3974	0.3663	0.3638	0.3854	0.3645
6	9.9647	9.8306	9.2782	9.2204	9.6716	0.4104	0.4055	0.3695	0.3672	0.3893	0.3657
7	11.9049	9.9356	9.3779	9.3124	9.5819	0.4298	0.4537	0.3898	0.3871	0.4105	0.3753
8	10.2351	10.118	9.5729	9.5039	10.0643	0.4257	0.4213	0.3885	0.3857	0.4096	0.3751
9	9.9235	9.8195	9.2920	9.2254	9.7570	0.4154	0.4117	0.3792	0.3764	0.3992	0.3553
10	10.3209	10.2166	9.6864	9.6229	10.1644	0.4186	0.4144	0.3849	0.3824	0.4045	0.3744
11	10.3688	10.2508	9.7951	9.7111	10.1833	0.4356	0.4307	0.4012	0.3978	0.4181	0.3779
12	10.0927	9.9685	9.5175	9.4497	9.9275	0.4036	0.3990	0.3718	0.3692	0.3889	0.3350
13	9.9658	9.8174	9.3086	9.2441	9.7480	0.4132	0.4075	0.3767	0.3741	0.3954	0.3665
14	10.2273	10.0872	9.5821	9.5135	10.0253	0.4151	0.4096	0.3803	0.3776	0.3986	0.3742
15	9.9007	9.7804	9.2745	9.2159	9.7173	0.1770	0.1752	0.1582	0.1572	0.1664	0.3629
16	10.1099	9.9575	9.4473	9.3961	9.8939	0.4133	0.4075	0.3773	0.3753	0.3962	0.3709
17	9.7460	9.6420	9.1074	9.0312	9.5781	0.4111	0.4072	0.3731	0.3699	0.3932	0.3279
18	10.0712	9.9197	9.3896	9.3206	9.8615	0.4070	0.4016	0.3704	0.3676	0.3903	0.3697
19	9.6156	9.4846	9.0011	8.9386	9.4243	0.3908	0.3859	0.3591	0.3566	0.3772	0.3581
20	10.3622	10.2212	9.7397	9.6671	10.1775	0.4168	0.4116	0.3835	0.3806	0.4019	0.3769
21	10.2167	10.0795	9.5857	9.5188	10.0218	0.4095	0.4045	0.376	0.3733	0.3942	0.3750
22	10.0846	9.9420	9.4157	9.3481	9.8832	0.4111	0.4057	0.3763	0.3736	0.3961	0.3680

ตาราง 48 (ต่อ)

ตัวอย่างจำลอง	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	9.7316	9.6526	9.1323	9.0769	9.5792	0.4059	0.4023	0.3718	0.3695	0.3902	0.3509
24	10.2521	10.0887	9.5674	9.4904	10.0257	0.4358	0.4295	0.3956	0.3924	0.4159	0.3789
25	9.8463	9.7199	9.2346	9.1732	9.6796	0.401	0.3959	0.3669	0.3645	0.3852	0.3651
26	10.0286	9.9152	9.7018	9.6386	10.2336	0.4013	0.3966	0.4019	0.3993	0.4244	0.3641
27	10.0600	9.9406	9.418	9.3427	9.8736	0.4135	0.4089	0.3795	0.3764	0.3986	0.3749
28	10.1949	10.0704	9.5799	9.5185	10.0215	0.4076	0.4033	0.3742	0.3718	0.3926	0.3761
29	9.9161	9.7689	9.2683	9.2025	9.7025	0.406	0.4006	0.3694	0.3668	0.3879	0.3573
30	9.9593	9.8618	9.3263	9.2670	9.8065	0.4103	0.4067	0.3732	0.3709	0.3933	0.3710
รวม	10.0939	9.9152	9.4171	9.3509	9.8587	0.4047	0.4014	0.3701	0.3675	0.3889	0.3646

