

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการศึกษาเชิงประจักษ์ (Empirical Study) มีวัตถุประสงค์เพื่อเปรียบเทียบคุณภาพของการเทียบมาตราเชิงเส้นตรงเมื่อใช้แบบสอบร่วมภายในที่มีความยาวต่างกัน ในแบบสอบวัดผลสัมฤทธิ์ มีวิธีดำเนินการวิจัยดังนี้

แบบแผนการวิจัย

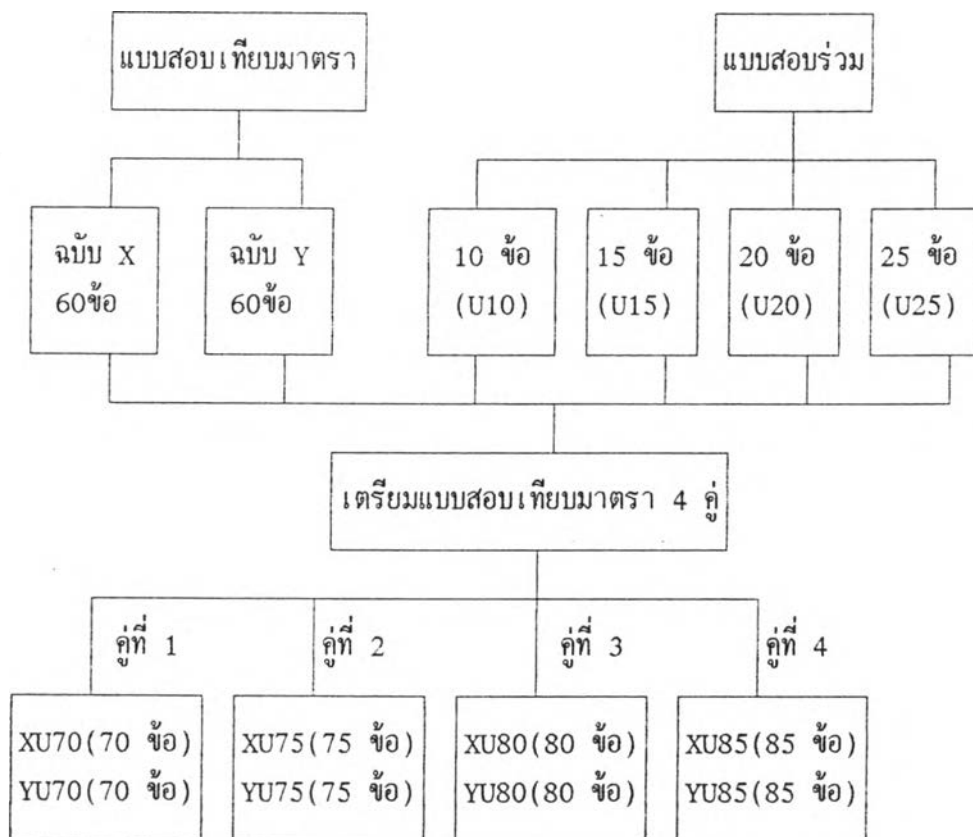
การเทียบมาตราครั้งนี้ นอกแบบรวบรวมข้อมูลโดยใช้แบบสอบร่วม (Anchor test) กับกลุ่มตัวอย่างสุ่ม 2 กลุ่ม แต่ละกลุ่มทำแบบสอบเพียง 1 ชุด แต่ละชุดของแบบสอบประกอบด้วยแบบสอบร่วมภายในที่แปรผันไปตามขนาดความยาว 4 ขนาด คือ 10 15 20 และ 25 ข้อ สรุปได้ดังภาพที่ 4

กลุ่มตัวอย่าง	แบบสอบ		
	แบบสอบเทียบมาตรา (X)	แบบสอบเทียบมาตรา (Y)	แบบสอบร่วม 4 ขนาด (U)
P ₁	/	-	/
P ₂	-	/	/

ภาพที่ 4: รูปแบบการรวบรวมข้อมูลในการเทียบมาตราโดยใช้แบบสอบร่วม (Anchor test-Random groups Design) (Petersen et al 1989: 244)

เนื่องจากแบบสอบเทียบมาตรฐานทั้งสองชุด เป็นแบบสอบที่สร้างขึ้นในแต่ละปีการศึกษา ค่าความเที่ยงของแบบสอบทั้งสองฉบับแตกต่างกัน ดังนั้น ผู้วิจัยจึงเลือกใช้แบบแผนการรวบรวมข้อมูลของแองกอฟรูปแบบ 3B (Angoff 1984:109) ซึ่งใช้กับแบบสอบเทียบมาตรฐานที่มีความเที่ยงไม่เท่ากัน โดยแบบสอบที่จะเทียบมาตรฐานมีทั้งหมด 4 คู่ แต่ละคู่มีความยาวไม่เท่ากัน ดังรายละเอียดในภาพที่ 5 และประเมินคุณภาพของการเทียบมาตรฐานจาก

1. ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานเชิงเส้นตรง คำนวณจากสูตรในแบบแผนการรวบรวมข้อมูลของแองกอฟ รูปแบบ 3B (Angoff 1984: 106) (จากสมการที่ 2 ในหน้า 60) เป็นการประเมินเชิงทฤษฎีซึ่งเป็นเกณฑ์ภายใน
2. ค่าดัชนีความแตกต่าง (C) ซึ่งเป็นความคลาดเคลื่อนรวม ได้จากการวิเคราะห์กลุ่มสอบทานผล ตามแนวของโคเลนและวิทนี (Kolen and Whitney 1982)



ภาพที่ 5: การจัดเตรียมแบบสอบสำหรับเทียบมาตรฐาน



ประชากรและกลุ่มตัวอย่าง

ประชากร

เนื่องจากการวิจัยครั้งนี้ผู้วิจัยได้ออกแบบ โดยใช้แบบสอบถามสัมฤทธิ์ผลปลายภาควิชา คณิตศาสตร์ชั้นมัธยมศึกษาปีที่สอง (ค203) ของกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดขอนแก่น กลุ่มที่ 2 เป็นแบบสอบถามมาตรา โดยใส่แบบสอบถามที่ผู้วิจัยสร้างขึ้นรวมอยู่ด้วย ดังนั้น ประชากรที่ใช้ในการวิจัยครั้งนี้ จึงเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2534 ของโรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษา จังหวัดขอนแก่น กลุ่มที่ 2 ซึ่งมีทั้งหมด 12 โรงเรียน จำนวนนักเรียนรวมทั้งสิ้น 3,018 คน (ฝ่ายสถิติกองแผนงาน กรมสามัญศึกษา 2533)

กลุ่มตัวอย่าง

กลุ่มตัวอย่างในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2534 ของโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา จังหวัดขอนแก่น กลุ่มที่ 2 ประมาณ ขนาดกลุ่มตัวอย่างโดยใช้ตารางยามาเน (Yamane อ้างถึงใน อุทุมพร จามรมาน 2533: 30) ที่ระดับความเชื่อมั่น 95% ความคลาดเคลื่อน $\pm 5\%$ ได้กลุ่มตัวอย่างจำนวน 353 คน

แต่การออกแบบการเทียบมาตราจะต้องใช้กลุ่มตัวอย่างเทียบมาตรา 2 กลุ่ม แต่ละกลุ่มไม่ควรมีขนาดต่ำกว่า 400 คน (Brennan and Kolen 1987: 285) จึงใช้กลุ่มตัวอย่าง เทียบมาตรารวมทั้งสิ้น 810 คน และกลุ่มสอบทานผลซึ่งใช้ประเมินคุณภาพในการเทียบมาตราซึ่ง มีขนาดไม่น้อยกว่า 10% ของกลุ่มตัวอย่างเทียบมาตรา (Kolen and Whitney 1982) อีก จำนวน 117 คน โดยมีขั้นตอนในการสุ่ม ดังต่อไปนี้

1. สุ่มอำเภอ

1.1 สุ่มโรงเรียนมัธยมศึกษาในจังหวัดขอนแก่น กลุ่มที่ 2 มีโรงเรียนอยู่ในอำเภอต่าง ๆ จากทั้งหมด 7 อำเภอ ดังนี้

1.1.1	อำเภอชุมแพ	มีจำนวน	3	โรงเรียน
1.1.2	อำเภอภูเวียง	มีจำนวน	3	โรงเรียน
1.1.3	อำเภอหนองเรือ	มีจำนวน	2	โรงเรียน
1.1.4	กิ่งอำเภอพระยืน	มีจำนวน	1	โรงเรียน
1.1.5	กิ่งอำเภอภูผาม่าน	มีจำนวน	1	โรงเรียน
1.1.6	อำเภอสีชมพู	มีจำนวน	1	โรงเรียน
1.1.7	อำเภอบ้านฝาง	มีจำนวน	1	โรงเรียน

