



บทที่ 3

วิธีดำเนินการวิจัย

การวิจัย เพื่อ เปรียบเทียบประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก 3 พารามิเตอร์ ระหว่างวิธีแมกซ์ิมไลค์ลิสต์ วิธีฮิวริสติก และวิธีของเบย์ ในแบบสอขวดผลสัมฤทธิ์ และแบบสอบความถนัด มีขั้นตอนในการดำเนินการวิจัยดังนี้

ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการวิจัยครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2531 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ จังหวัดจันทบุรี จำนวน 2,125 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยคือ นักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2531 สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ในจังหวัดจันทบุรี ซึ่งทำการสุ่มตัวอย่างแบบแบ่งชั้น (Stratified Sampling) โดยมีขั้นตอนดังต่อไปนี้

1. สํารวจขนาดประชากรในแต่ละอำเภอ จากทั้งหมด 7 อำเภอในจังหวัดจันทบุรี ดังนี้

1.1	อำเภอเมือง	จำนวนนักเรียน	902 คน
1.2	อำเภอขลุง	จำนวนนักเรียน	246 คน
1.3	อำเภอท่าใหม่	จำนวนนักเรียน	478 คน

1.4	อำเภอโป่งน้ำร้อน	จำนวนนักเรียน	77 คน
1.5	อำเภอมะขาม	จำนวนนักเรียน	140 คน
1.6	อำเภอแหลมสิงห์	จำนวนนักเรียน	103 คน
1.7	กิ่งอำเภอสอยดาว	จำนวนนักเรียน	179 คน

2. คำนวณหาขนาดที่พอดีของกลุ่มตัวอย่างที่ใช้ในการวิจัยโดยใช้สูตร

$$n = \frac{N}{1 + N e^2} \quad (\text{Taro Yamane 1967 : 14})$$

เมื่อ n แทน ขนาดของกลุ่มตัวอย่าง

N แทน ขนาดประชากร

e แทน ความคลาดเคลื่อนที่จะยอมรับได้ (ผู้วิจัยยอมให้เกิดความคลาดเคลื่อนได้ 5 เปอร์เซ็นต์)

จากจำนวนนักเรียน 2,125 คน สามารถคำนวณหาขนาดกลุ่มตัวอย่างที่มีขนาดพอดีได้ 337 คน แต่เนื่องจากผู้วิจัยต้องการให้เหมาะสมกับรูปแบบการวิเคราะห์โดยแบบจำลองโลจิสติก 3 พารามิเตอร์ ในอันที่จะทำให้ค่าพารามิเตอร์ที่ได้ในแต่ละวิธีประมาณค่า มีความคงที่มากขึ้น กลุ่มผู้วิจัยจึงใช้ขนาดของกลุ่มตัวอย่างประมาณ 1,000 คน (Warm 1978 : 109) ซึ่งมากกว่ากลุ่มตัวอย่างที่มีขนาดพอดี ที่ผู้วิจัยคำนวณได้จากสูตรข้างต้น

3. คำนวณหาขนาดของกลุ่มตัวอย่างที่พอดีของประชากรในแต่ละอำเภอ โดยใช้สูตร ดังนี้ (Freund 1984 : 239 - 240)

$$n_i = \frac{N_i \cdot n}{N}$$

เมื่อ n_i	แทน	ขนาดของกลุ่มตัวอย่างที่พหุคูณในแต่ละอำเภอ
N_i	แทน	ขนาดของประชากรในแต่ละอำเภอ
n	แทน	ขนาดของกลุ่มตัวอย่างทุกระดับ
N	แทน	ขนาดของประชากรทั้งหมด

จากการคำนวณได้กลุ่มตัวอย่างในแต่ละอำเภอ ดังนี้

3.1	อำเภอเมือง	จำนวนนักเรียน	425	คน
3.2	อำเภอขลุง	จำนวนนักเรียน	116	คน
3.3	อำเภอท่าใหม่	จำนวนนักเรียน	225	คน
3.4	อำเภอโป่งน้ำร้อน	จำนวนนักเรียน	36	คน
3.5	อำเภอมะขาม	จำนวนนักเรียน	66	คน
3.6	อำเภอแหลมสิงห์	จำนวนนักเรียน	48	คน
3.7	กิ่งอำเภอสอยดาว	จำนวนนักเรียน	84	คน

4. หลังจากได้จำนวนกลุ่มตัวอย่างในแต่ละอำเภอแล้ว จึงทำการสุ่มกลุ่มตัวอย่างด้วยวิธีการสุ่มอย่างง่าย โดยใช้ห้องเรียนเป็นหน่วยการสุ่ม กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ ผู้วิจัยได้แบ่งออกเป็น 2 กลุ่มคือ กลุ่มที่ 1 เป็นกลุ่มที่ใช้ในการทดลองสอบ (try out) เพื่อจะได้ศึกษาปัญหาต่าง ๆ เช่น การดำเนินการสอบ เวลาที่กำหนดให้นักเรียนตอบ เพื่อศึกษาคุณภาพของแบบสอบและเพื่อความสมบูรณ์ของกระบวนการสอบในขั้นดำเนินการจริง ซึ่งจำนวนนักเรียนที่ใช้ในการทดลองสอบ มีทั้งหมด 97 คน เป็นนักเรียนโรงเรียนบางกะจะ อำเภอเมือง จังหวัดจันทบุรี จำนวน 2 ห้องเรียน

กลุ่มที่ 2 เป็นกลุ่มนักเรียนที่ใช้ในการวิจัย เพื่อนำผลการสอบไปวิเคราะห์หาค่าพารามิเตอร์ของข้อสอบ และของผู้เข้าสอบ ด้วยวิธีแมกซิมัมไลค์ลิตูด วิธีฮิวริสติก และวิธีของเบย์ แต่เนื่องจากการวิจัยครั้งนี้ต้องทำการเก็บข้อมูลผลการตอบแบบสอบ 2 ครั้ง ใน 2 แบบสอบจากนักเรียนแต่ละคน ทำให้มีโอกาสนักเรียนบางคนจะเข้าสอบไม่ครบ 2 ครั้ง ผู้วิจัยจึงต้องสุ่มตัวอย่างเกินจากที่กำหนดไว้เพื่อป้องกันปัญหาที่อาจเกิดขึ้นดังกล่าว ซึ่งได้จำนวนนักเรียนที่เป็นกลุ่มตัวอย่างทั้งหมด 1,123 คน ดังแสดงไว้ในตารางที่ 4