ตาราง 49 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านตัวเลข จากตัวอย่างขนาด 800 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	10.1664	10.0195	9.5234	9.4591	9.9577	0.4117	0.4117	0.3815	0.3789	0.4001	0.4117
2	10.2289	10.1073	9.6187	9.6187	10.0663	0.4142	0.4096	0.3813	0.3813	0.4000	0.4142
3	9.9562	9.8265	9.3330	9.2692	9.7745	0.4097	0.4147	0.3798	0.3772	0.3989	0.4097
4	9.8775	9.7603	9.2507	9.1830	9.6984	0.4007	0.3961	0.3676	0.3649	0.3860	0.4007
5	9.9661	9.8660	9.2965	9.2286	9.8100	0.4021	0.3983	0.3670	0.3644	0.3880	0.4021
6	9.8742	9.7484	9.2640	9.1989	9.6827	0.3944	0.3899	0.3629	0.3603	0.3802	0.3944
7	10.0396	9.9221	9.4021	9.3393	9.8563	0.4020	0.3977	0.3685	0.3660	0.3871	0.4020
8	10.1410	9.9712	9.4805	9.4194	9.9247	0.4105	0.4043	0.3745	0.3721	0.3935	0.4105
9	10.0562	9.9287	9.4296	9.3681	9.8811	0.4116	0.4069	0.3762	0.3738	0.3952	0.4116
10	10.2447	10.0807	9.5617	9.4925	10.0423	0.4214	0.4148	0.3848	0.3820	0.4052	0.4214
11	10.0325	9.8683	9.3873	9.3244	9.8158	0.4172	0.4111	0.3808	0.3782	0.3995	0.4172
12	9.9132	9.8070	9.2537	9.1909	9.7511	0.4097	0.4055	0.3734	0.3709	0.3944	0.4097
13	9.8999	9.7677	9.2712	9.1912	9.7171	0.4058	0.4008	0.3713	0.3681	0.3900	0.4058
14	9.9914	9.8710	9.3464	9.2757	9.8120	0.4096	0.4049	0.3749	0.3721	0.3945	0.4096
15	9.8715	9.7493	9.2574	9.1783	9.6895	0.4141	0.4094	0.3792	0.3760	0.3979	0.4141
16	10.0247	9.9067	9.4297	9.3611	9.8671	0.4043	0.3999	0.3727	0.3700	0.3907	0.4043
17	9.7435	9.6417	9.1372	9.0715	9.5912	0.4022	0.3981	0.3688	0.3662	0.3878	0.4022
18	10.0338	9.8892	9.4054	9.3375	9.8351	0.4127	0.4071	0.3775	0.3748	0.3956	0.4127
19	10.0506	9.913	9.4329	9.3702	9.8601	0.4165	0.4109	0.3815	0.3789	0.3996	0.4165
20	9.7835	9.6563	9.1738	9.1738	9.6060	0.3942	0.3897	0.3617	0.3617	0.3796	0.3942
21	9.8737	9.7459	9.2564	9.1903	9.6960	0.4028	0.3981	0.3686	0.3660	0.3870	0.4028
22	9.8546	9.7409	9.2316	9.1709	9.6800	0.4125	0.4080	0.3774	0.3749	0.3966	0.4125

ตาราง 49 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	9.8903	9.7755	9.2244	9.1640	9.7086	0.4127	0.4083	0.3784	0.3759	0.3966	0.4127
24	9.7984	9.6632	9.1867	9.1180	9.6121	0.4022	0.3970	0.3686	0.3658	0.3868	0.4022
25	10.1263	10.0040	9.4967	9.4244	9.9573	0.4152	0.4104	0.3794	0.3765	0.3984	0.4152
26	10.2488	10.0905	9.8904	9.8252	10.0308	0.4122	0.4060	0.3829	0.3804	0.3910	0.4122
27	9.7313	9.6424	9.1070	9.0367	9.5932	0.4033	0.3998	0.3694	0.3666	0.3899	0.4033
28	10.0618	9.9205	9.4247	9.3628	9.8777	0.4080	0.4027	0.3737	0.3712	0.3927	0.4080
29	10.023	9.9093	9.3690	9.2981	9.8543	0.4143	0.4099	0.3786	0.3757	0.3993	0.4143
30	10.0464	9.9264	9.4127	9.3411	9.8687	0.4209	0.4162	0.3846	0.3817	0.4041	0.4209
รวม	9.9850	9.8573	9.3618	9.2994	9.8039	0.409	0.4046	0.3749	0.3724	0.3935	0.4090