ผู้วิจัยจึงสุ่มแบบ Cluster ตามขนาดอำเภอ โดยมีอำเภอที่มี 3 โรงเรียน มีอยู่ 2 อำเภอ อำเภอที่มี 2 โรงเรียน มี 1 อำเภอ อำเภอที่มี 1 โรงเรียน 4 อำเภอ แต่เนื่องจากโรงเรียนในอำเภอสีชมพู และอำเภอบ้านฝางเป็นโรงเรียนในโครงการนำร่องทดลองใช้หลักสูตรที่ปรับปรุงใหม่ จึงตัดออกไป เหลืออำเภอที่มี 1 โรงเรียน 2 อำเภอ จึงสุ่มอำเภอที่มี 3 โรงเรียน 1 อำเภอ อำเภอที่มี 1 โรงเรียน 1 อำเภอ รวมอำเภอที่มี 2 โรงเรียน อีก 1 อำเภอ ได้อำเภอและโรงเรียนที่เป็นกลุ่มตัวอย่าง ดังต่อไปนี้

อำเภอชุมแพ: โรงเรียนชุมแพศึกษา โรงเรียนจตุรมิตรวิทยาคาร โรงเรียนโนนหันวิทยายน

อำเภอหนองเรือ: โรงเรียนหนองเรือวิทยา โรงเรียนจรเข้วิทยายน

กิ่งอำเภอพระยืน: โรงเรียนประชารัฐวิทยาเสริม

2. สุ่มห้องเรียน

หลังจากได้โรงเรียนซึ่งเป็นกลุ่มตัวอย่างในแต่ละอำเภอแล้ว จึงทำการสุ่มห้องเรียนในแต่ละโรงเรียน ด้วยวิธีการสุ่มอย่างง่าย ตามสัดส่วนจำนวนนักเรียนในแต่ละโรงเรียน โดยกลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ผู้วิจัยแบ่งออกเป็น 3 กลุ่ม คือ

กลุ่มที่ 1 เป็นกลุ่มที่ใช้ในการทดลองสอบ (Try Out) เพื่อจะได้ศึกษาคุณภาพของแบบสอบและเพื่อความสมบูรณ์ของกระบวนการสอบในขั้นดำเนินการจริง เวลาที่กำหนดให้นักเรียนตอบ ซึ่งในการนำแบบสอบของกลุ่มโรงเรียนไปทดลองสอบ ใช้นักเรียนจำนวน 69 คน เป็นนักเรียนโรงเรียนประชารัฐวิทยาเสริม จำนวน 2 ห้องเรียน

กลุ่มที่ 2 เป็นกลุ่มที่ใช้ในการวิจัยเพื่อนำผลการสอบไปวิเคราะห์เทียบมาตรฐานคะแนนของแบบสอบ ด้วยวิธีการเทียบมาตรฐานเชิงเส้นตรง แต่เนื่องจากการวิจัยครั้งนี้ต้องการเก็บข้อมูลโดยใช้แบบสอบ 2 ฉบับ ซึ่งนักเรียนทำแบบสอบคนละ 1 ฉบับ เพื่อให้ได้นักเรียนที่สอบแบบสอบแต่ละฉบับมีความสามารถทางการเรียนใกล้เคียงกัน จึงสุ่มนักเรียนโดยแจกแบบสอบสลับกันไป ได้จำนวนนักเรียนที่เป็นกลุ่มตัวอย่างเทียบมาตรฐานทั้งหมด 831 คน สอบแบบสอบฉบับที่หนึ่ง (ฉบับ X) 426 คน สอบแบบสอบฉบับที่สอง (ฉบับ Y) 405 คน จึงสุ่มผู้สอบฉบับที่หนึ่งออก 21 คน เพื่อให้ได้จำนวนผู้สอบในแบบสอบแต่ละฉบับจำนวนเท่ากัน ดังแสดงไว้ในตารางที่ 1

กลุ่มที่ 3 เป็นกลุ่มที่ใช้ในการวิจัย เพื่อนำผลการสอบไปวิเคราะห์ตรวจสอบคุณภาพของวิธีการเทียบมาตรฐานเชิงเส้นตรง โดยกลุ่มสอบทานผลแต่ละคนจะต้องทำแบบสอบทั้ง 2 ฉบับ ทำให้มีโอกาสที่นักเรียนบางคนจะทำแบบสอบไม่ครบทั้ง 2 ฉบับ ผู้วิจัยจึงสุ่มตัวอย่างเกินจากที่กำหนดไว้ เพื่อป้องกันปัญหาที่อาจเกิดขึ้นดังกล่าว ซึ่งได้จำนวนนักเรียนที่เป็นกลุ่มสอบทานผล 136 คน เป็นนักเรียนโรงเรียนโนนหันวิทยายน จำนวน 4 ห้องเรียน มีนักเรียนที่สอบไม่ครบทั้ง 2 ฉบับ จำนวน 19 คน เหลือนักเรียนเป็นกลุ่มสอบทานผล จำนวน 117 คน ดังแสดงไว้ในตารางที่ 1