ตารางที่ 4 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย จำแนกตามอำเภอและโรงเรียน

อำเภอ	โรงเรียน	จำนวน	
		ห้องเรียน	นักเรียน (คน)
เมือง	เบญจมาชุตีศ	4	194
	ศรียานุสรณ์	5	243
ขลุง	ขลุงรัชดาภิเษก	3	135
ท่าใหม่	ท่าใหม่พูลสวัสดิ์ราษฎร์นุกูล	4	164
	ศรีรัตนราษฎร์เคราะห์	1	35
	มัธยมวัดเขาสุกิม	1	37
	นายายอามพิทยาคม	1	39
โป่งน้ำร้อน	โป่งน้ำร้อนพิทยาคม	1	28
มะขาม	มะขามสรรเสริญ	2	77
แหลมสิงห์	แหลมสิงห์พิทยาคม	2	76
สอยดาว	สอยดาววิทยา	3	95
รวม		27	1,123

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นแบบสอบ 2 ฉบับ คือ

1. แบบสอบผลสัมฤทธิ์วิชาภาษาไทย (การใช้ภาษาไทย) เป็นแบบเลือกตอบ ที่มี 5 ตัวเลือก จำนวน 40 ข้อ ใช้เวลาทำ 40 นาที สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของสำนักทดสอบทางการศึกษาและจิตวิทยา มศว ประสานมิตร สังกัดทบวงมหาวิทยาลัย ซึ่งมีค่าความยากของข้อสอบตั้งแต่ 0.14 - 0.82 ค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.18 - 0.61 มีค่าความเที่ยงที่คำนวณโดยสูตรคูเคอร์ริชาร์ดสันที่ 21 เท่ากับ 0.7567 มีความตรงตามเนื้อเรื่อง และมีเกณฑ์ปกติปลายปีระดับประเทศซึ่งกลุ่มผู้สร้างได้ดำเนินการสร้างแบบสอบตามลำดับขั้น ดังนี้

ขั้นที่ 1 จัดตั้งคณะกรรมการสร้างแบบสอบมาตรฐาน

ขั้นที่ 2 กำหนดวัตถุประสงค์ของโครงการสร้างแบบสอบมาตรฐาน

ขั้นที่ 3 กำหนดขอบข่ายงาน และวางแผนดำเนินการ ซึ่งมีขั้นตอนที่

ต่อเนื่อง ดังนี้

- 3.1 ประชุมคณะกรรมการประจำกลุ่มวิชาภาษาไทย เพื่อวิเคราะห์หลักสูตร กำหนดโครงสร้างของวิชาภาษาไทยฉบับการใช้ภาษาไทยกำหนดจำนวนข้อและ เวลาสอบและออกข้อสอบ
- 3.2 ประชุมคณะกรรมการเพื่อวิจารณ์ข้อสอบ และเลือกข้อสอบที่วิจารณ์แล้ว เข้าสู่ชุด และส่งพิมพ์ต้นฉบับ
- 3.3 กำหนดกลุ่มตัวอย่างเพื่อออกทดสอบภาคสนาม เพื่อทดลองข้อสอบครั้งที่ 1 นำผลมาวิเคราะห์รายข้อ หาค่าความยาก ค่าอำนาจจำแนก และความเที่ยง
- 3.4 ปรับปรุงแก้ไขข้อสอบจากผลการวิเคราะห์จากข้อ 3.3 แล้วนำไปทดลองสอบใหม่อีก 3-4 ครั้ง แต่ละครั้งวิเคราะห์ผลแล้วปรับปรุงจนมั่นใจว่าเป็นแบบสอบที่มีคุณภาพสูง ทั้งค่าความยาก ค่าอำนาจจำแนก และความเที่ยง

- 3.5 สร้างเกณฑ์ปกติ (Norm) ของแบบสอบจากกลุ่มตัวอย่างของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ซึ่งเลือกมาจากประชากรทั้งประเทศ
- 3.6 ศึกษาความตรงของแบบสอบ โดยหาสหสัมพันธ์ ระหว่างคะแนนจากการสอบ แบบสอบที่สร้างขึ้น กับ ผลสัมฤทธิ์ทางการเรียนภาษาไทย ด้านการใช้ภาษาไทยของนักเรียนในโรงเรียน
- (อรุณศรี กุมุท 2525 : 67-82)

2. แบบสอบความถนัดจำแนกด้านการใช้ภาษาไทย เป็นแบบเลือกตอบที่มี 5 ตัวเลือก จำนวน 80 ข้อ ใช้เวลา 40 นาที สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของ อัมพิกา สุริยินทร์ ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่มีค่าความยากของข้อสอบตั้งแต่ 0.13 - 0.89 ค่าอำนาจจำแนกของข้อสอบตั้งแต่ 0.06-0.59 มีค่าความเที่ยงที่คำนวณโดยสูตรคูเดอรีชาร์ดสันที่ 20 เท่ากับ 0.863 มีค่าความตรงร่วมสมัย 0.503 และมีเกณฑ์ปกติเปอร์เซ็นต์ไทล์ ซึ่งผู้สร้างได้ดำเนินการพัฒนาแบบสอบตามลำดับขั้นดังนี้

ขั้นที่ 1 การสร้างแบบสอบ

- 1.1 ตั้งวัตถุประสงค์ ของการสร้างแบบสอบ
- 1.2 ศึกษาเอกสารที่เกี่ยวข้องกับการสร้างแบบสอบ
- 1.3 กำหนดลักษณะ และเนื้อหาของแบบสอบ
- 1.4 ปรับปรุงแบบสอบ