ตาราง 50 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) จากแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากกลุ่มตัวอย่างขนาด 300 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	12.9829	12.9323	12.3118	12.2645	12.7878	0.5390	0.5371	0.5033	0.5013	0.5238	0.4618
2	12.7306	12.6397	12.1338	12.0922	12.5895	0.4695	0.4661	0.4442	0.4427	0.4618	0.4528
3	12.8366	12.8030	12.2063	12.1551	12.7160	0.5066	0.5050	0.4767	0.4747	0.4970	0.4779
4	13.7794	13.7024	13.0609	13.0148	13.6088	0.5366	0.5334	0.5005	0.4988	0.5218	0.4925
5	13.3273	13.2583	12.6168	12.5645	13.2117	0.5132	0.5099	0.4796	0.4776	0.5020	0.4774
6	13.6752	13.5661	12.9859	12.9329	13.4594	0.5284	0.5236	0.4977	0.4957	0.5164	0.4843
7	14.0494	13.9523	13.1454	13.0919	13.7071	0.5305	0.5264	0.5139	0.5118	0.5362	0.4912
8	13.6458	13.5983	13.0937	13.0444	13.6380	0.5584	0.5559	0.4977	0.4958	0.5187	0.4937
9	13.5123	13.4708	12.5598	12.5078	13.1305	0.5130	0.5111	0.4857	0.4837	0.5080	0.4789
10	12.4274	12.3892	12.4154	12.3635	12.8804	0.4850	0.4831	0.5209	0.5187	0.5401	0.4439
11	13.6964	13.6264	12.8882	12.8394	13.4203	0.5293	0.5260	0.4844	0.4825	0.5045	0.4925
12	13.3136	13.2319	12.4521	12.4101	12.9319	0.5130	0.5095	0.4846	0.4830	0.5039	0.4678
13	13.3134	13.2379	12.3711	12.3242	12.8405	0.5117	0.5085	0.4700	0.4682	0.4880	0.4746
14	13.3639	13.2840	12.2455	12.2085	12.7278	0.5069	0.5036	0.4513	0.4499	0.4694	0.4796
15	13.0776	12.9679	12.6765	12.6161	13.1656	0.5067	0.5020	0.4976	0.4953	0.5168	0.4624
16	13.1436	13.0721	12.2442	12.1921	12.7273	0.5034	0.5003	0.4637	0.4617	0.4822	0.4733
17	12.7410	12.6822	12.6044	12.5464	13.0715	0.5431	0.5402	0.5050	0.5027	0.5238	0.4548
18	13.2536	13.1664	12.5413	12.5045	13.0114	0.5130	0.5095	0.4756	0.4742	0.4942	0.4698
19	13.1142	13.0077	12.5241	12.4722	13.0177	0.5014	0.4973	0.4776	0.4756	0.4973	0.4671
20	13.5973	13.5177	12.7186	12.6771	13.1816	0.5399	0.5363	0.4843	0.4827	0.5025	0.4837
21	13.2728	13.1919	12.6512	12.6051	13.0697	0.5110	0.5076	0.4705	0.4687	0.4864	0.4640
22	13.3728	13.3213	12.5903	12.5417	13.1090	0.5349	0.5326	0.4828	0.4810	0.5031	0.4720

ตาราง 50 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	13.8182	13.7781	12.6736	12.6141	13.2742	0.5713	0.5691	0.4782	0.476	0.5006	0.4887
24	13.2499	13.1739	12.2621	12.2098	12.7788	0.5425	0.5389	0.5023	0.5002	0.5241	0.4819
25	13.3409	13.2549	12.5272	12.4882	13.0571	0.5293	0.5258	0.4797	0.4782	0.5010	0.4808
26	12.6944	12.6010	12.2286	12.1712	12.6846	0.5057	0.5017	0.4792	0.4770	0.4971	0.4586
27	12.7609	12.7315	12.7705	12.7143	13.3016	0.4986	0.4971	0.5050	0.5028	0.5260	0.4596
28	13.9877	13.9712	13.0124	12.9642	13.6287	0.5316	0.5305	0.4960	0.4941	0.5196	0.4950
29	13.1911	13.1247	12.5870	12.5489	13.0808	0.4933	0.4905	0.4797	0.4782	0.4988	0.4760
30	14.4749	14.4133	13.1766	13.1378	13.7069	0.5361	0.5333	0.4856	0.4842	0.5054	0.5060
รวม	13.3248	13.2556	12.6092	12.5606	13.1172	0.5201	0.5171	0.4858	0.4839	0.5057	0.4754

