

519.544

ร 637 ก

ร 3

การศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์
ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติก แบบหนึ่ง สอง สาม พารามิเตอร์
และขนาดกลุ่มตัวอย่างต่างกัน

19 S.A. 2539

ปริญาณิพนธ์
ของ
รุ่งฤดี ทรัพย์นันทิ

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

พฤษภาคม 2539

ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

ท 126210

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญาบัตรฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาการศึกษามหาบัณฑิต
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

คณะกรรมการควบคุม

.....ประธาน

(รองศาสตราจารย์อังคณา สายยศ)

.....กรรมการ

(อาจารย์ ดร. ชูศักดิ์ ชัมภลลิขิต)

คณะกรรมการสอบ

.....ประธาน

(รองศาสตราจารย์อังคณา สายยศ)

.....กรรมการ

(อาจารย์ ดร. ชูศักดิ์ ชัมภลลิขิต)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(รองศาสตราจารย์ชูศรี วงศ์รัตนะ)

บัณฑิตวิทยาลัยอนุมัติให้รับปริญญาบัตรฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิตวิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(ดร.ศิริภา พูลสุวรรณ)

วันที่ 8 เดือน พฤษภาคม พ.ศ. 2539

ประกาศคุณูปการ

ปริญญาโทพนธบัณนี้สำเร็จลุล่วงด้วยดี เพราะได้รับความกรุณาจาก รองศาสตราจารย์ อังคณา สายยศ อาจารย์ ดร.ชูศักดิ์ ชัมภลิขิต รองศาสตราจารย์ชูศรี วงศ์รัตนะ ที่ให้ค่าปรึกษา แนะนำ ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบคุณ ผู้อำนวยการ ผู้ช่วยผู้อำนวยการฝ่ายวิชาการ หัวหน้าหมวดวิชาคณิตศาสตร์ และคณะอาจารย์ของสถานศึกษาที่เป็นกลุ่มตัวอย่าง และขอขอบใจนักเรียนทุกคนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบคุณผู้เชี่ยวชาญ ดังมีรายนามในภาคผนวก ที่ให้ความช่วยเหลือและให้คำแนะนำในการสร้างเครื่องมือที่ใช้ในการวิจัย

ขอขอบคุณพี่ ๆ เพื่อน ๆ น้อง ๆ ทุกคนที่ให้ความช่วยเหลือและเป็นกำลังใจ ในการทำปริญญาโทพนธบัณนี้

คุณค่าและประโยชน์ของปริญญาโทพนธบัณนี้ ขอมอบเป็นเครื่องบูชาพระคุณของบิดา มารดา ครู-อาจารย์ และผู้มีพระคุณของผู้วิจัยทุกท่าน

รุ่งฤดี ทรัพย์นิธิ

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
จุดมุ่งหมายของการศึกษาค้นคว้า	7
ความสำคัญของการศึกษาค้นคว้า	7
ขอบเขตของการศึกษาค้นคว้า	8
นิยามศัพท์เฉพาะ	9
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	11
ทฤษฎีการตอบสนองข้อสอบ	11
ความคลาดเคลื่อนมาตรฐาน	21
การวิเคราะห์ความสอดคล้องกับโมเดล	24
แบบทดสอบและกลุ่มตัวอย่าง	25
งานวิจัยที่เกี่ยวข้อง	26
3 วิธีดำเนินการศึกษาค้นคว้า	34
ประชากร	34
กลุ่มตัวอย่าง	34
เครื่องมือที่ใช้ในการศึกษาค้นคว้า	35
วิธีดำเนินการสร้างเครื่องมือที่ใช้ในการศึกษาค้นคว้า	35
การเก็บรวบรวมข้อมูล	42
การวิเคราะห์ข้อมูล	42
4 ผลการวิเคราะห์ข้อมูล	44
การนำเสนอผลการวิเคราะห์ข้อมูล	44
ผลการวิเคราะห์ข้อมูล	44

5	สรุปผล อภิปราย และข้อเสนอแนะ	83
	จุดมุ่งหมายของการศึกษาค้นคว้า	83
	กลุ่มตัวอย่าง	83
	เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล	83
	วิธีดำเนินการเก็บรวบรวมข้อมูล	83
	การวิเคราะห์ข้อมูล	84
	สรุปผลการวิเคราะห์ข้อมูล	84
	อภิปรายผล	87
	ข้อเสนอแนะ	92
	บรรณานุกรม	93
	ภาคผนวก	97
	ประวัติผู้วิจัย	121

บัญชีตาราง

ตาราง		หน้า
1	โอกาสที่จะเกิดของแบบการตอบสนอง	14
2	โรงเรียนที่ได้จากการสุ่ม จำแนกตามเขต โรงเรียน จำนวนห้องเรียน และจำนวนนักเรียน	35
3	ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคิดข้อสอบ ที่ไม่สอดคล้องกับโมเดลออกแล้ว	45
4	ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคิดข้อสอบ ที่ไม่สอดคล้องกับโมเดลออกแล้ว	47
5	ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคิดข้อสอบ ที่ไม่สอดคล้องกับโมเดลออกแล้ว	51
6	ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วย โมเดล 1 พารามิเตอร์ ก่อนและหลังจากคิดผู้สอบที่ไม่สอดคล้องกับ โมเดลออกแล้ว	56
7	ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วย โมเดล 2 พารามิเตอร์ ก่อนและหลังจากคิดผู้สอบที่ไม่สอดคล้องกับ โมเดลออกแล้ว	57

8	ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐาน ในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วย โอมเดล 3 พารามิเตอร์ ก่อนและหลังจากคัดผู้สอบที่ไม่สอดคล้องกับ โอมเดลออกแล้ว	58
9	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	59
10	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 พารามิเตอร์ เป็นรายคู่	60
11	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	61
12	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	61
13	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า อำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 พารามิเตอร์ เป็นรายคู่	62
14	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	63
15	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	63

16	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ เป็นรายคู่	64
17	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	64
18	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	65
19	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ เป็นรายคู่	66
20	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	67
21	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ เป็นรายคู่	67
22	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	68
23	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ เป็นรายคู่	69
24	การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชุก ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250 คน	70

35	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่	78
36	การเปรียบเทียบความแตกต่างของค่าเฉลี่ยความคลาดเคลื่อนในการ ประมาณค่าความสามารถ ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 500 คน	78
37	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่	79
38	การเปรียบเทียบความแตกต่างของค่าเฉลี่ยความคลาดเคลื่อนในการ ประมาณค่าความสามารถ ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 750 คน	80
39	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่	80
40	การเปรียบเทียบความแตกต่างของค่าเฉลี่ยความคลาดเคลื่อนในการ ประมาณค่าความสามารถ ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 1,000 คน	81
41	การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่	82
42	ค่าไอเกนขององค์ประกอบที่มีค่ามากกว่า 1.000	99
43	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	101
44	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	102

45	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	104
46	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	108
47	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	109
48	ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการ ประมาณค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน	111
49	ค่าความยาก ค่าอำนาจจำแนก ของข้อสอบวิชาคณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม	116
50	รายละเอียดเนื้อหาและจุดประสงค์เชิงพฤติกรรม วิชาคณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ.2533)	118

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 แสดงโคงลักษณะข้อสอบของข้อสอบ 2 ข้อ	16
2 แสดงการแจกแจงความสามารถของผู้สอบ 2 กลุ่ม	16
3 แสดงโคงลักษณะข้อสอบโมเดลหนึ่งพารามิเตอร์	18
4 แสดงโคงลักษณะข้อสอบโมเดลสองพารามิเตอร์	19
5 แสดงโคงลักษณะข้อสอบโมเดลสามพารามิเตอร์	20
6 แสดงความสัมพันธ์ระหว่างความคลาดเคลื่อนมาตรฐานในการประมาณค่า และอินฟอร์เมชันของแบบทดสอบ	23
7 แผนผังแสดงขั้นตอนการดำเนินการสร้างแบบทดสอบ	36
8 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ของข้อสอบ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	46
9 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ของข้อสอบ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	49
10 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ของข้อสอบ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	50
11 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ของข้อสอบ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	53
12 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ของข้อสอบ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	54

13	แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา ของข้อสอบ จากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้อง กับโอมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโอมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม	55
14	องค์ประกอบที่มีค่าไอเกนมากกว่า 1.0000	99

ภูมิหลัง

การวัดผลทางการศึกษาและจิตวิทยานั้น มักเป็นการสอบวัดคุณลักษณะภายใน (Latent Traits) ที่ไม่สามารถสังเกตได้โดยตรงเหมือนการวัดทางกายภาพ เช่น ความสามารถในการใช้เหตุผล ความรู้ความสามารถทางด้านวิชาการ เป็นต้น การใช้แบบทดสอบเป็นสิ่งเข้าเพื่อกระตุ้นให้บุคคลแสดงพฤติกรรมออกมา แล้วอนุมานจากพฤติกรรมนั้น นำไปสู่การหาความสามารถที่แท้จริง ซึ่งเป็นคุณลักษณะภายในของบุคคล (สงวน ลักษณะ. 2525 : 47) นักจิตวิทยาและนักวัดผลได้พยายามสร้างทฤษฎีต่าง ๆ ขึ้น เพื่อให้ผลการวัดเป็นตัวแทนของคุณลักษณะหรือความสามารถของบุคคลอย่างแท้จริง ทฤษฎีการวัดที่สำคัญและเป็นที่น่าสนใจในปัจจุบันได้แก่ ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory : CTT) และทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) หรือทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT)

นับเป็นเวลาอันยาวนานที่นักวัดผลได้ใช้ทฤษฎีการวัดแบบดั้งเดิม ในการสร้างแบบทดสอบและแปลความหมายของคะแนน ด้วยเหตุที่ทฤษฎีง่ายต่อการสร้างความเข้าใจและสามารถประยุกต์ใช้กับการสอบวัดผลสัมฤทธิ์ และการวัดทางด้านความถนัด แต่ทฤษฎีนี้มีข้อบกพร่องหลายประการ (Hambleton. 1989 : 147 - 148) คือ

1. ค่าสถิติของข้อทดสอบ ได้แก่ ค่าความยากและค่าอำนาจจำแนก มีค่าไม่คงที่ขึ้นอยู่กับกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ ถ้ากลุ่มตัวอย่างมีลักษณะต่างกัน ค่าความยากและค่าอำนาจจำแนก จะมีค่าไม่เท่ากัน ความไม่เป็นอิสระจากกลุ่มตัวอย่าง ทำให้ประโยชน์ของการใช้ค่าสถิติในการพัฒนาแบบทดสอบและประยุกต์ใช้ในการอื่นลดลง
2. คะแนนจากการสังเกตและค่าของคะแนนจริง ขึ้นอยู่กับแบบทดสอบ ซึ่งค่าทั้งสองจะสูงหรือต่ำ เกิดจากการเปลี่ยนแปลงของค่าความยาก ถ้าใช้แบบทดสอบหลายฉบับที่วัดความสามารถเดียวกัน แต่ความยากแตกต่างกัน จะนำคะแนนสอบคนละฉบับมาเปรียบเทียบกันไม่ได้
3. การจัดข้อสอบให้สอดคล้องกับระดับความสามารถของผู้สอบ (Tailored Testing) แบบทดสอบที่ประกอบด้วยข้อสอบที่มีความยากแตกต่างกัน จะไม่มีประโยชน์ในการเปรียบเทียบระหว่างผู้สอบ ทฤษฎีการวัดแบบดั้งเดิมไม่ได้ให้สารสนเทศ (Information) ในผลของการสอบที่ได้จากแบบทดสอบที่ตรงกับระดับความสามารถของผู้สอบ

4. ข้อตกลงเกี่ยวกับความคลาดเคลื่อนในการวัดของผู้สอบทุกคนมีค่าเท่ากัน ซึ่งไม่สมเหตุสมผลในความเป็นจริง ความคลาดเคลื่อนในการวัดของแบบทดสอบที่ยาก ผู้สอบที่มีความสามารถต่ำ จะมีค่ามากกว่าผู้สอบที่มีความสามารถปานกลางและความสามารถสูง แสดงว่าความแปรปรวนของความคลาดเคลื่อนในการวัดของผู้สอบแต่ละคน น่าจะไม่เท่ากัน

5. นิยามของแบบทดสอบคู่ขนาน แบบสอบที่คู่ขนานกันอย่างแท้จริงจะเลือกค่าความเชื่อมั่น แต่ในทางปฏิบัติไม่สามารถสร้างแบบทดสอบให้คู่ขนานกันได้อย่างแท้จริง จะมีผลให้ความไม่คงที่ของค่าประมาณความเชื่อมั่นของแบบทดสอบ ความไม่คงที่ของค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบ และความไม่คงที่ของความยาวของแบบทดสอบที่ต้องการที่จะให้ได้ค่าความเชื่อมั่นที่พอใจ

จากข้อบกพร่องที่กล่าวถึงข้างต้น สามารถช่วยเบาบางไปได้บ้างจากกรอบของทฤษฎีการตอบสนองข้อสอบ (Hambleton, 1989 : 151) และระยะเวลากว่า 20 ปี ที่ผู้สร้างแบบทดสอบจำนวนมากได้นำทฤษฎีการตอบสนองข้อสอบ มาใช้ในการพัฒนาแบบทดสอบมากกว่ารูปแบบการวัดแบบดั้งเดิม จะเห็นได้จากการใช้โมเดลโลจิสติกแบบ 1, 2 และ 3 พารามิเตอร์มากขึ้นในแบบทดสอบที่ให้คะแนนแบบ 0,1 ในการพัฒนาแบบทดสอบและการคัดเลือกข้อสอบ การศึกษาความลำเอียงของข้อสอบ การประยุกต์ใช้คอมพิวเตอร์เพื่อการเรียนการสอน ตลอดจนการเปรียบเทียบคะแนนและรายงานผลคะแนนสอบ ในปัจจุบันนี้ทั้งผู้ผลิตแบบทดสอบ หน่วยงานทางการศึกษา รวมทั้งโรงเรียนขนาดใหญ่ ได้ใช้ทฤษฎีนี้ในการสร้างแบบทดสอบ (Hambleton, Jones and Rogers, 1993 : 143) ทฤษฎีการตอบสนองข้อสอบมีข้อตกลงเบื้องต้นที่สำคัญคือ (Hambleton, 1989 : 149)

1. ผลการสอบของผู้สอบได้มาจากแบบทดสอบที่วัดความสามารถหรือคุณลักษณะเดียว
2. ความสัมพันธ์ระหว่างผลการสอบแต่ละข้อ กับความสามารถที่วัดโดยแบบทดสอบ

สามารถอธิบายได้ด้วยฟังก์ชันลักษณะข้อคำถาม (Item characteristic Function)

ทฤษฎีนี้จะให้ค่าความน่าจะเป็นของความสามารถของผู้สอบ ที่ระดับการตอบต่าง ๆ ผู้สอบที่มีความสามารถสูงกว่า มีความน่าจะเป็นที่จะตอบคำถามได้ถูกต้องมากกว่า ผู้สอบที่มีความสามารถต่ำกว่า ฟังก์ชันคุณลักษณะข้อสอบ หรือโค้งลักษณะข้อสอบ (ICC) ในโมเดลของแบบทดสอบที่มีมิติเดียว และสามารถอธิบายได้ด้วยโมเดลการตอบข้อสอบแบบ 1, 2 และ 3 พารามิเตอร์ ซึ่งในปัจจุบันนี้ใช้ในการออกแบบและวิจัยในการทดสอบทางการศึกษาและจิตวิทยา ความแตกต่างที่สำคัญระหว่างโมเดลนี้อยู่ที่รูปแบบทางคณิตศาสตร์ ที่ใช้อธิบายโค้งลักษณะข้อสอบ (Item Characteristic Curve) ความแตกต่างที่สำคัญอีกอย่างหนึ่งคือ วิธีการตรวจให้

คะแนนของการตอบรายข้อ โมเดลที่สำคัญประกอบด้วยลักษณะดังนี้

1. โลจิสติกโมเดลแบบหนึ่งพารามิเตอร์ (One-parameter Logistic Model) มีพารามิเตอร์สำหรับแสดงลักษณะของข้อสอบหนึ่งค่า คือ ค่าความยากของข้อสอบ (Item Difficulty : b) โดยถือว่าข้อสอบทุกข้อมีค่าอำนาจจำแนก (Item discrimination : a) เท่ากันหมด และมีค่าการเดา (Guessing Parameter : c) เป็นศูนย์
2. โลจิสติกโมเดลแบบสองพารามิเตอร์ (Two-parameter Logistic Model) มีพารามิเตอร์สำหรับแสดงลักษณะของข้อสอบสองค่า คือ ค่าอำนาจจำแนกของข้อสอบ (a) และค่าความยากของข้อสอบ (b)
3. โลจิสติกโมเดลแบบสามพารามิเตอร์ (Three-parameter Logistic Model) มีพารามิเตอร์สำหรับแสดงลักษณะของข้อสอบสามค่า คือ ค่าอำนาจจำแนกของข้อสอบ (a) ค่าความยากของข้อสอบ (b) และค่าการเดา (c) ซึ่งเป็นค่าโอกาสที่ผู้สอบที่มีความสามารถต่ำจะทำข้อสอบถูก

การนำทฤษฎีการตอบข้อสอบไปใช้มีขั้นตอนดังนี้ (1) เก็บรวบรวมข้อมูลและเลือกโมเดลที่เหมาะสม (2) นำข้อมูลผลการสอบไปวิเคราะห์เพื่อหาค่าพารามิเตอร์ต่าง ๆ ตามโมเดลที่เลือกไว้ ในการเลือกโมเดลของทฤษฎีการตอบข้อสอบไปใช้ ต้องพิจารณาความสอดคล้องระหว่างโมเดลกับลักษณะของข้อมูล เช่น โมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ ควรใช้กับข้อมูลจากแบบทดสอบที่ข้อสอบทุกข้อมีค่าอำนาจจำแนก (a) เท่ากัน และค่าโอกาสในการตอบถูกหรือค่าการเดา (c) มีค่าเท่ากับ 0.0 หรือโมเดลโลจิสติกแบบ 2 พารามิเตอร์ ควรใช้กับข้อมูลจากแบบสอบที่ข้อสอบทุกข้อมีค่าการเดา (c) เท่ากับ 0.0 เป็นต้น

✱ โดยปกติแล้วไม่จำเป็นที่จะเป็นการวัดหรือการประมาณค่าใด ๆ ก็ตาม จะต้องมีความคลาดเคลื่อนเกิดขึ้นเสมอ โดยเฉพาะอย่างยิ่งในการวัดหรือการประมาณค่าความสามารถหรือคุณลักษณะที่ซ่อนเร้นอยู่ภายใน (Traits) ทางด้านการศึกษาและจิตวิทยาก็จะมีความคลาดเคลื่อนค่อนข้างสูง ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ เป็นค่าสถิติที่สำคัญแสดงถึงประสิทธิภาพการประมาณค่าของข้อสอบจากแบบทดสอบที่วิเคราะห์ด้วยทฤษฎีการตอบสองข้อสอบ (IRT) ความคลาดเคลื่อนมาตรฐานที่สำคัญคือ

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ (Standard Error of Estimated Item Parameters) ลักษณะของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ เป็นประเด็นที่สำคัญในการพัฒนาแบบทดสอบ ค่าสถิตินี้ได้จากส่วนกลับของเมตริกซ์ของอินฟอर्मชัน โดยใช้สมการของนิวตัน-ราฟสัน (Newton-Raphson

Equation) ซึ่งเป็นผลมาจากกระบวนการในการประมาณค่าพารามิเตอร์ข้อสอบ ในทางปฏิบัติ ค่าความคลาดเคลื่อนมาตรฐานนี้ เป็นผลจากการคำนวณขั้นสุดท้ายของกระบวนการประมาณค่า ด้วยวิธีแมกซิมัมไลกิลิฮูด

ไทซ์เซนและไวเนอร์ (Thissen and Wainer, 1982) ได้ศึกษาความคลาดเคลื่อนของพารามิเตอร์ข้อสอบ สำหรับโมเดลหนึ่ง สอง และสามพารามิเตอร์ ภายใต้ข้อตกลงของการทราบค่าความสามารถของผู้สอบ (θ) ซึ่งมีการแจกแจงแบบปกติ $n(0,1)$ โดยการใช้โมเดล 2 พารามิเตอร์ และโมเดล 3 พารามิเตอร์ เมื่อ $c=0$ พบว่าถึงแม้ว่าค่าสถิติของ a และ b จะเท่ากัน แต่เมตริกซ์ของอินฟอर्मชันจะไม่เท่ากัน เมตริกซ์ของโมเดลสามพารามิเตอร์ยังคงมีความสัมพันธ์กับพารามิเตอร์ c สำหรับข้อสอบที่ง่ายและมีค่าอำนาจจำแนกต่ำ ความคลาดเคลื่อนมาตรฐานภายใต้โมเดล 2 พารามิเตอร์ ประมาณอย่างคร่าว ๆ เท่ากับ .09 สำหรับโมเดล 2 พารามิเตอร์และโมเดล 3 พารามิเตอร์ เมื่อ $c=0$ ความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์ข้อสอบจะไม่เท่ากัน และที่เห็นได้ชัดคือ ความคลาดเคลื่อนมาตรฐานของความยาก (b) ของข้อสอบ ภายใต้รหัสโมเดล จะมีความคงที่น้อยกว่าค่าที่ได้จากโมเดล 2 และ 3 พารามิเตอร์

จากนั้นยังพบอีกว่า โมเดลสามพารามิเตอร์ ดีน้อยกว่าโมเดล 1 และ 2 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ภายใต้โมเดลสามพารามิเตอร์ จะยอมรับได้เพียงค่ากลางของพิสัยของค่า b และ a เมื่อค่า c ต่ำ สรุปว่าการใช้วิธีการประมาณค่าแบบแมกซิมัมไลกิลิฮูด สำหรับโมเดลสามพารามิเตอร์จะให้ผลไม่ตรงกับในทางปฏิบัติ ในบางกรณี หรือไม่เช่นนั้นจะต้องใช้กลุ่มตัวอย่างที่มีขนาดใหญ่มาก เพื่อให้ข้อจำกัดลดลง เดอกรูเจอร์ (De Gruijter, D.N.M., 1984 : 269 - 272) ทำการทดสอบด้วยวิธีเดียวกันนี้ พบว่าการใช้การแจกแจงขั้นแรก (prior distribution) ของค่า c จะลดความคลาดเคลื่อนมาตรฐานของทั้งสามพารามิเตอร์

ข้อจำกัดหนึ่งของการศึกษาความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ ข้อสอบที่กล่าวข้างต้นนั้น อยู่บนพื้นฐานของข้อสอบรายข้อและการทราบค่า θ ลอร์ดและวิงเกอร์สกี (Lord and Wingersky, 1985 : 69 - 88) ศึกษาเมตริกซ์ความแปรปรวนและความแปรปรวนร่วมแบบสุ่มของการประมาณค่าแบบแมกซิมัมไลกิลิฮูด เมื่อทุกพารามิเตอร์ไม่ทราบค่า โดยศึกษาความคลาดเคลื่อนมาตรฐานของโมเดลสามพารามิเตอร์ โดยกำหนดพารามิเตอร์ในการทดสอบแบบจำลองกับข้อสอบ 45 และ 90 ข้อ โดยสุ่มซ้ำ (replicate) ค่าพารามิเตอร์ข้อสอบ 15 ข้อ ใช้ความสามารถ (θ) ที่มีการแจกแจงแบบปกติและแบบเท่ากัน

(uniform) กับกลุ่มตัวอย่าง 2 ขนาด คือ 1,500 คน และ 6,000 คน มีการรายงานค่าความคลาดเคลื่อนมาตรฐานของข้อสอบ 15 ข้อ ที่ใช้เป็นฐานเท่านั้น และต้องการค่าพารามิเตอร์ความยากของข้อสอบที่คัดเลือกไว้ โดยมีค่าเฉลี่ยเท่ากับ 0 และมีความแตกต่างระหว่างกลุ่มข้อสอบที่คัดเลือกไว้เท่ากับ 1 (scale unit) เมตริกซ์นี้เรียกว่า สเกลหลัก (capital scale) มีการรายงานความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ A, B และ C ของเมตริกซ์นี้ด้วย

2. ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ (The Standard Error of Estimate : SEE(θ)) คือค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนของค่าประมาณความสามารถ ซึ่งแสดงด้วยสมการดังต่อไปนี้

$$SEE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (\text{Warm. 1978 : 77})$$

ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (SEE(θ)) จะแปรผันตรงกันข้ามกับค่าสารสนเทศ (Information) ของแบบทดสอบ ณ ระดับความสามารถ ($I(\theta)$) นั้น เมื่อค่าสารสนเทศของแบบทดสอบ ณ ระดับความสามารถใดมีค่ามากแล้ว ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า (SEE(θ)) ณ ระดับความสามารถนั้น จะมีค่าน้อย และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าจะมีค่าน้อยที่สุด ณ ระดับที่ค่าสารสนเทศของแบบทดสอบมีค่าสูงสุด นั่นคือแบบทดสอบฉบับนั้นสามารถประมาณค่าความสามารถได้อย่างถูกต้องแม่นยำ ณ ระดับความสามารถนั้น ค่าเฉลี่ยของค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า จะมีความสัมพันธ์กับค่าความเชื่อมั่นในทฤษฎีการทดสอบแบบดั้งเดิม (r_{xx}) เมื่อคะแนนอยู่ในรูปคะแนนมาตรฐาน ที่มีค่าความเบี่ยงเบนมาตรฐานเท่ากับ 1.0 ซึ่งแสดงด้วยสมการดังต่อไปนี้

$$r_{xx} = 1 - \frac{SEE(\theta)^2}{\sigma^2} \quad (\text{Warm. 1978 : 77})$$

นั่นก็หมายความว่า แบบทดสอบที่มีค่าความเชื่อมั่นสูงอาจจะเป็นแบบทดสอบที่ไม่ค่อยดีนักต่อการนำไปใช้ เนื่องจากค่าสารสนเทศของแบบทดสอบนั้นจะมีค่าต่ำ ณ ระดับความสามารถที่ไม่พึงปรารถนา เช่นเดียวกับแบบทดสอบที่มีค่าความเชื่อมั่นต่ำ อาจจะเป็นแบบสอบที่ดี เมื่อแบบ

ทดสอบนั้นมีสารสนเทศสูงตามต้องการ ดังนั้นค่าความเชื่อมั่นของแบบทดสอบจะนำไปสู่ความเข้าใจผิดในแบบทดสอบนั้น ๆ ได้ (Warm. 1978 : 77)

ในการสร้างแบบทดสอบที่ประกอบด้วยข้อสอบที่ผ่านการวิเคราะห์ด้วยทฤษฎีการตอบสนองข้อสอบ (IRT) มักจะมีค่าความแปรปรวนเกี่ยวกับขนาดของกลุ่มผู้สอบ และความยาวของแบบทดสอบที่เหมาะสม เพื่อให้มีประสิทธิภาพในการประมาณค่าความสามารถได้ใกล้เคียงความจริงที่สุด ปัจจัยที่ส่งผลต่อประสิทธิภาพในการประมาณค่า ได้แก่ การเลือกใช้โมเดลที่เหมาะสม วิธีการประมาณค่า ลักษณะการแจกแจงความสามารถของกลุ่มผู้สอบ ค่าพารามิเตอร์ที่ต้องการประมาณค่า และจุดประสงค์ของการนำโมเดลไปใช้ (Hambleton. 1989 : 171)

นักวิจัยหลายท่านได้ทำการศึกษาและเสนอแนะขนาดของกลุ่มตัวอย่าง และความยาวของแบบทดสอบที่เหมาะสม เช่น ไรท์และสโตน (Wright and Stone. 1979) เสนอให้ใช้ข้อสอบ 20 ข้อ และกลุ่มตัวอย่างขนาด 200 คน สำหรับการวิเคราะห์โดยใช้โมเดลแบบหนึ่งพารามิเตอร์ ฮิวลิน, ลิสแซคและดราสโกว (Hulin, Lissak and Drasgow. 1982 : 249 - 260) เสนอให้ใช้ข้อสอบอย่างน้อยที่สุด 30 ข้อ และกลุ่มตัวอย่างตั้งแต่ 500 คนขึ้นไป สำหรับโมเดลโลจิสติกแบบสองพารามิเตอร์ สวามินาธานและจีฟฟอร์ด (Swaminathan and Gifford. 1983 : 9 - 30) พบว่าแบบทดสอบที่มีข้อสอบ 20 ข้อ และผู้สอบ 1,000 คน เหมาะสมกับโมเดลโลจิสติก และเพื่อให้ได้ค่าประมาณที่ดี ควรใช้ข้อสอบ 80 ข้อ

จากการศึกษางานวิจัยเกี่ยวกับทฤษฎี IRT พบว่านักการศึกษาส่วนใหญ่มักจะใช้กลุ่มตัวอย่างที่มีขนาดใหญ่และแบบทดสอบที่มีจำนวนมากข้อ กับการวิเคราะห์ด้วยโมเดลโลจิสติกแบบ 2 และ 3 พารามิเตอร์ เนื่องจากพบว่ากลุ่มตัวอย่างไม่ต่ำกว่า 1,000 คน และแบบทดสอบที่มีจำนวนมากข้อจะได้ค่าประมาณพารามิเตอร์ที่ดี แต่ในทางปฏิบัติ การพัฒนาแบบทดสอบการใช้กลุ่มตัวอย่างที่มีขนาดใหญ่ มักจะประสบอุปสรรคหลายด้าน เช่น นักเรียนมักเบื่อหน่ายกับการตอบแบบทดสอบที่มีจำนวนมากข้อ สิ้นเปลืองงบประมาณ เวลา ตลอดจนอุปสรรคในการเข้าถึงกลุ่มตัวอย่าง เป็นต้น ผู้วิจัยจึงมีความสนใจที่จะศึกษาถึงแบบทดสอบที่มีประสิทธิภาพในการประมาณค่าความสามารถได้ใกล้เคียงความเป็นจริงที่สุด โดยศึกษาจากความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถ ของแบบทดสอบที่มีจำนวนข้อน้อย ด้วยนัยที่ว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ (SEE) มีความสัมพันธ์กับค่าพารามิเตอร์ของข้อสอบ และมีความสัมพันธ์ในทางตรงกันข้ามกับค่าสารสนเทศของแบบทดสอบ (Test Information) (Hambleton. 1989 : 164 ; Warm. 1978 : 77) นั่นคือสารสนเทศของแบบทดสอบ

๗ ระดับความสามารถโคมีค่ามากแล้ว ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า ๗ ระดับความสามารถนั้นจะต่ำ โดยมุ่งศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ และความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถ ภายใต้อิมเพลโตโลจิสติกแบบ 1, 2 และ 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่างที่มีขนาดต่าง ๆ กัน ขนาดของกลุ่มตัวอย่างที่ต่างกันนี้ น่าจะส่งผลถึงค่าพารามิเตอร์ที่ได้จากการประมาณค่าด้วยอิมเพลโตโลจิสติกแบบเดียวกัน หรืออิมเพลโตโลจิสติกที่ต่างกันก็ได้ ซึ่งจะเป็ประโยชน์ในการสร้างข้อสอบที่มีคุณภาพโดย ใช้กลุ่มตัวอย่างที่มีขนาดเหมาะสมได้

จุดมุ่งหมายของการศึกษาค้นคว้า

เพื่อศึกษาความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ที่เป็นผลจากการประมาณค่าด้วยอิมเพลโตโลจิสติก แบบ 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากันโดยมีจุดมุ่งหมายของการศึกษาค้นคว้า ดังต่อไปนี้

1. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน ที่วิเคราะห์ด้วยอิมเพลโตโลจิสติกแบบเดียวกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน ที่วิเคราะห์ด้วยอิมเพลโตโลจิสติกแบบเดียวกัน
3. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างการวิเคราะห์ด้วยอิมเพลโตโลจิสติกแบบต่างกัน ที่มีจำนวนกลุ่มตัวอย่างเท่ากัน
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ระหว่างการวิเคราะห์ด้วยอิมเพลโตโลจิสติกแบบต่างกัน ที่มีจำนวนกลุ่มตัวอย่างเท่ากัน

ความสำคัญของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้ เป็นการศึกษาถึงความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ภายใต้อิมเพลโตโลจิสติกแบบ 1, 2 และ 3 พารามิเตอร์ เมื่อใช้กลุ่มตัวอย่างที่มีจำนวนต่างกัน ผลของการศึกษาจะทำให้ได้ข้อสอบที่มีคุณภาพ โดยใช้กลุ่มตัวอย่างที่เหมาะสม ที่วิเคราะห์ด้วยอิมเพลโตโลจิสติก และเป็นแนวทางในการพัฒนาแบบทดสอบ โดยใช้อิมเพลโตโลจิสติกต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 เขตบางกะปิ ลาดพร้าว บึงกุ่ม ลาดกระบัง หนองจอกและมีนบุรี กรุงเทพมหานคร จำนวนโรงเรียน 23 โรงเรียน จำนวนห้องเรียน 367 ห้องเรียน จำนวนนักเรียน 12,029 คน

2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนโรงเรียน 10 โรงเรียน จำนวนห้องเรียน 52 ห้องเรียน จำนวนนักเรียน 2,500 คน ซึ่งได้มาจากการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) โดยใช้โรงเรียนเป็นหน่วยการสุ่ม

ตัวแปรที่ใช้ในการศึกษาค้นคว้าครั้งนี้มีตัวแปรดังนี้

1. ตัวแปรต้น ได้แก่

1.1 โมเดลการตอบสนองข้อสอบ 3 รูปแบบ คือ

1.1.1 โมเดลโลจิสติกแบบ 1 พารามิเตอร์

1.1.2 โมเดลโลจิสติกแบบ 2 พารามิเตอร์

1.1.3 โมเดลโลจิสติกแบบ 3 พารามิเตอร์

1.2 จำนวนกลุ่มตัวอย่างที่ต่างกัน แบ่งออกเป็น 4 กลุ่มดังนี้

1.2.1 จำนวน 250 คน

1.2.2 จำนวน 500 คน

1.2.3 จำนวน 750 คน

1.2.4 จำนวน 1,000 คน

2. ตัวแปรตาม ได้แก่

2.1 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ
ด้านต่อไปนี้

2.1.1 พารามิเตอร์ด้านความยากของข้อสอบ

2.1.2 พารามิเตอร์ด้านอำนาจจำแนกของข้อสอบ

2.1.3 พารามิเตอร์ด้านการเดาของข้อสอบ

2.2 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ

นิยามศัพท์เฉพาะ

1. พารามิเตอร์ หมายถึง คุณสมบัติเชิงสถิติของข้อสอบรายข้อ ที่ใช้อธิบายความสัมพันธ์ระหว่างคุณสมบัติของผู้สอบและโอกาสในการตอบข้อสอบ ได้แก่

1.1 ค่าความยาก (b) หมายถึง ความยากของข้อสอบตามโมเดลโลจิสติก ซึ่งแสดงตำแหน่งของโค้งลักษณะข้อสอบบนมาตราความสามารถ (Ability Scale) ที่ค่าความน่าจะเป็นในการตอบถูกมีค่า 0.50 เมื่อไม่มีการเดา

1.2 ค่าอำนาจจำแนก (a) หมายถึง ค่าอำนาจจำแนกของข้อสอบตามโมเดลโลจิสติก ซึ่งมีค่าเป็นสัดส่วนกับค่าความชันของโค้งลักษณะข้อสอบ (Item Characteristic Curve) ณ จุดเปลี่ยนโค้งข้อสอบ

1.3 ค่าการเดา (c) หมายถึง ค่าความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำมีโอกาสจะทำข้อสอบข้อที่ i ได้ถูก เป็นตำแหน่งที่โค้งลักษณะข้อสอบ (ICC) ตัดแกนตั้ง โดยที่มีจุดตัดแกนตั้งสูงจะมีโอกาสของการเดาได้ถูกสูงกว่าโค้งที่มีจุดตัดแกนตั้งต่ำ

2. โมเดลโลจิสติก หมายถึง รูปแบบทางคณิตศาสตร์ที่ใช้อธิบายโค้งลักษณะข้อสอบ (ICC) โดยมีรูปแบบต่างกันตามจำนวนพารามิเตอร์ที่ใช้อธิบายโค้งลักษณะข้อสอบ ได้แก่ โมเดลดังนี้

2.1 โมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ หรือราส์โมเดล เป็นรูปแบบคณิตศาสตร์ที่มีค่าพารามิเตอร์แสดงลักษณะความยากของข้อสอบ

2.2 โมเดลโลจิสติกแบบสองพารามิเตอร์ เป็นรูปแบบคณิตศาสตร์ที่มีค่าพารามิเตอร์แสดงลักษณะอำนาจจำแนกของข้อสอบและลักษณะความยากของข้อสอบ

2.3 โมเดลโลจิสติกแบบสามพารามิเตอร์ เป็นรูปแบบคณิตศาสตร์ที่มีค่าพารามิเตอร์แสดงลักษณะอำนาจจำแนกของข้อสอบ ลักษณะความยากของข้อสอบ และลักษณะการเดาของผู้สอบ

3. ค่าความสามารถของผู้สอบ (Examinee's ability : θ) หมายถึงระดับความรู้ความเข้าใจในเนื้อหาวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม โดยกำหนดให้หน่วยของ θ ที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความเบี่ยงเบนมาตรฐานเท่ากับ 1

4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ (Standard Error of Estimation) คือ ค่าประมาณพารามิเตอร์ที่เบี่ยงเบนไปจากค่าที่แท้จริงของค่าพารามิเตอร์ ซึ่งจำแนกออกเป็น 2 ส่วน คือ

4.1 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบคือ ค่าประมาณพารามิเตอร์ข้อสอบที่เบี่ยงเบนไปจากค่าที่แท้จริงของข้อสอบ ได้แก่ พารามิเตอร์ ความยาก (b) พารามิเตอร์อำนาจจำแนก (a) และพารามิเตอร์การเดา (c)

4.2 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ (Standard Error of Ability Estimation) คือค่าประมาณพารามิเตอร์ความสามารถที่เบี่ยงเบนไปจากค่าความสามารถที่แท้จริง

5. สารสนเทศของข้อสอบ (Item Information Function) หมายถึง ความแม่นยำในการประมาณค่าความสามารถจริงจากผลการตอบข้อสอบข้อนั้น ๆ ซึ่งมีค่าเท่ากับ อัตราส่วนของกำลังสองของความชันโค้งลักษณะข้อสอบ ต่อค่าความแปรปรวนของข้อสอบข้อนั้น

6. สารสนเทศของแบบทดสอบ (Test Information Function) หมายถึง ผลรวมของสารสนเทศของข้อสอบทุก ๆ ข้อ ในแบบทดสอบนั้น ซึ่ง ณ ระดับความสามารถใดที่สารสนเทศของแบบทดสอบมีค่าสูง แสดงว่าแบบทดสอบนั้นมีคุณภาพดีมากในการประมาณค่าความสามารถที่ระดับนั้น

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการนำเสนอเอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยเรื่องนี้ ผู้วิจัยได้แบ่งการนำเสนอออกเป็นขั้นตอนดังนี้

1. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)
2. ความคลาดเคลื่อนมาตรฐาน (Standard Error)
3. การวิเคราะห์ความสอดคล้องกับโมเดล
4. แบบทดสอบและกลุ่มตัวอย่าง
5. งานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT)

ทฤษฎีการตอบสนองข้อสอบ มีชื่อที่เรียกอย่างอื่นคือ ทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) หรือทฤษฎีโค้งลักษณะข้อสอบ (Item Characteristic Curve Theory) หรือทฤษฎีการทดสอบแนวใหม่ (Modern Test Theory) ทฤษฎีนี้อธิบายความสัมพันธ์ระหว่างพฤติกรรมที่สังเกตได้ของผู้สอบกับคุณลักษณะที่ไม่สามารถสังเกตได้ หรือความสามารถของผู้สอบ ซึ่งความสัมพันธ์นี้เรียกว่า โค้งลักษณะข้อสอบนั่นเอง (Hambleton. 1989 : 149)

ความสัมพันธ์ระหว่างปริมาณของค่าสังเกต (คะแนนสอบ) และปริมาณที่ไม่สามารถสังเกตได้ (ความสามารถ) อธิบายด้วยฟังก์ชันทางคณิตศาสตร์ ฟังก์ชันนี้อยู่บนพื้นฐานข้อตกลงที่เกี่ยวกับคะแนนจากการทดสอบ โมเดลการตอบข้อสอบมีรูปแบบที่ต่างกันในความแตกต่างของข้อตกลงเกี่ยวกับคะแนนจากการทดสอบ ซึ่งโมเดลที่มีข้อตกลงที่สมเหตุสมผลกว่าจะสามารถลดปัญหาการทดสอบในทางปฏิบัติได้ดีกว่า

แมคโดนัลด์ (McDonald. 1982 : 379 - 396) ได้กำหนดกรอบการจำแนกโมเดลการตอบข้อสอบ โดยจำแนกดังนี้

1. โมเดลความเป็นมิติเดียว (Unidimensional) และหลายมิติ (Multidimensional)
2. โมเดลเชิงเส้นตรง (Linear) และโมเดลเชิงเส้นโค้ง (nonlinear)
3. โมเดลการตอบสนองแบบสองทาง (Dichotomous) และโมเดลการตอบสนองแบบหลายทาง (Polychotomous)

ในปัจจุบันนี้นักการศึกษาและนักวัดผล ได้ให้ความสนใจเป็นอย่างมากต่อการศึกษา
โมเดลความเป็นมิติเดียว โมเดลเชิงเส้นโค้ง และโมเดลการตอบสนองแบบสองทาง

ข้อตกลงเบื้องต้น

ในแต่ละโมเดลของทฤษฎีการตอบสนองข้อสอบ จะมีการกำหนดข้อตกลงที่ต่างกันไป
เกี่ยวกับลักษณะของข้อมูล และฟังก์ชันคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้
กับตัวแปรที่ไม่สามารถสังเกตได้ ข้อตกลงเบื้องต้นที่สำคัญ ได้แก่

1. มิติของแบบทดสอบ (Dimensionality) โดยหลักการของทฤษฎีคุณลักษณะแฝง
สมมุติให้ความสามารถของบุคคลมีอยู่ทั้งหมด k อย่าง ซึ่งความสามารถแต่ละอย่างนี้ต่างก็ส่งผล
ต่อการตอบสนองข้อสอบ (Test Performance) ต่าง ๆ ที่รวมกันเป็นแบบทดสอบ ข้อตกลง
เบื้องต้นของโมเดลการตอบข้อสอบส่วนมากจะกำหนดให้เพียงความสามารถหรือคุณลักษณะเดียว
ก็เพียงพอที่จะอธิบายคะแนนจากการสอบและความสัมพันธ์ระหว่างข้อสอบ โดยสมมุติให้ความ
สามารถแฝงเดียวนี้อ้างอิงไปสู่ความเป็นมิติเดียว (Unidimensional) (Hambleton,
1989 : 150) อย่างไรก็ตาม แบบทดสอบอาจจะไม่เป็นมิติเดียวเนื่องจากมีปัจจัยอื่นที่ส่งผล
ต่อการตอบข้อสอบของผู้สอบ เช่น ความรู้เกี่ยวกับการตอบข้อสอบ บุคลิกภาพ แรงจูงใจ
ความวิตกกังวล และความรู้ความจำเรื่องอื่น ๆ แต่ข้อตกลงเบื้องต้นข้อนี้ก็ไม่ได้เข้มงวดนัก
เพียงแต่ให้มีลักษณะเด่น (Dominant) ที่จะวัดในองค์ประกอบใดองค์ประกอบหนึ่งก็ใช้ได้

วิธีการตรวจสอบว่าแบบสอบนั้นวัดในมิติเดียวหรือไม่นั้น สามารถทำได้โดยการวิเคราะห์
องค์ประกอบ (Factor Analysis) แล้วสังเกตค่าไอเกน (Eigen Value) ค่าสูงสุด ว่า
แตกต่างจากค่าอื่น ๆ อย่างชัดเจนหรือไม่ (Hambleton and Thumb. 1973 : 195 -
211 ; อ้างอิงมาจาก ต่าย เชิญณี. 2534 : 23)

2. ความเป็นอิสระในการตอบข้อสอบ (local Independence) หมายถึงการตอบ
ข้อสอบแต่ละข้อของผู้สอบ มีความเป็นอิสระต่อกันเชิงสถิติ นั่นคือ การตอบข้อสอบข้อใดข้อหนึ่งถูก
หรือผิด จะไม่มีผลต่อการตอบข้ออื่น ๆ ด้วย มีเพียงความสามารถของผู้สอบและลักษณะของข้อสอบ
เท่านั้นที่ส่งผลต่อคะแนนสอบ ข้อตกลงเบื้องต้นของความเป็นอิสระในการตอบข้อสอบ สามารถ
กล่าวในเชิงคณิตศาสตร์ได้ว่า หมายถึงโอกาสที่จะตอบข้อสอบถูกต้องทั้งหมดจะเท่ากับผลคูณของโอกาส
การตอบข้อสอบถูกต้องเป็นรายข้อ ซึ่งสามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ดังนี้ (Hambleton,
1989 : 150)

$$\begin{aligned}
 & \text{Prob}(U_1=u_1, U_2=u_2, \dots, U_n=u_n / \theta) \\
 &= (P_1^{u_1} Q_1^{1-u_1})(P_2^{u_2} Q_2^{1-u_2}) \dots (P_n^{u_n} Q_n^{1-u_n}) \\
 &= \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i}
 \end{aligned}$$

เมื่อ U = ลักษณะของการตอบ

P = ความน่าจะเป็นของการตอบถูก

Q = $1-P$ (ความน่าจะเป็นของการตอบผิด)

จากสมการ ความน่าจะเป็นจะขึ้นอยู่กับความสามารถ (ability) ซึ่งโดยหลักการแล้วความสามารถของบุคคลจะมีหลายมิติ (Multidimensional) ในความเป็นมิติเดียวของโมเดลการตอบสนองข้อสอบ คือ โอกาสของแบบการตอบสนองที่สมบูรณ์ของผู้สอบจะเท่ากับความน่าจะเป็นซึ่งสัมพันธ์กับคะแนนของผู้สอบแต่ละคน เช่น เมื่อพิจารณาข้อสอบ 3 ข้อของผู้สอบ ณ ระดับความสามารถ θ_u สมมติ $P_1(\theta_u) = .8$, $P_2(\theta_u) = .5$ และ $P_3(\theta_u) = .4$ ถ้าเป็นไปตามข้อตกลงนี้ ความน่าจะเป็นของค่าสังเกตเท่ากับ $2^3 = 8$ ซึ่งแบบการตอบสนองที่เป็นไปได้ในกลุ่มตัวอย่างของผู้สอบ ณ ระดับความสามารถ θ_u ดังแสดงในตาราง 1

ตาราง 1 โอกาสที่จะเกิดของแบบการตอบสนอง

แบบการตอบสนอง	โอกาสที่จะเกิด
0 0 0	$Q_1 Q_2 Q_3 = .2 \times .5 \times .6 = .06$
1 0 0	$P_1 Q_2 Q_3 = .8 \times .5 \times .6 = .24$
0 1 0	$Q_1 P_2 Q_3 = .2 \times .5 \times .6 = .06$
0 0 1	$Q_1 Q_2 P_3 = .2 \times .5 \times .4 = .04$
1 1 0	$P_1 P_2 Q_3 = .8 \times .5 \times .6 = .24$
1 0 1	$P_1 Q_2 P_3 = .8 \times .5 \times .4 = .16$
0 1 1	$Q_1 P_2 P_3 = .2 \times .5 \times .4 = .04$
1 1 1	$P_1 P_2 P_3 = .8 \times .5 \times .4 = .04$

จากตาราง 1 ข้อตกลงเบื้องต้นของความเป็นอิสระในการตอบข้อสอบ เป็นข้อตกลงที่สำคัญของโมเดลการตอบสนองข้อสอบทุกโมเดล ในทางปฏิบัติสำหรับระดับความสามารถนั้น ๆ ความน่าจะเป็นของแบบการตอบสนองต่าง ๆ อยู่บนพื้นฐานของข้อตกลงเบื้องต้นในความเป็นอิสระในการตอบข้อสอบ (Local Independence) ซึ่งสามารถคำนวณได้และสามารถเปรียบเทียบอัตราส่วนที่ปรากฏของแบบการตอบสนองเดียวกัน ในกลุ่มตัวอย่างที่ระดับความสามารถเดียวกัน ข้อสังเกตเกี่ยวกับความสัมพันธ์ของข้อตกลงของความเป็นอิสระในการตอบข้อสอบในการที่ความสามารถที่มีมิติเดียวเกี่ยวข้องกับข้อตกลงเกี่ยวกับความเป็นมิติเดียว คือ

1. เมื่อแบบทดสอบวัดเพียงความสามารถเดียว สำหรับผู้สอบที่มีความสามารถระดับหนึ่งนั้น การตอบข้อสอบจะมีความเป็นอิสระในเชิงสถิติ ถ้าการตอบสนองไม่มีความเป็นอิสระในเชิงสถิติ ณ ระดับความสามารถนั้น ผู้สอบบางคน ณ ระดับความสามารถนั้นจะมีคะแนนสอบที่คาดหวังสูงกว่าผู้สอบคนอื่น ณ ระดับความสามารถเดียวกัน นั่นคือ สารสนเทศของผู้สอบจะอยู่บนความสามารถที่มากกว่าหนึ่งความสามารถ ซึ่งจะต้องมีการตรวจสอบถึงคะแนนสอบเหล่านั้น สิ่งนี้จะปรากฏชัดเจนในการฝ่าฝืนข้อตกลงเบื้องต้นของความเป็นมิติเดียว

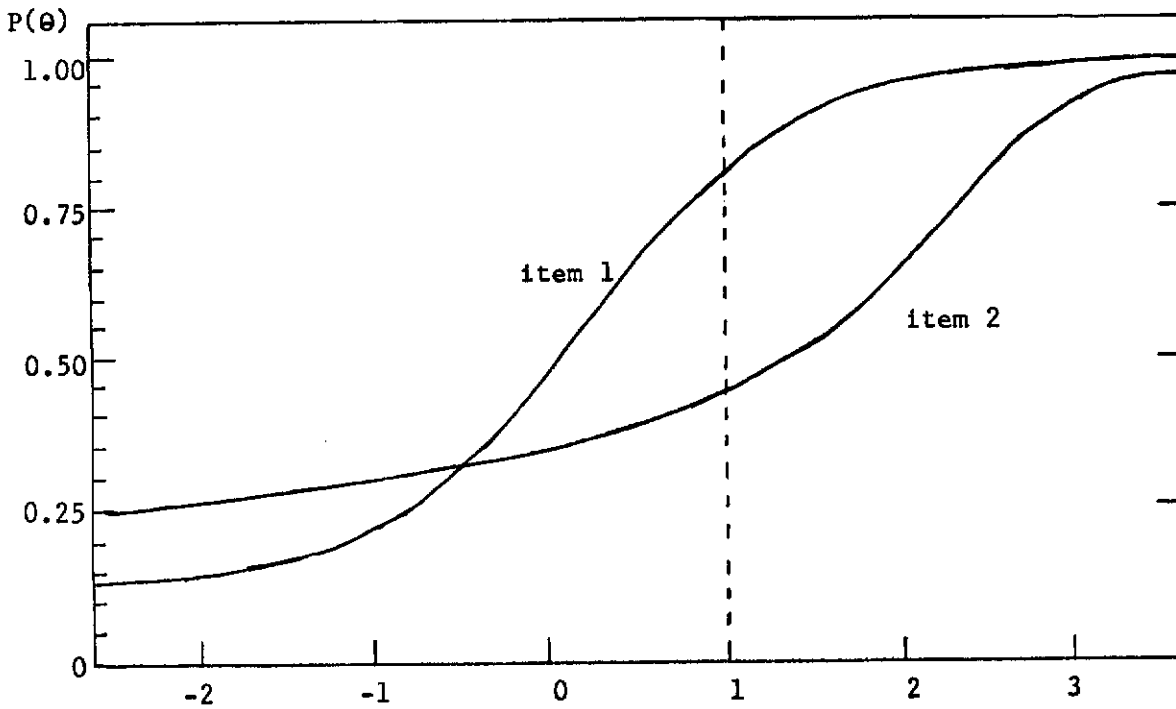
2. ข้อตกลงของความเป็นอิสระในการตอบ หมายถึงการตอบสนองข้อสอบจะมีความเป็นอิสระในเชิงสถิติ สำหรับผู้สอบ ณ ระดับความสามารถที่คงที่ ดังนั้นเพียงความสามารถเดียว

เท่านั้น ก็เพียงพอที่จะสามารถตรวจสอบความสัมพันธ์ระหว่างข้อสอบในแบบทดสอบนั้น

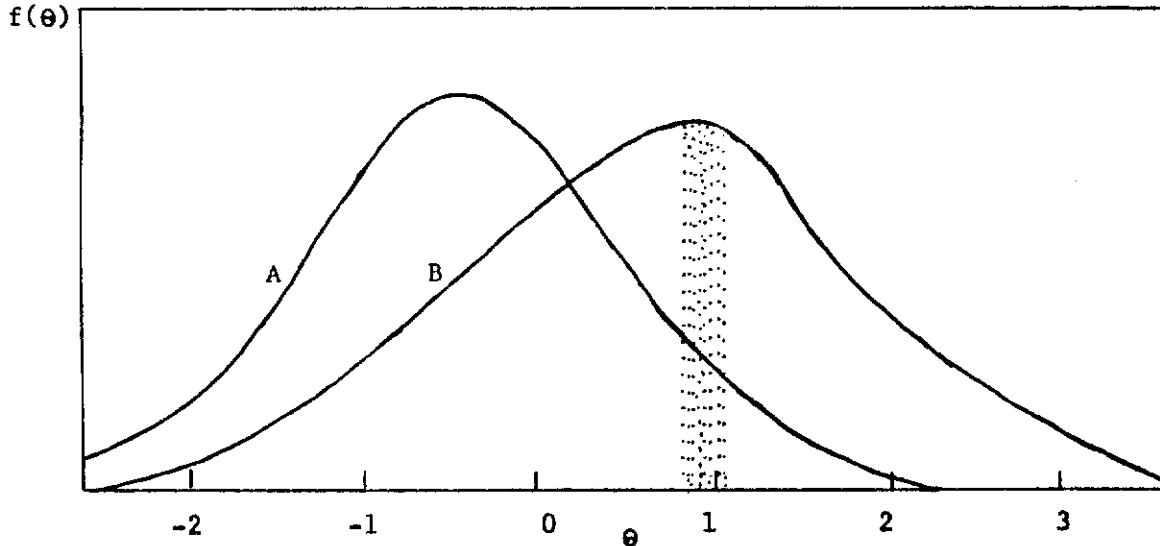
ความเป็นอิสระในการตอบข้อสอบ มีความสัมพันธ์กับความเป็นมิติเดียวของแบบทดสอบ คือ ถ้าแบบทดสอบฉบับใดมีความเป็นมิติเดียว แบบทดสอบฉบับนั้นจะมีความเป็นอิสระเฉพาะที่ในการตอบข้อสอบ (Hambleton, 1989 : 151) และในทางปฏิบัติความเป็นอิสระในการตอบข้อสอบ เป็นสิ่งที่ยากที่จะตรวจสอบได้จึงมักจะไม่มีการตรวจสอบ (Warm, 1978 : 101)

3. โค้งลักษณะข้อสอบ (Item Characteristic Curve : ICC) โค้งลักษณะข้อสอบคือฟังก์ชันคณิตศาสตร์ ที่แสดงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบข้อนั้นได้ถูกต้อง ($P_i(\theta)$) กับความสามารถ (θ) เป็นฟังก์ชันถดถอยของคะแนนรายข้อลงบนความสามารถ นั่นคือความน่าจะเป็นที่ผู้สอบแต่ละคนจะตอบข้อสอบรายข้อได้ถูกต้อง ขึ้นอยู่กับโค้งลักษณะข้อสอบ (ICC) ในแต่ละโมเดลที่ใช้ และความน่าจะเป็นนี้คือความเป็นอิสระของการแจกแจงความสามารถของผู้สอบ ซึ่งไม่ขึ้นอยู่กับการจัดกลุ่มของผู้สอบในประชากรที่มีความสามารถระดับเดียวกัน ตามหลักการของทฤษฎีนี้ หมายถึงภายในข้อจำกัดการเปลี่ยนแปลงของการแจกแจงความสามารถของกลุ่มตัวอย่างในทางปฏิบัติ จะไม่มีผลกระทบต่อค่าพารามิเตอร์ของข้อสอบ (Hambleton, 1989 : 151)

เนื่องจากความน่าจะเป็นที่ผู้สอบแต่ละคน จะตอบข้อสอบถูกต้องขึ้นอยู่กับรูปแบบของโค้งลักษณะข้อสอบ (Item Characteristic Curve) เท่านั้น ความน่าจะเป็นนี้คือความเป็นอิสระของการแจกแจงความสามารถของผู้สอบในประชากร ดังนั้นโอกาสที่จะตอบข้อสอบถูกต้องจะไม่ขึ้นอยู่กับการจัดกลุ่มของผู้สอบ ซึ่งมีระดับความสามารถเดียวกัน คุณสมบัติของความไม่เปลี่ยนแปลงของโค้งลักษณะข้อสอบ (ICC) คือลักษณะสำคัญที่ต้องการของโมเดลการตอบสนองข้อสอบ ความไม่เปลี่ยนแปลงนี้มีความสำคัญในการทดสอบที่เหมาะสมกับบุคคล (adaptive testing) การสร้างคลังข้อสอบ (Item Banking) และการประยุกต์ใช้ในการวัดผลการศึกษาต่าง ๆ ที่สำคัญคือ การเปลี่ยนแปลงของการแจกแจงความสามารถ ของกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ข้อสอบรายข้อ ในทางปฏิบัติจะไม่มีผลกระทบต่อค่าพารามิเตอร์ของข้อสอบ ลักษณะสำคัญของความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบ พิจารณาได้จากภาพประกอบ 1 และ 2 ดังต่อไปนี้



ภาพประกอบ 1 แสดงโค้งลักษณะข้อสอบของข้อสอบ 2 ข้อ



ภาพประกอบ 2 แสดงการแจกแจงความสามารถของผู้สอบ 2 กลุ่ม

จากภาพประกอบ 1 และ 2 แสดงให้เห็นโค้งลักษณะข้อสอบ (Item Characteristic Curve) ของข้อสอบ 2 ข้อ และการแจกแจงของความสามารถของกลุ่มผู้สอบ 2 กลุ่ม เมื่อข้อมูลสอดคล้องกับโมเดลการตอบสนองข้อสอบ จะได้ข้อสอบที่มีโค้งลักษณะ

ข้อสอบ (ICC) เดียวกันโดยประมาณ โดยไม่คำนึงถึงการแจกแจงของความสามารถของกลุ่มตัวอย่าง ที่ใช้ประมาณค่าพารามิเตอร์ของข้อสอบ โค้งลักษณะข้อสอบ (ICC) ระบบถึงโอกาสที่ผู้สอบซึ่งมีความสามารถ ณ ระดับนั้น จะตอบข้อสอบได้ถูกต้อง แต่โอกาสนี้ไม่ได้ขึ้นอยู่กับกลุ่มตัวอย่างที่มีความสามารถระดับนั้น

4. โมเดลของทฤษฎีการตอบสนองข้อสอบ (Item Response Model)

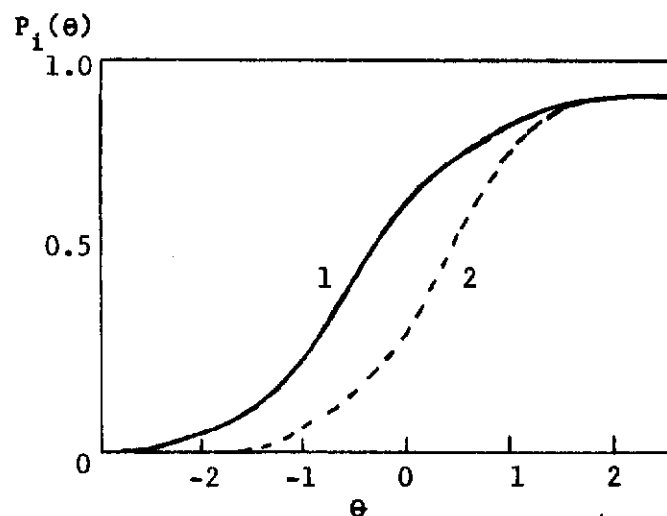
ปัจจุบันโมเดลการตอบสนองข้อสอบ ได้มีการนำมาใช้อย่างมากในการออกแบบและวิเคราะห์แบบทดสอบทางการศึกษาและจิตวิทยา ซึ่งโมเดลต่าง ๆ มีความแตกต่างกันในรูปแบบทางคณิตศาสตร์ของโค้งลักษณะข้อสอบ (Item Characteristic Curve) และความแตกต่างในวิธีการให้คะแนนของการตอบข้อสอบ โมเดลที่สำคัญมีดังนี้ (Hambleton, 1989 : 153 - 157)

4.1 โมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ (One-Parameter Logistic Model or Rasch Model) โมเดลการตอบข้อสอบ ซึ่งพัฒนาขึ้นโดย จอร์จ ราสช์ (Georg Rasch) นักคณิตศาสตร์ชาวเดนมาร์ก ได้เผยแพร่โมเดลนี้ในปี ค.ศ.1960 และเป็นที่สนใจในหมู่นักพัฒนาแบบทดสอบ ในปี ค.ศ.1967 ศาสตราจารย์ ไวท์ (Wright) และคณะได้ทำการฝึกสอนในหลายภาคเรียน โดยกล่าวถึงความแข็งแกร่งและการประยุกต์ใช้ราสช์โมเดล ความรวดเร็วในการพัฒนาและความง่ายต่อการประยุกต์ใช้ด้วยเหตุนี้ในปัจจุบัน ราสช์โมเดล จึงเป็นที่นิยมนำไปใช้จากผู้ปฏิบัติโดยทั่วไป โดยมีรูปแบบทางคณิตศาสตร์ดังนี้

$$P_i(\theta) = \frac{e^{Da(\theta_a - b_i)}}{1 + e^{Da(\theta_a - b_i)}} \quad (i=1,2,\dots,n)$$

จากสมการนี้คือรูปแบบของ ราสช์โมเดล เมื่อพิจารณาว่าเป็นกรณีเฉพาะ (special case) ของโมเดลโลจิสติกแบบสามพารามิเตอร์ โดยมีค่าพารามิเตอร์ของการเดา (c) เท่ากับ 0 และข้อสอบทุกข้อมีค่าอำนาจจำแนก (a) เท่ากัน นั่นคือสามารถกล่าวได้ว่าเป็นกรณีหนึ่งของโมเดลโลจิสติกแบบหนึ่งพารามิเตอร์นั่นเอง ในความหมายนี้ ราสช์โมเดล ก็คือกรณีเฉพาะของโมเดลโลจิสติกแบบสามพารามิเตอร์ของเบียร์นบอม โดยมีข้อตกลงที่เพิ่มเข้ามาของค่าอำนาจจำแนกที่เท่ากัน และไม่มีการเดาคำตอบถูกในผู้สอบที่มีความสามารถต่ำ ซึ่งในทางปฏิบัติจะไม่พบในการสอบแบบเลือกตอบ (Divgi, 1986 : 283 - 298 ; Traub, 1983 : 57 - 70)

ถึงแม้ว่า ราชส์โมเดล จะมองดูเหมือนกับว่าเป็นกรณีเฉพาะของโมเดลโลจิสติกแบบ 2 และ 3 พารามิเตอร์ แต่ก็ยังคงมีคุณสมบัติบางอย่างที่เป็นที่สนใจของผู้ใช้เพราะโมเดลนี้มีพารามิเตอร์น้อย ซึ่งง่ายแก่การนำไปใช้ ปัญหาจากการประมาณค่าด้วยราชส์โมเดล มีน้อยกว่าโมเดลอื่น ๆ และคุณสมบัติของจุดประสงค์เฉพาะ ซึ่งทำให้มีการประมาณค่าที่แยกจากกันอย่างสมบูรณ์ของข้อสอบและความสามารถ (Rasch. 1966 : 45 - 57) นั่นคือพารามิเตอร์ความสามารถ สามารถประมาณค่าได้โดยปราศจากความลำเอียง และมีความเป็นอิสระของข้อสอบที่เลือกมาจากข้อสอบที่สอดคล้องกับโมเดล เช่นเดียวกับพารามิเตอร์ข้อสอบ ซึ่งสามารถประมาณค่าได้โดยปราศจากความลำเอียง และเป็นอิสระจากความสามารถของกลุ่มตัวอย่างที่สอดคล้องกับโมเดล โด่งลักษณะข้อสอบโมเดลหนึ่งพารามิเตอร์ มีลักษณะดังต่อไปนี้ (Hambleton. 1989 : 154)