ขั้นที่ 2 การเลือกกลุ่มตัวอย่างประชากร

ขั้นที่ 3 การดำเนินการสอบ และการเก็บรวบรวมข้อมูล

ขั้นที่ 4 การวิเคราะห์ข้อมูล

(อัมพิกา สุริยินทร์ 2522 : 32-37)

ซึ่งจากการศึกษากระบวนการสร้าง และพัฒนาแบบสอบทั้ง 2 ฉบับ ประกอบกับ
คุณภาพด้านต่าง ๆ ของแบบสอบแล้ว ผู้วิจัยจึงถือว่า แบบสอบทั้ง 2 ฉบับนี้ เป็นแบบสอบที่มี
คุณภาพอยู่ในระดับที่เชื่อถือได้

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล โดยนำหนังสือรับรองจากบัณฑิตวิทยาลัยไป
ติดต่อขอความร่วมมือกับสำนักงานศึกษาธิการจังหวัดจันทบุรี ให้แจ้งไปยังโรงเรียนที่เป็น
กลุ่มตัวอย่าง ที่ใช้ในการทดลอง เครื่องมือและเก็บข้อมูลจริง จากนั้นผู้วิจัยได้นำหนังสือขอ
ความร่วมมือในการวิจัยไปยังโรงเรียนเหล่านั้น ซึ่งมีวิธีการดำเนินการเก็บรวบรวมข้อมูลดังนี้

1. ขั้นการทดลองเครื่องมือที่ใช้ในการวิจัย

การทดลองทำโดยนำแบบสอบ ทั้ง 2 ฉบับ คือ แบบสอบผลสัมฤทธิ์วิชา
ภาษาไทย (การใช้ภาษาไทย) และแบบสอบความถนัดจำแนกด้านการใช้ภาษาไทย ไปสอบ
กับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนบางกะจะ อำเภอเมือง จังหวัดจันทบุรี จำนวน
97 คน โดยมีจำนวนนักเรียนที่ตอบแบบสอบแต่ละฉบับเป็น 49 คน และ 48 คน ตามลำดับ
เพื่อเป็นการชั่งชั่งการดำเนินการสอบ รวมทั้งการศึกษาเวลา และปัญหาต่าง ๆ เพื่อ
นำมาแก้ไขในการดำเนินการสอบจริงให้มีความสมบูรณ์ และมีคุณภาพมากยิ่งขึ้น สำหรับการ
ศึกษาเรื่องเวลาที่ใช้ในการทำข้อสอบนั้น ได้ตรวจสอบว่า เวลาที่กำหนดให้นักเรียนใช้ตอบ
แบบสอบ นักเรียนทำเสร็จทันเวลาหรือไม่ ทั้งนี้เพราะ ข้อตกลงเบื้องต้นของทฤษฎีการตอบ
สนองต่อข้อสอบนั้น กำหนดว่า ต้องไม่มีเงื่อนไขของความเร็ว หรือเวลาเข้ามาเกี่ยวข้อง
หรือมีอิทธิพลต่อการตอบสนองต่อข้อสอบ ผู้เข้าสอบทุกคนต้องมีโอกาสทำถึงข้อสุดท้าย
(Hambleton and Swaminathan 1985 : 30) จึงต้องกำหนดเวลาให้นักเรียน
ทำแบบสอบ เท่าที่จะมีเวลาตอบข้อสอบทุกข้อ ผลจากการทดลองสอบ ปรากฏว่า นักเรียน
ทำได้ทันตามเวลาที่กำหนดไว้ โดยเฉพาะแบบสอบความถนัด แม้จะมีจำนวนถึง 80 ข้อ แต่
นักเรียนส่วนใหญ่ทำเสร็จตามเวลาที่กำหนดไว้ เช่นเดียวกัน ดังนั้นผู้วิจัยกำหนดเวลาให้

นักเรียนตอบแบบสอบทั้งสองฉบับเท่ากัน คือ 40 นาที แล้วนำผลการสอบมาวิเคราะห์ หาค่าความเที่ยง ด้วยสูตรคูเคอร์ริชาร์ดสัน สูตรที่ 20 ซึ่งได้ค่าความเที่ยงของแบบสอบผลสัมฤทธิ์เท่ากับ 0.7176 และ แบบสอบความถนัด เท่ากับ 0.8395 ค่าความเที่ยงของแบบสอบที่ได้นี้มีค่าใกล้เคียงกับค่าความเที่ยงของแบบเดิม ก่อนที่จะนำมาทดลองใช้ คือ 0.7576 และ 0.863 ตามลำดับ ดังนั้นถือว่า แบบสอบผลสัมฤทธิ์ และแบบสอบความถนัดนี้ เมื่อนำไปใช้กับกลุ่มตัวอย่างประชากรนักเรียนชั้นมัธยมศึกษาปีที่ 3 จังหวัดจันทบุรี ปีการศึกษา 2531 ยังคงเป็นแบบสอบที่มีความเที่ยงอยู่ในระดับที่เชื่อถือได้

2. ขั้นตอนการเก็บข้อมูลจากกลุ่มตัวอย่างที่ใช้ในการวิจัย

2.1 ในการสอบผู้วิจัยได้ติดต่ออาจารย์ผู้สอน เพื่อชี้แจงวัตถุประสงค์ของการวิจัย และนัดหมายเรื่อง วัน เวลา และสถานที่สอบ ซึ่งนักเรียนแต่ละคน จะต้องทำการตอบแบบสอบ ทั้ง 2 ฉบับ โดยได้ดำเนินการสอบแบบสอบความถนัด ก่อนสอบปลายภาค ประมาณ 2 เดือน และดำเนินการสอบแบบสอบวัดผลสัมฤทธิ์ ก่อนสอบปลายภาค 1-2 สัปดาห์ เพื่อให้เป็นช่วงเวลาที่เหมาะสมในอันที่จะทำให้แบบสอบมีความตรงเชิงทำนาย และความตรงร่วมสมัยตามลำดับ