ตาราง 51 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาด 500 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	12.5623	12.4954	12.0024	11.9538	12.4500	0.5143	0.5113	0.4680	0.4661	0.4853	0.4555
2	12.8580	12.8230	12.9224	12.8746	13.4567	0.5141	0.5124	0.5029	0.5010	0.5240	0.4643
3	13.5346	13.4885	12.7733	12.7300	13.2555	0.5186	0.5151	0.4928	0.4912	0.5119	0.4874
4	12.9915	12.9496	11.9918	11.9408	12.4204	0.5138	0.5118	0.4827	0.4807	0.5003	0.4652
5	12.9912	12.9113	12.8746	12.8253	13.3624	0.5199	0.5164	0.4939	0.4920	0.5128	0.4743
6	13.3226	13.2484	12.2248	12.1755	12.7446	0.5258	0.5223	0.4821	0.4802	0.5027	0.4860
7	13.5045	13.4244	12.9497	12.9093	13.4310	0.5253	0.5216	0.4897	0.4882	0.5083	0.4915
8	13.1398	13.0479	12.3819	12.3173	12.8630	0.5282	0.5243	0.4841	0.4816	0.5032	0.4732
9	13.0804	13.0134	12.3175	12.2758	12.8366	0.5117	0.5088	0.4866	0.4850	0.5076	0.4705
10	13.3806	13.3082	12.6624	12.6128	13.1739	0.5209	0.5176	0.4945	0.4925	0.5144	0.4801
11	13.1495	13.1013	12.816	12.7724	13.3646	0.5245	0.5219	0.4941	0.4924	0.5153	0.4762
12	13.4011	13.3495	12.5225	12.4753	12.9840	0.5310	0.5285	0.4956	0.4937	0.5145	0.4861
13	12.9660	12.9002	12.4573	12.4095	12.9554	0.5246	0.5215	0.4814	0.4795	0.5013	0.4705
14	12.8140	12.7654	12.7604	12.7173	13.2571	0.4881	0.4862	0.4914	0.4897	0.5110	0.4510
15	13.3754	13.308	12.4821	12.4416	12.9882	0.5244	0.5216	0.4928	0.4912	0.5129	0.4858
16	12.9728	12.9043	12.7326	12.6853	13.2738	0.5138	0.5107	0.4980	0.4961	0.5192	0.4654
17	13.3044	13.2301	12.336	12.2858	12.8134	0.5113	0.5080	0.4926	0.4906	0.5119	0.4788
18	12.8029	12.7290	12.1944	12.1397	12.6765	0.5032	0.4998	0.4600	0.4579	0.4786	0.4624
19	12.6360	12.5616	12.71	12.6715	13.2390	0.5175	0.5142	0.4919	0.4904	0.5130	0.4561
20	13.9971	13.9345	12.3526	12.3145	12.8403	0.5679	0.5647	0.4832	0.4818	0.5027	0.5064
21	12.8800	12.8295	12.6459	12.5973	13.1844	0.5047	0.5024	0.4815	0.4797	0.5021	0.4648
22	12.9543	12.8760	12.1943	12.1453	12.6518	0.5290	0.5254	0.4739	0.4720	0.4918	0.4659

ตาราง 51 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	13.5484	13.4876	12.1337	12.0927	12.6208	0.5207	0.5180	0.4838	0.4822	0.5035	0.4899
24	12.9845	12.8993	11.9698	11.9203	12.4813	0.4919	0.4884	0.4828	0.4808	0.5040	0.4651
25	12.5782	12.5163	13.3132	13.2629	13.8723	0.4981	0.4950	0.5316	0.5296	0.5541	0.4532
26	13.5702	13.5207	11.9559	11.9124	12.4224	0.5335	0.5312	0.4699	0.4682	0.4885	0.4882
27	13.3819	13.3045	12.2697	12.2186	12.7592	0.5213	0.5179	0.4760	0.4740	0.4951	0.4797
28	13.4795	13.4049	12.3632	12.3148	12.8204	0.5231	0.5197	0.4954	0.4935	0.5141	0.4844
29	12.7677	12.6980	12.8905	12.8435	13.4260	0.5173	0.5140	0.4916	0.4898	0.5124	0.4606
30	12.5869	12.5142	12.3814	12.3329	12.8306	0.5009	0.4976	0.4647	0.4629	0.4820	0.4462
รวม	13.1172	13.0515	12.4861	12.4390	12.9819	0.5180	0.5149	0.4870	0.4852	0.5066	0.4728