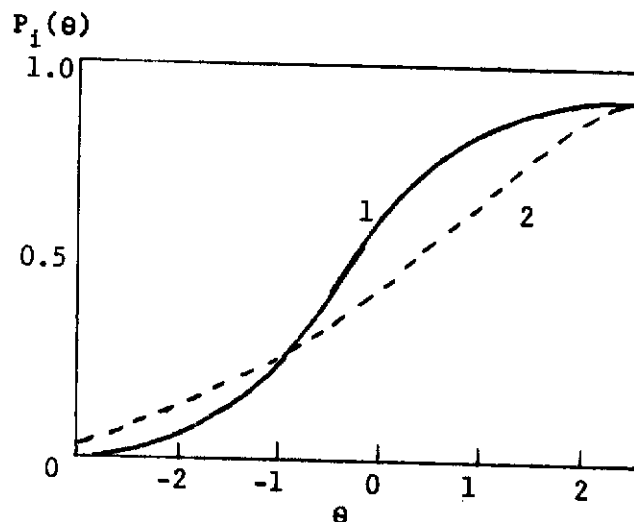
ภาพประกอบ 3 แสดงโด่งลักษณะข้อสอบโมเดลหนึ่งพารามิเตอร์

4.2 โมเดลโลจิสติกแบบสองพารามิเตอร์ (Two-Parameter Logistic Model) เบิร์นบอม (Birnbaum. 1968 : 399-400) ได้เสนอโด่งลักษณะข้อสอบ (Item Characteristic Curves) โดยใช้ฟังก์ชันโลจิสติกแบบสองพารามิเตอร์ ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i=1,2,\dots,n)$$

จากรูปแบบสมการดังกล่าว ค่า $P_i(\theta)$, a_i และ b_i มีความหมายเช่นเดียวกันกับของโมเดลปกติสะสม D คือ ค่าคงที่ (Scaling factor) เสนอโดย ลอร์ด เบิร์นบอมและคนอื่นๆ ซึ่ง

นำมาใช้แปลความหมายพารามิเตอร์ของโลจิสติกโมเดล คล้ายกับโมเดลปกติสะสมแบบสองพารามิเตอร์ แต่เป็นที่นิยมมากกว่า เนื่องจากโมเดลโลจิสติกสามารถคำนวณได้ง่ายกว่าโมเดลปกติสะสม และแสดงให้เห็นว่าเมื่อ $D=1.7$ ค่า $P_i(\theta)$ ของโมเดลปกติสะสมและโมเดลโลจิสติก จะมีค่าแตกต่างกันน้อยกว่า .01 สำหรับทุกค่าความสามารถ (Birnbaum. 1968 : 399 ; Hambleton. 1979 : 155) โด่งลักษณะข้อสอบโมเดลสองพารามิเตอร์ มีลักษณะดังต่อไปนี้ (Hambleton. 1989 : 155)



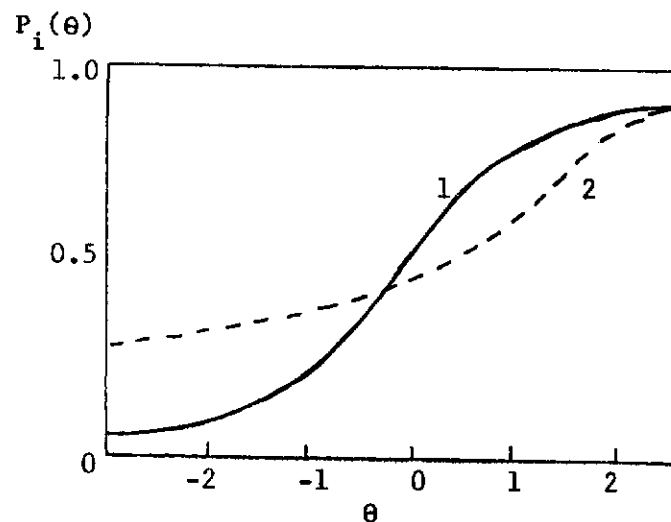
ภาพประกอบ 4 แสดงโด่งลักษณะข้อสอบโมเดลสองพารามิเตอร์

4.3 โมเดลโลจิสติกแบบสามพารามิเตอร์ (Three-Parameter Logistic Model) โมเดลโลจิสติกแบบสามพารามิเตอร์นี้ได้จาก โมเดลแบบสองพารามิเตอร์ โดยการเพิ่มพารามิเตอร์ตัวที่ 3 คือ ค่าการเดา (c) รูปแบบทางคณิตศาสตร์ของโด่งโลจิสติกแบบสามพารามิเตอร์ แสดงดังนี้

$$P_i(\theta) = c_i + (1-c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i=1, 2, \dots, n)$$

จากรูปแบบคณิตศาสตร์ดังกล่าว ค่าของ $P_i(\theta)$, b_i , a และ D มีความหมายเช่นเดียวกับ โมเดลแบบสองพารามิเตอร์ โดยปรับแก้ค่า b_i บนสเกลความสามารถ ด้วยความน่าจะเป็นของการตอบถูกเท่ากับ $(1+c_i)/2$ ซึ่งมีความน่าจะเป็นมากกว่า .50 ส่วนพารามิเตอร์ของ

การเดา (c) คือแอมบิโกล่าของโด่งสารสนเทศข้อสอบ (ICC) แสดงถึงความน่าจะเป็นที่ผู้สอบที่มีความสามารถต่ำ ตอบข้อสอบได้ถูกต้อง การรวมค่า c เข้าในโมเดลก็เพื่อที่จะอธิบายการตอบข้อสอบของผู้สอบที่มีความสามารถต่ำในแบบทดสอบเลือกตอบ โด่งลักษณะข้อสอบโมเดลสามพารามิเตอร์ มีลักษณะดังนี้ (Hambleton. 1989 : 154)



ภาพประกอบ 5 แสดงโด่งลักษณะข้อสอบโมเดลสามพารามิเตอร์

จากภาพโด่งลักษณะข้อสอบโมเดลสามพารามิเตอร์ ค่า b จะบอกตำแหน่ง ณ จุดซึ่งโด่งลักษณะข้อสอบ (ICC) มีความชันมากที่สุดบนสเกลความสามารถ ณ จุดนี้จะมีความสัมพันธ์กับความน่าจะเป็นของการตอบข้อสอบถูกต้อง เท่ากับ $(1+c_i)/2$ (เท่ากับ $.50$ เมื่อ $c_i = .00$) ความชันของ ICC ณ ตำแหน่ง b_i มีค่าเท่ากับ $.425a_i(1-c_i)$ เมื่อ a_i คือค่าอำนาจจำแนก ผลลัพธ์นี้ได้จากการดิฟเฟอเรนเชียล (differentiate) นิพจน์ทางคณิตศาสตร์ของ ICC ในส่วนของ θ โดยกำหนดค่าให้ $\theta = b$

2. ความคลาดเคลื่อนมาตรฐาน (Standard Error)

การพัฒนาแบบทดสอบให้มีความเที่ยงตรงในการประมาณค่าความสามารถที่แท้จริงได้มากที่สุดนั้น จะต้องคำนึงถึงการศึกษาความคลาดเคลื่อนจากการประมาณค่าพารามิเตอร์ด้วย เนื่องจากการประมาณค่าพารามิเตอร์มีความคลาดเคลื่อนมาก นั่นคือความเที่ยงตรงในการประมาณค่าความสามารถจะลดน้อยลง ทฤษฎีการตอบสนองข้อสอบศึกษาความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ และความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถ ในเวลาเดียวกัน

2.1 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ (Standard Error of Estimated Item Parameters) ลักษณะของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ เป็นประเด็นที่สำคัญในการพัฒนาแบบทดสอบ แต่เนื่องจากที่ผ่านมาการศึกษาถึงขนาดของความคลาดเคลื่อนมาตรฐานนี้ไม่มากนัก เนื่องจากในทางปฏิบัติสูตรการคำนวณมีความยุ่งยาก ในปัจจุบันโปรแกรมที่นิยมใช้ในการวิเคราะห์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ด้วย และในการศึกษารังนี้ได้ใช้โปรแกรม BILOG 3 ในการวิเคราะห์ข้อสอบตามโมเดลโลจิสติกแบบ 1, 2 และ 3 พารามิเตอร์ ซึ่งประมาณค่าด้วยวิธี Marginal Maximum Likelihood โดยจะให้ค่าประมาณพารามิเตอร์ข้อสอบได้เที่ยงตรงกว่าค่าประมาณที่ได้จากโปรแกรม LOGIST (Yen. 1985 ; อ้างอิงมาจาก Baker. 1987 : 133) และใช้วิธีคำนวณค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าพารามิเตอร์ข้อสอบ ด้วยวิธี EM (Em algorithm) (Bock and Aitkin. 1981 : 443 - 459)

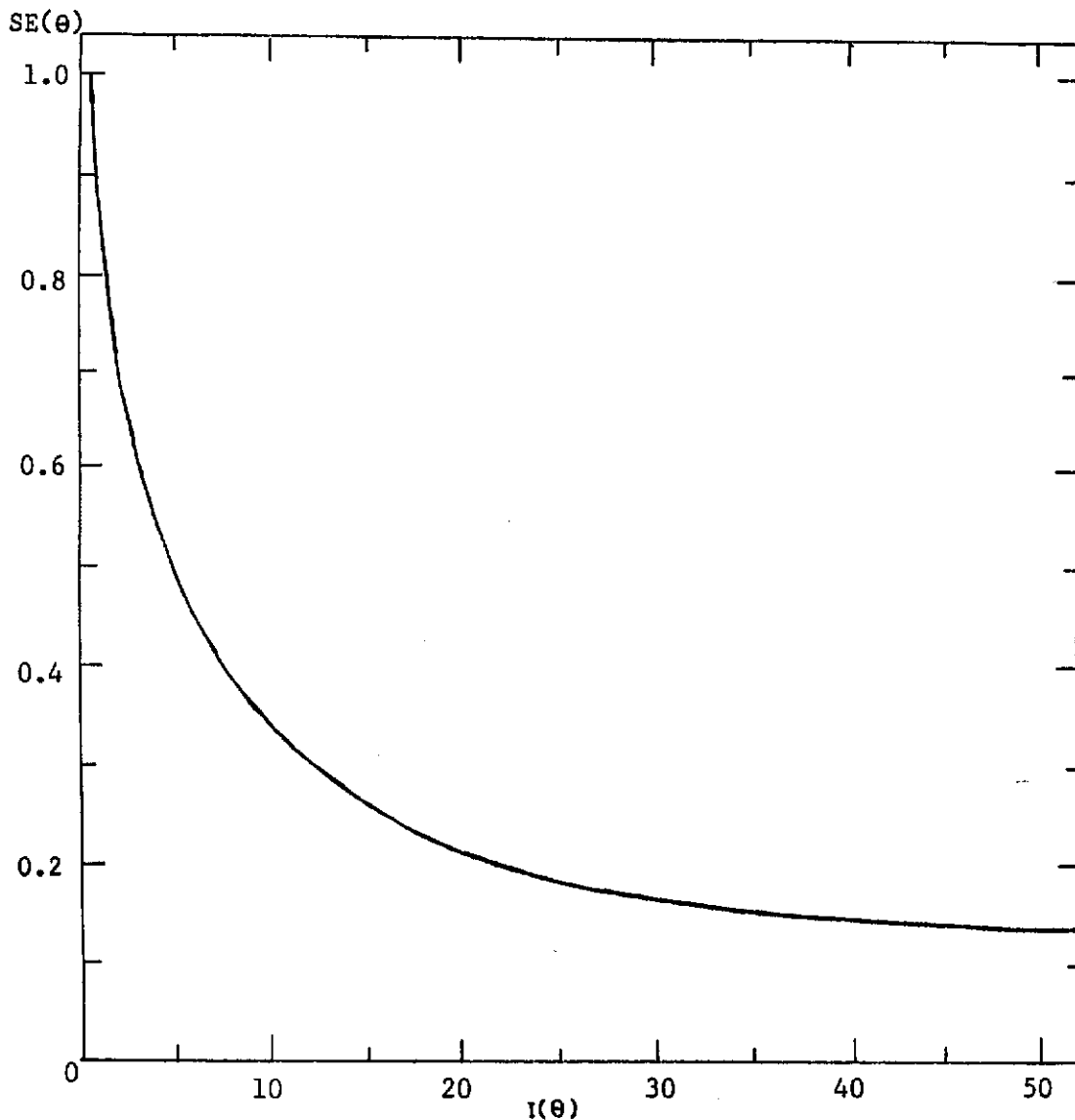
2.2 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถ (Standard Error of Estimate: SEE(θ)) การศึกษาค่าสารสนเทศของแบบทดสอบ ณ ระดับความสามารถ θ จะมีความสัมพันธ์ในทางกลับกันกับความเที่ยงตรงของค่าประมาณความสามารถ ณ ระดับนั้น ดังสมการต่อไปนี้

$$SE(\theta_0) = \frac{1}{\sqrt{I(\theta_0)}}$$

เมื่อ $SE(\theta_0)$ คือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่า นั่นคือ เมื่อประมาณค่าความสามารถ ด้วยวิธีแมกซิมัมไลค์ฮูด (Maximum Likelihood Estimate) จะพบช่วงของความเชื่อมั่นของการแปลค่าประมาณความสามารถ ณ ระดับต่าง ๆ กัน ในกรอบของทฤษฎีการตอบสนองข้อสอบ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า ($SE(\theta)$) สามารถนำไปใช้ประโยชน์ในลักษณะคล้ายกับค่าความคลาดเคลื่อนมาตรฐานในการวัด (Standard error of Measurement) ในทฤษฎีการวัดแบบดั้งเดิม (Classical Theory) แต่ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่า จะขึ้นอยู่กับแต่ละระดับของความสามารถ ในขณะที่ค่าความคลาดเคลื่อนมาตรฐานในการวัด (SE_{\dots}) มีเพียงค่าเดียว ซึ่งเป็นค่าสถิติของกลุ่มตัวอย่าง และนำไปใช้ในลักษณะเดียวกันกับผู้สอบทุกคนในกลุ่มตัวอย่าง (พจนานุกรม อักษรสารพ. 2528 : 67 ; Hambleton. 1989 : 164)

ความคลาดเคลื่อนมาตรฐานในการประมาณค่า $SE(\theta)$ เป็นค่าเบี่ยงเบนมาตรฐานของการแจกแจงปกติ ซึ่งสัมพันธ์กับค่าประมาณความสามารถของผู้สอบที่มีความสามารถ θ_0 การประมาณค่าจะดีที่สุดเมื่อสารสนเทศไม่ต่ำเกินไป

ขนาดของความคลาดเคลื่อนมาตรฐาน ขึ้นอยู่กับ (1) จำนวนข้อสอบในแบบทดสอบ โดยค่าความคลาดเคลื่อนมาตรฐานค่ามีความสัมพันธ์กับแบบทดสอบที่มีขนาดยาว (2) คุณภาพของข้อสอบ โดยทั่วไปความคลาดเคลื่อนมาตรฐานจะมีขนาดเล็กลงเมื่อข้อสอบมีค่าอำนาจจำแนกสูง ซึ่งการตอบข้อสอบถูกเกิดจากความสามารถของผู้สอบโดยไม่มีโอกาส (3) ความสอดคล้องระหว่างความยากของข้อสอบกับความสามารถของผู้สอบ เมื่อแบบทดสอบที่ประกอบด้วยข้อสอบที่มีค่าความยากเท่ากับค่าความสามารถของผู้สอบ ความสัมพันธ์ระหว่างความคลาดเคลื่อนมาตรฐานในการประมาณค่าและสารสนเทศของแบบทดสอบแสดงได้ดังต่อไปนี้ (Hambleton. 1989 : 156)



ภาพประกอบ 6 แสดงความสัมพันธ์ระหว่างความคลาดเคลื่อนมาตรฐานในการประมาณค่า และสารสนเทศของแบบทดสอบ

จากภาพประกอบ 6 แสดงให้เห็นความสัมพันธ์ระหว่างค่าความคลาดเคลื่อนมาตรฐาน และสารสนเทศของแบบทดสอบ จะเห็นได้อย่างชัดเจนว่า ขนาดของความคลาดเคลื่อนมาตรฐาน จะมีความคงที่อย่างรวดเร็ว เมื่อสารสนเทศของแบบทดสอบมีค่าสูงกว่า 25 ซึ่งจะมีผลกระทบ เพียงเล็กน้อยจากขนาดของความคลาดเคลื่อนในการประมาณค่าความสามารถ

ในการศึกษาครั้งนี้ คำนวณค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่าพารามิเตอร์ข้อสอบ ผู้วิจัยใช้โปรแกรม BILOG ซึ่งคำนวณด้วยวิธี EM (Em algorithm) (Bock and Aitkin. 1981 : 443 - 459)

3. การวิเคราะห์ความสอดคล้องกับโมเดล (The Analysis of Fit)

ทฤษฎีการตอบสนองข้อสอบ สามารถนำมาใช้ในการแก้ปัญหาการวัดและการทดสอบ แต่อย่างไรก็ตามผลจากการวิเคราะห์ด้วยโมเดลการตอบสนองข้อสอบ ไม่อาจเชื่อมั่นในค่าที่ประมาณได้ในทันที เนื่องจากอาจมีข้อมูลที่ไม่สอดคล้องกับโมเดลปะปนอยู่ ประโยชน์จากโมเดลการตอบสนองข้อสอบสามารถเป็นจริงขึ้นได้ เมื่อพบว่ามีความสอดคล้องระหว่างโมเดลกับข้อมูล การที่ข้อมูลสอดคล้องกับโมเดลอย่างไม่สมบูรณ์จะทำให้การประมาณค่าพารามิเตอร์ของข้อสอบ และพารามิเตอร์ความสามารถ ได้ค่าที่ไม่คงที่ (Hambleton. 1989 : 172)

ความสอดคล้องของข้อสอบได้จากการตรวจสอบความแตกต่างระหว่างความน่าจะเป็นแบบโพสทีเรีย (Posterior Probability) ของการตอบถูก ณ ระดับความสามารถนั้น กับความน่าจะเป็น ณ ระดับความสามารถเดียวกัน ความแตกต่างนี้เรียกว่า "Standardized posterior residuals" ซึ่งสามารถแสดงได้ดังสมการ (Mislevy and Bock. 1990 : 10)

$$d_{ik} = \frac{\sum_1^s w_{1k} [x_{1j} - P_j(x_k)]}{\left\{ \sum_1^s w_{1k} [x_{1j} - P_j(x_k)]^2 \right\}^{\frac{1}{2}}}$$

เมื่อ

$$w_{1k} = \frac{r_1 P(x_1/x_k)}{P(x_1)}$$

r_1 = ความถี่ของแบบการตอบสนอง x_1

$P(x_1)$ = Marginal probability ของแบบการตอบสนอง 1

x_k = Quadrature point

x_{1j} = แบบการตอบสนอง 1 ของข้อสอบข้อที่ j

$P_j(x_k)$ = ความน่าจะเป็น ณ Quadrature point ของข้อที่ j

s = จำนวนค่าสังเกตที่มีแบบการตอบสนองแตกต่างกัน

N = จำนวนผู้สอบ

ส่วนนี้ความสอดคล้องพิจารณาจากค่าพารามิเตอร์รากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (population root mean square) ของความเบี่ยงเบนโพสทีเรีย (posterior deviates) จากสูตร

$$\text{RMS}(d'_1) = \left[\frac{\sum_k^q N_k d_{jk}^2}{\sum_k^q N_k} \right]^{1/2}$$

เมื่อ N_k = ค่าคาดหวังของจำนวนผู้สอบ ณ quadrature point

ค่าส่วนที่เหลือโพสทีเรีย (posterior residual) มีความสัมพันธ์อย่างสูงกับค่าสถิติไคสแควร์ (chi-square : χ^2) นักการศึกษาบางท่านจึงทำการทดสอบความสอดคล้องโดยใช้สถิติไคสแควร์ เนื่องจากได้มีการจัดคะแนนของผู้สอบออกเป็นกลุ่มตามความสามารถในระดับเดียวกัน ในกระบวนการประมาณค่าพารามิเตอร์ จึงเป็นการง่ายที่จะใช้ค่าสถิติไคสแควร์ทดสอบความสอดคล้องกับโมเดล

ในการศึกษาครั้งนี้ผู้วิจัยตรวจสอบความสอดคล้องของข้อสอบและผู้สอบ กับโมเดลโดยใช้เกณฑ์ค่าสถิติไคสแควร์ ณ ระดับนัยสำคัญทางสถิติ .01

4. แบบทดสอบและกลุ่มตัวอย่าง

คำถามเกี่ยวกับความยาวของแบบทดสอบและขนาดของกลุ่มตัวอย่างที่เหมาะสมสำหรับการประมาณค่าพารามิเตอร์เป็นสิ่งที่ตอบได้ยาก การที่จะกำหนดความยาวของแบบทดสอบและขนาดของกลุ่มตัวอย่าง ขึ้นอยู่กับหลายปัจจัย ดังนี้ (Hambleton, 1989 : 171)

1. การเลือกใช้โมเดล IRT โดยหลักการแล้วโมเดลที่มีพารามิเตอร์มากกว่าจะต้องใช้กลุ่มตัวอย่างที่มีขนาดใหญ่กว่า
2. การเลือกใช้วิธีการประมาณค่าพารามิเตอร์และการประยุกต์ใช้ โดยหลักการแล้ววิธีการประมาณค่าแบบเบย์ (Bayesian Method) จะให้ค่าประมาณพารามิเตอร์ดีกว่าวิธีการประมาณค่าแบบแมกซิมัมไลค์ลิฮูด (Maximum Likelihood Method) ในแบบทดสอบที่มีขนาดสั้น และกลุ่มตัวอย่างขนาดเล็ก

3. การแจกแจงของค่าความสามารถ (ability) ในกลุ่มผู้สอบที่เป็นตัวอย่าง โดยหลักการกลุ่มตัวอย่างที่เป็นวิวิธพันธุ์ (heterogeneous) จะให้ค่าพารามิเตอร์ที่เป็นที่พอใจมากกว่ากลุ่มตัวอย่างที่เป็นเอกพันธ์ (homogeneous)

4. ค่าพารามิเตอร์ที่แท้จริงของข้อสอบหรือความสามารถ หรือทั้งสองค่าที่จะประมาณค่า โดยหลักการแล้วจะใช้แบบทดสอบที่ยาวและกลุ่มตัวอย่างขนาดใหญ่มาประมาณค่าพารามิเตอร์ทั้งสอง

5. ความสำคัญของโปรแกรมประยุกต์ IRT ค่าความคลาดเคลื่อนที่ยอมรับได้ของโปรแกรมประยุกต์ที่มีขนาดเล็กกว่า แบบทดสอบที่ยาวกว่า และกลุ่มตัวอย่างที่มีขนาดใหญ่กว่า เป็นสิ่งที่จำเป็น

5. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับความยาวของแบบทดสอบ ขนาดกลุ่มตัวอย่างที่เหมาะสมและความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ มีดังต่อไปนี้

งานวิจัยที่เกี่ยวข้องภายในประเทศ

บุญชิต รอดแก้ว (2533) ได้ศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์ด้วยโมเดลโลจิสติกสามพารามิเตอร์กับกลุ่มตัวอย่างที่มีลักษณะต่างกันคือ กลุ่มสูง กลาง ต่ำ และกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน 3 กลุ่ม คือ กลุ่ม 1,052 คน กลุ่ม 2,104 คน และกลุ่ม 3,155 คน โดยศึกษากับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ผลการศึกษาพบว่า

1. ค่าอำนาจจำแนกของข้อสอบของกลุ่มผู้สอบที่มีระดับความสามารถแตกต่างกัน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าอำนาจจำแนกของกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความยากของข้อสอบทั้งของกลุ่มผู้สอบที่มีระดับความสามารถแตกต่างกันและกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ค่าการเดาของข้อสอบทั้งของกลุ่มผู้สอบที่มีระดับความสามารถแตกต่างกันและกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

เวรดี อินทสระ (2530) ทำการเปรียบเทียบความคลาดเคลื่อนของการเทียบมาตรฐาน ระหว่างรูปแบบอิงทฤษฎีการตอบสนองกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ โดยใช้กลุ่มตัวอย่างจำนวน 2,823 คน ผลการศึกษาพบว่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน (SEE) ของรูปแบบอิงทฤษฎีการตอบสนองข้อสอบ มีค่าน้อยกว่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ และเมื่อทดสอบค่าเฉลี่ยของคะแนนความคลาดเคลื่อนของการเทียบมาตรฐานทั้งสองรูปแบบพบว่าคะแนนความคลาดเคลื่อนแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนความเพียงพอของความคลาดเคลื่อน การเทียบมาตรฐานรูปแบบอิงทฤษฎีการตอบสนองข้อสอบอยู่ในระดับน่าพอใจ และการเทียบมาตรฐานรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบอยู่ในระดับปานกลาง

บัญชา แสนทวี (2530) ได้ประยุกต์ใช้รูปแบบของราชส์ในการออกแบบโครงสร้างสนเทศของแบบทดสอบ ด้วยการจัดข้อสอบที่มีระดับความยากง่าย ให้เหมาะสมกับระดับความสามารถของนักเรียน พบว่ากลุ่มที่มีความสามารถต่ำและสูงเมื่อทดสอบความเหมาะสมกับความความสามารถของตนเอง ได้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถแตกต่างจากแบบทดสอบที่เหมาะสมกับกลุ่มความสามารถอื่น ๆ อย่างมีนัยสำคัญทางสถิติที่ระดับ ทั้งสามระดับ ได้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถแตกต่างอย่างไม่มีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบเมื่อกลุ่มความสามารถต่ำ สอบแบบทดสอบที่เหมาะสมกับความสามารถแล้วมีค่าสูงกว่าค่าความเชื่อมั่นที่ได้จากแบบทดสอบอื่น ๆ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 กลุ่มความสามารถปานกลางสอบแบบทดสอบที่เหมาะสมกับความสามารถแล้ว มีค่าความเชื่อมั่นสูงกว่าที่ได้จากแบบทดสอบที่เหมาะสมกับความสามารถต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนกลุ่มความสามารถสูงสอบแบบทดสอบที่เหมาะสมกับความสามารถสูง ได้ค่าความเชื่อมั่นสูงกว่าที่ได้จากแบบทดสอบที่เหมาะสมกับความสามารถต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และความเชื่อมั่นจากการสอบแบบทดสอบที่เหมาะสมกับความสามารถสูงนี้ มีความแตกต่างจากความเชื่อมั่นที่ได้จากการสอบแบบทดสอบที่เหมาะสมกับความสามารถปานกลาง อย่างไม่มีนัยสำคัญทางสถิติ

นันทิยา พึ่งคำ (คำข เชิงฉ. 2534 ; อ้างอิงมาจาก นันทิยา พึ่งคำ.

2531 : 57 - 58) ได้ศึกษาเปรียบเทียบคุณภาพของการทดสอบแบบ ซีเอที (CAT) กับแบบดั้งเดิม (Conventional) ในการวัดความสามารถด้านคำศัพท์ ภาษาอังกฤษ โดยทำการศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 120 คน การทดสอบแบบ ซีเอที ใช้กลุ่มข้อสอบ (Item Pool) จำนวน 361 ข้อ โดยมีค่าความยากอยู่ระหว่าง -.90552 ถึง 3.61860

มีค่าอำนาจจำแนกอยู่ระหว่าง 0.80334 ถึง 2.0000 และมีค่าการเดาอยู่ระหว่าง 0.04421 ถึง 0.29886 การประมาณค่าความสามารถของผู้สอบโดยใช้วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข (Conditional Maximum Likelihood) ส่วนการทดสอบแบบดั้งเดิมใช้ข้อสอบทั้งหมด 40 ข้อ โดยมีค่าความยากอยู่ระหว่าง -0.76488 ถึง 2.97884 มีค่าอำนาจจำแนกอยู่ระหว่าง 0.94069 ถึง 2.0000 และมีค่าการเดาอยู่ระหว่าง 0.3462 ถึง 0.28620 ผลการศึกษาพบว่า การทดสอบแบบซีเอที กับแบบดั้งเดิม มีความเที่ยงตรงเชิงเกณฑ์สัมพัทธ์ไม่แตกต่างกัน แต่การทดสอบแบบซีเอที สามารถลดจำนวนข้อสอบลงได้มากกว่าครึ่งหนึ่งและความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากการทดสอบแบบซีเอที มีค่าต่ำกว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบจากการทดสอบแบบดั้งเดิม อย่างมีนัยสำคัญทางสถิติ

คำย ชัยงฉ (2534) ศึกษาเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (SEE) จากแบบทดสอบเทเลอ์รูปปริามิตที่มีรูปแบบจำนวนขึ้น และวิธีการให้คะแนน ที่แตกต่างกัน โดยใช้กลุ่มทดลอง 32 กลุ่ม กลุ่มละ 1,000 คน ในแต่ละกลุ่มทำการทดลองซ้ำ 40 ครั้ง ทาค่าโอกาสในการตอบข้อสอบถูก ด้วยโมเดลโลจิสติกแบบหนึ่งพารามิเตอร์ โดยใช้รูปแบบการทดสอบเทเลอ์แบบขนาดขึ้นคงที่และแบบขนาดขึ้นแปรผัน มีจำนวนขึ้นของการทดสอบ 4 แบบ คือ แบบที่มีจำนวน 7 ขึ้น 10 ขึ้น 13 ขึ้น และ 16 ขึ้น ให้คะแนนในการทดสอบ 4 วิธีคือ ให้คะแนนตามค่าความยากของข้อที่ยากที่สุดที่ทำถูก ให้คะแนนตามค่าเฉลี่ยความยากของข้อที่ทำถูก ให้คะแนนโดยใช้วิธีของเบส์ และให้คะแนนโดยใช้วิธีความเป็นไปได้สูงสุดแบบมีเงื่อนไข พบว่าการทดสอบรูปปริามิตแบบขนาดขึ้นแปรผันที่มี 16 ขึ้น ให้คะแนนโดยใช้วิธีการของเบส์ จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถต่ำสุด เมื่อเพิ่มจำนวนขึ้นของรูปแบบทั้งสอง ทำให้ค่าความเชื่อมั่นและค่าความเที่ยงตรงเพิ่มขึ้น ค่าความคลาดเคลื่อนมาตรฐานในการวัดลดลง และพบว่าการให้คะแนนโดยใช้วิธีของเบส์ กับจำนวนขึ้นตั้งแต่ 10 ขึ้นขึ้นไป จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบต่ำสุด

งานวิจัยที่เกี่ยวข้องต่างประเทศ

สวามินาธานและจีฟฟอร์ด (1983) ทำการศึกษาข้อมูลจำลองในการตรวจสอบความเที่ยงตรงและความลำเอียงของค่าประมาณพารามิเตอร์ของข้อสอบ ภาสได้โมเดลแบบสามพารามิเตอร์ โดยใช้โปรแกรม LOGIST ใช้แบบทดสอบที่มีความยาว 10, 15 และ 20 ข้อ

และใช้ขนาดของกลุ่มตัวอย่าง 50, 200 และ 1,000 คน พบว่าจำนวนข้อสอบและขนาดกลุ่มตัวอย่างเพิ่มขึ้น เส้นสหสัมพันธ์การคาดหมายของค่าประมาณความยากและอำนาจจำแนก มีความแตกต่างเล็กน้อยจากค่าพารามิเตอร์ที่ประมาณเอียงเป็นมุม 45 องศา นั่นคือข้อมูลเชิงประจักษ์ทำให้ทราบว่ามีความเที่ยงตรงในการประมาณค่าความยากและอำนาจจำแนก ด้วยโมเดลแบบสามพารามิเตอร์