2.2 ผู้วิจัยนำแบบสอบทั้ง 2 ฉบับ ไปทำการทดสอบนักเรียน จำนวน 1,123 คน จากโรงเรียนที่แสดงไว้ในตารางที่ 4 ส่วนการบริหารการสอบ ผู้วิจัยได้ขอความร่วมมือจากอาจารย์ผู้สอน ให้ช่วยในการดำเนินการสอบ

3. ขั้นตอนการบันทึกข้อมูล

หลังจากได้ดำเนินการสอบแล้ว นำกระดาษคำตอบมาจัดระบบซึ่งจะคัดเลือกเฉพาะผลการสอบของนักเรียนที่ทำการตอบแบบสอบทั้ง 2 ฉบับ และในแต่ละฉบับต้องตอบข้อสอบไม่ต่ำกว่า 1 ใน 3 ของจำนวนข้อสอบทั้งหมด ซึ่งเป็นเงื่อนไขของกระบวนการประมาณค่าในโปรแกรมโลจิสต์ (Wingersky 1982 : 8) จากการคัดเลือกทำให้เหลือจำนวนกระดาษคำตอบของนักเรียนที่เข้าสอบ ไว้ทำการประมาณค่าพารามิเตอร์จำนวนทั้งสิ้น 1,028 คน นำมาตรวจให้คะแนน และบันทึกลงในกระดาษลงรหัส แล้วจัดเตรียมข้อมูลลงเทป เพื่อนำไปวิเคราะห์ต่อไป

4. การเก็บข้อมูลคะแนนเกณฑ์

ผู้วิจัยติดต่อขอคะแนนสอบปลายภาควิชาภาษาไทย ด้านการใช้ภาษาไทย ของนักเรียนที่เป็นกลุ่มตัวอย่าง จากครูผู้สอน เพื่อนำมาเป็นคะแนนเกณฑ์

อนึ่ง แบบสอบปลายภาควิชาภาษาไทย ด้านการใช้ภาษาไทยระดับชั้นมัธยมศึกษาปีที่ 3 ซึ่งโรงเรียนมัธยมศึกษาตอนต้นทุกโรงเรียน ในจังหวัดจันทบุรี ใช้แบบสอบ เดียวกันนั้น เป็นแบบสอบวัดผลสัมฤทธิ์ที่สร้างโดยกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดจันทบุรี และมีกระบวนการสร้างข้อสอบที่เป็นระบบคือ

1. ตั้งคณะกรรมการออกข้อสอบ ซึ่งมาจากครูที่ทำการสอนรายวิชาภาษาไทย ชั้นมัธยมศึกษาปีที่ 3 โรงเรียนละ 1 คน
2. กำหนดตารางวิเคราะห์เนื้อหา
3. สร้างข้อสอบตามตารางวิเคราะห์เนื้อหา (ซึ่งขั้นตอนทั้ง 3 นี้ ต้องแล้วเสร็จก่อนเปิดภาคเรียนที่ 2)
4. นำข้อสอบรวมเป็นฉบับ แล้วนำไปสอบนักเรียนชั้นมัธยมศึกษาปีที่ 3 ทั้งจังหวัด โดยทำการสอบ 2 ครั้ง คือ
 - 4.1 สอบก่อนสอน (Pre - Test)
 - 4.2 สอบปลายภาคเรียน (Post - Test)
5. ทำการวิเคราะห์หาคุณภาพของข้อสอบเป็นรายข้อ โดยใช้ค่าสถิติในแนวอิงเกณฑ์ คือ ค่าความยาก และ ค่าดัชนีความไว (Sensitivity Index) (กรมวิชาการ 2530 : 152 - 156 อ้างถึงใน กลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดจันทบุรี 2531 : 25 - 30) และคัดเลือกข้อสอบที่มีคุณภาพดีไว้ใช้ในปีต่อไป

ด้วยเหตุนี้ ผู้วิจัย จึงถือว่า แบบสอบวัดผลสัมฤทธิ์วิชาภาษาไทย ด้านการใช้ภาษาไทย ของกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดจันทบุรีนี้ มีคุณภาพเชื่อถือได้ ส่วนคะแนนผลการสอบปลายภาคในวิชาภาษาไทยของนักเรียนนั้น เป็นคะแนนที่นักเรียนทำแบบสอบอย่างเต็มความสามารถ เพราะเป็นการสอบในสถานการณ์จริง และคะแนนสอบมีผลต่อการจบและไม่จบหลักสูตรมัธยมศึกษาตอนต้น ดังนั้นผู้วิจัยจึงถือว่า คะแนนสอบปลายภาควิชา

ภาษาไทย ด้านการใช้ภาษา ของนักเรียนที่เป็นกลุ่มตัวอย่าง เกิดจากกระบวนการวัดที่มีคุณภาพ สามารถนำมาเป็นคะแนนเกณฑ์ของการวิจัยครั้งนี้ได้

การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลสำหรับการวิจัยนี้ ผู้วิจัยทำการวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ของสถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย โดยใช้โปรแกรมสำเร็จรูปได้แก่ โปรแกรม Logist 5 version 2.5 และโปรแกรม SPSS x นอกจากนี้ยังใช้คอมพิวเตอร์ภาษาฟอร์แทรน (FORTRAN) ในการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ และวิธีอีวีรอสติก รวมทั้งใช้คำนวณค่าฟังก์ชันสารสนเทศของแบบสอบ และค่าสถิติต่าง ๆ ที่ใช้ในการทดสอบสมมุติฐาน ซึ่งขั้นตอนต่าง ๆ ในการวิเคราะห์ข้อมูลมีดังนี้

1. การตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว และความเป็นอิสระในการตอบข้อสอบ ตามข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองต่อข้อสอบได้กำหนดไว้ว่าแบบสอบที่จะใช้วิเคราะห์ด้วยแบบจำลองต่าง ๆ ในกรอบทฤษฎีนี้ จะต้องมียุทธศาสตร์การวัดเพียงมิติเดียวและความเป็นอิสระในการตอบข้อสอบ ผู้วิจัยจึงทำการตรวจสอบคุณสมบัติดังกล่าว โดยนำคะแนนผลการสอบของแบบสอบวัดผลสัมฤทธิ์ และแบบสอบความถนัด มาทำการวิเคราะห์ตัวประกอบของแบบสอบแต่ละฉบับ โดยใช้โปรแกรม SPSS x การวิเคราะห์ตัวประกอบนี้ ใช้การวิเคราะห์ตัวประกอบสำคัญ (Principle Component) แล้วหมุนแกนด้วยวิธีแวนิแมกซ์ (Varimax) ซึ่งการตรวจสอบคุณสมบัติการวัดเพียงมิติเดียว ก็จะพิจารณาจากค่าไอเกน (Eigen Value) ถ้าหากค่าไอเกนของตัวประกอบที่ 1 ของแบบสอบมีค่าสูงกว่าค่าไอเกนของตัวประกอบที่ 2 อย่างมาก และค่าไอเกนของตัวประกอบที่ 2 มีค่าสูงกว่าค่าไอเกนของตัวประกอบถัดไปเพียงเล็กน้อย จึงจะกล่าวได้ว่า แบบสอบนั้น มีคุณสมบัติการวัดเพียงมิติเดียว (Lord 1980 : 21) และการตรวจสอบคุณสมบัติความเป็นอิสระในการตอบข้อสอบก็ได้พิจารณาว่าถ้าแบบสอบมีคุณสมบัติการวัดเพียงมิติเดียว ก็จะมีคุณสมบัติความเป็นอิสระในการตอบข้อสอบด้วย (Hambleton and Swaminathan 1985 : 24)

2. การประมาณค่าพารามิเตอร์ อำนาจจำแนก (a_j) ค่าความยาก (b_j) ค่าการเดา (c_j) ของข้อสอบ และค่าความสามารถของผู้เข้าสอบ (θ_i) โดยวิธีประมาณค่าที่แตกต่างกัน 3 วิธี ดังนี้

2.1 การประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้เข้าสอบจากผลการสอบของกลุ่มตัวอย่างในแบบสอบแต่ละฉบับ ด้วยวิธีแมกซิมัมไลค์ลิฮูด โดยใช้โปรแกรม Logist 5 version 2.5 ของสถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นการวิเคราะห์ข้อสอบรายข้อตามแบบจำลองโลจิสติกที่มีพารามิเตอร์ 3 ตัว โดยสามารถเขียนเป็นรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ ได้ดังนี้

$$P_j(\theta) = c_j + (1 - c_j) \frac{e^{Da_j(\theta - b_j)}}{1 + e^{Da_j(\theta - b_j)}} \quad (j = 1, 2, \dots, m)$$

- เมื่อ $P_j(\theta)$ = ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ θ ตอบข้อสอบข้อ j ถูก
 θ = ระดับความสามารถของผู้สอบ
 a_j = ค่าอำนาจจำแนกของข้อสอบข้อที่ j
 b_j = ค่าความยากของข้อสอบข้อที่ j
 c_j = ค่าการเดาของข้อสอบข้อที่ j
 D = a scaling factor มีค่าเท่ากับ 1.7
 e = ค่าคงที่ มีค่าเท่ากับ 2.71828.....

2.2 การประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของผู้เข้าสอบ จากผลการสอบของกลุ่มตัวอย่าง ในแบบสอบแต่ละฉบับด้วยวิธีฮิวริสติก โดยใช้โปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรน ซึ่งสามารถเขียนรูปแบบความสัมพันธ์เชิงคณิตศาสตร์ของค่าพารามิเตอร์ ได้ดังนี้ (Warm 1987 : 51; Jensen 1976 cited by Hambleton and Swaminathan 1985 : 146)

$$a_j = \frac{D_j (P_j Q_j)^{1/2}}{[(KR-20)(1-C_j)^2 Y_j - D_j P_j Q_j]^{1/2}} \quad j = 1, 2, \dots, m$$

$$b_j = \frac{Y_j Z_j (KR-20)^{1/2} (1-C_j)}{D_j (P_j Q_j)^{1/2}} \quad j = 1, 2, \dots, m$$

c_j = "สัดส่วนของผู้ตอบถูกที่อยู่ที่ปลายล่างของโค้งที่เกิดจากการพล็อตจำนวนคนตอบถูกสะสมในแต่ละระดับคะแนนรวม หลังจากหักคะแนนของข้อที่กำลังหาค่าพารามิเตอร์ ออกจากคะแนนรวมทั้งฉบับก่อน"

เมื่อ a_j = ค่าอำนาจจำแนกของข้อที่ j

b_j = ค่าความยากของข้อที่ j

c_j = ค่าการเดาของข้อที่ j

C_j = ค่าการเดาของข้อที่ j ในทฤษฎีการทดสอบแบบดั้งเดิม
(สัดส่วนกลับของจำนวนตัวเลือกในข้อที่ j)

D_j = ค่าอำนาจจำแนกของข้อที่ j ในทฤษฎีการทดสอบแบบดั้งเดิม
(Point biserial correlation = $(\bar{X}_R - \bar{X}_T) / SD_x (P'_j / Q'_j)^{1/2}$)

P'_j = ค่าความยากของข้อที่ j ในทฤษฎีการทดสอบแบบดั้งเดิม
(สัดส่วนของผู้ตอบถูก ต่อผู้ตอบทั้งหมดในข้อที่ j)