ตาราง 52 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านภาษา จากตัวอย่างขนาด 800 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	12.9233	12.8252	12.3358	12.2928	12.7697	0.5180	0.5140	0.4880	0.4863	0.5057	0.4632
2	12.9005	12.8365	12.2938	12.2501	12.7724	0.5184	0.5153	0.4883	0.4865	0.5075	0.4572
3	12.6283	12.7129	12.1003	12.0584	12.5700	0.5183	0.4926	0.4669	0.4653	0.4854	0.4578
4	12.8484	12.7909	12.2312	12.1782	12.7304	0.4963	0.5041	0.4772	0.4751	0.4967	0.4555
5	13.1287	13.0783	12.5214	12.4813	13.0175	0.5070	0.5231	0.4939	0.4924	0.5136	0.4693
6	13.4172	13.3253	12.7976	12.7531	13.2659	0.5257	0.5243	0.4986	0.4969	0.5170	0.4825
7	12.8849	12.8159	12.2844	12.2427	12.7597	0.5285	0.5172	0.4887	0.4870	0.5082	0.4660
8	13.2713	13.1882	12.6568	12.6138	13.1336	0.5204	0.5201	0.4921	0.4904	0.5112	0.4766
9	13.1833	13.1206	12.5442	12.5009	13.0417	0.5238	0.5219	0.4941	0.4924	0.5143	0.4694
10	13.2057	13.1338	12.5904	12.5450	13.0832	0.5248	0.5163	0.4887	0.4870	0.5082	0.4725
11	13.0765	12.9999	12.4869	12.4413	12.9302	0.5195	0.5250	0.4970	0.4952	0.5151	0.4715
12	13.1341	13.0535	12.5302	12.4899	12.9974	0.5285	0.5038	0.4790	0.4774	0.4971	0.4719
13	13.1146	13.0440	12.5704	12.5199	13.0149	0.5073	0.5103	0.4858	0.4838	0.5034	0.4661
14	13.1034	13.0185	12.504	12.4572	12.9718	0.5134	0.4954	0.4720	0.4702	0.4901	0.4692
15	12.9700	12.8987	12.3338	12.2897	12.8346	0.4989	0.5033	0.4766	0.4749	0.4962	0.4663
16	13.1631	13.1019	12.5832	12.5352	13.0554	0.5065	0.5170	0.4912	0.4893	0.5097	0.4704
17	12.8044	12.7329	12.3265	12.2747	12.8403	0.5201	0.5063	0.4819	0.4799	0.5021	0.4613
18	13.2975	13.2299	12.2669	12.2168	12.6785	0.5096	0.5275	0.4819	0.4800	0.4984	0.4726
19	13.0516	12.9853	12.6592	12.6163	13.1818	0.5307	0.5275	0.4983	0.4966	0.5194	0.4746
20	12.7583	12.6939	12.4486	12.4024	12.9375	0.5308	0.4979	0.4990	0.4972	0.5187	0.4642
21	13.1292	13.0548	12.1591	12.1112	12.6320	0.5008	0.5143	0.4722	0.4704	0.4911	0.4701
22	13.0337	12.9721	12.5068	12.4550	13.0022	0.5177	0.5202	0.4876	0.4856	0.5073	0.4751

ตาราง 53 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	13.3237	13.2572	12.4450	12.3992	12.9202	0.5352	0.5321	0.4928	0.4909	0.5119	0.4903
24	13.1978	13.1063	12.7022	12.6485	13.1974	0.6303	0.5263	0.5049	0.5027	0.5250	0.4829
25	13.1681	13.0923	12.5976	12.5525	13.0494	0.5241	0.5207	0.4996	0.4978	0.5183	0.4702
26	13.0587	12.9899	12.5649	12.5108	13.0311	0.5275	0.5243	0.4941	0.4920	0.5129	0.4687
27	12.9676	12.8914	12.4372	12.3873	12.9274	0.5134	0.5099	0.4958	0.4938	0.5158	0.4735
28	12.8911	12.8264	12.2968	12.2473	12.8431	0.5279	0.5247	0.4810	0.4791	0.5028	0.4670
29	13.0123	12.9390	12.2846	12.2357	12.7616	0.5226	0.5192	0.4949	0.4930	0.5144	0.4740
30	12.9360	12.8748	12.4063	12.3574	12.8893	0.5153	0.5122	0.4913	0.4894	0.5108	0.4669
รวม	13.0528	12.9863	12.4489	12.4022	12.9280	0.5222	0.5156	0.4884	0.4866	0.5076	0.4699