ฮิวลิน ลิสแซค และดราสโกว (Hulin, Lissak and Drassgow. 1982) ได้ศึกษาค่าพารามิเตอร์ข้อสอบภายใต้โมเดลสองพารามิเตอร์และโมเดลสามพารามิเตอร์ โดยศึกษาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (root mean square : RMS) พบว่าความยาวของแบบทดสอบและขนาดของกลุ่มตัวอย่างขึ้นอยู่กับโมเดล สำหรับโมเดลสองพารามิเตอร์ เมื่อใช้แบบทดสอบ 30 ข้อ และกลุ่มตัวอย่าง 500 คน จะให้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่พอใจ ในขณะที่โมเดลสามพารามิเตอร์ต้องใช้ข้อสอบ 60 ข้อ และผู้สอบ 1,000 คน และพบว่าเมื่อประมาณค่าด้วยโมเดลสองพารามิเตอร์ โดยใช้กลุ่มตัวอย่างมากกว่า 500 คนแล้ว ค่าประมาณพารามิเตอร์ข้อสอบและพารามิเตอร์ข้อสอบ จะมีความสัมพันธ์กันสูงกว่าเมื่อประมาณค่าด้วยโมเดลสามพารามิเตอร์ โดยสหสัมพันธ์ระหว่างค่าประมาณอำนาจจำแนกและค่าอำนาจจำแนก ($r_{b\theta}$) มีค่าประมาณ .9 และสำหรับโมเดลสามพารามิเตอร์มีค่าประมาณ .5 ส่วนค่าสหสัมพันธ์ระหว่างค่าประมาณความยากและค่าความยาก ($r_{b\mu}$) นั้น สำหรับโมเดลสองพารามิเตอร์จะมีค่าสูงกว่า .94 ในทุกขนาดกลุ่มตัวอย่างและทุกความยาวของแบบทดสอบ ส่วนโมเดลสามพารามิเตอร์เมื่อใช้กลุ่มตัวอย่าง ขนาด 1,000 คน และข้อสอบจำนวน 30 ข้อ จะมีค่าสหสัมพันธ์ ($r_{b\mu}$) สูงกว่า .94

คูเบียค (Kubiak. 1986) ทำการศึกษาความเที่ยงตรงของวิธีการประมาณค่าแบบแมกซ์ิมัมไลกิลิฮูด (Maximum Likelihood) ในการประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ความสามารถ ด้วยโมเดลโลจิสติกแบบ 1, 2 และ 3 พารามิเตอร์ และทำการเปรียบเทียบระหว่างค่าประมาณความสามารถของผู้สอบจากการประมาณค่าแบบแมกซ์ิมัมไลกิลิฮูดกับคะแนนรวมจากแบบทดสอบตามทฤษฎีแบบดั้งเดิม (Classical Test Theory) สำหรับโมเดล 1, 2 และ 3 พารามิเตอร์ ใช้สถานการณ์จำลอง (Simulation) ด้วยกลุ่มตัวอย่างขนาด 200, 500, 1,000 และ 2,000 คน ข้อสอบจำนวน 15, 30 และ 60 ข้อ ทำการตรวจสอบความเที่ยงตรงของค่าประมาณค่าพารามิเตอร์และค่าพารามิเตอร์ ด้วยสหสัมพันธ์ระหว่างค่าประมาณพารามิเตอร์ ค่าพารามิเตอร์ และค่าเฉลี่ยรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง

(average root mean squared error) ผลจากการศึกษาชี้ให้เห็นว่า ขนาดของกลุ่มตัวอย่างและความยาวของแบบทดสอบที่เหมาะสม จะขึ้นอยู่กับโมเดลการตอบสนองข้อสอบและลักษณะของการแจกแจงของความสามารถ โมเดลแบบ 3 พารามิเตอร์ต้องการกลุ่มผู้สอบขนาดใหญ่และข้อสอบมากข้อ เพื่อที่จะสามารถเปรียบเทียบกับโมเดล 1 และ 2 พารามิเตอร์ได้อย่างเที่ยงตรง ค่าประมาณพารามิเตอร์ความสามารถจากโมเดล 3 พารามิเตอร์ จะมีความสัมพันธ์กับพารามิเตอร์ความสามารถ ต่ำกว่าความสัมพันธ์ระหว่างกับคะแนนรวมจากการทดสอบกับพารามิเตอร์ความสามารถ

โยส์ (Yoes. 1994) ทำการเปรียบเทียบวิธีการประมาณค่า 3 วิธี คือ แมกซ์ิมัม-ไลค์ลิฮูด (Joint maximum likelihood) เบย์ส์ (Bayesian) และมาร์จินัลแมกซ์ิมัมไลค์ลิฮูด (Marginal maximum likelihood) โดยตรวจสอบความเที่ยงตรงของกระบวนการประมาณค่าพารามิเตอร์ข้อสอบด้วยโปรแกรมโลจิสต์ (LOGIST) ไบล็อก (BILOG) และแอสแคล (ASCAL) โดยใช้วิธีมอนติ คาร์โล (monte carlo simulation) ตัวแปรอิสระ คือ ขนาดกลุ่มตัวอย่าง 250, 500, 1,000 และ 2,000 คน ความยาวของแบบทดสอบ 15, 25, 50, 75, 100 ข้อ การแจกแจงของความสามารถแบบปกติ (Normal) แบบเท่ากัน (Uniform) แบบเบ้ทางลบ (negatively skewed) และคุณลักษณะของแบบทดสอบ (Test characteristics) ระหว่างค่ามอดเรท (Moderate) กับค่าอำนาจจำแนกเฉลี่ยสูง และระหว่างการแจกแจงของค่าความยากของข้อสอบแบบเท่ากัน (Uniform) กับแบบปกติ (Normal) โดยการนำปัจจัยเหล่านี้มาทำการประมาณค่าพารามิเตอร์ 240 ครั้ง โดยกระทำทั้ง 3 กระบวนการ

ผลจากการวิเคราะห์ได้ทำการตรวจสอบค่าพารามิเตอร์ข้อสอบแต่ละค่าโดยใช้

- 1) ค่าสหสัมพันธ์ 2) ความลำเอียง และ 3) ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean squared error) ตรวจสอบฟังก์ชันการตอบสนองข้อสอบ (IRF) และสารสนเทศของแบบทดสอบ โดยใช้ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (RMSE) ตรวจสอบพารามิเตอร์ความสามารถ (๑) โดยใช้ 1) ค่าสหสัมพันธ์ 2) ความลำเอียง 3) ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (RMSE) และทำการวิเคราะห์ความแปรปรวนแบบแฟกทอเรียล (Factorial Analysis of Variance) ด้วยการแปลงค่า Log ของดัชนีรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง เพื่อชี้ให้เห็นถึงผลกระทบและปฏิสัมพันธ์ระหว่างปัจจัยที่นำมาใช้ในการวิเคราะห์

ผลที่ได้รับชี้ให้เห็นว่า กระบวนการประมาณค่าแบบมาร์จินัลแมกซิมัโมไลลียูด โดยใช้โปรแกรม BILOG ใช้ในการกำหนดค่า (generate value) ได้ดีกับแบบทดสอบที่มีจำนวนข้อน้อยและกลุ่มตัวอย่างขนาดเล็ก จากแบบการประมาณค่าทั้ง 3 วิธี ค่าพารามิเตอร์อำนาจจำแนก (a) มีค่าสูงกว่าค่าประมาณ (overestimated) ส่วนค่าการเดา (c) ได้ค่าประมาณไม่ดีนัก แต่สามารถประมาณค่าพารามิเตอร์ความยาก (b) ได้ดีพอ ๆ กัน สำหรับค่าสหสัมพันธ์ค่าความลำเอียง และค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ระหว่างค่าประมาณพารามิเตอร์ข้อสอบชี้ให้เห็นว่าวิธีการประมาณค่าทั้ง 3 วิธี จะให้ค่าที่ใกล้เคียงกัน ผลจากการวิเคราะห์ความแปรปรวน แสดงว่าขนาดของกลุ่มตัวอย่างและความยาวของแบบทดสอบ เป็นปัจจัยสำคัญในการประมาณค่าพารามิเตอร์ ส่วนลักษณะการแจกแจงของความสามารถ ไม่มีผลกระทบต่อการบวนการในการประมาณค่า

ลอร์ด (วิรัช วรรณรัตน์. 2533 : 41 ; อ้างอิงมาจาก Lord. 1983) ได้ศึกษาข้อมูลในการตรวจสอบความลำเอียงของฐานโค้ง (asymptotic biases) ในการประมาณค่าพารามิเตอร์แบบสามพารามิเตอร์ โดยใช้จำนวนข้อสอบ 90 ข้อ จำนวนตัวอย่าง 2,995 คน โดยใช้แบบทดสอบความถนัดด้านภาษาพบว่า ในกรณีความยากง่ายของข้อสอบนั้น เมื่อข้อสอบง่ายและยากง่ายปานกลางนั้น ความลำเอียงมีค่าเป็นลบ แต่ถ้าข้อสอบยากความลำเอียงมีค่าเป็นบวก ส่วนในการประมาณค่าอำนาจจำแนกนั้นความลำเอียงมีค่าเป็นบวก และการประมาณค่าโอกาสการเดานั้นความลำเอียงมีค่าเป็นลบ โดยลอร์ดสรุปว่า ในลักษณะทั่วไปถ้าค่าพารามิเตอร์ของข้อสอบมีค่าความคลาดเคลื่อนมาตรฐาน (standard error) มากแล้ว จะมีความลำเอียงมากด้วย

ลอร์ดและวิงเกอร์สกี (Lord and Wingersky. 1985 : 69 - 88) สรุปว่าจำนวนของข้อสอบและขนาดกลุ่มตัวอย่างที่นำมาใช้ในการศึกษา ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ A, B และ C จะเปลี่ยนแปลงไปในทางกลับกันกับขนาดของกลุ่มตัวอย่าง แต่จะมีผลกระทบเล็กน้อยจากการเปลี่ยนแปลงจำนวนข้อสอบ การใช้กลุ่มตัวอย่างที่มีการกระจายของความสามารถ (θ) แบบเรคแทงกิวลา (rectangular) จะให้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ น้อยกว่าค่าที่ได้จากการใช้กลุ่มตัวอย่างที่มีการแจกแจงของความสามารถเป็นแบบปกติ (Normal) ซึ่งใช้ข้อสอบมากเป็นสองเท่าสำหรับค่า A และ C ที่มีค่าต่ำ ซึ่งเป็นผลจากการใช้เกณฑ์ $B-2/A$ น้อยกว่า 1 มีความคลาดเคลื่อนมาตรฐานที่คำนวณด้วยข้อมูลที่ได้จากผู้สอบที่มีการแจกแจงความสามารถแบบเรคแทงกิวลา จะมีค่าต่ำใกล้เคียงความคลาดเคลื่อนมาตรฐานที่คำนวณด้วยข้อมูลที่มีการแจกแจงของความ

สามารถแบบปกติ ซึ่งมีขนาดของกลุ่มตัวอย่างใหญ่เป็น 4 เท่า

เดอ กรูจ์เตอร์ (De Gruijter. 1985 : 247 - 249) ศึกษาความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ความยากของข้อสอบ โดยใช้รหัสโมเดล พบว่าขนาดของความคลาดเคลื่อนมาตรฐานที่ได้จากการศึกษาครั้งนี้ และขนาดของความคลาดเคลื่อนมาตรฐานซึ่งเป็นผลจากการศึกษาของ ริงเกอร์สกีและลอร์ด (1984) ขึ้นอยู่กับวิธีการระบุการแก้ปัญหา โดยยึดเมตริกซ์ของพารามิเตอร์ความสามารถ กรณีของรหัสโมเดลในการยึดเมตริกซ์พารามิเตอร์ความสามารถโดยกำหนดให้ค่าเฉลี่ยของความยาก (b) เท่ากัน จะให้ค่าความคลาดเคลื่อนมาตรฐานที่แตกต่างจากการกำหนดให้ความยากของข้อสอบแต่ละข้อมีค่าเท่ากับ 0

เบเกอร์ (Baker. 1987 : 111 - 138) ได้ศึกษาเอกสารทางทฤษฎีที่เกี่ยวข้องคุณสมบัติของค่าประมาณพารามิเตอร์ สรุปว่าโมเดลโลจิสติกแบบ 1 พารามิเตอร์ และ 2 พารามิเตอร์ มีโดเมนลักษณะข้อสอบ (ICC) ดีกว่าแบบ 3 พารามิเตอร์ นอกจากนี้ค่าความคลาดเคลื่อนมาตรฐานที่ได้จากการใช้โมเดล 1 พารามิเตอร์ และ 2 พารามิเตอร์ จะมีค่าน้อยกว่าแบบ 3 พารามิเตอร์ ทั้งในทางทฤษฎีและข้อมูลเชิงประจักษ์

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง จะเห็นว่าขนาดของกลุ่มตัวอย่างจะส่งผลต่อค่าประมาณพารามิเตอร์และความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ และนักการศึกษาก็ยังคงศึกษาถึงคุณลักษณะของค่าประมาณพารามิเตอร์ โดยใช้กลุ่มตัวอย่างที่มีขนาดต่าง ๆ กัน ทำให้ผู้วิจัยมีความสนใจที่จะศึกษาถึงผลกระทบจากขนาดของกลุ่มตัวอย่างต่าง ๆ กัน เมื่อประมาณค่าพารามิเตอร์ด้วยโมเดลโลจิสติกต่างกันว่าจะให้ค่าประมาณค่าพารามิเตอร์และความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์เป็นเช่นใด เพื่อให้ได้ข้อสอบที่มีคุณภาพและเป็นแนวทางในการพัฒนาแบบทดสอบที่มีคุณภาพต่อไป

สมมติฐานของการวิจัย

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของข้อสอบ ระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน มีค่าแตกต่างกัน เมื่อวิเคราะห์ด้วยโมเดลโลจิสติกแบบเดียวกัน
2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน มีค่าแตกต่างกัน เมื่อวิเคราะห์ด้วยโมเดลโลจิสติกแบบเดียวกัน
3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ของข้อสอบ ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติก แบบต่างกัน เมื่อใช้กลุ่มตัวอย่างจำนวนเท่ากัน มีค่าแตกต่างกัน
4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติก แบบต่างกัน เมื่อใช้กลุ่มตัวอย่างจำนวนเท่ากัน มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 ประกอบด้วย เขตบางกะปิ ลาดพร้าว บึงกุ่ม ลาดกระบัง หนองจอกและมีนบุรี จำนวนโรงเรียน 17 โรงเรียน จำนวนห้องเรียน 239 ห้องเรียน จำนวนนักเรียน 10,371 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนโรงเรียน 10 โรงเรียน จำนวนห้องเรียน 52 ห้องเรียน จำนวนนักเรียน 2,500 คน ซึ่งสุ่มมาด้วยวิธีการสุ่มแบบแบ่งชั้น โดยมีเขตเป็นชั้น และโรงเรียนเป็นหน่วยการสุ่ม ซึ่งมีรายละเอียดดังตารางที่ 2

ตาราง 2 โรงเรียนที่ได้จากการสุ่ม จำแนกตามเขต โรงเรียน จำนวนห้องเรียน
และจำนวนนักเรียน

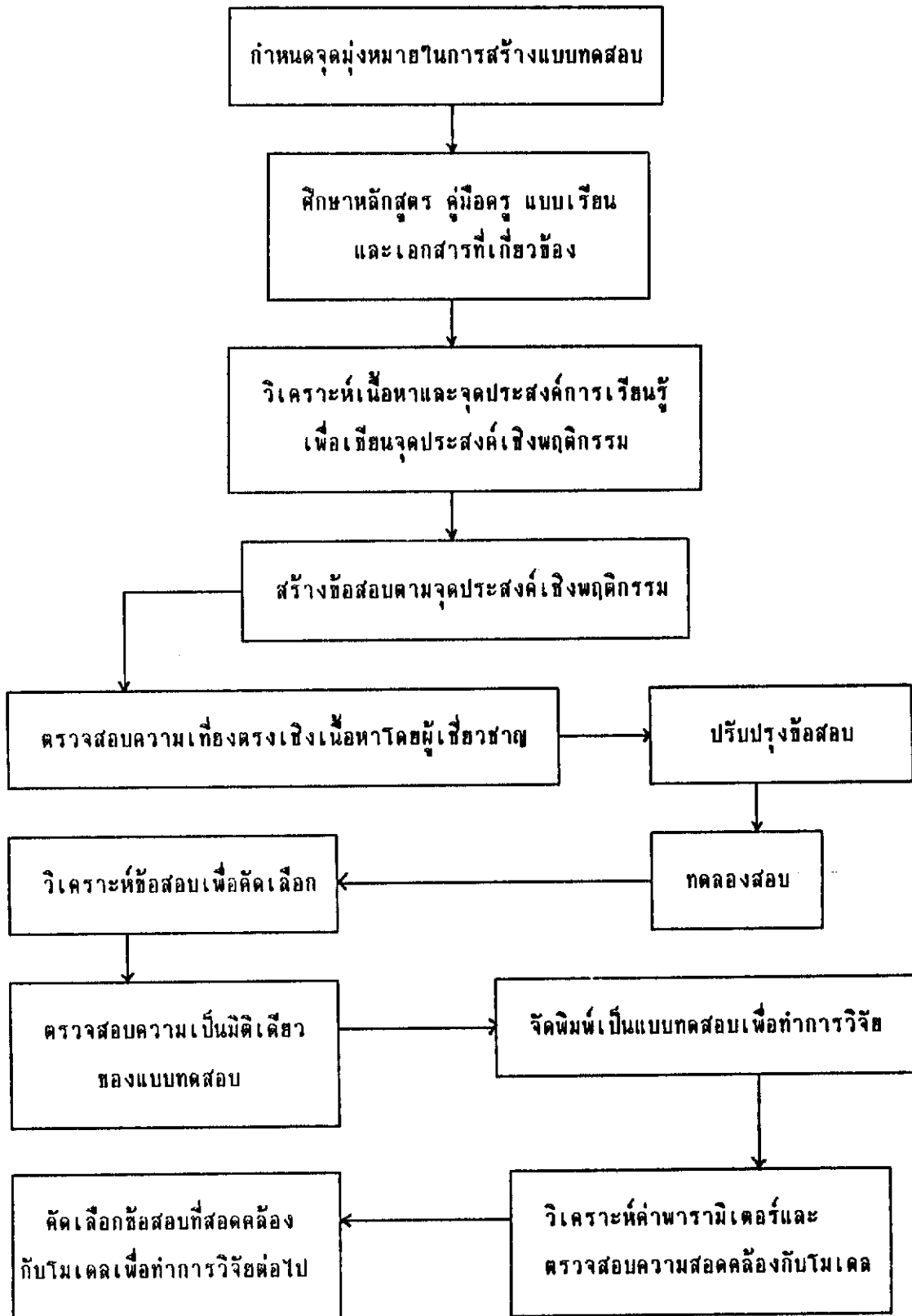
เขต	โรงเรียน	จำนวนห้องเรียน	จำนวนนักเรียน
บางกะปิ	1. บางกะปิ	8	397
	2. มีศยมวัดบึงทองหลาง	5	218
มีนบุรี	3. เศรษฐบุศรป่าเพ็ญ	6	295
	4. สตรีเศรษฐบุศรป่าเพ็ญ	5	233
ลาดกระบัง	5. เทพศิรินทร์ร่มเกล้า	6	295
	6. รัตนโกสินทร์สมโภช ลาดกระบัง	5	219
หนองจอก	7. วัดหนองจอก	4	211
ลาดพร้าว	8. ลาดปลาเค้าพิทยาคม	5	257
บึงกุ่ม	9. เดเรียนอุดมศึกษาน้อมเกล้า	3	117
	10. นวมินราชูทิศ กรุงเทพมหานคร	5	258
รวม		52	2,500

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนชนิด 5
ตัวเลือก วิชาคณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม ซึ่งผู้วิจัยสร้างขึ้นเอง
ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) จำนวน 20 ข้อ

วิธีดำเนินการสร้างแบบทดสอบ

การสร้างแบบทดสอบ ผู้วิจัยดำเนินการสร้างแบบทดสอบตามแผนผังดังนี้



ภาพประกอบ 7 แผนผังแสดงขั้นตอนการดำเนินการสร้างแบบทดสอบ

จากแผนภูมิค่าเป็นการสร้างแบบทดสอบ มีรายละเอียดดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ เพื่อสร้างแบบทดสอบไว้ใช้ในการวิจัย
2. ศึกษาเนื้อหาจากหลักสูตร คู่มือครู แบบเรียนและเอกสารที่เกี่ยวข้องในวิชา
คณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม
3. วิเคราะห์เนื้อหาและจุดประสงค์การเรียนรู้ เพื่อเขียนจุดประสงค์เชิงพฤติกรรม
ในแต่ละหน่วยการเรียนรู้ ซึ่งปรากฏว่าได้จุดประสงค์ 15 จุดประสงค์ ดังรายละเอียด
ในภาคผนวก ง
4. สร้างข้อสอบเลือกตอบชนิด 5 ตัวเลือก โดยสร้างตามจุดประสงค์เชิงพฤติกรรม
ทั้งหมด ซึ่งสร้างข้อสอบในจุดประสงค์ที่ 1-2 จุดประสงค์ละ 5 ข้อ จุดประสงค์ที่ 3,4,5,6,
9,10,11 และ 12 จุดประสงค์ละ 2 ข้อ จุดประสงค์ที่ 1,2,7 และ 15 จุดประสงค์ละ 3 ข้อ
จุดประสงค์ที่ 8,13 และ 14 จุดประสงค์ละ 4 ข้อ รวมข้อสอบทั้งหมด 40 ข้อ
5. นำข้อสอบที่สร้างขึ้น 40 ข้อ ไปให้ผู้เชี่ยวชาญซึ่งมีความรู้ในเนื้อหาและการเขียน
ข้อสอบในวิชาคณิตศาสตร์ เป็นอย่างดี จำนวน 5 คน เป็นผู้ตรวจความเที่ยงตรงเชิงเนื้อหา
(Content Validity) โดยพิจารณาความสอดคล้องระหว่างข้อสอบแต่ละข้อกับจุดประสงค์
เชิงพฤติกรรม และข้อสอบที่ถือว่าวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรม จะต้องได้รับความ
เห็นชอบจากผู้เชี่ยวชาญไม่น้อยกว่า 3 คน จึงถือว่าข้อสอบข้อนั้นใช้ได้

- ตัวอย่าง** แบบประเมินความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม
- คำชี้แจง** ให้ท่านพิจารณาว่าข้อสอบต่อไปนี้ข้อใดตรงตามจุดประสงค์เชิงพฤติกรรมที่กำหนดไว้หรือไม่ โดยพิจารณาให้คะแนนดังนี้
- 1 = แน่ใจว่าข้อสอบวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรม
- 0 = ไม่แน่ใจว่าข้อสอบวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรม
- 1 = แน่ใจว่าข้อสอบไม่ได้วัดตรงตามจุดประสงค์เชิงพฤติกรรม

จุดประสงค์เชิงพฤติกรรม	ข้อสอบ	คะแนนการพิจารณา		
		1	0	-1
เมื่อกำหนดจำนวนในรูปทศนิยมมาให้ นักเรียนสามารถบอกค่าประจำหลักของเลขโดดที่อยู่ในรูปทศนิยมตำแหน่งต่าง ๆ ได้	<p>0. ค่าของ 4×10 ตรงกับค่าของเลข 4 ของจำนวนในข้อใด</p> <p>1. 4.582</p> <p>2. 42.825</p> <p>3. 178.466</p> <p>4. 322.145</p> <p>5. 429.037</p> <p>00. ค่าของ $6 \times \frac{1}{10^4}$ ตรงกับค่าของเลข 6 ของจำนวนในข้อใด</p> <p>1. 6.3618</p> <p>2. 11.8946</p> <p>3. 13.0462</p> <p>4. 661.7140</p> <p>5. 6134.5321</p>			

จากผลการตรวจสอบความเที่ยงตรงเชิงเนื้อหาของผู้เชี่ยวชาญ ปรากฏว่าข้อสอบทั้ง 40 ข้อ มีความสอดคล้องกับจุดประสงค์เชิงพฤติกรรม

6. ปรับปรุงข้อสอบตามที่คุณเชี่ยวชาญได้พิจารณาไว้ เพื่อให้ได้ข้อสอบจำนวน 40 ข้อ

7. นำแบบทดสอบไปทดลองสอบกับกลุ่มตัวอย่างที่ใช้ในการสร้างเครื่องมือ ประมาณ 300 คน โดยเป็นกลุ่มตัวอย่างคนละกลุ่มกับที่ใช้ในการวิจัย โดยวิธีการสุ่มตัวอย่างแบบง่าย มีห้องเรียนเป็นหน่วยการสุ่ม ซึ่งได้แก่โรงเรียนสวนกุหลาบวิทยาลัย โรงเรียนสตรีวัดอัปสรสวรรค์ โรงเรียนราชประชานิธิ โดยสุ่มนักเรียนมาโรงเรียนละ 100 คน รวมทั้งสิ้นมีนักเรียน 300 คน

8. นำผลการสอบไปวิเคราะห์หาคุณภาพของข้อสอบ ตามทฤษฎีการทดสอบแบบดั้งเดิม เพื่อวิเคราะห์ข้อสอบเป็นรายข้อ โดยพิจารณาค่าความยากและค่าอำนาจจำแนก ทั้งตัวถูกและตัวลวง ด้วยโปรแกรม Item Analysis เพื่อคัดเลือกข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ คือ มีค่าความยาก อยู่ระหว่าง .20 ถึง .80 และมีค่าอำนาจจำแนก ตั้งแต่ .20 ขึ้นไป

สำหรับในการคัดเลือกข้อสอบครั้งนี้ สามารถคัดเลือกข้อสอบได้เพียง 12 ข้อ ที่มีค่าความยากตั้งแต่ .38 ถึง .78 และค่าอำนาจจำแนกตั้งแต่ .27 ถึง .59 จึงได้ข้อสอบไม่ครบตามจำนวนที่ต้องการเพื่อนำมาใช้ในการวิจัย

9. จากการพิจารณาผลการวิเคราะห์คุณภาพข้อสอบ พบว่าข้อสอบที่สร้างตามจุดประสงค์เชิงพฤติกรรม จุดประสงค์ที่ 2 และจุดประสงค์ที่ 3 มีค่าความยากและค่าอำนาจจำแนก ไม่อยู่ในเกณฑ์ที่กำหนด ผู้วิจัยจึงปรับปรุงและสร้างข้อสอบเพิ่มเติม โดยเพิ่มข้อสอบในจุดประสงค์ที่ 1, 9, 11 และ 15 จุดประสงค์ละ 3 ข้อ จุดประสงค์ที่ 13 และ 14 จุดประสงค์ละ 4 ข้อ จุดประสงค์ที่ 7 และ 8 จุดประสงค์ละ 5 ข้อ รวมข้อสอบทั้งหมด 40 ข้อ

10. นำข้อสอบที่สร้างขึ้น 40 ข้อ ไปให้ผู้เชี่ยวชาญชุดเดิมพิจารณา แล้วปรับปรุงข้อสอบตามที่คุณเชี่ยวชาญได้พิจารณาไว้ทั้ง 40 ข้อ

11. นำแบบทดสอบไปทดลองสอบกับกลุ่มตัวอย่างที่ใช้ในการสร้างเครื่องมือ ประมาณ 300 คน โดยเป็นกลุ่มตัวอย่างคนละกลุ่มกับการทดลองครั้งแรก และคนละกลุ่มกับที่ใช้ในการวิจัยโดยวิธีการสุ่มตัวอย่างแบบง่าย มีห้องเรียนเป็นหน่วยการสุ่ม โรงเรียนดังกล่าว ได้แก่ โรงเรียนสวนกุหลาบวิทยาลัย โรงเรียนสตรีวัดอัปสรสวรรค์ โรงเรียนฤทธิยะวารณาลัย โดยสุ่มนักเรียนมาโรงเรียนละ 100 คน รวมทั้งสิ้นมีนักเรียน 300 คน

12. นำผลการสอบไปวิเคราะห์หาคุณภาพของข้อสอบ ตามทฤษฎีการทดสอบแบบดั้งเดิม เพื่อวิเคราะห์ข้อสอบเป็นรายข้อ โดยโปรแกรมวิเคราะห์ข้อสอบ เพื่อคัดเลือกข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ที่กำหนด สำหรับในการคัดเลือกข้อสอบครั้งนี้ สามารถคัดเลือกข้อสอบได้ 20 ข้อ ที่มี

ค่าความยากตั้งแต่ .43 ถึง .73 และค่าอำนาจจำแนกตั้งแต่ .32 ถึง .68 รายละเอียดดังตาราง 49 ในภาคผนวก ค

13. นำผลจากข้อ 12 มาตรวจสอบความเป็นมิติเดียวของแบบทดสอบ โดยการวิเคราะห์องค์ประกอบ (Factor Analysis) โดยใช้โปรแกรม SPSSPC⁺ ซึ่งสกัดองค์ประกอบโดยวิเคราะห์องค์ประกอบหลัก (Principle Component Analysis : PC) แล้วหมุนแกนด้วยวิธีแวนิแมกซ์ (Varimax) ผลการวิเคราะห์ได้องค์ประกอบที่มีค่าไอเกนมากกว่า 1.0000 อยู่ 4 องค์ประกอบ โดยมีค่าไอเกนขององค์ประกอบที่ 1, 2, 3 และ 4 ดังนี้ 4.125, 1.368, 1.219 และ 1.069 ตามลำดับ และปรากฏว่าค่าไอเกนขององค์ประกอบที่ 1 มีค่าสูงกว่าองค์ประกอบที่ 2 อย่างมาก และสามารถอธิบายความแปรปรวนของคะแนนรวมทั้งหมดได้ 20.6 เปอร์เซ็นต์ ขณะที่องค์ประกอบที่ 2 และองค์ประกอบต่อไปมีค่าใกล้เคียงกัน และสามารถอธิบายความแปรปรวนของคะแนนรวมต่างกันไม่มากนัก ดังนั้นจึงสรุปได้ว่าแบบทดสอบฉบับนี้ มีคุณสมบัติการวัดเพียงมิติเดียวโดยประมาณ (Hambleton, 1985 : 21) ดังตาราง 42 และภาพประกอบ 14 ในภาคผนวก ก

14. นำข้อสอบที่คัดเลือกไว้ จัดพิมพ์เป็นรูปเล่ม แล้วไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างจำนวน 2,500 คน ดังแสดงในตาราง 1

15. นำข้อมูลจากกลุ่มตัวอย่างสุ่มเข้ากลุ่ม 4 กลุ่มคือ กลุ่ม 250 คน 500 คน 750 คน และ 1,000 คน มาวิเคราะห์หาค่าพารามิเตอร์ข้อสอบ และพารามิเตอร์ความสามารถ ตามทฤษฎีการตอบสนองข้อสอบ ด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ จากนั้นวิเคราะห์ความสัมพันธ์ของพารามิเตอร์ข้อสอบและพารามิเตอร์ความสามารถ กับโมเดล แล้วคัดเลือกข้อสอบและข้อสอบที่ไม่สอดคล้องกับโมเดลออก คงไว้เฉพาะข้อสอบและผู้สอบที่สอดคล้องกับโมเดล เพื่อนำไปศึกษาความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ต่อไป

ค่าพารามิเตอร์ความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา (c) และค่าพารามิเตอร์ความสามารถ (θ) ที่สอดคล้องกับโมเดล 1, 2 และ 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ปรากฏดังต่อไปนี้

กลุ่มตัวอย่าง 250 คน ค่าพารามิเตอร์ข้อสอบที่สอดคล้องกับโมเดล 1 พารามิเตอร์ ได้แก่ ค่าความยาก มีค่าระหว่าง -0.994 ถึง 1.226 มีข้อสอบที่สอดคล้องกับโมเดล จำนวน 17 ข้อ และพารามิเตอร์ความสามารถของผู้สอบที่สอดคล้องกับโมเดล มีค่าอยู่ระหว่าง -1.142 ถึง 3.530 ค่าพารามิเตอร์ข้อสอบที่สอดคล้องกับโมเดล 2 พารามิเตอร์ ได้แก่ ค่าความยาก มีค่าระหว่าง -0.954 ถึง 1.916 ค่าอำนาจจำแนก มีค่าระหว่าง 0.192 ถึง 0.713

ได้แก่ ค่าความชาก มีค่าระหว่าง -1.563 ถึง 1.355 มีข้อสอบที่สอดคล้องกับโมเดล จำนวน 11 ข้อ และพารามิเตอร์ความสามารถของผู้สอบที่สอดคล้องกับโมเดล มีค่าอยู่ระหว่าง -1.218 ถึง 3.355 ค่าพารามิเตอร์ข้อสอบที่สอดคล้องกับโมเดล 2 พารามิเตอร์ ได้แก่ ค่าความชาก มีค่าระหว่าง -0.083 ถึง 3.430 ค่าอำนาจจำแนก มีค่าระหว่าง 0.192 ถึง 0.704 จำนวนข้อสอบที่สอดคล้องกับโมเดล จำนวน 14 ข้อ ค่าพารามิเตอร์ความสามารถของผู้สอบที่สอดคล้องกับโมเดล มีค่าระหว่าง -1.627 ถึง 3.862 และค่าพารามิเตอร์ข้อสอบที่สอดคล้องกับโมเดล 3 พารามิเตอร์ ได้แก่ ค่าความชาก มีค่าระหว่าง -2.184 ถึง 3.403 ค่าอำนาจจำแนก มีค่าระหว่าง 0.289 ถึง 1.853 ค่าการเดา มีค่าระหว่าง 0.127 ถึง 0.272 มีข้อสอบที่สอดคล้องกับโมเดล จำนวน 14 ข้อ ค่าพารามิเตอร์ความสามารถของผู้สอบที่สอดคล้องกับโมเดล มีค่าระหว่าง -3.945 ถึง 3.732