P_j = $(P'_j - C_j) / (1 - C_j)$

Q'_j = $1 - P'_j$

Q_j = $1 - P_j$

$KR-20$ = ค่าความเที่ยงของแบบสอบ ที่คำนวณโดยสูตรคูเดอริชาร์ดสันสูตรที่ 20

Z_j = ค่าของคะแนนมาตรฐาน ที่มีพื้นที่ใต้โค้งปกติด้านขวามือ เท่ากับ P_j

Y_j = ค่าความสูงของโค้งปกติ ณ จุด Z_j

\bar{X}_R = คะแนนเฉลี่ยของคะแนนรวมของกลุ่มที่ทำข้อสอบข้อ j ถูก

\bar{X}_i = คะแนนเฉลี่ยของคะแนนรวมของผู้เข้าสอบทุกคน

SD_x = ค่าเบี่ยงเบนมาตรฐานของคะแนนรวมของผู้เข้าสอบทุกคน

ส่วนการประมาณค่าความสามารถของบุคคลใช้วิธี Conditional Maximum Likelihood Estimation of Ability ซึ่งเป็นการประมาณค่าอย่างมีเงื่อนไข ภายหลังจากทราบค่าพารามิเตอร์ของข้อสอบแต่ละข้อแล้ว โดยขบวนการนิวตันกราฟสัน และ กำหนดค่าความสามารถเริ่มต้นสำหรับผู้สอบคนที่ a เป็น

$$\theta_{0a} = \ln [r_a / (n-r_a)] = \theta_0$$

เมื่อ r_a = จำนวนข้อสอบที่ผู้สอบคนที่ a ตอบถูก

n = จำนวนข้อสอบที่ผู้สอบคนที่ a ทำทั้งหมด

\ln = natural logarithm

จะได้ค่า θ ของการทำซ้ำครั้งที่ $m+1$ นั้นคือ

$$\theta_{m+1} = \theta_m - h_m$$

$$\text{โดยที่ } h_m = \left[\frac{\partial \ln L}{\partial \theta_a} \right] / \left[\frac{\partial^2 \ln L}{\partial \theta_a^2} \right]$$

$$\frac{\partial \ln L}{\partial \theta_a} = D \sum_{j=1}^n a_j (U_{ja} - P_{ja}) (P_{ja} - c_j) / [P_{ja} (1 - c_j)]$$

$$\frac{\partial^2 \ln L}{\partial \theta_a^2} = D^2 \sum_{j=1}^n a_j (P_{ja} - c_j) (U_{ja} - c_j - P_{ja}) Q_{ja} / [P_{ja} (1 - c_j)^2]$$

- เมื่อ U_{ja} = ค่าผลการตอบ (ถูกเป็น 1 ผิดเป็น 0) ของผู้ตอบคนที่ a
 P_{ja} = โอกาสที่ผู้ตอบคนที่ a ทำข้อสอบข้อ j ถูกต้อง
 $= c_j + (1-c_j)/(1+\exp(-Da_j(\theta_a-b_j)))$
 D = a scaling factor มีค่าเท่ากับ 1.7
 θ_a = ค่าความสามารถของผู้สอบคนที่ a
 a_j = ค่าอำนาจจำแนกของข้อที่ j ที่ประมาณโดยวิธีชิวริสติก
 b_j = ค่าความยากของข้อที่ j ที่ประมาณโดยวิธีชิวริสติก
 c_j = ค่าการเดาของข้อที่ j ที่ประมาณโดยวิธีชิวริสติก

ซึ่งเมื่อ $|h_m| = .001$ จึงสิ้นสุดการประมาณค่าความสามารถของบุคคลที่ a
 (Hambleton and Swaminathan 1985 : 81-88)

2.3 การประมาณค่าพารามิเตอร์ของข้อสอบ และค่าความสามารถของ
 ผู้เข้าสอบ จากผลการสอบของกลุ่มตัวอย่างในแบบสอบแต่ละฉบับด้วยวิธีของเบย์ ดำเนินการ
 ดังนี้

กำหนดค่าเริ่มต้นที่จะใช้ในการประมาณค่าพารามิเตอร์ โดย

$$\theta^{(0)} = \ln \left[\frac{q_1}{1-q_1} \right]$$

$$b_j^{(0)} = \frac{Z_j}{D_j}$$

$$a_j^{(0)} = \frac{D_j}{(1-D_j)^{1/2}}$$

$$c_j^{(0)} = \frac{1}{m_j}$$

$$q_i = \frac{\sum_{j=1}^n u_{ij}}{n} \quad , \quad P_j = \frac{\sum_{i=1}^N u_{ij}}{N}$$

- เมื่อ u_{ij} = ผลการตอบข้อที่ j ของนักเรียนคนที่ i
 Z_j = ค่าของคะแนนมาตรฐานที่พื้นที่ใต้โค้งปกติด้านขวามีค่าเท่ากับ P_j
 D_j = ค่า Point - biserial Correlation ระหว่างคะแนนรายข้อ กับคะแนนรวม
 m_j = จำนวนตัวเลือกของข้อที่ j

และทำการประมาณค่าพารามิเตอร์ด้วยวิธีการหาค่าซ้ำ ๆ จนได้ค่าคงที่ (Iterative) ดังนี้

ขั้นที่ 1 ประมาณค่าความสามารถของผู้สอบ

$$\theta_i^{(k+1)} = \theta_i^{(k)} - \frac{g[\theta_i^{(k)}]}{h[\theta_i^{(k)}]}$$

ขั้นที่ 2 ประมาณค่าอำนาจจำแนก a ค่าความยาก b และค่าการเดา c ของข้อสอบ

$$a_j^{(k+1)} = a_j^{(k)} - \frac{g[a_j^{(k)}]}{h[a_j^{(k)}]}$$

$$b_j^{(k+1)} = b_j - \frac{g[b_j^{(k)}]}{h[b_j^{(k)}]}$$

$$c_j^{(k+1)} = c_j - \frac{g[c_j^{(k)}]}{h[c_j^{(k)}]}$$

เมื่อ $\theta_i^{(k)}$ = ค่าความสามารถของคนที i ที่ได้จากการประมาณครั้งที่ k

$a_j^{(k)}$ = ค่าอำนาจจำแนกของข้อที่ j ที่ได้จากการประมาณครั้งที่ k

$b_j^{(k)}$ = ค่าความยากของข้อที่ j ที่ได้จากการประมาณครั้งที่ k

$c_j^{(k)}$ = ค่าการเดาของข้อที่ j ที่ได้จากการประมาณครั้งที่ k

$$g[\theta_i^{(k)}] = D \sum_{j=1}^n a_j (P_{ij} - c_j) (U_{ij} - P_{ij}) / P_{ij} (1 - c_j) - \theta_i$$