ตาราง 54 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มตัวอย่างขนาด 300 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	10.4797	10.2447	10.1204	10.0925	10.1875	0.4132	0.4050	0.3908	0.3898	0.3952	0.3769
2	10.1824	9.9373	9.8835	9.8524	9.9512	0.3930	0.3859	0.3733	0.3721	0.3774	0.3713
3	9.5048	9.3708	9.1237	9.0998	9.3316	0.3717	0.3670	0.3412	0.3403	0.3501	0.3699
4	11.4509	11.2428	11.0916	11.0497	11.2588	0.4378	0.4310	0.4154	0.4139	0.4221	0.4101
5	10.5021	10.3422	10.1499	10.1195	10.2964	0.4163	0.4086	0.3943	0.3931	0.4007	0.3875
6	10.5219	10.4075	10.1868	10.1592	10.3666	0.4201	0.4159	0.3984	0.3973	0.4064	0.3785
7	10.2856	10.1155	9.9788	9.9392	10.0878	0.4118	0.4054	0.3897	0.3881	0.3950	0.3719
8	10.5282	10.3170	10.2039	10.1577	10.2644	0.4284	0.4204	0.4064	0.4045	0.4105	0.3911
9	10.8653	10.7188	10.4365	10.4140	10.6255	0.4183	0.4131	0.3929	0.3921	0.4011	0.3935
10	10.4699	10.3032	9.0641	9.0407	9.1781	0.4196	0.4131	0.3411	0.3403	0.3464	0.3842
11	10.4354	10.218	10.0508	10.0249	10.1427	0.4023	0.3945	0.3798	0.3788	0.3842	0.3818
12	10.7064	10.4743	10.3561	10.3267	10.4478	0.4164	0.4084	0.3923	0.3912	0.3979	0.3834
13	10.6826	10.5348	10.5728	10.5344	10.4204	0.409	0.4036	0.3860	0.3846	0.3841	0.3816
14	10.6553	10.4990	10.2535	10.2246	10.4239	0.4154	0.4101	0.3902	0.3891	0.3984	0.3989
15	10.6938	10.5388	10.3365	10.3077	10.5248	0.4119	0.4060	0.3868	0.3857	0.3944	0.3829
16	10.6024	10.4273	10.2693	10.2464	10.3769	0.4179	0.4117	0.3948	0.3939	0.4005	0.3988
17	10.7645	10.5644	10.4066	10.3722	10.5224	0.4132	0.4061	0.3915	0.3902	0.3973	0.3961
18	10.5209	10.3542	10.1817	10.1333	10.3021	0.4188	0.4127	0.3965	0.3946	0.4025	0.3834
19	10.5130	10.3498	10.1915	10.1638	10.2639	0.4151	0.4088	0.3930	0.3919	0.3968	0.3839
20	10.8005	10.5542	10.4426	10.4131	10.4972	0.4313	0.4216	0.4080	0.4068	0.4114	0.3946
21	10.2121	10.0227	9.8664	9.8453	9.9817	0.4015	0.3942	0.3797	0.3788	0.3848	0.3718
22	10.0949	9.9431	9.7648	9.7416	9.8935	0.3979	0.3923	0.3769	0.3760	0.3830	0.3675

ตาราง 54 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	10.6104	10.4370	10.2092	10.1823	10.3474	0.4109	0.4049	0.3886	0.3876	0.3955	0.3898
24	10.5563	10.3555	10.1433	10.0977	10.2729	0.4432	0.4353	0.4131	0.4112	0.4196	0.3900
25	10.6391	10.4833	10.3021	10.2631	10.4493	0.4168	0.4114	0.3921	0.3906	0.3989	0.3894
26	10.3985	10.2240	10.2201	10.1935	10.3587	0.4172	0.4110	0.3969	0.3959	0.4025	0.3839
27	10.3499	10.1766	9.9925	9.9707	10.1417	0.4091	0.4027	0.3868	0.3860	0.3940	0.3787
28	10.6872	10.5549	10.3562	10.3351	10.5337	0.4047	0.4001	0.3851	0.3844	0.3926	0.3900
29	10.4774	10.3112	10.0770	10.0459	10.2163	0.4143	0.4081	0.3912	0.3900	0.3978	0.3841
30	11.2239	11.0055	10.8842	10.8557	10.9938	0.4261	0.4181	0.4044	0.4033	0.4096	0.4093
รวม	10.5472	10.3676	10.1706	10.1401	10.2886	0.4141	0.4076	0.3892	0.3881	0.3950	0.3858