16. นำข้อสอบที่สอดคล้องกับโมเดล 1, 2 และ 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ไปทำการวิจัยต่อไป

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูลดังนี้

1. วางแผนในการดำเนินการสอบ โดยการติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียน ของโรงเรียนที่เป็นกลุ่มตัวอย่างในการวิจัย เพื่อกำหนดวัน เวลา ในการดำเนินการสอบ
2. นำแบบทดสอบที่จัดพิมพ์ไว้แล้ว ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัย จำนวน 2,500 คน
3. นำผลการสอบในข้อ 2 มาจัดเตรียมข้อมูลเพื่อวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ ได้แก่ ค่าความชาก (b) ค่าอำนาจจำแนก (a) ค่าการเดา (c) และค่าความสามารถของผู้สอบ (θ) ตามทฤษฎีการตอบสนองข้อสอบแบบโลจิสติกโมเดล โดยใช้โปรแกรม BILOG 3

การวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการวิเคราะห์โดยใช้เครื่องคอมพิวเตอร์ ซึ่งมีรายละเอียดดังนี้

1. ตรวจสอบความเป็นมิติเดียวของแบบทดสอบ (Unidimension) โดยนำข้อมูลที่
ได้จากการตอบมาวิเคราะห์องค์ประกอบ (Factor Analysis) โดยใช้โปรแกรม SPSS/PC
โดยสกัดองค์ประกอบด้วยวิธีวิเคราะห์องค์ประกอบหลัก (Principle Component
Analysis : PC) แล้วหมุนแกนด้วยวิธีแวนแมกซ์ (Varimax) ซึ่งการตรวจสอบคุณสมบัติการ
วัดเพียงมิติเดียวของแบบทดสอบด้วยวิธีการวิเคราะห์องค์ประกอบ จะใช้หลักการพิจารณาจาก
ค่าไอเกน (Eigen Value) ถ้าค่าไอเกนขององค์ประกอบที่ 1 มีค่าสูงกว่าค่าไอเกนของ
องค์ประกอบที่ 2 อย่างมาก และค่าไอเกนขององค์ประกอบที่ 2 มีค่าใกล้เคียงกับองค์
ประกอบที่เหลือเพียงเล็กน้อย ก็แสดงว่าแบบทดสอบนั้นมีคุณสมบัติการวัดเพียงมิติเดียว (Warm.
1978 : 104 ; Lord. 1980 : 21)
2. วิเคราะห์หาคุณภาพของข้อสอบเป็นรายข้อ จากผลการสอบของกลุ่มตัวอย่าง
ที่ใช้ในการวิจัย โดยวิเคราะห์หาค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่าการเดา
(c) ของข้อสอบเป็นรายข้อ ตามทฤษฎีการตอบข้อคำถาม (Item Response Theory)
โดยใช้โลจิสติกโมเดลแบบ 1, 2 และ 3 พารามิเตอร์ วิเคราะห์ด้วยโปรแกรม BILOG 3
3. วิเคราะห์ความสอดคล้องกับโมเดล (Fit) โดยพิจารณาค่าสถิติไคสแควร์
แสดงความสอดคล้อง ระหว่างข้อสอบกับโมเดล (Item Fit) และระหว่างผู้สอบกับโมเดล
(Person Fit)
4. วิเคราะห์ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ
โดยใช้โปรแกรม BILOG 3
5. วิเคราะห์ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความ
สามารถ (Standard Error of Estimate : SEE(θ)) โดยใช้โปรแกรม BILOG 3
6. ทดสอบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารา-
มิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีขนาดต่างกัน ในแต่ละโมเดลโดย
ใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว (One - way ANOVA) ด้วยโปรแกรม SPSS
7. ทดสอบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารา-
มิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบ ระหว่างโมเดล 1, 2 และ 3 พารามิเตอร์ ที่มีกลุ่ม
ตัวอย่างขนาดเดียวกัน โดยใช้การวิเคราะห์ความแปรปรวนทางเดียว (One - way ANOVA)
ด้วยโปรแกรม SPSS
8. ทดสอบความแตกต่างของค่าพารามิเตอร์อำนาจจำแนก ระหว่างโมเดล 2 และ
3 พารามิเตอร์ ภายในกลุ่มตัวอย่างเดียวกัน โดยใช้การทดสอบ t-test ด้วย โปรแกรม SPSS

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการนำเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยจะนำเสนอใน 5 ตอน ดังต่อไปนี้

ตอนที่ 1 ผลการประมาณค่าพารามิเตอร์ผู้สอบและพารามิเตอร์ข้อสอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 4 กลุ่ม โดยแยกพิจารณา ดังนี้

1.1 ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าข้อสอบ แต่ละกลุ่มตัวอย่างและแต่ละแบบการประมาณค่า

1.2 ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถผู้สอบ แต่ละกลุ่มตัวอย่างและแต่ละแบบการประมาณค่า

ตอนที่ 2 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล

ตอนที่ 3 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล

ตอนที่ 4 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดล ในแต่ละกลุ่มตัวอย่าง

ตอนที่ 5 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดล ในแต่ละกลุ่มตัวอย่าง

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ผลการประมาณค่าพารามิเตอร์ผู้สอบและพารามิเตอร์ข้อสอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 4 กลุ่ม คือ กลุ่ม 250 คน 500 คน 750 คน และ 1,000 คน

ผู้วิจัยนำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการวิจัย จำนวน 2,500 คน แล้วนำข้อมูลที่ได้มาสุ่มเข้ากลุ่มตัวอย่าง 4 กลุ่ม คือ 250 คน 500 คน 750 คน และ 1,000 คน โดยวิธีการสุ่มอย่างเป็นระบบ (Systematic Random Sampling) จากนั้นนำมาวิเคราะห์ผลด้วยโปรแกรม BILOG 3 โมเดล 1, 2 และ 3 พารามิเตอร์ กับกลุ่มตัวอย่างทุกกลุ่ม ผลการวิเคราะห์ได้ค่าประมาณพารามิเตอร์ข้อสอบ พารามิเตอร์ผู้สอบ (๑) และค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ซึ่งผู้วิจัยได้นำเสนอตารางเปรียบเทียบค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ก่อน

และหลังจากคัดข้อสอบและผู้สอบที่ไม่สอดคล้องกับโมเดลออก

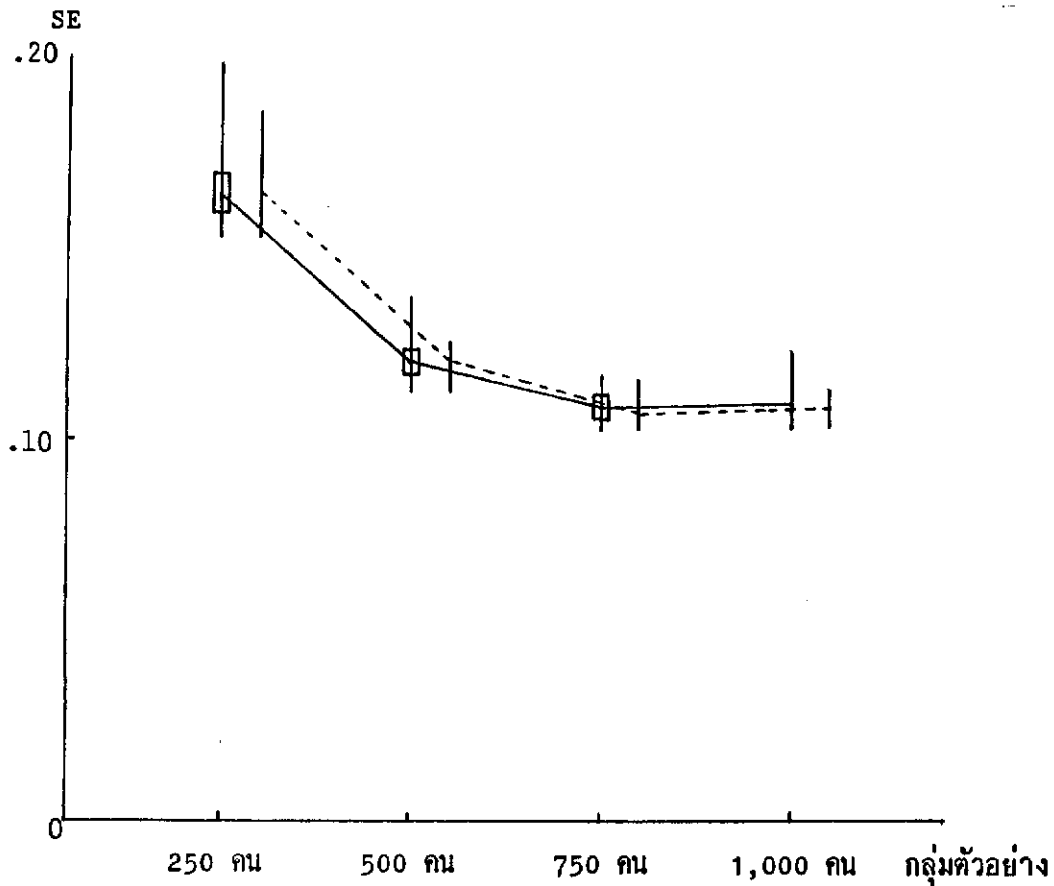
1.1 ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ แต่ละกลุ่มตัวอย่าง ในแต่ละโมเดล ปรากฏดังตาราง 3 ถึงตาราง 8 และ ภาพประกอบ 8 ถึงภาพประกอบ 13 ดังต่อไปนี้

ตาราง 3 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว

กลุ่มตัวอย่าง	ก่อนคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก			คัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว		
	SE _e	\bar{X}	S.D.	SE _e	\bar{X}	S.D.
250 คน	0.152 ถึง 0.198	0.163	0.010	0.152 ถึง 0.175	0.164	0.000
500 คน	0.112 ถึง 0.137	0.120	0.005	0.112 ถึง 0.125	0.120	0.000
750 คน	0.102 ถึง 0.116	0.108	0.006	0.102 ถึง 0.115	0.106	0.000
1,000 คน	0.103 ถึง 0.123	0.109	0.000	0.103 ถึง 0.113	0.108	0.000

จากตาราง 3 ผลการวิเคราะห์แสดงให้เห็นว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ของกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ก่อนคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก มีค่าอยู่ระหว่าง 0.152 ถึง 0.198, 0.112 ถึง 0.137, 0.102 ถึง 0.116 และ 0.103 ถึง 0.123 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.163, 0.120, 0.108, 0.089 ตามลำดับ และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.010, 0.005, 0.006 และ 0.000 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าความยาก จะมีแนวโน้มลดลง และเมื่อพิจารณาผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก หลังจากคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว พบว่ามีค่าอยู่ระหว่าง 0.152 ถึง 0.175, 0.112 ถึง 0.125, 0.102 ถึง 0.115 และ 0.103 ถึง 0.113 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.164, 0.120, 0.106, 0.088 ตามลำดับ และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.000 ทั้ง 4 กลุ่มตัวอย่าง แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าความยาก จะมีแนวโน้มลดลง และมีความเที่ยงตรงในการประมาณค่าพารามิเตอร์มากขึ้น เพื่อให้เห็นภาพได้ชัดเจนยิ่งขึ้น ผู้วิจัยได้นำความคลาดเคลื่อนมาตรฐานในการ

ประมาณค่าพารามิเตอร์ความยาก สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนการคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกและหลังจากคัดข้อสอบที่ไม่สอดคล้องกับโมเดล แสดงดังภาพประกอบ 8



— ก่อนคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก
 --- หลังคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก

ภาพประกอบ 8 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ก่อนคัดข้อสอบที่ไม่สอดคล้องกับโมเดล และหลังจากคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

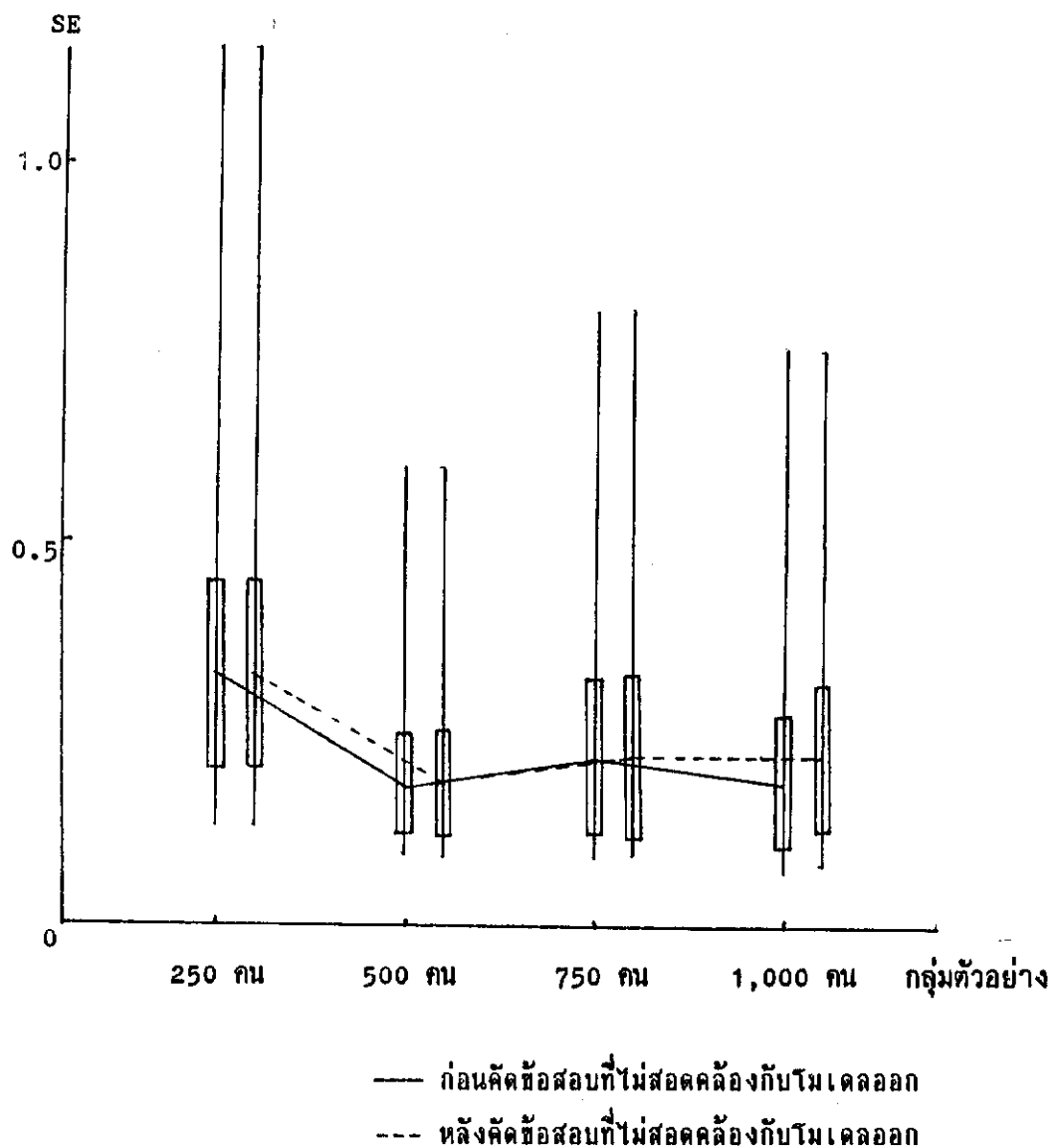
จากภาพประกอบ 8 แสดงว่าสำหรับโอมเคล 1 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบ จะมีค่าลดลง และการกระจายของความคลาดเคลื่อนมีค่าลดลงเช่นเดียวกัน ซึ่งสอดคล้องกับการศึกษาของ เอร์รี่ (Urry. 1974 : 253 - 269)

ตาราง 4 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จาก การประมาณค่าด้วยโอมเคล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคิดข้อสอบที่ไม่สอดคล้องกับโอมเคลออกแล้ว

กลุ่มตัวอย่าง	ก่อนคิดข้อสอบที่ไม่สอดคล้องกับโอมเคลออก						หลังคิดข้อสอบที่ไม่สอดคล้องกับโอมเคลออกแล้ว					
	SE _x	\bar{Y}	S.D.	SE _x	\bar{Y}	S.D.	SE _x	\bar{Y}	S.D.	SE _x	\bar{Y}	S.D.
250 คน	0.129 ถึง 1.154	0.324	0.239	0.058 ถึง 0.167	0.108	0.029	0.129 ถึง 1.154	0.324	0.239	0.058 ถึง 0.167	0.108	0.032
500 คน	0.088 ถึง 0.598	0.181	0.132	0.053 ถึง 0.105	0.081	0.014	0.088 ถึง 0.598	0.182	0.134	0.053 ถึง 0.105	0.081	0.000
750 คน	0.089 ถึง 0.807	0.218	0.192	0.043 ถึง 0.083	0.064	0.011	0.089 ถึง 0.807	0.241	0.210	0.043 ถึง 0.077	0.062	0.000
1,000 คน	0.069 ถึง 0.764	0.182	0.170	0.043 ถึง 0.071	0.057	0.009	0.077 ถึง 0.764	0.217	0.192	0.043 ถึง 0.070	0.053	0.000

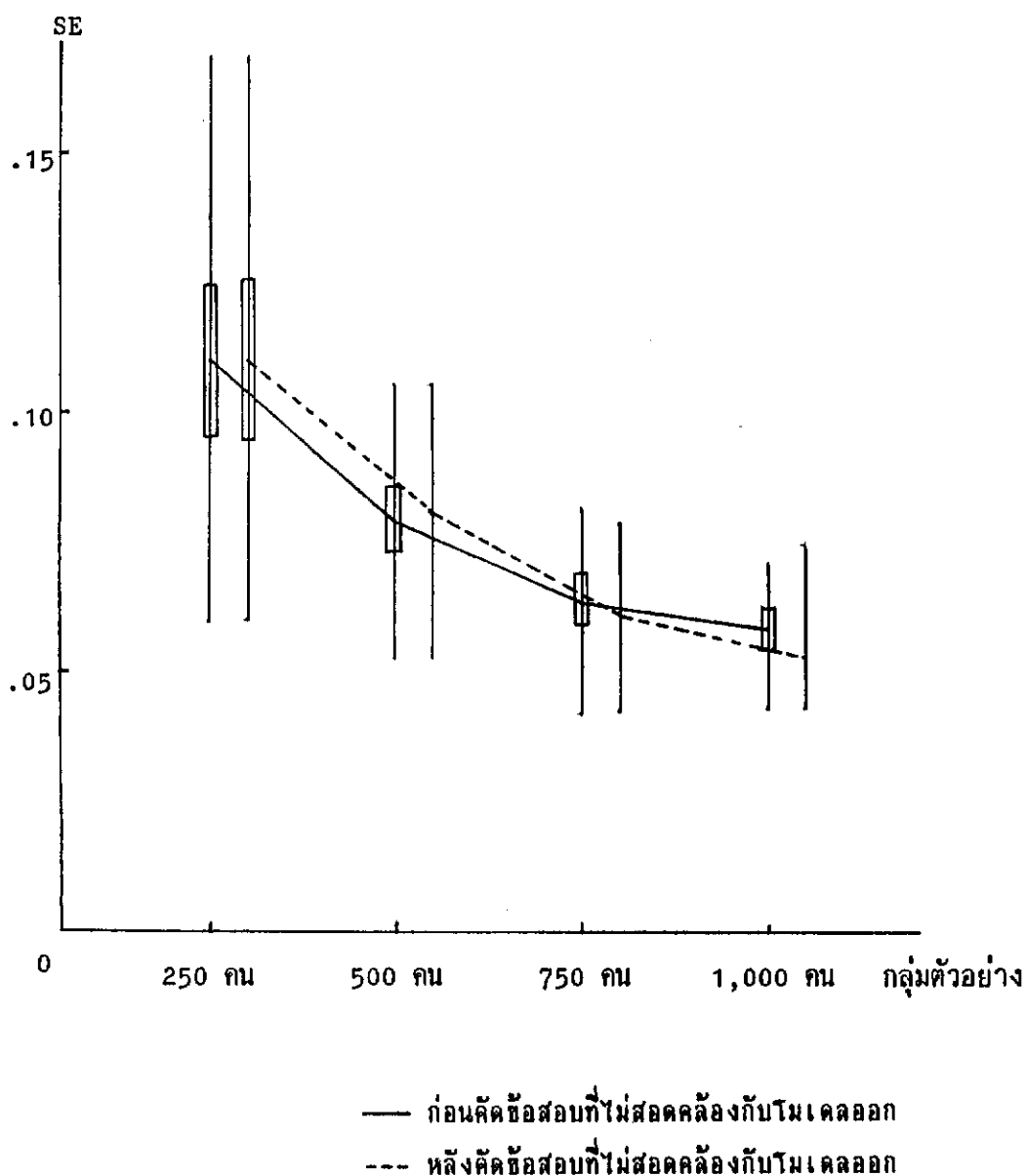
จากตาราง 4 ผลการวิเคราะห์แสดงให้เห็นว่า คลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโอมเคล 2 พารามิเตอร์ ของกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากอยู่ระหว่าง 0.129 ถึง 1.154, 0.088 ถึง 0.598, 0.089 ถึง 0.807 และ 0.069 ถึง 0.764 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.324, 0.181, 0.218, 0.182 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.239, 0.132, 0.192, 0.170 ตามลำดับ ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.058 ถึง 0.167, 0.053 ถึง 0.105, 0.043 ถึง 0.083 และ 0.043 ถึง 0.071 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.108, 0.081, 0.064, 0.057 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.029, 0.014, 0.011, 0.170 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบ จะมีแนวโน้มลดลง เมื่อพิจารณาผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์หลังคิดข้อสอบที่ไม่สอดคล้องกับโอมเคลออกแล้วพบว่า ความคลาดเคลื่อนมาตรฐานในการ

ประมาณค่าความชาก มีค่าอยู่ระหว่าง 0.129 ถึง 1.154, 0.088 ถึง 0.598, 0.089 ถึง 0.807 และ 0.077 ถึง 0.764 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.324, 0.182, 0.241, 0.217 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.239, 0.134, 0.210, 0.192 ตามลำดับ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.058 ถึง 0.167, 0.053 ถึง 0.105, 0.043 ถึง 0.077 และ 0.043 ถึง 0.070 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.108, 0.081, 0.062, 0.053 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.032, 0.000, 0.000, 0.000 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ข้อสอบ จะมีแนวโน้มลดลง เพื่อให้เห็นภาพ ได้ชัดเจนยิ่งขึ้น ผู้วิจัยได้นำความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนการคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก และหลังจากคัดข้อสอบที่ไม่สอดคล้องกับโมเดล แสดงดังภาพประกอบ 9 และ ภาพประกอบ 10



ภาพประกอบ 9 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้องกับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

จากภาพประกอบ 9 แสดงว่าสำหรับโมเดล 2 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าความยาก จะมีค่าลดลง และการกระจายของความคลาดเคลื่อนมีค่าลดลงเช่นเดียวกัน ซึ่งสอดคล้องกับการศึกษาของ เอลรี่ (Urry, 1974 : 253 - 269)



ภาพประกอบ 10 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกของข้อสอบ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้องกับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

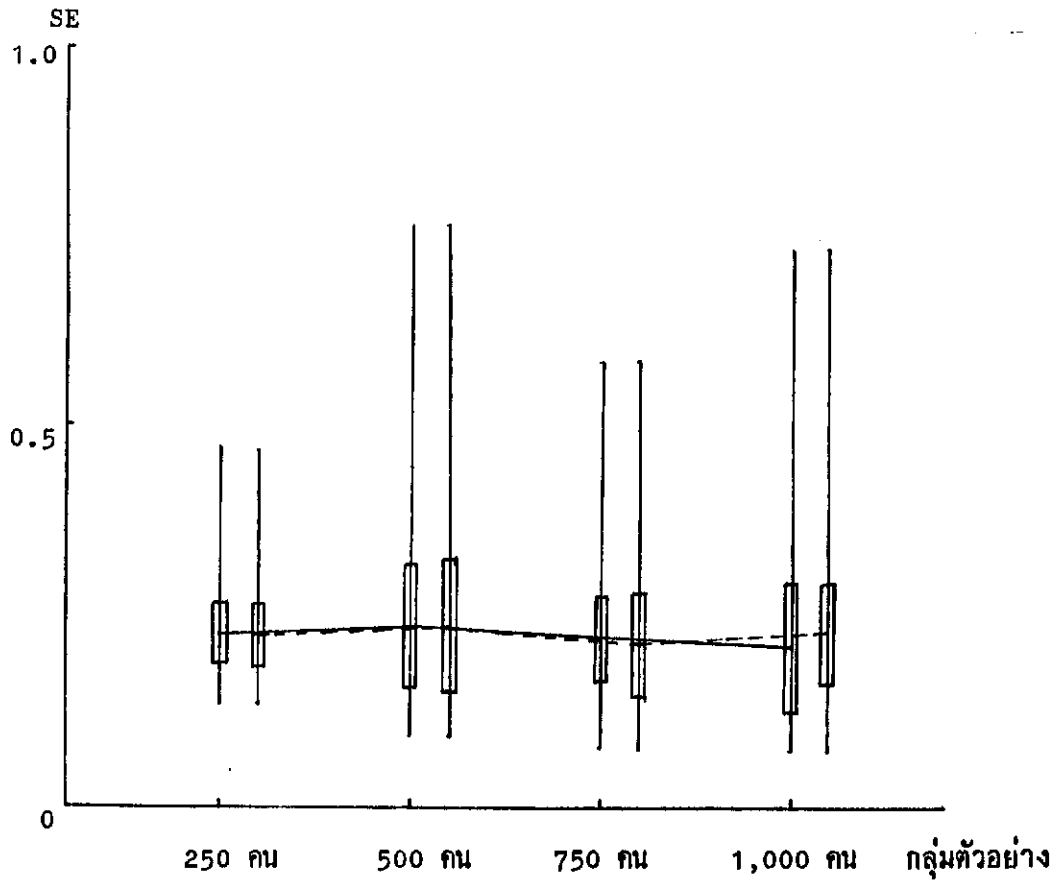
จากภาพประกอบ 10 แสดงว่าสำหรับโมเดล 2 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าอำนาจจำแนก จะมีค่าลดลง และการกระจายของความคลาดเคลื่อนมีค่าลดลงเช่นเดียวกัน ซึ่งสอดคล้องกับการศึกษาของ เออร์รี่ (Urry, 1974 : 253 - 269)

ตาราง 5 ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนและหลังคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว

กลุ่มตัวอย่าง	ก่อนคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก									
	$SE_{\bar{X}}$	\bar{X}	S.D.	$SE_{\bar{X}}$	\bar{X}	S.D.	$SE_{\bar{X}}$	\bar{X}	S.D.	
250 คน	0.133 ถึง 0.468	0.224	0.079	0.108 ถึง 1.366	0.441	0.287	0.032 ถึง 0.098	0.058	0.021	
500 คน	0.089 ถึง 0.764	0.234	0.160	0.084 ถึง 0.493	0.263	0.122	0.026 ถึง 0.095	0.053	0.020	
750 คน	0.074 ถึง 0.584	0.220	0.113	0.074 ถึง 0.546	0.254	0.122	0.020 ถึง 0.099	0.054	0.025	
1,000 คน	0.072 ถึง 0.735	0.209	0.170	0.060 ถึง 0.448	0.210	0.095	0.021 ถึง 0.098	0.050	0.022	
กลุ่มตัวอย่าง	คัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว									
	$SE_{\bar{X}}$	\bar{X}	S.D.	$SE_{\bar{X}}$	\bar{X}	S.D.	$SE_{\bar{X}}$	\bar{X}	S.D.	
250 คน	0.133 ถึง 0.468	0.225	0.077	0.108 ถึง 1.366	0.452	0.290	0.032 ถึง 0.098	0.057	0.000	
500 คน	0.089 ถึง 0.764	0.236	0.167	0.084 ถึง 0.493	0.275	0.122	0.026 ถึง 0.095	0.051	0.000	
750 คน	0.074 ถึง 0.584	0.215	0.134	0.074 ถึง 0.546	0.309	0.118	0.020 ถึง 0.099	0.043	0.000	
1,000 คน	0.072 ถึง 0.735	0.229	0.128	0.060 ถึง 0.448	0.331	0.095	0.021 ถึง 0.098	0.043	0.000	

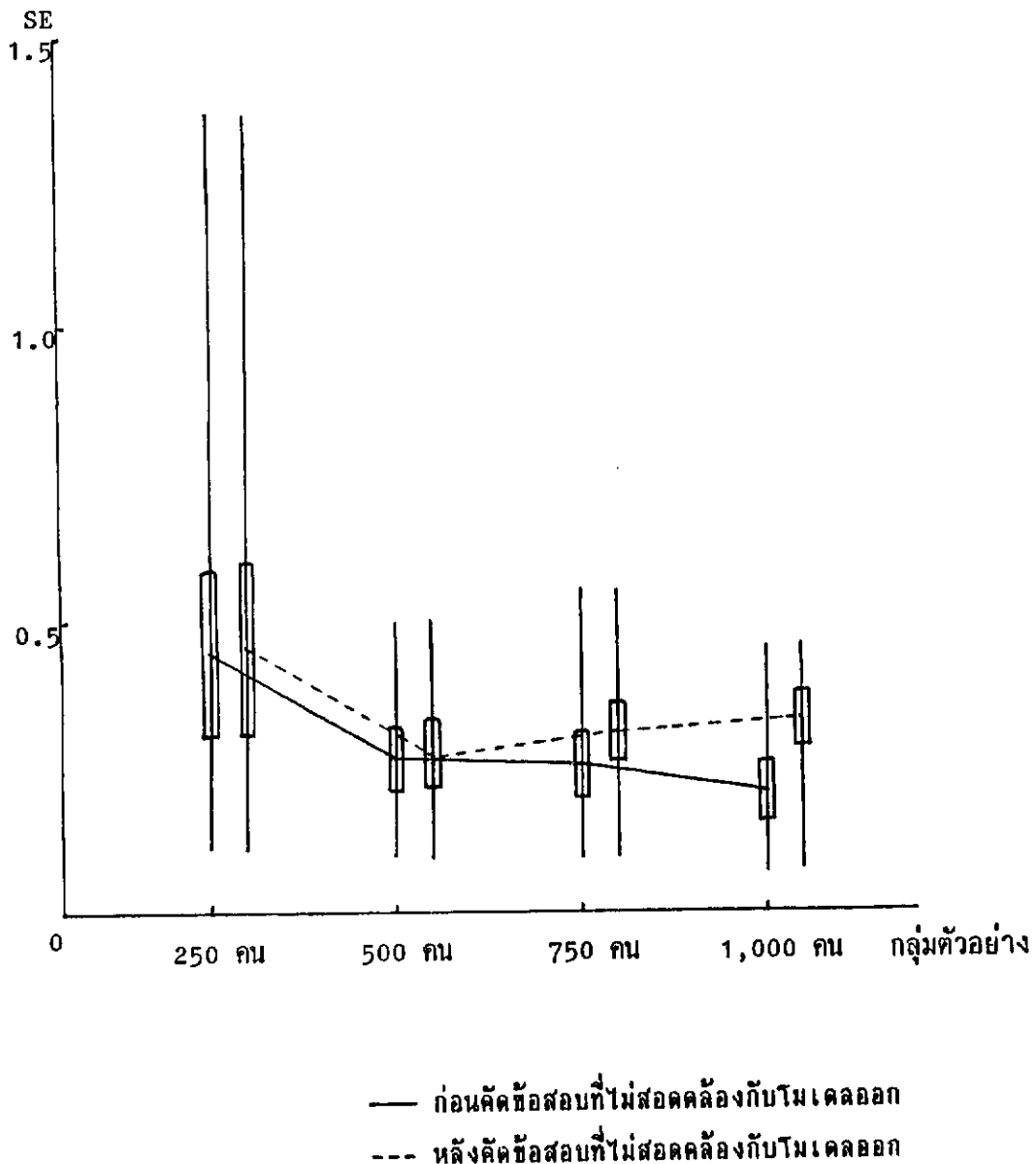
จากตาราง 5 ผลการวิเคราะห์แสดงให้เห็นว่า ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าอยู่ระหว่าง 0.133 ถึง 0.468, 0.089 ถึง 0.764, 0.074 ถึง 0.584 และ 0.072 ถึง 0.735 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.224, 0.234, 0.220, 0.209 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.079, 0.160, 0.113, 0.170 ตามลำดับ ความคลาดเคลื่อนมาตรฐาน