$$h[\theta_j^{(k)}] = D^2 \sum_{j=1}^n a_j^2 (P_{ij} - c_j) (U_{ij} - P_{ij})^2 \theta_{ij} / P_{ij}^2 (1 - c_j)^2 - 1$$

$$g[a_j^{(k)}] = D \sum_{i=1}^N (\theta_i - b_j) (P_{ij} - c_j) (U_{ij} - P_{ij}) / P_{ij} (1 - c_j)$$

$$+ (V_j - 1) / a_j - a_j / W_j$$

$$h[a_j^{(k)}] = D^2 \sum_{i=1}^N (\theta_i - b_j)^2 (P_{ij} - c_j) (U_{ij} c_j - P_{ij}^2) Q_{ij} / P_{ij} (1 - c_j)^2$$

$$- (V_j - 1) / a_j - 1 / W_j$$

$$g[b_j^{(k)}] = -D \sum_{i=1}^N a_j (P_{ij} - c_j) (U_{ij} - P_{ij}) / P_{ij} (1 - c_j)$$

$$h[b_j^{(k)}] = D^2 \sum_{i=1}^N a_j^2 (P_{ij} - c_j) (U_{ij} c_j - P_{ij}^2) Q_{ij}^2 / P_{ij} (1 - c_j)^2$$

$$g[c_j^{(k)}] = D \sum_{i=1}^N (U_{ij} - P_{ij}) / P_{ij} (1 - c_j) + S_j / c_j - t_j / (1 - c_j)$$

$$h[c_j^{(k)}] = D^2 \sum_{i=1}^N [U_{ij}(2P_{ij}-1) - P_{ij}^2] / P_{ij}^2 (1-c_j)^2$$

$$-s_j / c_j^2 + t_j / (1-c_j)^2$$

U_{ij} = ผลการตอบข้อ j ของผู้สอบคนที่ i

D = a scaling factor ซึ่งมีค่า = 1.7

P_{ij} = $c_j + (1-c_j) / (1 + \exp(-Da_j(\theta_i - b_j)))$

v_j = degree of freedom ของ a_j จาก prior distribution

w_j = Scaling factor ของ a_j จาก Prior distribution

s_j = ค่าพารามิเตอร์ตัวที่หนึ่ง ของ c_j จาก prior distribution

t_j = ค่าพารามิเตอร์ตัวที่สอง ของ c_j จาก prior distribution

การประมาณค่าพารามิเตอร์ จะดำเนินการตามขั้นที่หนึ่ง และขั้นที่สอง ซ้ำ ๆ จนกว่าจะได้ค่าคงที่ ค่าอำนาจจำแนก a_j ค่าความยาก b_j ค่าการเดา c_j และค่าความสามารถ θ_i ที่ได้ในครั้งสุดท้าย คือ ค่าที่ประมาณได้ จากวิธีของเบย์ เมื่อได้มีการกำหนดการแจกแจงเริ่มแรกของค่าอำนาจจำแนกด้วยการแจกแจงแบบไคว (Chi distribution) กำหนดการแจกแจงเริ่มแรกของค่าการเดาด้วยการแจกแจงแบบเบต้า (Beta distribution) (Swaminatham and Gifford 1986 : 589-601)

3. การคำนวณค่าฟังก์ชันสารสนเทศของแบบสอบแต่ละฉบับ จากการประมาณค่าพารามิเตอร์แต่ละวิธี ณ ระดับความสามารถต่าง ๆ โดยการนำค่าพารามิเตอร์ของข้อสอบในแบบสอบแต่ละฉบับ จากการประมาณค่าแต่ละวิธี มาคำนวณหาค่าฟังก์ชันสารสนเทศของข้อสอบในแบบสอบแต่ละฉบับ ณ ระดับความสามารถต่าง ๆ ซึ่งได้แบ่งระดับความสามารถออกเป็น 31 ระดับ คือ -3.0, -2.8, -2.6, -2.4, -2.2, -2.0, -1.8, -1.6, -1.4, -1.2, -1.0, -0.8, -0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8 และ 3.0 ซึ่งใช้สูตรในการคำนวณดังนี้

ค่าฟังก์ชันสารสนเทศสารสนเทศแบบสอบ ณ ระดับความสามารถ θ_j ($I(\theta_j)$)

$$I(\theta_j) = \sum_{i=1}^n \frac{P_i'(\theta_j)^2}{P_i(\theta_j)Q_i(\theta_j)}$$

เมื่อ i แทน ข้อสอบข้อที่ $1, 2, \dots, n$

$$P_i'(\theta_j) = \frac{1.7 a_i (1-c_i)}{e^{1.7a_i(\theta_j - b_i)} + 2 + e^{-1.7a_i(\theta_j - b_i)}}$$

$$P_i(\theta_j) = \frac{c_i + e^{1.7a_i(\theta_j - b_i)}}{1 + e^{1.7a_i(\theta_j - b_i)}}$$

$$Q_i(\theta_j) = \frac{1 - c_i}{1 + e^{1.7a_i(\theta_j - b_i)}}$$

4. การคำนวณค่าประสิทธิภาพเชิงสัมพัทธ์ของแบบสอบ ระหว่างวิธีประมาณค่าพารามิเตอร์ 2 วิธีใด ๆ ณ ระดับความสามารถต่าง ๆ โดยนำค่าฟังก์ชันสารสนเทศของแบบสอบแต่ละฉบับ ที่ได้จากการประมาณค่าแต่ละวิธี ณ ระดับความสามารถต่าง ๆ มาคำนวณซึ่งใช้สูตรในการคำนวณ ดังนี้

ค่าประสิทธิภาพเชิงสัมพัทธ์ของแบบสอบฉบับหนึ่ง (A) ที่ประมาณค่าพารามิเตอร์ด้วยวิธี y เปรียบเทียบกับ แบบสอบฉบับนั้น ที่ประมาณค่าพารามิเตอร์ด้วยวิธี x ณ ระดับความสามารถ θ_j ($REA(y, x)$)