ตาราง 54 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มตัวอย่างขนาด 500 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	10.4388	10.2887	10.1189	10.0833	10.2623	0.4177	0.4115	0.3944	0.3931	0.4004	0.3803
2	10.5687	10.3889	10.2372	10.2078	10.3408	0.4221	0.4154	0.3987	0.3976	0.4042	0.3786
3	10.2179	10.0423	10.1116	10.0817	10.1946	0.4108	0.4046	0.3807	0.3796	0.3851	0.3740
4	10.6126	10.4327	10.2931	10.2688	10.4051	0.4142	0.408	0.3925	0.3916	0.3984	0.3815
5	10.1370	9.9601	9.8092	9.7773	9.9080	0.3936	0.3876	0.3688	0.3676	0.3735	0.3662
6	10.1403	9.9805	9.8245	9.7969	9.9266	0.4117	0.4058	0.3859	0.3849	0.3909	0.3702
7	10.4374	10.2717	10.0443	10.0443	10.2385	0.4082	0.4026	0.3814	0.3814	0.3901	0.3802
8	10.3903	10.2403	10.1275	10.1275	10.1222	0.4273	0.4217	0.4014	0.4014	0.4044	0.3844
9	10.7151	10.5530	10.3641	10.3387	10.5324	0.4255	0.4194	0.4022	0.4012	0.4099	0.3935
10	10.4249	10.2102	10.0432	10.014	10.1684	0.4139	0.4057	0.3883	0.3871	0.3943	0.3827
11	10.4072	10.2200	10.0579	10.0268	10.1558	0.4182	0.4112	0.3948	0.3935	0.4000	0.3723
12	10.7496	10.5386	10.4177	10.3864	10.4952	0.4266	0.4187	0.4044	0.4032	0.4091	0.3862
13	10.2855	10.0814	9.9797	9.9464	10.0564	0.4108	0.4036	0.3885	0.3872	0.3930	0.3759
14	10.3954	10.1895	10.0560	10.0311	10.1632	0.4082	0.4006	0.3868	0.3858	0.3924	0.3840
15	11.0382	10.8482	10.6957	10.6639	10.8079	0.4270	0.4204	0.4016	0.4004	0.4072	0.3994
16	10.7149	10.5388	10.4352	10.4124	10.5261	0.4135	0.4071	0.3939	0.3930	0.3985	0.3808
17	10.2054	10.0477	10.0260	9.9987	9.8701	0.4048	0.3989	0.3792	0.3782	0.3771	0.3730
18	10.6629	10.4639	10.3022	10.2795	10.4164	0.4228	0.4157	0.3970	0.3962	0.4030	0.3883
19	10.5329	10.3425	10.1721	10.1481	10.2924	0.4088	0.4021	0.3869	0.3860	0.3929	0.3816
20	10.8314	10.6515	10.4928	10.4648	10.6054	0.4251	0.4180	0.4024	0.4013	0.4076	0.3979
21	10.3380	10.1886	9.9939	9.9729	10.1434	0.4087	0.4036	0.3853	0.3844	0.3925	0.3796
22	10.2333	10.0520	9.8817	9.8496	9.9929	0.4018	0.3950	0.3795	0.3783	0.3849	0.3748

ตาราง 54 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	10.4317	10.2800	10.0899	10.0640	10.2481	0.4076	0.4020	0.3848	0.3838	0.3918	0.3768
24	10.7270	10.5734	10.4063	10.3828	10.5253	0.4277	0.4221	0.4050	0.4041	0.4109	0.3941
25	10.4267	10.2519	10.1016	10.0688	10.2109	0.4065	0.4000	0.3863	0.3850	0.3916	0.3801
26	10.2518	10.0975	9.9253	9.8962	10.0481	0.4282	0.4222	0.4034	0.4022	0.4098	0.3803
27	10.6372	10.4506	10.3117	10.2799	10.4208	0.4204	0.4135	0.3989	0.3977	0.4045	0.3857
28	10.4838	10.3059	10.1031	10.0851	10.2664	0.4160	0.4093	0.3911	0.3904	0.3986	0.3824
29	10.0341	9.8818	9.6943	9.6704	9.8267	0.4080	0.4021	0.3823	0.3813	0.3883	0.3681
30	10.4067	10.2498	10.0779	10.0513	10.2256	0.4039	0.3980	0.3826	0.3816	0.3890	0.3789
รวม	10.4625	10.2874	10.1398	10.1140	10.2619	0.4146	0.4082	0.3910	0.3900	0.3969	0.3811

ตาราง 55 ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการวัดอย่างมีเงื่อนไขของคะแนนแปลงรูปแบบตำแหน่งเปอร์เซ็นไทล์ (PR) และแบบคะแนนปกติ (Z) ของแบบทดสอบวัดความถนัดทางการเรียนด้านเหตุผล จากกลุ่มตัวอย่างขนาด 800 คน