ฐานในการประมาณค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.108 ถึง 1.366, 0.084 ถึง 0.493, 0.074 ถึง 0.546 และ 0.060 ถึง 0.448 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.441, 0.263, 0.254, 0.210 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.287, 0.122, 0.122, 0.095 ตามลำดับ และความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา มีค่าอยู่ระหว่าง 0.032 ถึง 0.098, 0.026 ถึง 0.095, 0.020 ถึง 0.099 และ 0.021 ถึง 0.098 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.058, 0.053, 0.054, 0.050 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.021, 0.020, 0.025, 0.022 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์จะมีแนวโน้มลดลง และเมื่อพิจารณาผลการวิเคราะห์หลังคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าอยู่ระหว่าง 0.133 ถึง 0.468, 0.089 ถึง 0.764, 0.074 ถึง 0.584 และ 0.072 ถึง 0.735 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.225, 0.236, 0.215, 0.229 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.077, 0.167, 0.134, 0.128 ตามลำดับ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าอยู่ระหว่าง 0.108 ถึง 1.366, 0.084 ถึง 0.493, 0.074 ถึง 0.546 และ 0.060 ถึง 0.448 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.452, 0.275, 0.309, 0.331 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 0.290, 0.122, 0.118, 0.095 ตามลำดับ และความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา มีค่าอยู่ระหว่าง 0.032 ถึง 0.098, 0.026 ถึง 0.095, 0.020 ถึง 0.099 และ 0.021 ถึง 0.098 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.057, 0.051, 0.043, 0.043 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.000 ถึง 4 กลุ่มตัวอย่าง แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ ส่วนใหญ่จะมีแนวโน้มลดลง และการกระจายมีค่าลดลงเช่นเดียวกัน เพื่อให้เห็นภาพได้ชัดเจนยิ่งขึ้น ผู้วิจัยได้นำความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ก่อนการคัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก และหลังจากคัดข้อสอบที่ไม่สอดคล้องกับโมเดล แสดงดังภาพประกอบ 11 ถึง ภาพประกอบ 13



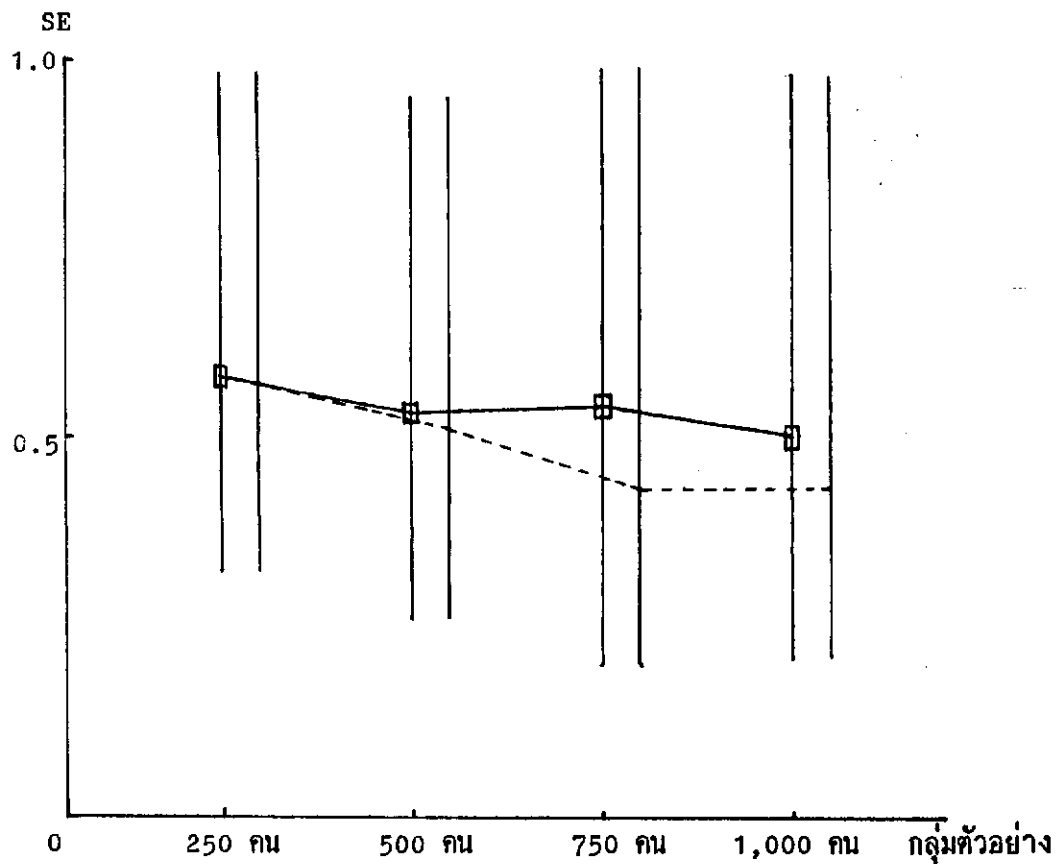
ภาพประกอบ 11 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนฝึกซ้อมที่ไม่สอดคล้องกับโมเดล และหลังจากฝึกซ้อมที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

จากภาพประกอบ 11 แสดงว่าสำหรับโมเดล 3 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าความยาก จะมีค่าเพิ่มขึ้น และการกระจายของความคลาดเคลื่อนมีค่าเพิ่มขึ้นเช่นเดียวกัน แต่เพิ่มขึ้นอย่างไม่ชัดเจนนัก



ภาพประกอบ 12 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกของข้อสอบจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนคิดข้อสอบที่ไม่สอดคล้องกับโมเดล และหลังจากคิดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

จากภาพประกอบ 12 แสดงว่าสำหรับโมเดล 3 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าอำนาจจำแนก จะมีค่าลดลง และการกระจายของความคลาดเคลื่อนมีค่าลดลงเช่นเดียวกัน ซึ่งสอดคล้องกับการศึกษาของ เฮอร์รี่ (Urry, 1974 : 253 - 269)



— ก่อนตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก
 --- หลังตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออก

ภาพประกอบ 13 แสดงความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดาของข้อสอบจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนตัดข้อสอบที่ไม่สอดคล้องกับโมเดล และหลังจากตัดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว ของกลุ่มตัวอย่าง 4 กลุ่ม

จากภาพประกอบ 13 แสดงว่าสำหรับโมเดล 3 พารามิเตอร์ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าการเดา จะมีค่าลดลง และการกระจายของความคลาดเคลื่อนมีค่าลดลงเช่นเดียวกัน ซึ่งสอดคล้องกับการศึกษาของ เอร์รี่ (Urry, 1974 : 253 - 269)

1.2 ค่าสถิติพื้นฐานของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ แต่ละกลุ่มตัวอย่าง ในแต่ละโมเดล ปรากฏผลดังตาราง 6 ถึง ตาราง 8 ดังต่อไปนี้

ตาราง 6 ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ก่อนและหลังจากตัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว

กลุ่มตัวอย่าง	ก่อนตัดผู้สอบที่ไม่สอดคล้องกับโมเดลออก			ตัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว		
	SE	\bar{X}	S.D.	SE	\bar{X}	S.D.
250 คน	0.805 ถึง 999.000	12.896	108.893	0.805 ถึง 1.779	0.896	0.161
500 คน	0.814 ถึง 999.000	6.940	77.154	0.855 ถึง 1.898	0.951	0.148
750 คน	0.938 ถึง 999.000	11.710	102.584	1.004 ถึง 1.628	1.094	0.100
1,000 คน	0.881 ถึง 999.000	14.977	117.314	0.948 ถึง 1.539	1.036	0.105

จากตาราง 6 ผลการวิเคราะห์แสดงให้เห็นว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (SE_0) ของกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ก่อนตัดผู้สอบที่ไม่สอดคล้องกับโมเดลออก มีค่าอยู่ระหว่าง 0.805 ถึง 999.000, 0.814 ถึง 999.000, 0.938 ถึง 999.000 และ 0.881 ถึง 999.000 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 12.896, 6.940, 11.710 และ 14.977 ตามลำดับ และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 108.893, 77.154, 102.584 และ 117.314 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างใหญ่ขึ้น ความคลาดเคลื่อนเฉลี่ยส่วนใหญ่จะมีค่าสูงขึ้น และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบหลังตัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว มีค่าอยู่ระหว่าง 0.805 ถึง 1.779, 0.855 ถึง 1.898, 1.004 ถึง 1.628, 0.948 ถึง 1.539 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 0.896, 0.951, 1.094 และ 1.036 ตามลำดับ ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.161, 0.148, 0.100 และ 0.105 ตามลำดับ แสดงว่าเมื่อกลุ่ม

ตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าเฉลี่ยจะสูงขึ้น แต่การกระจายของความคลาดเคลื่อนจะลดลง

ตาราง 7 ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วยโอมเคล 2 พารามิเตอร์ ก่อนและหลังจากคัดผู้สอบที่ไม่สอดคล้องกับโอมเคลออกแล้ว

กลุ่มตัวอย่าง	ก่อนคัดผู้สอบที่ไม่สอดคล้องกับโอมเคลออก			คัดผู้สอบที่ไม่สอดคล้องกับโอมเคลออกแล้ว		
	SE	\bar{X}	S.D.	SE	\bar{X}	S.D.
250 คน	0.747 ถึง 999.000	16.852	125.490	0.926 ถึง 1.907	1.017	0.145
500 คน	0.755 ถึง 999.000	12.883	108.787	0.807 ถึง 2.018	0.914	0.176
750 คน	0.886 ถึง 999.000	12.992	108.738	0.970 ถึง 1.913	1.059	0.118
1,000 คน	0.829 ถึง 999.000	16.931	125.292	0.939 ถึง 1.970	1.029	0.130

จากตาราง 7 ผลการวิเคราะห์ที่แสดงให้เห็นว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (SE_0) ของกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลมาจากการประมาณค่าด้วยโอมเคล 1 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ก่อนคัดผู้สอบที่ไม่สอดคล้องกับโอมเคลออก มีค่าอยู่ระหว่าง 0.747 ถึง 999.000, 0.755 ถึง 999.000, 0.886 ถึง 999.000 และ 0.829 ถึง 999.000 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 16.852, 12.883, 12.992 และ 16.931 ตามลำดับ และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 125.490, 108.787, 108.738 และ 125.292 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างใหญ่ขึ้น ความคลาดเคลื่อนเฉลี่ยส่วนใหญ่จะมีค่าสูงขึ้น และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบหลังคัดผู้สอบที่ไม่สอดคล้องกับโอมเคลออกแล้ว มีค่าอยู่ระหว่าง 0.926 ถึง 1.907, 0.807 ถึง 2.018, 0.970 ถึง 1.913, 0.939 ถึง 1.970 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 1.017, 0.914, 1.059 และ 1.029 ตามลำดับ ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.145, 0.176, 0.118 และ 0.130 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาด

ใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าเฉลี่ยจะสูงขึ้น แต่การกระจายของความคลาดเคลื่อนจะลดลง

ตาราง 8 ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ที่เป็นผลจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ก่อนและหลังจากคัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว

กลุ่มตัวอย่าง	ก่อนคัดผู้สอบที่ไม่สอดคล้องกับโมเดลออก			คัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว		
	SE	\bar{X}	S.D.	SE	\bar{X}	S.D.
250 คน	0.456 ถึง 999.000	77.695	264.805	0.479 ถึง 7.830	1.252	1.222
500 คน	0.455 ถึง 999.000	59.320	233.415	0.505 ถึง 7.091	1.354	1.179
750 คน	0.513 ถึง 999.000	43.178	198.642	0.563 ถึง 9.645	1.800	1.905
1,000 คน	0.518 ถึง 999.000	50.367	215.450	0.621 ถึง 6.303	1.916	1.432

จากตาราง 8 ผลการวิเคราะห์แสดงให้เห็นว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ (SE_u) ของกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ก่อนคัดผู้สอบที่ไม่สอดคล้องกับโมเดลออก มีค่าอยู่ระหว่าง 0.456 ถึง 999.000, 0.455 ถึง 999.000, 0.513 ถึง 999.000 และ 0.518 ถึง 999.000 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 77.695, 59.320, 43.178 และ 50.367 ตามลำดับ และมีค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 264.805, 233.415, 198.642 และ 215.450 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างใหญ่ขึ้น ความคลาดเคลื่อนเฉลี่ยส่วนใหญ่จะมีค่าลดลง และมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบหลังคัดผู้สอบที่ไม่สอดคล้องกับโมเดลออกแล้ว มีค่าอยู่ระหว่าง 0.479 ถึง 7.830, 0.505 ถึง 7.091, 0.563 ถึง 9.645, 0.621 ถึง 6.303 ตามลำดับ มีค่าเฉลี่ยเท่ากับ 1.252, 1.354, 1.800 และ 1.916 ตามลำดับ ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.222, 1.179, 1.905 และ 1.432 ตามลำดับ แสดงว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนในการประมาณค่าเฉลี่ยจะสูงขึ้น และการกระจายของความคลาดเคลื่อนมีค่าสูงขึ้นเช่นเดียวกัน

ตอนที่ 2 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มผู้สอบ 250 คน 500 คน 750 คน และ 1,000 คน ที่เป็นผลมาจากการประมาณค่าพารามิเตอร์ด้วยโมเดล 1, 2 และ 3 พารามิเตอร์

ในการวิเคราะห์ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบนั้น ผู้วิจัยได้คัดเลือกข้อสอบที่ไม่สอดคล้องกับโมเดลออก และทำการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่า โดยนำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เปรียบเทียบระหว่างกลุ่มผู้สอบ 250, 500, 750 และ 1,000 คน ซึ่งแต่ละกลุ่มได้ตัดผู้ที่ไม่สอดคล้องกับโมเดลออกแล้ว ในแต่ละโมเดลโดยการวิเคราะห์ความแปรปรวน ปรากฏผลดังต่อไปนี้

2.1 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากของข้อสอบที่ประมาณค่าพารามิเตอร์ด้วยโมเดล 1 พารามิเตอร์ เมื่อคัดข้อที่ไม่สอดคล้องกับโมเดลออกแล้ว ได้ข้อสอบเฉพาะที่สอดคล้องกับโมเดล จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ดังตาราง 9

ตาราง 9 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.042	.014	850.314*
ภายในกลุ่ม	52	.001	.000	
รวมทั้งหมด	55	0.028		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 9 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ จากกลุ่มตัวอย่าง

250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเชฟเฟ่ ดังแสดงในตาราง 9

ตาราง 10 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	.164	.120	.106	.108
250 คน	-	.044*	.058*	.056*
500 คน		-	.014*	.012*
750 คน			-	.002*
1,000 คน				-

จากตาราง 10 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500, 750, 1,000 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750, 1,000 คน และกลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่

2.2 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและค่าอำนาจจำแนก ของข้อสอบที่ประมาณค่าพารามิเตอร์ด้วยโมเดล 2 พารามิเตอร์ เมื่อคัดข้อที่ไม่สอดคล้องกับโมเดลออกแล้ว ได้ข้อสอบเฉพาะที่สอดคล้องกับโมเดล จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ดังตาราง 11 ถึงตาราง 13

ตาราง 11 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.212	.071	1.784
ภายในกลุ่ม	65	2.562	.039	
รวมทั้งหมด	68	2.774		

จากตาราง 11 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 12 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.032	.011	31.071*
ภายในกลุ่ม	65	.022	.000	
รวมทั้งหมด	68	.054		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 12 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮาก ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเซฟเฟ่ ดังแสดงในตาราง 13

ตาราง 13 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	.108	.081	.062	.053
250 คน	-	.027*	.046*	.055*
500 คน		-	.019	.028*
750 คน			-	.009
1,000 คน				-

จากตาราง 13 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500, 750, 1,000 คน และกลุ่มตัวอย่าง 500 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มตัวอย่าง 750 คน กับกลุ่ม 500 คน และ 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2.3 เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮาก ค่าอำนาจจำแนก และค่าการเดา ของข้อสอบที่ประมาณค่าพารามิเตอร์ด้วยโอมเดล 3 พารามิเตอร์ เมื่อคิดข้อที่ไม่สอดคล้องกับโอมเดลออกแล้ว ได้ข้อสอบเฉพาะที่สอดคล้องกับโอมเดล จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน ดังตาราง 14 ถึงตาราง 17

ตาราง 14 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชุก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.003	.001	.052
ภายในกลุ่ม	60	1.337	.022	
รวมทั้งหมด	63	1.340		

จากตาราง 14 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชุก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 15 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.504	.168	4.850*
ภายในกลุ่ม	60	2.079	.035	
รวมทั้งหมด	63	2.583		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 15 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชุก ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเชฟเฟ่ ดังแสดงในตาราง 16

ตาราง 16 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	.452	.275	.309	.331
250 คน	-	.177*	.143	.121*
500 คน		-	.034	.056
750 คน			-	.022
1,000 คน				-

จากตาราง 16 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500 คน และกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกนั้นมีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 17 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	.002	.001	1.410
ภายในกลุ่ม	60	.026	.000	
รวมทั้งหมด	63	.028		

จากตาราง 17 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ตอนที่ 3 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ (๑) ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ที่ประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์

ในการวิเคราะห์ผลการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ผู้วิจัยได้คัดเลือกผู้สอบที่ไม่สอดคล้องกับโมเดลออก และทำการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มตัวอย่าง ในแต่ละโมเดล โดยนำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เปรียบเทียบระหว่างกลุ่มผู้สอบ 250, 500, 750 และ 1,000 คน ในแต่ละโมเดล โดยการวิเคราะห์ความแปรปรวน ปรากฏผลดังตาราง 18 ถึงตาราง 23

ตาราง 18 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	9.299	3.100	217.630*
ภายในกลุ่ม	2196	31.278	.014	
รวมทั้งหมด	2199	40.577		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 18 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ จาก

กลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเซฟเฟ่ ดังแสดงในตาราง 19

ตาราง 19 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	.896	.951	1.094	1.036
250 คน	-	.055*	.198*	.140*
500 คน		-	.143*	.085*
750 คน			-	.058*
1,000 คน				-

จากตาราง 19 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500, 750, 1,000 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750, 1,000 คน และกลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่

ตาราง 20 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	9.821	1.940	98.990*
ภายในกลุ่ม	2164	32.416	.020	
รวมทั้งหมด	2167	48.237		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 20 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเซฟเฟ่ ดังแสดงในตาราง 21

ตาราง 21 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	1.017	.914	1.059	1.029
250 คน	-	.103*	.042*	.012
500 คน		-	.145*	.115*
750 คน			-	.030*
1,000 คน				-

จากตาราง 21 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500, 750 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750, 1,000 คน และกลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มตัวอย่าง 250 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 22 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	3	125.160	41.720	17.612*
ภายในกลุ่ม	1848	4377.684	2.369	
รวมทั้งหมด	1851	4502.844		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 22 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 250, 500, 750 และ 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่มีขนาดกลุ่มตัวอย่างต่างกันอย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์เป็นรายคู่ด้วยค่าสถิติการทดสอบเชฟเฟ่ ดังแสดงในตาราง 23

ตาราง 23 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ เป็นรายคู่

กลุ่มตัวอย่าง	250 คน	500 คน	750 คน	1,000 คน
	1.252	1.354	1.800	1.916
250 คน	-	.102	.548*	.664*
500 คน		-	.446*	.562*
750 คน			-	.116
1,000 คน				-

จากตาราง 23 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 750, 1,000 คน และกลุ่มตัวอย่าง 500 คน กับกลุ่ม 750, 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตอนที่ 4 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างค่าที่ประมาณจากโมเดล 1, 2 และ 3 พารามิเตอร์ ในกลุ่มผู้สอบที่มีจำนวนเท่ากัน

ในการวิเคราะห์ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบนั้น ผู้วิจัยได้นำข้อสอบที่สอดคล้องกับโมเดลมาทำการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่า โดยนำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เปรียบเทียบระหว่างโมเดล จากการวิเคราะห์ภายในแต่ละกลุ่มผู้สอบ 250, 500, 750 และ 1,000 คน ซึ่งแต่ละกลุ่มได้คัดเลือกผู้ที่ไม่สอดคล้องกับโมเดลออกแล้ว ปรากฏผลดังต่อไปนี้

4.1 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างค่าที่ประมาณด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ เมื่อวิเคราะห์จากกลุ่มตัวอย่าง 250 คน

ผู้วิจัยนำข้อสอบที่สอดคล้องกับโอมเดล มาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความยาก ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ โดยการวิเคราะห์ความแปรปรวน และเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์อำนาจจำแนก โดยการทดสอบ t-test ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลดังตาราง 24 ถึงตาราง 26

ตาราง 24 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ระหว่างค่าที่ประมาณด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	.252	.126	5.532*
ภายในกลุ่ม	53	1.205	.023	
รวมทั้งหมด	55	1.457		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 24 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากที่เป็นผลจากการประมาณค่าด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลจากการประมาณค่าด้วยโอมเดลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติการทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 25

ตาราง 25 การเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮาก ระหว่างค่าที่ประมาณด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	.164	.324	.225
1 พารามิเตอร์	-	.160 [*]	.061
2 พารามิเตอร์		-	.099
3 พารามิเตอร์			-

จากตาราง 25 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 พารามิเตอร์ และโอมเดล 2 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความฮาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ กับโอมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 26 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ระหว่างค่าที่ประมาณด้วยโอมเดล 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250 คน

แบบประมาณค่า	\bar{x}	s	n	t
2 พารามิเตอร์	.108	.029	20	-5.15 [*]
3 พารามิเตอร์	.452	.290	19	

จากตาราง 26 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 250 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.2 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ข้อสอบระหว่างค่าที่ประมาณด้วยโสมเดล 1, 2 และ 3 พารามิเตอร์ เมื่อวิเคราะห์จากกลุ่ม ตัวอย่าง 500 คน

ผู้วิจัยนำข้อสอบที่สอดคล้องกับโสมเดล มาเปรียบเทียบความคลาดเคลื่อน มาตรฐานในการประมาณค่าพารามิเตอร์ความยาก ระหว่างค่าที่เป็นผลมาจากการประมาณค่า ด้วยโสมเดล 1, 2 และ 3 พารามิเตอร์ โดยการวิเคราะห์ความแปรปรวน และเปรียบเทียบ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์อำนาจจำแนก โดยการทดสอบ t-test ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโสมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลดังตาราง 27 ถึงตาราง 29

ตาราง 27 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความยาก ระหว่างค่าที่ประมาณด้วยโสมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 500 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	.116	.058	3.550*
ภายในกลุ่ม	50	.814	.016	
รวมทั้งหมด	52	.930		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 27 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความยากที่เป็นผลจากการประมาณค่าด้วยโสมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่ม ตัวอย่าง 500 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ย ของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลจากการประมาณค่า ด้วยโสมเดลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติ การทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 28

ตาราง 28 การเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ระหว่างค่าที่ประมาณด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	.120	.182	.236
1 พารามิเตอร์	-	.062	.116*
2 พารามิเตอร์		-	.054
3 พารามิเตอร์			-

จากตาราง 28 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 พารามิเตอร์ และโอมเดล 3 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 พารามิเตอร์ กับโอมเดล 1 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตาราง 29 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ระหว่างค่าที่ประมาณด้วยโอมเดล 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 500 คน

แบบประมาณค่า	\bar{x}	s	n	t
2 พารามิเตอร์	.081	.014	19	-6.69*
3 พารามิเตอร์	.275	.122	18	

จากตาราง 29 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 500 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.3 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ข้อสอบระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เมื่อวิเคราะห์จากกลุ่ม ตัวอย่าง 750 คน

ผู้วิจัยนำข้อสอบที่สอดคล้องกับโมเดล มาเปรียบเทียบความคลาดเคลื่อน มาตรฐานในการประมาณค่าพารามิเตอร์ความยาก ระหว่างค่าที่เป็นผลมาจากการประมาณค่า ด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ โดยการวิเคราะห์ความแปรปรวน และเปรียบเทียบ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์อำนาจจำแนก โดยการทดสอบ t-test ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลดังตาราง 30 ถึงตาราง 31

ตาราง 30 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความยาก ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 750 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	.134	.067	2.907
ภายในกลุ่ม	38	.878	.023	
รวมทั้งหมด	40	1.012		

จากตาราง 30 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่า ความยาก ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ในกลุ่ม ตัวอย่าง 750 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 31 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ระหว่างค่าที่ประมาณด้วยโมเดล 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 750 คน

แบบประมาณค่า	\bar{x}	s	n	t
2 พารามิเตอร์	.062	.010	16	-7.66*
3 พารามิเตอร์	.309	.116	13	

จากตาราง 31 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 750 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

4.4 เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ข้อสอบระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เมื่อวิเคราะห์จากกลุ่มตัวอย่าง 1,000 คน

ผู้วิจัยนำข้อสอบที่สอดคล้องกับโมเดล มาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความยาก ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ โดยการวิเคราะห์ความแปรปรวน และเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์อำนาจจำแนก โดยการทดสอบ t-test ระหว่างค่าที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลดังตาราง 32 ถึงตาราง 33

ตาราง 32 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ระหว่างค่าที่ประมาณด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	.146	.073	2.617
ภายในกลุ่ม	36	1.004	.028	
รวมทั้งหมด	38	1.450		

จากตาราง 32 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ในกลุ่มตัวอย่าง 1,000 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .05

ตาราง 33 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ระหว่างค่าที่ประมาณด้วยโมเดล 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 1,000 คน

แบบประมาณค่า	\bar{x}	s	n	t
2 พารามิเตอร์	.053	.008	14	-5.86 ^{**}
3 พารามิเตอร์	.222	.108	14	--

จากตาราง 33 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ จากกลุ่มตัวอย่าง 1,000 คน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตอนที่ 5 ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ (๑) ระหว่างค่าที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ในกลุ่มตัวอย่างที่มีขนาดเท่ากัน

ในการวิเคราะห์ผลการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ผู้วิจัยได้ทำการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดล ในแต่ละกลุ่มตัวอย่างที่เท่ากัน โดยนำค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เปรียบเทียบระหว่างโมเดลในกลุ่มผู้สอบ 250, 500, 750 และ 1,000 คน ซึ่งแต่ละกลุ่มได้คัดเลือกผู้ที่ไม่สอดคล้องกับโมเดลออกแล้ว ปรากฏผลดังตาราง 34 ถึงตาราง 41

ตาราง 34 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 250 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	13.897	6.948	13.896*
ภายในกลุ่ม	636	318.004	.500	
รวมทั้งหมด	638	331.900		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 34 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ของกลุ่มตัวอย่าง 250 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติการทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 35

ตาราง 35 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	.896	1.017	1.252
1 พารามิเตอร์	-	.121	.356*
2 พารามิเตอร์		-	.235*
3 พารามิเตอร์			-

จากตาราง 35 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ กับโอมเดล 1 และโอมเดล 2 พารามิเตอร์ มีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีความนัยสำคัญทางสถิติ

ตาราง 36 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถผู้สอบ ระหว่างค่าที่ประมาณค่าด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 500 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	50.340	25.170	53.603*
ภายในกลุ่ม	1283	602.451	.470	
รวมทั้งหมด	1285	652.791		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 36 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ของกลุ่มตัวอย่าง 500 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติการทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 37

ตาราง 37 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	.951	.914	1.354
1 พารามิเตอร์	-	.037	.403*
2 พารามิเตอร์		-	.440*
3 พารามิเตอร์			-

จากตาราง 37 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ กับโมเดล 1 และโมเดล 2 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

ตาราง 38 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างค่าที่ประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 750 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	222.045	111.023	95.126*
ภายในกลุ่ม	1942	2266.537	1.167	
รวมทั้งหมด	1944	2488.582		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 38 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ของกลุ่มตัวอย่าง 750 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติการทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 39

ตาราง 39 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	1.094	1.059	1.800
1 พารามิเตอร์	-	.035	.706*
2 พารามิเตอร์		-	.741*
3 พารามิเตอร์			-

จากตาราง 39 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเคล 3 พารามิเตอร์ กับ โอมเคล 1 และโอมเคล 2 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเคล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

ตาราง 40 การเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ ระหว่างค่าที่ประมาณค่าด้วยโอมเคล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 1,000 คน

แหล่งความแปรปรวน	df	SS	MS	F
ระหว่างกลุ่ม	2	350.906	175.453	325.682*
ภายในกลุ่ม	2347	1264.387	.539	
รวมทั้งหมด	2349	1615.293		

* มีนัยสำคัญที่ระดับ .05

จากตาราง 40 แสดงว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโอมเคล 1, 2 และ 3 พารามิเตอร์ ของกลุ่มตัวอย่าง 1,000 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่ามีค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโอมเคลต่างกัน อย่างน้อย 1 คู่ มีค่าต่างกัน ซึ่งทำการวิเคราะห์แบบรายคู่ด้วยค่าสถิติการทดสอบแบบเซฟเฟ่ ดังแสดงในตาราง 41

ตาราง 41 การเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1, 2 และ 3 พารามิเตอร์ เป็นรายคู่

แบบการประมาณค่า	1 พารามิเตอร์	2 พารามิเตอร์	3 พารามิเตอร์
	1.036	1.029	1.916
1 พารามิเตอร์	-	.007	.880*
2 พารามิเตอร์		-	.887*
3 พารามิเตอร์			-

จากตาราง 41 แสดงว่าค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 3 พารามิเตอร์ กับโอมเดล 1 และโอมเดล 2 พารามิเตอร์ มีค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโอมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

จุดมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน ที่วิเคราะห์ด้วยโมเดลโลจิสติกแบบเดียวกัน
2. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถผู้สอบระหว่างกลุ่มผู้สอบที่มีจำนวนไม่เท่ากัน ที่วิเคราะห์ด้วยโมเดลโลจิสติกแบบเดียวกัน
3. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติกแบบต่างกัน ที่มีจำนวนกลุ่มตัวอย่างเท่ากัน
4. เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถผู้สอบระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติกแบบต่างกัน ที่มีจำนวนกลุ่มตัวอย่างเท่ากัน

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 1 ปีการศึกษา 2538 ของโรงเรียนสังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่ 5 กรุงเทพมหานคร จำนวนนักเรียน 2,500 คน ซึ่งเลือกมาโดยวิธีการสุ่มแบบแบ่งชั้น

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 เรื่องทศนิยม (ค 101) ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2531 (ปรับปรุงพุทธศักราช 2521) เป็นแบบเลือกตอบชนิด 5 ตัวเลือก จำนวน 20 ข้อ ที่ผู้วิจัยสร้างขึ้น

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้นำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มตัวอย่าง โดยดำเนินการตามขั้นตอน ดังนี้

1. วางแผนในการดำเนินการสอบ โดยติดต่อขอความร่วมมือในการนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มตัวอย่าง ตลอดจนกำหนดวันและเวลาในการทดสอบ
2. นำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่าง
3. นำข้อมูลที่เกี่ยวข้องมาวิเคราะห์ด้วยโปรแกรม BILOG และทดสอบสมมติฐานการวิจัย

การวิเคราะห์ข้อมูล

ผู้วิจัยได้ดำเนินการวิเคราะห์ข้อมูล ดังนี้

1. วิเคราะห์องค์ประกอบ (Factor Analysis: PC) แล้วหมุนแกนด้วยวิธีแวนริแมกซ์ (Varimax) โดยใช้โปรแกรม SPSS
2. วิเคราะห์ความสัมพันธ์กับโมเดล
3. หาค่าประมาณพารามิเตอร์และค่าสถิติพื้นฐาน ของค่าพารามิเตอร์และความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์
4. เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล
5. เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล
6. เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดล ในแต่ละกลุ่มผู้สอบ
7. เปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดล ในแต่ละกลุ่มผู้สอบ

สรุปผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลสรุปได้ดังต่อไปนี้

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ หลังจากคิดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้นความคลาดเคลื่อนมาตรฐานเฉลี่ยของการประมาณค่าพารามิเตอร์ข้อสอบส่วนใหญ่มีค่าลดลง และการกระจายของความคลาดเคลื่อน มีค่าลดลงเช่นเดียวกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ความสามารถของผู้สอบ หลังจากคิดข้อสอบที่ไม่สอดคล้องกับโมเดลออกแล้ว เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความคลาดเคลื่อนมาตรฐานเฉลี่ยของการประมาณค่าพารามิเตอร์ผู้สอบส่วนใหญ่มีค่าสูงขึ้น ส่วนการกระจายของความคลาดเคลื่อน มีค่าลดลง
2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชก มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่ เมื่อประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชก ระหว่างกลุ่มตัวอย่าง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500 คน 750 คน และ

1,000 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05 ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 500 คน กับกลุ่ม 750 คน และกลุ่ม 750 คน กับกลุ่ม 1,000 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และเมื่อประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ปรากฏว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ระหว่างกลุ่มตัวอย่าง มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500 คน และ 1,000 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 นอกนั้นมีค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเดา ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากผลการศึกษาสรุปได้ว่า เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มตัวอย่างมีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ สำหรับค่าความยากมีค่าไม่แตกต่างกัน ส่วนค่าอำนาจจำแนก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง เมื่อประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าข้อสอบ สำหรับค่าความยากและค่าการเดา มีค่าไม่แตกต่างกัน ส่วนค่าอำนาจจำแนกมีค่าแตกต่างกันในบางกลุ่มตัวอย่าง และปรากฏว่ากลุ่มตัวอย่างที่นำมาใช้วิเคราะห์แล้วให้ค่าประมาณพารามิเตอร์ที่ดีที่สุดคือ กลุ่มตัวอย่าง 750 คน

3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ระหว่างกลุ่มผู้สอบ 250 คน 500 คน 750 คน และ 1,000 คน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ทุกคู่ เมื่อประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500 คน และ 750 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750 คน และ 1,000 คน กลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนกลุ่มตัวอย่าง 250 คน กับกลุ่ม 1,000 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถจากกลุ่มตัวอย่าง 250 คน กับกลุ่ม 750 คน และ 1,000 คน กลุ่มตัวอย่าง 500 คน กับกลุ่ม 750 คน และ 1,000 คน มีค่าแตกต่างกันอย่างมีนัยสำคัญทาง

สถิติที่ระดับ .05 ส่วนกลุ่มตัวอย่าง 250 คน กับกลุ่ม 500 คน กลุ่มตัวอย่าง 750 คน กับกลุ่ม 1,000 คน มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

จากผลการศึกษารูปได้ว่า เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ ค่าเฉลี่ย ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ มีค่าแตกต่างกัน ระหว่างกลุ่ม ตัวอย่างทั้ง 4 กลุ่ม และเมื่อประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ ปรากฏผล เช่นเดียวกันคือ ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ แตกต่างกันในบางกลุ่มตัวอย่าง และปรากฏว่ากลุ่มตัวอย่างที่นำมาใช้วิเคราะห์แล้วให้ค่า ประมาณพารามิเตอร์ที่ดีที่สุดคือ กลุ่มตัวอย่าง 750 คน

4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ที่สอบ เมื่อวิเคราะห์ จากกลุ่มตัวอย่าง 250 คน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ แตกต่างจากโมเดล 2 พารามิเตอร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณ ค่าความชาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ กับโมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าเฉลี่ยความคลาดเคลื่อน มาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อวิเคราะห์จากกลุ่ม ตัวอย่าง 500 คน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชาก ที่เป็นผล มาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ กับโมเดล 3 พารามิเตอร์ มีค่าแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณ ค่าความชาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ กับโมเดล 1 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าเฉลี่ยความคลาดเคลื่อน มาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อวิเคราะห์จากกลุ่ม ตัวอย่าง 750 คน ค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความชาก ที่เป็น ผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มี นัยสำคัญทางสถิติ ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างมี นัยสำคัญทางสถิติที่ระดับ .05 เมื่อวิเคราะห์จากกลุ่มตัวอย่าง 1,000 คน ค่าเฉลี่ยความ คลาดเคลื่อนมาตรฐานในการประมาณค่าความชาก ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าเฉลี่ยความ คลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ที่เป็นผลมาจากการประมาณค่าด้วย

โมเดล 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากผลการศึกษารูปได้พบว่า ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดลทั้ง 3 มีค่าแตกต่างกันในแต่ละกลุ่มตัวอย่างค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันระหว่างโมเดลในทุกขนาดกลุ่มตัวอย่าง ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ในบางกลุ่มตัวอย่างมีค่าแตกต่างกัน และปรากฏว่ากลุ่มตัวอย่างที่นำมาใช้วิเคราะห์แล้วให้ค่าประมาณพารามิเตอร์ที่ดีที่สุดคือ กลุ่มตัวอย่าง 750 คน

5. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ เมื่อวิเคราะห์จากกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน พบว่าค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ กับโมเดล 1 และ 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนค่าเฉลี่ยความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของผู้สอบ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ กับโมเดล 2 พารามิเตอร์ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ในทุกขนาดกลุ่มตัวอย่าง และปรากฏว่ากลุ่มตัวอย่างที่นำมาใช้วิเคราะห์แล้วให้ค่าประมาณพารามิเตอร์ที่ดีที่สุดคือ กลุ่มตัวอย่าง 750 คน

จากผลการศึกษารูปได้พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ มีค่าแตกต่างกันในบางโมเดล ในทุกขนาดกลุ่มตัวอย่าง โดยโมเดล 3 พารามิเตอร์ มีค่าแตกต่างจากโมเดล 1 และ 2 พารามิเตอร์

อภิปรายผล

1. การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล พบว่าเมื่อทำการเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ สำหรับความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ระหว่างกลุ่มตัวอย่างทั้ง 4 กลุ่ม มีความแตกต่างกัน สอดคล้องกับสมมติฐานข้อที่ 1 ซึ่งความแตกต่างนี้น่าจะเป็นผลมาจากขนาดของกลุ่มตัวอย่างที่ไม่เท่ากัน จากการศึกษาของ ลอร์ดและวิงเกอร์สกี (Lord and Wingersky, 1985 : 69 - 88) กล่าวว่าค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ข้อสอบ จะแปรไปในทางตรงกันข้ามกับรากที่สองของขนาดกลุ่มตัวอย่าง แต่จะเปลี่ยนแปลงไปน้อยมากเมื่อจำนวนข้อสอบเปลี่ยนไปซึ่งสอดคล้องกับการศึกษานี้ เมื่อพิจารณาค่าความคลาดเคลื่อนมาตรฐาน พบว่ามีค่าลดลงเมื่อขนาด

กลุ่มตัวอย่างใหญ่ขึ้น สำหรับโมเดล 2 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกของกลุ่ม 250 คน และ 500 คน แตกต่างจากกลุ่ม 750 คน และ 1,000 คน นอกนั้นความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 1 เนื่องจากขนาดของกลุ่มตัวอย่างจะส่งผลต่อความสัมพันธ์ระหว่างค่าอำนาจจำแนกกับค่าประมาณอำนาจจำแนก ซึ่งจะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกเช่นเดียวกัน ทำให้ความคลาดเคลื่อนมาตรฐานจากกลุ่มตัวอย่าง 250 คน มีค่าแตกต่างไปจากกลุ่มอื่น ๆ สอดคล้องกับการศึกษาของ ฮิวลิน ลิสแซค และ ดราสโกว์ (Hulin, Lissak and Drasgow. 1982 : 249 - 260) ซึ่งศึกษาค่าพารามิเตอร์ข้อสอบภายใต้โมเดล 2 และ 3 พารามิเตอร์ โดยพิจารณาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อน (Root mean square) เป็นเกณฑ์ พบว่าสหสัมพันธ์ระหว่างค่าอำนาจจำแนกและค่าประมาณอำนาจจำแนก จะมีค่าเท่ากับ .9 เมื่อกลุ่มตัวอย่างมีขนาดมากกว่า 500 คน สำหรับโมเดล 2 พารามิเตอร์ และค่าสหสัมพันธ์จะมีค่ามากกว่า .9 เมื่อกลุ่มตัวอย่างมีขนาดมากกว่า 1,000 คน และข้อสอบมีมากกว่า 30 ข้อ ในทำนองเดียวกันกับความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 500 คน มีค่าแตกต่างจากกลุ่ม 1,000 คน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก จากกลุ่มตัวอย่างทั้ง 4 กลุ่ม มีค่าไม่แตกต่างกัน น่าจะเนื่องมาจากค่าประมาณความยากจะเปลี่ยนแปลงไปเพียงเล็กน้อยแม้กลุ่มตัวอย่างจะมีขนาดเปลี่ยนแปลงไป ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 1 สำหรับโมเดล 3 พารามิเตอร์ พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่มตัวอย่าง 250 คน มีค่าแตกต่างจากกลุ่ม 500 คน และ 1,000 คน อย่างมีนัยสำคัญทางสถิติ นอกนั้นความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 1 ความแตกต่างนี้เนื่องมาจากขนาดของกลุ่มตัวอย่าง ที่ส่งผลต่อความสัมพันธ์ระหว่างค่าอำนาจจำแนกและค่าประมาณอำนาจจำแนก และจากการที่ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากกลุ่ม 1,000 คน แตกต่างจากกลุ่มอื่น ๆ นั้น น่าจะเป็นเพราะโมเดล 3 พารามิเตอร์ จะประมาณค่าอำนาจจำแนกได้ใกล้เคียงกับค่าที่แท้จริง เมื่อใช้กลุ่มตัวอย่างตั้งแต่ 1,000 คน ขึ้นไป ซึ่งจากการศึกษาของ ฮิวลิน ลิสแซค และ ดราสโกว์ (Hulin, Lissak and Drasgow. 1982 : 249 - 260) กล่าวว่าสำหรับโมเดล 3 พารามิเตอร์ เมื่อใช้กลุ่มตัวอย่าง มากกว่าหรือเท่ากับ 1,000 คน และข้อสอบมากกว่า 30 ข้อแล้ว สหสัมพันธ์ระหว่างค่าอำนาจจำแนกและค่าประมาณอำนาจจำแนก จะมีค่ามากกว่า .9 และประมาณ .5 เมื่อกลุ่มตัวอย่างประมาณ 500 คน นั่นคือจะส่งผลต่อความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนกด้วย ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 1 และความไม่แตกต่างนี้

น่าจะเป็นผลมาจากค่าอำนาจจำแนกที่เปลี่ยนแปลงไป ซึ่งส่งผลต่อค่าความยากและความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเคา มีค่าไม่แตกต่างกันในระหว่างกลุ่มตัวอย่าง อาจจะเป็นเนื่องจากค่าการเคาจะมีค่าสูงหรือต่ำ เมื่อความยากของข้อสอบมีค่ามากหรือน้อยด้วย จึงมีผลให้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าการเคาไม่แตกต่างกันในระหว่างกลุ่มตัวอย่าง ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 1

2. การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มผู้สอบ ในแต่ละโมเดล พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ที่เป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ มีค่าแตกต่างกัน ซึ่งสอดคล้องกับสมมติฐานข้อที่ 2 ทั้งนี้ น่าจะเนื่องมาจากขนาดของกลุ่มตัวอย่างที่ไม่เท่ากัน เพราะการเพิ่มหรือลดขนาดของกลุ่มตัวอย่างในงานวิจัยนี้ ไม่ได้เพิ่มหรือลดในเชิงสัมพันธ์แบบจำลอง (stimulate) แต่เพิ่มหรือลดจากข้อมูลจริง จึงทำให้เกิดอิทธิพลขึ้นได้อย่างชัดเจน จากการศึกษาของ Hambleton และ Swaminathan (1985) พบว่าขนาดของความสามารถ จะสูงขึ้นเมื่อขนาดของกลุ่มตัวอย่างใหญ่ขึ้น ซึ่งความสามารถที่มีขนาดสูงขึ้น จะมีผลในทางกลับกันกับความคลาดเคลื่อนมาตรฐาน ที่จะมีค่าลดลงเมื่อพิจารณาในแต่ละโมเดล พบว่าโมเดล 1 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ระหว่างกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน มีค่าแตกต่างกัน สอดคล้องกับสมมติฐานข้อ 2 ซึ่งความแตกต่างที่เกิดขึ้นนี้เป็นผลจากความแตกต่างของความคลาดเคลื่อนในการประมาณค่าความยากที่เป็นผลมาจากการประมาณค่าด้วย โมเดล 1 พารามิเตอร์ สำหรับโมเดล 2 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถของกลุ่มตัวอย่าง 500 คน และ 750 คน มีค่าแตกต่างจากกลุ่ม 250 คน และ 1,000 คน อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 2 ในส่วนของความแตกต่างนั้น น่าจะเนื่องมาจากความแตกต่างของค่าอำนาจจำแนกของแต่ละกลุ่มตัวอย่าง ซึ่งจากการศึกษาของ แฮมเบิลตัน (Hambleton, 1989 : 164) กล่าวว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ จะมีขนาดเล็กลงเมื่อค่าอำนาจจำแนกมีค่าสูง และค่าอำนาจจำแนกก็ขึ้นอยู่กับขนาดของกลุ่มตัวอย่าง ในทำนองเดียวกันสำหรับโมเดล 3 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ของกลุ่ม 250 คน และ 500 คน มีค่าแตกต่างจากกลุ่ม 750 คน และ 1,000 คน ซึ่งเป็นผลมาจากค่าอำนาจจำแนกที่แตกต่างกัน นอกจากนี้ค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 2

3. การเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดล ในแต่ละกลุ่มตัวอย่าง ปรากฏว่าเมื่อพิจารณาในแต่ละกลุ่มตัวอย่าง

ผลการเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ซึ่งเป็นผลมาจากการประมาณค่าด้วยโมเดล 1, 2 และ 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่างทั้ง 4 กลุ่ม พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก ระหว่างโมเดล 2 และ 3 พารามิเตอร์ มีความแตกต่างกันทุกขนาดกลุ่มตัวอย่าง ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ กับโมเดล 2 พารามิเตอร์ และ 3 พารามิเตอร์ มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 3 เมื่อพิจารณาค่าเฉลี่ยของค่าอำนาจจำแนกจากทุกกลุ่มตัวอย่างแล้ว ค่าอำนาจจำแนกที่ประมาณค่าด้วยโมเดล 3 พารามิเตอร์ จะมีค่าสูงกว่าค่าอำนาจจำแนกที่ประมาณค่าด้วยโมเดล 2 พารามิเตอร์ และขณะเดียวกันค่าความคลาดเคลื่อนมาตรฐานก็มีค่าสูงขึ้นด้วย ซึ่งสอดคล้องกับการศึกษาของ เบเกอร์ (Baker. 1987 : 111 - 138) ที่สรุปว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ที่ประมาณค่าด้วยโมเดล 1 และ 2 พารามิเตอร์ จะมีค่าต่ำกว่าค่าที่ประมาณด้วยโมเดล 3 พารามิเตอร์ ในส่วนของความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก จากการประมาณค่าด้วยกลุ่มตัวอย่าง 250 คน มีความแตกต่างระหว่างโมเดล 1 กับ 2 พารามิเตอร์ และไม่แตกต่างจากโมเดล 3 พารามิเตอร์ ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 3 ความแตกต่างนี้อาจเนื่องมาจากความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก ที่ประมาณค่าด้วยโมเดล 2 พารามิเตอร์ มักจะมีค่าสูงกว่าอีก 2 โมเดล ในกลุ่มตัวอย่างที่มีขนาดต่ำกว่า 500 คน จึงทำให้มีความแตกต่างขึ้น กลุ่มตัวอย่าง 500 คน ความคลาดเคลื่อนมาตรฐานมีความแตกต่างกันระหว่างโมเดล 1 กับ 3 พารามิเตอร์ แต่ไม่แตกต่างจากโมเดล 2 พารามิเตอร์ ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 4 เมื่อพิจารณาข้อสอบที่สอดคล้องกับโมเดล พบว่าข้อสอบข้อที่ 14 และ 19 สอดคล้องกับโมเดล 1 พารามิเตอร์ แต่ยังคงมีความคลาดเคลื่อนในการประมาณค่าความยากสูงทั้ง 2 ข้อ เมื่อเทียบกับอีก 2 โมเดล จึงทำให้ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากแตกต่างกันได้ สำหรับกลุ่มตัวอย่าง 750 คน และ 1,000 คน ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากมีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 3 ความไม่แตกต่างกันนี้อาจเนื่องมาจากเมื่อกลุ่มตัวอย่างใหญ่ขึ้น ความเที่ยงตรงในการประมาณค่าความยากจะสูงขึ้น

4. การเปรียบเทียบ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ ผู้สอบ ระหว่างโมเดล ในแต่ละกลุ่มตัวอย่าง พบว่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ ซึ่งเป็นผลจากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ มีค่าแตกต่างจากโมเดล 1 และ 2 พารามิเตอร์ ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความสามารถ จากการประมาณค่าด้วยโมเดล 1 และ 2 พารามิเตอร์ มีค่าไม่แตกต่างกัน ซึ่งไม่สอดคล้องกับสมมติฐานข้อที่ 4 ในทุกกลุ่มตัวอย่าง ที่ปรากฏผลเช่นนี้อาจเนื่องมาจาก โมเดล

3 พารามิเตอร์ จะให้ผลการประมาณค่าพารามิเตอร์ที่เที่ยงตรงน้อยกว่าโมเดล 1 และ 2 พารามิเตอร์

5. ผลจากการทดสอบความสอดคล้องของข้อสอบกับโมเดล พบข้อสังเกตบางประการคือ

5.1 จำนวนข้อสอบที่ไม่สอดคล้องกับโมเดล ระหว่างโมเดล 1, 2 และ 3 พารามิเตอร์ พบว่าเมื่อวิเคราะห์จากโมเดลที่มีจำนวนพารามิเตอร์ที่ประมาณค่ามีมากขึ้น จำนวนข้อสอบที่ไม่สอดคล้องกับโมเดล จะมีจำนวนลดลง

5.2 จำนวนข้อสอบที่ไม่สอดคล้องกับโมเดล เมื่อใช้กลุ่มตัวอย่างขนาด 200 คน 500 คน 750 คน และ 1,000 คน พบว่าเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ขึ้น จำนวนข้อสอบที่ไม่สอดคล้องกับโมเดล จะมีมากขึ้นด้วยเช่นเดียวกัน

จากข้อสังเกตข้างต้นสามารถสรุปได้ว่า เมื่อข้อสอบสอดคล้องกับโมเดลที่ใช้ในการประมาณค่าแล้ว การประมาณค่าคุณลักษณะของข้อสอบจะเป็นอิสระจากขนาดของกลุ่มตัวอย่างได้มากขึ้น เมื่อใช้เฉพาะข้อสอบที่สอดคล้องกับโมเดลมาประมาณค่า

จากการศึกษาครั้งนี้ อาจสรุปได้ว่าเมื่อจะประมาณค่าพารามิเตอร์ ด้วยวิธีการประมาณค่าแบบ มาจินัลแมกซิมัมไลคิฮูด (Marginal Maximum Likelihood) แล้ว สำหรับการเลือกใช้โมเดล 1 พารามิเตอร์ สามารถใช้กลุ่มตัวอย่างขนาดเล็กในการวิเคราะห์ได้ จะทำให้ประหยัดทั้งเวลาและงบประมาณ สำหรับโมเดล 2 พารามิเตอร์ ควรใช้กลุ่มตัวอย่างตั้งแต่ 500 คนขึ้นไป จะให้ค่าพารามิเตอร์ที่เที่ยงตรงกว่าการใช้กลุ่มตัวอย่างขนาดเล็ก ส่วนการใช้โมเดล 3 พารามิเตอร์ ในการประมาณค่าพารามิเตอร์ข้อสอบนั้น สามารถใช้กลุ่มตัวอย่าง 750 คน และ 1,000 คน ได้ แต่กลุ่มตัวอย่าง 1,000 คน จะให้ค่าประมาณพารามิเตอร์ข้อสอบที่เที่ยงตรงมากกว่า แม้ค่าความสามารถที่ได้จากกลุ่ม 1,000 คน จะเที่ยงตรงน้อยกว่ากลุ่มตัวอย่างที่มีขนาดน้อยกว่า 500 คน แต่ก็ไม่มากนักก็อย่างชัดเจนโดยภาพรวมแล้ว โมเดล 1 และ 2 พารามิเตอร์ จะให้ค่าประมาณพารามิเตอร์ที่เที่ยงตรงกว่าโมเดล 3 พารามิเตอร์ และขนาดกลุ่มตัวอย่าง 750 คน เหมาะสมที่สุดที่จะนำมาใช้ในการวิเคราะห์เพื่อประมาณค่าพารามิเตอร์

ข้อเสนอแนะ

1. ผลจากการศึกษาค้างนี้ สำหรับการที่จะเลือกนำโมเดลใดมาใช้ประมาณค่าพารามิเตอร์นั้น ถ้าต้องการประมาณค่าความชุกเพียงค่าเดียว ก็ควรใช้โมเดล 1 พารามิเตอร์ เพราะจะให้ค่าประมาณความชุกที่ดีกว่าอีก 2 โมเดล เมื่อต้องการประมาณค่าอำนาจจำแนกด้วย ก็ควรใช้โมเดล 2 พารามิเตอร์ เพราะจะให้ค่าประมาณอำนาจจำแนก ได้เที่ยงตรงกว่าโมเดล 3 พารามิเตอร์ สำหรับขนาดของกลุ่มตัวอย่างที่เหมาะสมที่จะนำมาใช้ในการวิเคราะห์แล้วได้ค่าพารามิเตอร์ที่ดีที่สุดคือ กลุ่มตัวอย่างขนาด 750 คน

2. การศึกษาค้างนี้ใช้จำนวนข้อสอบคงที่ ในครั้งต่อไปอาจทำการศึกษาโดยใช้ข้อสอบหลายฉบับ ที่มีจำนวนข้อแตกต่างกัน เพื่อศึกษาถึงผลของจำนวนข้อสอบว่าจะส่งผลถึงความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์อย่างไร

3. อาจทำการศึกษาเปรียบเทียบระหว่างวิธีประมาณค่าแบบต่าง ๆ ว่าจะมีความเที่ยงตรงในการประมาณค่าพารามิเตอร์ แตกต่างกันได้ดีเพียงใด

มรดกทางวัฒนธรรม

บรรณานุกรม

- โกวิท ประวาลพุกษ์. "ความเข้าใจพื้นฐานเกี่ยวกับโด่งลักษณะข้อสอบ." วารสารการ
วัดผลการศึกษา. 4 : 55 - 56 ; กันยายน - สิงหาคม 2525.
- คำย เชื้องฉ. การศึกษาเปรียบเทียบประสิทธิภาพในการประมาณค่าความสามารถของ
ผู้สอบ จากการทดสอบเทเลอ์รูปปริมาตร ที่มีรูปแบบ จำนวนชั้น และวิธีการให้
คะแนนที่แตกต่างกัน โดยใช้วิธีมอนติ คาร์โล. ปรินญาณินพนธ์ กศ.ค. กรุงเทพฯ
: มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2534. อัดสำเนา.
- บัญญัติ แสนทวี. การประยุกต์รูปแบบราชสีในการออกแบบโด่งแสดงสารสนเทศของแบบ
ทดสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ตามระดับความสามารถของนักเรียนชั้น
ประถมศึกษาปีที่ 6. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย,
2530. อัดสำเนา.
- บุญชิต รอดแก้ว. การศึกษาความไม่แปรปรวนของค่าพารามิเตอร์ของข้อสอบที่วิเคราะห์
ด้วยโมเดลโลจิสติกสามพารามิเตอร์กับกลุ่มตัวอย่างที่มีลักษณะต่างกัน.
ปรินญาณินพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2533. อัดสำเนา.
- พจวจิต อินทสุวรรณ. ทฤษฎีการตอบข้อคำถาม. ม.ป.ท, 2528.
- เรวดี อินทสระ. การเปรียบเทียบความคลาดเคลื่อนของการเทียบระหว่างรูปแบบอิง
ทฤษฎีการตอบข้อสอบกับรูปแบบการใช้เทคนิคการวิเคราะห์องค์ประกอบ.
ปรินญาณินพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร,
2530. อัดสำเนา.
- วิรัช วรรณรัตน์. การเปรียบเทียบลักษณะการแจกแจงความถี่ของค่าความสามารถจริง
ที่ประมาณตามแบบดั้งเดิมกับแบบคุณลักษณะแฝง. วิทยานิพนธ์ ค.ค. กรุงเทพฯ
: จุฬาลงกรณ์มหาวิทยาลัย, 2533. อัดสำเนา.
- สงบ ลักษณะ. "ความสามารถกับการตอบข้อสอบ," วารสารการวัดผลการศึกษา.
4 : 47 - 54 ; พฤษภาคม - สิงหาคม 2525.
- สำเร้ง บุญเรืองรัตน์. ทฤษฎีการวัดและประเมินผลการศึกษา (ฉบับปรับปรุง).
กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทร
วิโรฒ ประสานมิตร, 2529.
- Bock, R.D., and Aitkin, M. "Marginal Maximum Likelihood Estimation
of Item Parameters : An application of an EM algorithm,"
Psychometrika. 46 : 443 - 459, 1981.

- De Gruijter, D.N.M. "A comment on Some Standard Errors in Item Response Theory". Psychometrika. 49 : 269 - 272 ; June, 1984.
- De Gruijter, D.N.M. A Note on The Asymptotic Variance-Covariance Matrix of Item Parameter Estimates in The Rasch Model,". Psychometrika. 50 : 247 - 249, 1985.
- Divgi, D.R. "Does The Rasch Model Really Work for Multiple Choice Items? Not if You Look Closely,". Journal of Educational Measurement. 23 : 283 - 298, 1986.
- Hambleton, Ronald K. "Principles and Selected Applications of Item Response Theory,". Educational Measurement. Third Edition. New York : American Council on Education and Macmillan Publishing Company A Division of Macmillan, Inc., 1989.
- Hambleton, Ronald K., Russell W. Jones and Rogers H. Jane, "Influence of Item Parameter Estimation Errors in Test Development,". Journal of Educational Measurement. 30 : 143 - 155 ; Summer, 1993.
- Hulin, C.L., Lissak, R.I., and Ddrassgow, F. "Recovery of Two- and Three-parameter Logistic Item Characteristic Curves : a Monte Carlo Study,". Applied Psychological Measurement, 6 : 249 - 260, 1982.
- Kubiak, Anna T. "Estimation of Item and Ability Parameters by Three Logistic Models of Item Response Theory," Dissertation Abstracts International. 47(7) : 2512-A ; 1987.
- Lord, Frederic M., and Novick, Melvin R. Statistical Theories of Mental Test Scores. Massachusetts : Addison-Wesley Publishers, 1980.
- McDonald, R.P. "Linear versus Non-linear Models in Item Response Theory,". Applied Psychological Measurement. 6 : 379 - 396, 1982.

- Thissen, David and Howard Wainer. "Some Standard Errors in Item Response Theory". Psychometrika. 47 : 397 - 412 ; December, 1982.
- Urry, V.W. "Approximations to item parameters of mental test models and their uses,". Educational and psychological Measurement. 34 : 253 - 269, 1974.
- Warm, T.A. A Primer of Item Response Theory. Oklahoma : U.S. Coast Guard Institute, 1978.
- Wright, B.D. and M.H. Stone. Best Test Design. Chicago, Illinois, MESA Press, 1979.
- Yoes, Michael Ernest. "A Comparison of The Effectiveness of Item Parameter Estimation Techniques Used With The Three-Paramete Logistic Item Response Theory Model," Dissertation Abstracts International. 54(12) : 6497-B ; June, 1994.

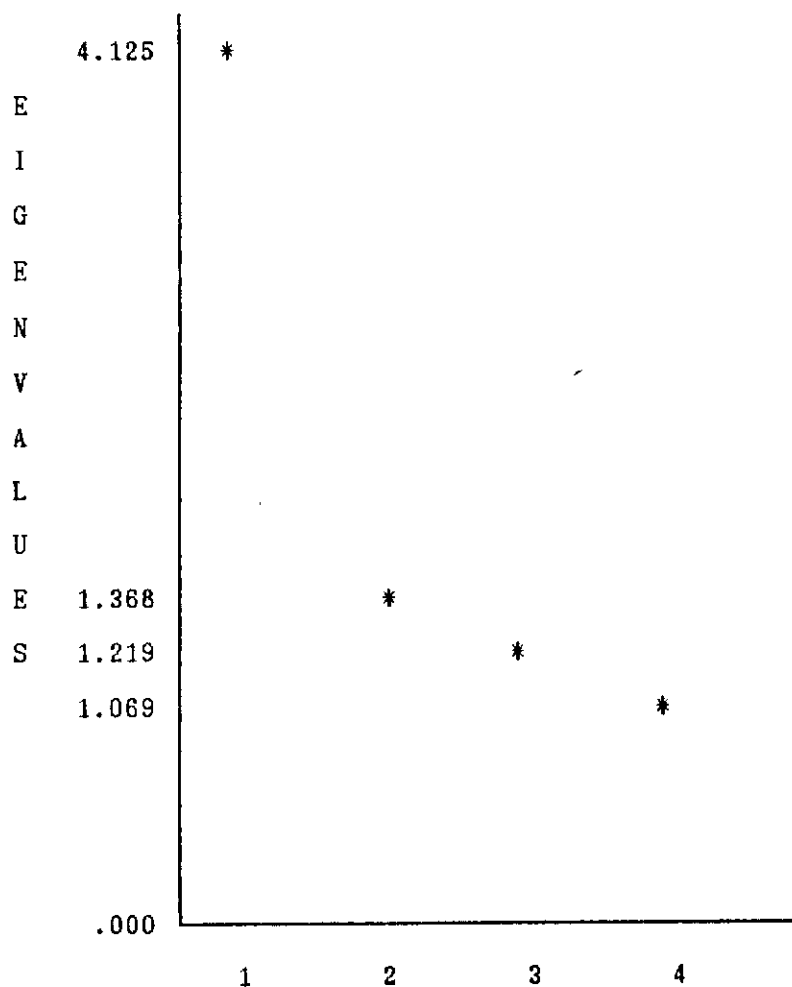
ภาคผนวก

ภาคผนวก ก
ผลการวิเคราะห์องค์ประกอบที่มีค่าไอเกิน มากกว่า 1
เพื่อตรวจสอบความเป็นมิติเดียวของแบบทดสอบ

ตาราง 42 ค่าไอเกินขององค์ประกอบที่มีค่ามากกว่า 1.0000

องค์ประกอบที่	ค่าไอเกิน	% ความแปรปรวน	% ความแปรปรวนสะสม
1	4.125	20.6	20.6
2	1.368	6.8	27.5
3	1.219	6.1	33.6
4	1.069	5.3	38.9

จากตาราง 42 สามารถนำค่าไอเกินมาเขียนกราฟความสัมพันธ์ โดยให้แกน x แทนจำนวนองค์ประกอบ และแกน y แทนค่าไอเกิน เพื่อการพิจารณาที่ชัดเจนยิ่งขึ้น ดังภาพประกอบ 14



ภาพประกอบ 14 องค์ประกอบที่มีค่าไอเกินมากกว่า 1.0000

ภาคผนวก ข

ค่าประมาณพารามิเตอร์ที่ข้อสอบและความคลาดเคลื่อนมาตรฐาน

ในการประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วย

โมเดล 1, 2 และ 3 พารามิเตอร์

สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

ตาราง 43 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณ
ค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง
250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน		500 คน		750 คน		1,000 คน	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.
1	-0.994	0.164	-1.137	0.124	-1.231	0.110	-1.563	0.113
2	-1.147	0.170	-0.880	0.121	-0.884	0.107	-1.433	0.113
3	0.279	0.154	0.358	0.114	0.231	0.102	0.395	0.105
4	-0.360	0.158	-0.333	0.116	-0.163	0.103	-0.654	0.106
5	-0.597	0.162	-0.564	0.117	-0.668	0.105	-1.016	0.109
6	0.574	0.162	0.744	0.120	0.720	0.108	0.449	0.105
7	0.587	0.160	0.769	0.119	0.908	0.106	0.808	0.107
8	0.410	0.152	0.759	0.119	0.547	0.102	0.522	0.103
9	0.774	0.161	0.981	0.123	1.292	0.109	1.071	0.108
10	0.829	0.175	0.717	0.122	1.065	0.111	0.908	0.112
11	-0.177	0.158	-0.051	0.116	0.037	0.103	-0.449	0.104
12	0.908	0.164	0.794	0.124	1.279	0.116	0.874	0.111
13	0.595	0.158	0.561	0.121	0.607	0.105	0.400	0.104
14	1.650	0.198	1.354	0.137	1.888	0.129	1.745	0.123
15	0.870	0.161	0.947	0.125	1.531	0.115	1.245	0.112
16	0.979	0.160	0.924	0.116	1.216	0.108	1.111	0.107
17	0.870	0.161	0.629	0.116	1.039	0.106	0.933	0.107
18	0.631	0.161	0.412	0.112	0.651	0.104	0.510	0.105
19	1.226	0.167	1.091	0.115	1.483	0.110	1.355	0.110
20	1.337	0.160	1.252	0.122	1.328	0.106	1.522	0.109
\bar{X}	0.462	0.163	0.466	0.120	0.644	0.108	0.581	0.089
S.D.	0.755	0.010	0.701	0.005	0.853	0.006	0.783	0.000