$REA(y, x) =$ ค่าฟังก์ชันสารสนเทศของแบบสอบ A ที่เป็นผลจากการประมาณด้วยวิธี y ณ ระดับความสามารถ θ_j / ค่าฟังก์ชันสารสนเทศของแบบสอบ A ที่เป็นผลจากการประมาณด้วยวิธี x ณ ระดับความสามารถ θ_j

5. การคำนวณค่าความตรงเชิง เกณฑ์สัมพัทธ์ของแบบสอบแต่ละประเภท ที่เป็น ความสัมพันธ์ระหว่างค่าความสามารถของบุคคลซึ่งประมาณได้ โดยวิธีแมกซิมัมไลค์ลิฮูด วิธี สิวริสติก หรือวิธีของเบย์ กับเกณฑ์ ตามกรอบทฤษฎีการสอบแบบดั้งเดิม ด้วยสูตรสัมประสิทธิ์ สหสัมพันธ์แบบเพียร์สัน ดังนี้

$$r_{xy} = \frac{N \sum xy - \sum x \sum y}{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}$$

เมื่อ r_{xy} คือ ค่าสัมประสิทธิ์ความตรงเชิง เกณฑ์สัมพัทธ์

X คือ ค่าความสามารถของบุคคลที่ได้จากการประมาณค่าด้วยวิธีใด ๆ จากผลของการตอบแบบสอบ ที่ต้องการหาความตรง

Y คือ คะแนนเกณฑ์

6. การคำนวณค่าความเที่ยงของแบบสอบแต่ละฉบับตามกรอบของทฤษฎีการทดสอบแบบดั้งเดิม โดยใช้การประมาณค่าความเที่ยงด้วยสูตร คูเคอร์ริชาร์ดสัน สูตรที่ 20 ในการคำนวณดังนี้ (Allen and Yen 1979 : 84)

$$r_{tt} = \frac{N}{N-1} \frac{(\sigma_t^2 - \sum PQ)}{\sigma_t^2}$$

เมื่อ N แทน จำนวนข้อสอบในแบบสอบ

P แทน สัดส่วนของการตอบถูกของข้อสอบแต่ละข้อ

Q แทน สัดส่วนของการตอบ ผิด ของข้อสอบแต่ละข้อ

Q² แทน ความแปรปรวนของแบบสอบ

7. การทดสอบสมมุติฐาน

7.1 ผู้วิจัยนำค่าประสิทธิภาพเชิงสัมพัทธ์ของแบบสอบ (A) มาทดสอบดังนี้
เมื่อ $RE_A (y, x) = 1$ แสดงว่า แบบสอบที่ประมาณค่า ด้วยวิธี y
และวิธี x มีคุณภาพเท่ากัน

$RE_A (y, x) > 1$ แสดงว่า แบบสอบที่ประมาณค่า ด้วยวิธี y
มีคุณภาพสูงกว่า เมื่อประมาณ
ด้วยวิธี x

$RE_A (y, x) < 1$ แสดงว่า แบบสอบที่ประมาณค่า ด้วยวิธี y
มีคุณภาพต่ำกว่า เมื่อประมาณ
ด้วยวิธี x

7.2 ผู้วิจัยนำค่าความตรงเชิงเกณฑ์สัมพัทธ์ของแบบสอบ ทั้งแบบสอบผล-
สัมฤทธิ์ และแบบสอบความถนัด ที่เป็นผลจากความสัมพันธ์ระหว่างค่าความสามารถของบุคคล
ที่ประมาณโดยวิธีแมกซิมัมไลค์ลิสตูด วิธีฮิวริสติก หรือวิธีของเบย์ กับ เกณฑ์ ซึ่งเป็นคุณลักษณะ
ที่สำคัญประการหนึ่งในการตรวจสอบคุณภาพของแบบสอบตามกรอบทฤษฎีการทดสอบแบบดั้งเดิม
มาทดสอบความแตกต่างโดยวิธีการแปลงให้อยู่ในรูปฟังก์ชันพิช เซอร์ซี และทดสอบ
ด้วยค่าไคสแควร์ (Chi square : χ^2) จากสูตรดังนี้ (Steel and Torrie 1981 : 281)

$$\chi^2 = \sum Z_{r1}^2 (N_1 - 3) - \frac{[\sum Z_{r1} (N_1 - 3)]^2}{\sum (N_1 - 3)}, \quad df = n - 1$$

เมื่อ i แทน วิธีประมาณค่า วิธีที่ 1, 2 และ 3

Z_{r1} แทน ค่าความตรงของแบบสอบที่เป็นผลจากการประมาณค่าแต่ละวิธี
ที่แปลงให้อยู่ในรูปของฟังก์ชันพิช เซอร์ซี

N_1 แทน จำนวนนักเรียนในแต่ละวิธีประมาณค่า

df แทน ชั้นของความเป็นอิสระ

n แทน จำนวนค่าของความตรงของแบบสอบ ที่นำมาทดสอบ

ถ้าพบว่า ค่าความตรงของแบบสอบที่ได้จากการประมาณค่าความสามารถแต่ละวิธี ที่นำมาทดสอบมีความแตกต่างกันอย่างมีนัยสำคัญ ก็นำมาทดสอบความแตกต่างของค่าความตรงของแบบสอบ ที่ได้จากการประมาณค่าความสามารถด้วยวิธีต่าง ๆ ครั้งละ 2 วิธี โดยการแปลงให้อยู่ในรูปฟังก์ชันพิช เซอร์ซี แล้วทดสอบด้วยค่าสัมประสิทธิ์พิช เซอร์ซี จากสูตรดังนี้ (Ferguson 1976 : 184)

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{1/(N_1-3) + 1/(N_2-3)}}$$

เมื่อ

Z_{r1}, Z_{r2} แทน ค่าความตรงของแบบสอบที่ได้จากการประมาณค่าแต่ละวิธี
ที่แปลงให้อยู่ในรูปของฟังก์ชันพิช เซอร์ซี

N_1, N_2 แทน จำนวนนักเรียนในการประมาณค่าแต่ละวิธี