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
1	10.1869	10.0283	9.8796	9.8540	9.9980	0.4006	0.3950	0.3802	0.3792	0.3859	0.3760
2	10.7914	10.6155	10.4163	10.3939	10.5203	0.4224	0.4157	0.4001	0.3992	0.4050	0.3924
3	10.4433	10.2628	10.2524	10.2254	10.3705	0.4239	0.4170	0.4013	0.4002	0.4071	0.3805
4	10.1546	9.9931	9.9664	9.9393	10.0885	0.4070	0.4010	0.4009	0.3998	0.4073	0.3720
5	10.309	10.1342	9.9223	9.8977	10.0249	0.4162	0.4096	0.3838	0.3828	0.3890	0.3792
6	10.5413	10.3691	10.0878	10.0676	10.2016	0.4324	0.4256	0.3926	0.3918	0.3982	0.3880
7	10.3983	10.188	10.1378	10.1130	10.2160	0.4115	0.4037	0.4088	0.4078	0.4136	0.3820
8	10.3628	10.1823	10.0537	10.0314	10.1658	0.4109	0.4042	0.3904	0.3895	0.3960	0.3829
9	10.4523	10.2464	10.0821	10.0590	10.1694	0.4159	0.4082	0.3899	0.3890	0.3946	0.3858
10	12.0173	11.8481	10.1435	10.1191	10.3129	0.4773	0.4715	0.3936	0.3927	0.4010	0.3873
11	10.1775	10.0201	11.0125	10.9758	11.1448	0.4046	0.3990	0.4464	0.4449	0.4535	0.3758
12	10.7173	10.5446	10.0900	10.0657	10.2126	0.4255	0.4190	0.3775	0.3766	0.3833	0.3915
13	10.4791	10.2882	10.2338	10.2055	10.3401	0.4257	0.4183	0.4044	0.4033	0.4098	0.3897
14	10.523	10.3581	10.1716	10.1503	10.2998	0.4177	0.4114	0.4004	0.3996	0.4068	0.3901
15	10.5547	10.3553	10.2115	10.1850	10.3054	0.4107	0.4034	0.3953	0.3943	0.4003	0.3834
16	10.9993	10.8131	10.4270	10.4011	10.1653	0.4236	0.4168	0.3896	0.3886	0.4032	0.3979
17	10.3204	10.1597	10.1764	10.1463	10.3062	0.4155	0.4094	0.3984	0.3972	0.4048	0.3724
18	10.6324	10.4508	9.8381	9.8137	9.9988	0.4139	0.4073	0.3825	0.3815	0.3895	0.3841
19	10.4945	10.3308	10.2201	10.1889	10.3523	0.4164	0.4103	0.3940	0.3928	0.4005	0.3864
20	10.2876	10.1112	10.0567	10.0314	10.1553	0.4204	0.4135	0.4003	0.3993	0.4055	0.3815
21	10.6042	10.4402	10.0672	10.038	10.2137	0.4195	0.4133	0.3908	0.3897	0.3975	0.3877
22	10.7303	10.5642	10.2700	10.2373	10.4389	0.4230	0.4168	0.3927	0.3915	0.4002	0.3908

ตาราง 55 (ต่อ)

ตัวอย่างที่	PR					Z					IRTSS
	BINSS	CBINSS	F-Q			BINSS	CBINSS	F-Q			
			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$			$F\text{-}Q_{R20}$	$F\text{-}Q_{RB}$	$F\text{-}Q_{CB}$	
23	10.5491	10.3662	10.3162	10.2923	10.4142	0.4148	0.4080	0.4016	0.4007	0.4067	0.3878
24	10.1036	9.9369	9.9636	9.9372	10.0652	0.4131	0.4069	0.3985	0.3975	0.4042	0.3717
25	10.6008	10.4440	9.9530	9.9221	10.1165	0.4210	0.4153	0.3831	0.3820	0.3909	0.3881
26	10.3993	10.2055	10.1788	10.1496	10.2525	0.4222	0.4147	0.3991	0.3979	0.4034	0.3870
27	10.6153	10.4102	10.1604	10.1359	10.2539	0.4289	0.4208	0.3973	0.3963	0.4022	0.3904
28	10.4089	10.2146	10.1658	10.1366	10.2728	0.4125	0.4054	0.4053	0.4042	0.4111	0.3847
29	10.6299	10.4602	10.1326	10.1125	10.2891	0.4240	0.4176	0.3893	0.3885	0.3965	0.3917
30	10.6813	10.4905	10.2894	10.2612	10.4309	0.4247	0.4171	0.3980	0.3969	0.4041	0.3954
รวม	10.5389	10.3611	10.1626	10.1362	10.2699	0.4199	0.4132	0.3962	0.3952	0.4024	0.3851

ประวัติย่อผู้วิจัย

ชื่อ – นามสกุล	นางสาวสุทธิวรรณ พีรศักดิ์โสภณ
สถานที่เกิด	อำเภอตาคลี จังหวัดนครสวรรค์
สถานที่อยู่ปัจจุบัน	138 ม. 4 บ.วังสารคาม อ.เมือง จ. มหาสารคาม 44000
ตำแหน่งหน้าที่การงานในปัจจุบัน	ผู้ช่วยศาสตราจารย์ ระดับ 8
สถานที่ทำงานปัจจุบัน	คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม อ. เมือง จ.มหาสารคาม 44000

ประวัติการศึกษา

พ.ศ. 2523	กศ.บ คณิตศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ พิษณุโลก
พ.ศ. 2529	กศ.ม การวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร กรุงเทพมหานคร
พ.ศ. 2546	กศ.ด การทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร กรุงเทพมหานคร