ตาราง 44 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณ
ค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง
250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน				500 คน			
	a	S.E.	b	S.E.	a	S.E.	b	S.E.
1	0.523	0.105	-0.954	0.235	0.809	0.089	-1.033	0.152
2	0.687	0.147	-0.926	0.180	0.631	0.088	-0.790	0.124
3	0.422	0.094	0.475	0.222	0.541	0.079	0.371	0.125
4	0.675	0.135	-0.242	0.136	0.729	0.095	-.239	0.088
5	0.713	0.149	-0.438	0.137	0.736	0.105	-.417	0.090
6	0.612	0.117	0.639	0.167	0.618	0.083	0.686	0.126
7	0.476	0.105	0.787	0.233	0.512	0.071	0.814	0.161
8	0.385	0.091	0.693	0.256	0.573	0.081	0.717	0.132
9	0.470	0.102	0.989	0.262	0.534	0.078	0.960	0.164
10	0.818	0.167	0.734	0.164	0.705	0.088	0.681	0.115
11	0.703	0.140	-0.096	0.129	0.644	0.084	-0.049	0.095
12	0.364	0.088	1.453	0.405	0.694	0.087	0.710	0.117
13	0.417	0.097	0.860	0.273	0.635	0.087	0.444	0.114
14	0.685	0.141	1.553	0.257	0.864	0.105	1.011	0.121
15	0.370	0.092	1.343	0.379	0.636	0.086	0.901	0.133
16	0.280	0.072	1.916	0.566	0.325	0.061	1.583	0.327
17	0.329	0.086	1.526	0.444	0.439	0.070	0.795	0.169
18	0.448	0.098	0.881	0.258	0.352	0.062	0.724	0.200
19	0.300	0.080	2.192	0.625	0.229	0.053	2.455	0.598
20	0.192	0.058	3.536	1.154	0.293	0.062	2.162	0.462
\bar{X}	0.493	0.108	0.846	0.324	0.565	0.081	0.624	0.181
S.D.	0.174	0.029	1.079	0.239	0.165	0.014	0.864	0.132

ข้อสอบ	750 RM				1,000 RM			
	a	S.E.	b	S.E.	a	S.E.	b	S.E.
1	0.595	0.076	-0.996	0.125	0.568	0.062	-0.948	0.109
2	0.674	0.083	-0.655	0.093	0.677	0.071	-0.753	0.083
3	0.499	0.064	0.259	0.101	0.537	0.058	0.552	0.094
4	0.543	0.070	-0.113	0.090	0.650	0.067	-0.208	0.069
5	0.583	0.077	-0.537	0.096	0.714	0.071	-0.449	0.069
6	0.668	0.076	0.588	0.089	0.578	0.061	0.603	0.089
7	0.456	0.061	0.965	0.155	0.483	0.056	0.938	0.129
8	0.447	0.060	0.614	0.124	0.429	0.052	0.762	0.127
9	0.352	0.059	1.691	0.283	0.388	0.050	1.334	0.186
10	0.592	0.070	0.934	0.121	0.704	0.070	0.829	0.089
11	0.548	0.069	0.066	0.090	0.555	0.058	-0.083	0.077
12	0.622	0.071	1.076	0.127	0.593	0.059	0.840	0.101
13	0.471	0.064	0.648	0.129	0.401	0.051	0.670	0.130
14	0.691	0.075	1.474	0.148	0.652	0.067	1.456	0.135
15	0.363	0.060	1.946	0.314	0.402	0.051	1.511	0.201
16	0.299	0.053	1.843	0.343	0.301	0.047	1.765	0.288
17	0.316	0.054	1.504	0.279	0.309	0.046	1.518	0.252
18	0.327	0.054	0.930	0.201	0.377	0.050	0.863	0.149
19	0.225	0.049	2.865	0.639	0.230	0.043	2.582	0.497
20	0.178	0.043	3.208	0.807	0.192	0.043	3.430	0.764
\bar{X}	0.472	0.064	0.915	0.218	0.487	0.057	0.861	0.182
S.D.	0.154	0.011	1.100	0.192	0.159	0.009	1.068	0.170

ตาราง 45 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ จากการประมาณค่าด้วยโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.448	0.108	-1.303	0.468	0.226	0.096
2	0.766	0.187	-1.121	0.307	0.235	0.098
3	1.211	0.388	0.716	0.204	0.262	0.064
4	0.993	0.227	-0.177	0.207	0.185	0.079
5	1.102	0.302	-0.253	0.214	0.219	0.089
6	1.050	0.241	0.552	0.163	0.125	0.053
7	1.239	0.392	0.864	0.194	0.224	0.057
8	0.942	0.288	0.884	0.245	0.227	0.068
9	1.142	0.297	0.959	0.194	0.182	0.052
10	2.304	1.366	0.625	0.133	0.152	0.042
11	0.821	0.169	-0.011	0.243	0.197	0.081
12	1.873	0.735	1.116	0.164	0.211	0.042
13	1.356	0.375	0.983	0.199	0.233	0.052
14	1.299	0.301	1.151	0.151	0.084	0.032
15	1.618	0.615	1.213	0.188	0.233	0.044
16	1.226	0.430	1.380	0.227	0.216	0.045
17	1.586	0.704	1.314	0.201	0.246	0.044
18	1.323	0.413	0.858	0.185	0.206	0.054
19	1.410	0.560	1.529	0.221	0.202	0.039
20	1.467	0.717	2.233	0.374	0.234	0.032
\bar{X}	1.259	0.441	0.676	0.224	0.205	0.058
S.D.	0.406	0.287	0.867	0.079	0.043	0.021

ข้อสอบ	500 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.393	0.084	-1.626	0.477	0.216	0.095
2	0.509	0.094	-0.932	0.344	0.217	0.094
3	1.705	0.493	0.872	0.129	0.282	0.041
4	0.799	0.134	-0.085	0.199	0.164	0.073
5	0.965	0.199	-0.188	0.203	0.203	0.084
6	1.029	0.243	0.897	0.151	0.162	0.050
7	1.807	0.443	0.972	0.113	0.195	0.034
8	0.827	0.191	0.936	0.182	0.154	0.056
9	1.061	0.281	1.163	0.152	0.171	0.044
10	1.294	0.294	0.858	0.125	0.163	0.042
11	0.652	0.113	0.116	0.221	0.153	0.069
12	1.091	0.225	0.862	0.134	0.133	0.043
13	1.586	0.414	0.813	0.122	0.230	0.042
14	1.965	0.441	0.972	0.089	0.092	0.026
15	-0.927	0.204	1.021	0.154	0.127	0.046
16	1.035	0.386	1.725	0.231	0.238	0.038
17	0.720	0.201	1.209	0.243	0.209	0.063
18	0.766	0.234	1.333	0.251	0.262	0.059
19	0.880	0.348	2.319	0.393	0.243	0.034
20	0.618	0.240	3.119	0.764	0.236	0.037
\bar{X}	1.031	0.263	0.818	0.234	0.192	0.053
S.D.	0.436	0.122	1.036	0.160	0.049	0.020

ข้อสอบ	750 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.331	0.074	-2.019	0.584	0.238	0.099
2	0.629	0.105	-0.675	0.268	0.214	0.090
3	0.945	0.230	0.698	0.196	0.262	0.064
4	0.713	0.162	0.324	0.283	0.269	0.088
5	0.768	0.152	-0.149	0.275	0.269	0.096
6	0.776	0.141	0.694	0.166	0.137	0.055
7	1.191	0.293	1.066	0.127	0.195	0.040
8	0.600	0.127	0.850	0.246	0.178	0.068
9	0.803	0.242	1.683	0.212	0.196	0.043
10	1.018	0.205	1.009	0.135	0.155	0.043
11	0.796	0.169	0.412	0.230	0.228	0.077
12	1.330	0.281	1.023	0.101	0.137	0.032
13	1.374	0.383	0.991	0.126	0.257	0.041
14	2.095	0.546	1.152	0.074	0.083	0.020
15	0.933	0.271	1.624	0.173	0.155	0.037
16	0.855	0.251	1.709	0.201	0.199	0.040
17	0.889	0.268	1.518	0.186	0.217	0.044
18	1.013	0.337	1.387	0.177	0.290	0.044
19	1.106	0.412	2.080	0.256	0.221	0.026
20	1.037	0.434	2.451	0.387	0.260	0.025
\bar{X}	0.960	0.254	0.891	0.220	0.208	0.054
S.D.	0.364	0.122	1.002	0.113	0.054	0.025

ข้อสอบ	1,000 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.289	0.060	-2.184	0.621	0.225	0.098
2	0.647	0.090	-0.860	0.235	0.189	0.085
3	1.026	0.224	0.818	0.139	0.219	0.049
4	0.767	0.145	0.047	0.244	0.245	0.086
5	0.722	0.102	-0.458	0.211	0.184	0.080
6	0.867	0.172	0.762	0.152	0.175	0.054
7	1.438	0.319	1.083	0.102	0.222	0.031
8	1.053	0.259	1.057	0.132	0.244	0.042
9	0.705	0.156	1.404	0.177	0.167	0.046
10	1.316	0.244	0.826	0.091	0.133	0.034
11	0.576	0.088	0.052	0.241	0.172	0.075
12	1.114	0.189	0.812	0.104	0.134	0.037
13	1.368	0.336	1.012	0.112	0.270	0.036
14	1.853	0.448	1.231	0.072	0.127	0.021
15	1.028	0.221	1.211	0.116	0.136	0.036
16	0.619	0.162	1.752	0.231	0.179	0.049
17	1.078	0.247	1.299	0.122	0.188	0.034
18	0.987	0.269	1.225	0.151	0.272	0.043
19	0.908	0.274	1.945	0.202	0.212	0.031
20	0.513	0.188	3.403	0.735	0.210	0.034
\bar{X}	0.944	0.210	0.822	0.209	0.195	0.050
S.D.	0.366	0.095	1.140	0.170	0.044	0.022

ตาราง 46 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณ
ค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 1 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง
250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน		500 คน		750 คน		1,000 คน	
	b	S.E.	b	S.E.	b	S.E.	b	S.E.
1	-0.994	0.164	-1.137	0.124	-1.231*	0.110*	-1.563	0.113
2	-1.147*	0.170*	-0.880	0.121	-0.884*	0.107*	-1.433*	0.113*
3	0.279	0.154	0.358	0.114	0.231	0.102	0.395	0.105
4	-0.360	0.158	-0.333	0.116	-0.163	0.103	-0.654*	0.106*
5	-0.597	0.162	-0.564	0.117	-0.668*	0.105*	-1.016*	0.109*
6	0.574	0.162	0.744	0.120	0.720*	0.108*	0.449*	0.105*
7	0.587	0.160	0.769*	0.119*	0.908	0.106	0.808	0.107
8	0.410	0.152	0.759	0.119	0.547	0.102	0.522	0.103
9	0.774	0.161	0.981	0.123	1.292	0.109	1.071	0.108
10	0.829	0.175	0.717	0.122	1.065*	0.111*	0.908*	0.112*
11	-0.177	0.158	-0.051*	0.116*	0.037	0.103	-0.449*	0.104*
12	0.908	0.164	0.794	0.124	1.279*	0.116*	0.874*	0.111*
13	0.595	0.158	0.561	0.121	0.607	0.105	0.400	0.104
14	1.650*	0.198*	1.354*	0.137*	1.888*	0.129*	1.745*	0.123*
15	0.870	0.161	0.947	0.125	1.531	0.115	1.245	0.112
16	0.979	0.160	0.924	0.116	1.216	0.108	1.111	0.107
17	0.870	0.161	0.629	0.116	1.039	0.106	0.933	0.107
18	0.631	0.161	0.412	0.112	0.651	0.104	0.510	0.105
19	1.226	0.167	1.091*	0.115*	1.483	0.110	1.355	0.110
20	1.337*	0.160*	1.252	0.122	1.328*	0.106*	1.522*	0.109*
\bar{X}	0.436	0.164	0.385	0.120	0.782	0.106	0.711	0.106
S.D.	0.161	0.000	0.716	0.000	0.559	0.000	0.648	0.000

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ตาราง 47 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 2 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน				500 คน			
	a	S.E.	b	S.E.	a	S.E.	b	S.E.
1	0.523	0.105	-0.954	0.235	0.609	0.089	-1.033	0.152
2	0.687	0.147	-0.926	0.180	0.631	0.088	-0.790	0.124
3	0.422	0.094	0.475	0.222	0.541	0.079	0.371	0.125
4	0.675	0.135	-0.242	0.136	0.729	0.095	-0.239	0.088
5	0.713	0.149	-0.438	0.137	0.736	0.105	-0.417	0.090
6	0.612	0.117	0.639	0.167	0.618	0.083	0.686	0.126
7	0.476	0.105	0.787	0.233	0.512*	0.071*	0.814*	0.161*
8	0.385	0.091	0.693	0.256	0.573	0.081	0.717	0.132
9	0.470	0.102	0.989	0.262	0.534	0.078	0.960	0.164
10	0.818	0.167	0.734	0.164	0.705	0.088	0.681	0.115
11	0.703	0.140	-0.096	0.129	0.644	0.084	-0.049	0.095
12	0.364	0.088	1.453	0.405	0.694	0.087	0.710	0.117
13	0.417	0.097	0.860	0.273	0.635	0.087	0.444	0.114
14	0.685	0.141	1.553	0.257	0.864	0.105	1.011	0.121
15	0.370	0.092	1.343	0.379	0.636	0.086	0.901	0.133
16	0.280	0.072	1.916	0.566	0.325	0.061	1.583	0.327
17	0.329	0.086	1.526	0.444	0.439	0.070	0.795	0.169
18	0.448	0.098	0.881	0.258	0.352	0.062	0.724	0.200
19	0.300	0.080	2.192	0.625	0.229	0.053	2.455	0.598
20	0.192	0.058	3.536	1.154	0.293	0.062	2.162	0.462
\bar{X}	0.493	0.108	0.846	0.324	0.568	0.081	0.614	0.182
S.D.	0.174	0.029	1.079	0.239	0.170	0.000	0.887	0.134

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

7	0.456	0.061	0.965	0.155	0.483*	0.056*	0.938*	0.129*
8	0.447	0.060	0.614	0.124	0.429	0.052	0.762	0.127
9	0.352	0.059	1.691	0.283	0.388	0.050	1.334	0.186
10	0.592	0.070	0.934	0.121	0.704	0.070	0.829	0.089
11	0.548	0.069	0.066	0.090	0.555	0.058	-0.083	0.077
12	0.622	0.071	1.076	0.127	0.593	0.059	0.840	0.101
13	0.471*	0.064*	0.648*	0.129*	0.401	0.051	0.670	0.130
14	0.691*	0.075*	1.474*	0.148*	0.652*	0.067*	1.456*	0.135*
15	0.363	0.060	1.946	0.314	0.402	0.051	1.511	0.201
16	0.299	0.053	1.843	0.343	0.301	0.047	1.765	0.288
17	0.316	0.054	1.504	0.279	0.309	0.046	1.518	0.252
18	0.327	0.054	0.930	0.201	0.377	0.050	0.863	0.149
19	0.225	0.049	2.865	0.639	0.230	0.043	2.582	0.497
20	0.178	0.043	3.208	0.807	0.192	0.043	3.430	0.764
\bar{X}	0.472	0.064	0.915	0.218	0.487	0.057	0.861	0.182
S.D.	0.154	0.011	1.100	0.192	0.159	0.009	1.068	0.170

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ตาราง 48 ค่าประมาณพารามิเตอร์ข้อสอบและความคลาดเคลื่อนมาตรฐานในการประมาณ
ค่าพารามิเตอร์ สำหรับข้อสอบที่สอดคล้องกับโมเดล 3 พารามิเตอร์ สำหรับกลุ่มตัวอย่าง
250 คน 500 คน 750 คน และ 1,000 คน

ข้อสอบ	250 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.448	0.108	-1.303	0.468	0.226	0.096
2	0.766	0.187	-1.121	0.307	0.235	0.098
3	1.211*	0.388*	0.716*	0.204*	0.262*	0.064*
4	0.993	0.227	-0.177	0.207	0.185	0.079
5	1.102	0.302	-0.253	0.214	0.219	0.089
6	1.050	0.241	0.552	0.163	0.125	0.053
7	1.239	0.392	0.864	0.194	0.224	0.057
8	0.942	0.288	0.884	0.245	0.227	0.068
9	1.142	0.297	0.959	0.194	0.182	0.052
10	2.304	1.366	0.625	0.133	0.152	0.042
11	0.821	0.189	-0.011	0.243	0.197	0.081
12	1.873	0.735	1.116	0.164	0.211	0.042
13	1.356	0.375	0.983	0.199	0.233	0.052
14	1.299	0.301	1.151	0.151	0.084	0.032
15	1.618	0.615	1.213	0.188	0.233	0.044
16	1.226	0.430	1.380	0.227	0.216	0.045
17	1.586	0.704	1.314	0.201	0.246	0.044
18	1.323	0.413	0.858	0.185	0.206	0.054
19	1.410	0.560	1.529	0.221	0.202	0.039
20	1.467	0.717	2.233	0.374	0.234	0.032
\bar{X}	1.273	0.452	0.773	0.225	0.206	0.057
S.D.	0.412	0.290	0.891	0.077	0.045	0.000

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ข้อสอบ	500 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.393	0.084	-1.626	0.477	0.216	0.095
2	0.509	0.094	-0.932	0.344	0.217	0.094
3	1.705	0.493	0.872	0.129	0.282	0.041
4	0.799	0.134	-0.085	0.199	0.164	0.073
5	0.965*	0.199*	-0.188*	0.203*	0.203*	0.084*
6	1.029	0.243	0.897	0.151	0.162	0.050
7	1.807	0.443	0.972	0.113	0.195	0.034
8	0.827	0.191	0.936	0.182	0.154	0.056
9	1.061	0.281	1.163	0.152	0.171	0.044
10	1.294	0.294	0.858	0.125	0.163	0.042
11	0.652*	0.113*	0.116*	0.221*	0.153*	0.069*
12	1.091	0.225	0.862	0.134	0.133	0.043
13	1.586	0.414	0.813	0.122	0.230	0.042
14	1.965	0.441	0.972	0.089	0.092	0.026
15	0.927	0.204	1.021	0.154	0.127	0.046
16	1.035	0.386	1.725	0.231	0.238	0.038
17	0.720	0.201	1.209	0.243	0.209	0.063
18	0.766	0.234	1.333	0.251	0.262	0.059
19	0.880	0.348	2.319	0.393	0.243	0.034
20	0.618	0.240	3.119	0.764	0.236	0.037
\bar{X}	1.056	0.275	0.913	0.236	0.194	0.051
S.D.	0.451	0.122	1.049	0.167	0.055	0.000

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ข้อสอบ	750 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.331	0.074	-2.019	0.584	0.238	0.099
2	0.629*	0.105*	-0.675*	0.268*	0.214*	0.090*
3	0.945	0.230	0.698	0.196	0.262	0.064
4	0.713*	0.162*	0.324*	0.283*	0.269*	0.088*
5	0.768*	0.152*	-0.149*	0.275*	0.269*	0.096*
6	0.776*	0.141*	0.694*	0.166*	0.137*	0.055*
7	1.191	0.293	1.066	0.127	0.195	0.040
8	0.600*	0.127*	0.850*	0.246*	0.178*	0.068*
9	0.803	0.242	1.683	0.212	0.196	0.043
10	1.018*	0.205*	1.009*	0.135*	0.155*	0.043*
11	0.796*	0.169*	0.412*	0.230*	0.228*	0.077*
12	1.330	0.281	1.023	0.101	0.137	0.032
13	1.374	0.383	0.991	0.126	0.257	0.041
14	2.095	0.546	1.152	0.074	0.083	0.020
15	0.933	0.271	1.624	0.173	0.155	0.037
16	0.855	0.251	1.709	0.201	0.199	0.040
17	0.889	0.268	1.518	0.186	0.217	0.044
18	1.013	0.337	1.387	0.177	0.290	0.044
19	1.106	0.412	2.080	0.256	0.221	0.026
20	1.037	0.434	2.451	0.387	0.260	0.025
\bar{X}	0.454	0.309	1.182	0.215	0.208	0.043
S.D.	0.373	0.118	1.074	0.134	0.055	0.000

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ข้อสอบ	1,000 คน					
	a	S.E.	b	S.E.	c	S.E.
1	0.289	0.060	-2.184	0.621	0.225	0.098
2	0.647	0.090	-0.860	0.235	0.189	0.085
3	1.026*	0.224*	0.818*	0.139*	0.219*	0.049*
4	0.767*	0.145*	0.047*	0.244*	0.245*	0.086*
5	0.722*	0.102*	-0.458*	0.211*	0.184*	0.080*
6	0.867*	0.172*	0.762*	0.152*	0.175*	0.054*
7	1.438	0.319	1.083	0.102	0.222	0.031
8	1.053*	0.259*	1.057*	0.132*	0.244*	0.042*
9	0.705	0.156	1.404	0.177	0.167	0.046
10	1.316	0.244	0.826	0.091	0.133	0.034
11	0.576	0.088	0.052	0.241	0.172	0.075
12	1.114*	0.189*	0.812*	0.104*	0.134*	0.037*
13	1.368	0.336	1.012	0.112	0.270	0.036
14	1.853	0.448	1.231	0.072	0.127	0.021
15	1.028	0.221	1.211	0.116	0.136	0.036
16	0.619	0.162	1.752	0.231	0.179	0.049
17	1.078	0.247	1.299	0.122	0.188	0.034
18	0.987	0.269	1.225	0.151	0.272	0.043
19	0.908	0.274	1.945	0.202	0.212	0.031
20	0.513	0.188	3.403	0.735	0.210	0.034
\bar{X}	0.987	0.331	0.673	0.527	0.280	0.056
S.D.	0.485	0.195	2.096	0.728	0.063	0.032

* ข้อสอบที่ไม่สอดคล้องกับโมเดล

ภาคผนวก ค

คำความชากและคำอำนาจจำแนกของแบบทดสอบ
วิชาคณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่อง ทศนิยม

ตาราง 49 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบ วิชาคณิตศาสตร์ (ค 101)
 ชั้นมัธยมศึกษาปีที่ 1 เรื่อง ทศนิยม

ข้อสอบ	ค่าความยาก	ค่าอำนาจจำแนก
1	.727	.525
2	.662	.576
3	.632	.657
4	.617	.552
5	.600	.541
6	.585	.460
7	.577	.528
8	.566	.463
9	.546	.343
10	.542	.576
11	.537	.477
12	.537	.518
13	.532	.562
14	.532	.676
15	.532	.461
16	.512	.408
17	.478	.321
18	.478	.515
19	.444	.496
20	.428	.345

ภาคผนวก ง

เนื้อหา จุดประสงค์เชิงพฤติกรรม

วิชาคณิตศาสตร์ (ค 101) ชั้นมัธยมศึกษาปีที่ 1 เรื่อง ทศนิยม

ตาราง 50 รายละเอียดเนื้อหาและจุดประสงค์เชิงพฤติกรรม วิชาคณิตศาสตร์ (ค 101)
 ชั้นมัธยมศึกษาปีที่ 1 เรื่อง ทศนิยม ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521
 (ฉบับปรับปรุง พ.ศ.2533)

เนื้อหา	จุดประสงค์เชิงพฤติกรรม
1. ค่าประจำหลักของทศนิยม	1. เมื่อกำหนดจำนวนในรูปทศนิยมมาให้ นักเรียนสามารถบอกค่าประจำหลักของเลขโดดที่อยู่ในรูปทศนิยมตำแหน่งต่าง ๆ ได้
2. การเปรียบเทียบทศนิยม	2. เมื่อกำหนดจำนวนในรูปทศนิยมตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถเปรียบเทียบว่าจำนวนเหล่านั้นได้
3. การบวกและการลบทศนิยม	3. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถบวกจำนวนเหล่านั้นได้
4. การคูณและการหารทศนิยม	4. เมื่อกำหนดโจทย์ปัญหาที่มีการบวกทศนิยมมาให้ นักเรียนสามารถหาคำตอบได้
5. การแทนเศษส่วนด้วยทศนิยม	5. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถลบจำนวนเหล่านั้นได้
	6. เมื่อกำหนดโจทย์ปัญหาที่มีการลบทศนิยม มาให้ นักเรียนสามารถหาคำตอบได้
	7. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถบอกและลบจำนวนเหล่านั้นได้
	8. เมื่อกำหนดโจทย์ปัญหาที่มีการบวกและการลบทศนิยมมาให้ นักเรียนสามารถหาคำตอบได้
	9. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถคูณจำนวนเหล่านั้นได้
	10. เมื่อกำหนดโจทย์ปัญหาที่มีการคูณทศนิยม มาให้ นักเรียนสามารถหาคำตอบได้
	11. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้ นักเรียนสามารถหารจำนวนเหล่านั้นได้

เนื้อหา	จุดประสงค์เชิงพฤติกรรม
	<p>12. เมื่อกำหนดโจทย์ปัญหาที่มีการหารทศนิยม มาให้นักเรียนสามารถหาคำตอบได้</p> <p>13. เมื่อกำหนดจำนวนในรูปทศนิยม ตั้งแต่ 2 จำนวนขึ้นไปมาให้นักเรียนสามารถคูณจำนวนเหล่านั้นได้</p> <p>14. เมื่อกำหนดโจทย์ปัญหาที่มีการคูณทศนิยมมาให้นักเรียนสามารถหาคำตอบได้</p> <p>15. เมื่อกำหนดจำนวนในรูปเศษส่วนมาให้นักเรียนสามารถเขียนให้อยู่ในรูปทศนิยมได้</p>

รายนามผู้เชี่ยวชาญที่พิจารณาคุณภาพของแบบทดสอบ

- | | |
|-------------------------|------------------|
| 1. รองศาสตราจารย์อังคณา | สายยศ |
| 2. อาจารย์ไทรรงค์ | เจนการ |
| 3. อาจารย์อรนุช | ฉัตรรัตน์พาณิชย์ |
| 4. อาจารย์จรีรัตน์ | นันทย์ทวีกุล |
| 5. อาจารย์ชลธิชา | สีลวิตตะ |
| 6. อาจารย์อรินทร์ | น่วมถนอม |

ประวัติย่อของผู้วิจัย

ชื่อ นางสาวรุ่งฤดี ชื่อสกุล ทรัพย์นิธิ
 เกิด วันที่ 21 กุมภาพันธ์ พุทธศักราช 2502
 สถานที่เกิด เขตสัมพันธวงศ์ กรุงเทพมหานคร
 สถานที่อยู่ปัจจุบัน 30/159 ถนนสุขุมวิท 1 คลองกุ่ม บึงกุ่ม กรุงเทพมหานคร
 ตำแหน่งหน้าที่ปัจจุบัน นักวิชาการสอบ
 สถานที่ทำงานปัจจุบัน สำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ

ประวัติการศึกษา

พ.ศ. 2521 มีธยมศึกษาตอนปลาย
 พ.ศ. 2526 บศ.บ (การเงินและการธนาคาร)
 จากมหาวิทยาลัยรามคำแหง
 พ.ศ. 2539 กศ.ม. (การวัดผลการศึกษา)
 จากมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

การศึกษาเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์
ระหว่างการวิเคราะห์ด้วยโมเดลโลจิสติก แบบหนึ่ง สอง สาม พารามิเตอร์
และขนาดกลุ่มตัวอย่างต่างกัน

ปริญญาโท

ของ

รุ่งฤดี ทรัพย์นิต

519.544

ธ 637ก 5.3

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

พฤษภาคม 2539

ลิขสิทธิ์ เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบ ระหว่างการวิเคราะห์ด้วยโมเดล โทจิสติก แบบหนึ่ง สอง สาม พารามิเตอร์ และขนาดกลุ่มตัวอย่าง 250 คน 500 คน 750 คน และ 1,000 คน โดยศึกษากับกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 1 เครื่องมือที่ใช้ในการวิจัยเป็นทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ แบบ 5 ตัวเลือก จำนวน 20 ข้อ การประมาณค่าพารามิเตอร์ข้อสอบและพารามิเตอร์ผู้สอบ ใช้โปรแกรมไบล็อก (BILOG) ด้วยวิธีการประมาณค่าแบบมาร์จินัลแมกซ์ิมัมไลกิลิฮูด (Marginal Maximum Likelihood)

ผลการวิจัยสรุปได้ดังนี้

1. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดล 2 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง เมื่อประมาณค่าด้วยโมเดล 3 พารามิเตอร์ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยากและการเดา มีค่าไม่แตกต่างกัน ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

2. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างกลุ่มตัวอย่างที่มีจำนวนไม่เท่ากัน เมื่อประมาณค่าด้วยโมเดล 1 พารามิเตอร์ มีค่าแตกต่างกัน เมื่อประมาณค่าด้วยโมเดล 2 และ 3 พารามิเตอร์ ปรากฏผลเช่นเดียวกันคือ ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

3. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ข้อสอบ ระหว่างโมเดล เมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน ความคลาดเคลื่อนมาตรฐานในการประมาณค่าอำนาจจำแนก มีค่าแตกต่างกัน ทุกขนาดกลุ่มตัวอย่าง ส่วนความคลาดเคลื่อนมาตรฐานในการประมาณค่าความยาก มีค่าแตกต่างกันในบางกลุ่มตัวอย่าง

4. ความคลาดเคลื่อนมาตรฐานในการประมาณค่าพารามิเตอร์ผู้สอบ ระหว่างโมเดล เมื่อวิเคราะห์ด้วยกลุ่มตัวอย่างที่มีจำนวนเท่ากัน มีค่าแตกต่างกัน จากทุกขนาดกลุ่มตัวอย่าง โดยโมเดล 3 พารามิเตอร์ แตกต่างจากโมเดล 1 และ 2 พารามิเตอร์

จากการศึกษาพบว่ากลุ่มตัวอย่างที่นำมาใช้ในการวิเคราะห์ และให้ค่าประมาณพารามิเตอร์ที่ดีที่สุด คือ กลุ่มตัวอย่าง 750 คน

A COMPARISON OF STANDARD ERROR OF PARAMETER ESTIMATION
BETWEEN ONE-, TWO- AND THREE-PARAMETER LOGISTIC MODEL
USING DIFFERENCE SAMPLE SIZE

AN ABSTRACT

BY

RUNGRUEDEE SUPNITI

Presented in partial fulfilment of the requirements for the Master
of Education Degree in Educational Measurement
at Srinakharinwirot University

May 1995

The purpose of this research were to compare the standard errors of items and abilities estimations between one-, two- and three-parameter logistic model with sample sizes of 250, 500, 750, and 1,000 examinees. The research instrument used were Mathematic achievement tests with five multiple-choices and test length of 20 items for Mathayomsuksa 1 students. Marginal Maximum Likelihood Procedure was used to estimate item and ability parameters.

The results of the research were summarized as follows.

1. The standard error of item parameter yielded by the one-parameter model was statistically significant differences by all sample sizes. Under the two-parameter model, the standard error of item difficulty was not significantly differences by all sample sizes, the standard error of item discrimination was statistically significant differences in some sample sizes. Under the three-parameter, the standard error of item difficulty and item guessing were not significantly difference by all sample sizes, the standard error of item discrimination was statistically significant difference by all sample sizes.

2. The standard error of ability parameter yielded by the one-parameter model was statistically significant differences by all sample sizes. Under the two- and three-parameter model, the standard error of ability parameter was statistically significant differences in some sample sizes.

3. The standard error of item discrimination yielded by one-, two-, and three-parameter model with the same sample size of all sample sizes were statistically significant differences. With respect to the standard error of item difficulty was statistically significant differences in some sample sizes.

4. The standard error of ability parameter yielded by one-, two-, and three-parameter model for the same sample size of all sample sizes were statistically significant differences, with

respect to the standard error of ability parameter yielded by the three-parameter differences from the others.

It was found that the sample size of 750 examinees yielded the most significant consistency parameter estimation.