ตารางที่ 1 รายชื่อโรงเรียนและจำนวนนักเรียน

โรงเรียน	จำนวนทั้งหมด		กลุ่มตัวอย่าง เทียบมาตรา		กลุ่มสอบทานผล	
	ห้องเรียน	นักเรียน	ห้องเรียน	นักเรียน	ห้องเรียน	นักเรียน
1. ชุมแพศึกษา	10	518	6	295	-	-
2. จตุรมิตรวิทยาคาร	6	271	5	213	-	-
3. หนองเรือวิทยา	6	262	5	207	-	-
4. จระเข้วิทยายน	5	166	3	116	-	-
5. โนนหันวิทยายน	6	257	-	-	4	136
รวม	33	1,474	19	831	4	136

เครื่องมือที่ใช้ในการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้ออกแบบ โดยใช้แบบสอบวัดผลสัมฤทธิ์ปลายภาค ที่กลุ่มโรงเรียนเป็นผู้สร้างมาเป็นแบบสอบเทียบมาตรา โดยใช้แบบสอบร่วมที่ผู้วิจัยสร้างขึ้นรวมอยู่ในแบบสอบเทียบมาตราทั้ง 2 ฉบับ ประกอบด้วยกลุ่มโรงเรียนที่ 2 จังหวัดขอนแก่นมีกระบวนการสร้างข้อสอบของกลุ่มโรงเรียน โดยมีแบบสอบในแต่ละปีการศึกษาที่ใช้จุดประสงค์ในการออกข้อสอบจุดประสงค์เดียวกัน แบบสอบดังกล่าวเคยใช้ในสถานการณ์สอบจริงมาแล้วผู้วิจัยจึงได้นำมาใช้ในการวิจัยครั้งนี้ ดังมีรายละเอียดเกี่ยวกับเครื่องมือที่ใช้ในการวิจัยต่อไปนี้

1. แบบสอบเทียบมาตราเป็นแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ภาคต้น (ค203) ปีการศึกษา 2532 และ 2533 ที่สร้างโดยกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดขอนแก่น กลุ่มที่ 2 ฉบับละ 60 ข้อ เป็นแบบสอบแบบเลือกตอบชนิด 4 ตัวเลือก แบบสอบทั้ง 2 ฉบับ วัดความรู้ครอบคลุมเนื้อหา 6 เรื่อง คือ จำนวนและตัวเลข ระบบจำนวนเต็ม พื้นที่ ความเท่ากันทุกประการ เศษส่วนและทศนิยม คุณสมบัติของรูปสามเหลี่ยมมุมฉากจากเนื้อหาหลักสูตรพุทธศักราช 2521 โดยมีกระบวนการสร้างที่เป็นระบบ คือ

1.1 ตั้งคณะกรรมการออกข้อสอบซึ่งมาจากครูที่ทำการสอนรายวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 โรงเรียนละ 1 คน

- 1.2 กำหนดตารางวิเคราะห์เนื้อหา (ตั้งรายละเอียดในภาคผนวก)
- 1.3 สร้างข้อสอบตามตารางวิเคราะห์เนื้อหา (ซึ่งขั้นตอนทั้ง 3 นี้ ต้องแล้วเสร็จก่อนเปิดภาคเรียนที่ 1)
- 1.4 ทำการวิเคราะห์คุณภาพของแบบสอบและค่าสถิติพื้นฐานของผู้สอบในแต่ละโรงเรียน

โดยจุดประสงค์ที่ใช้สร้างแบบสอบในปีการศึกษา 2532 และ 2533 เป็นจุดประสงค์เชิงพฤติกรรมชุดเดียวกัน แบบสอบทั้ง 2 ฉบับ จะใช้เป็นแบบสอบเทียบมาตรฐานฉบับ X (ปีการศึกษา 2532) และฉบับ Y (ปีการศึกษา 2533)

ได้นำแบบสอบทั้ง 2 ฉบับ ไปทดลองใช้กับ นักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนพระราชรัฐวิทยาเสริม 2 ห้อง จำนวน 69 คน (โดยแจกแบบสอบสลับกันให้นักเรียนแต่ละคน) ได้จำนวนผู้สอบในแบบสอบฉบับ X 32 คน ฉบับ Y 37 คน ได้ค่าความเที่ยงของแบบสอบสำหรับแบบสอบฉบับ X = .84 ฉบับ Y = .80

จากการศึกษากระบวนการสร้างและคุณภาพของแบบสอบ ด้านความเที่ยงแล้ว ผู้วิจัยพิจารณาว่าแบบสอบทั้ง 2 ฉบับ เป็นแบบสอบที่มีคุณภาพเพียงพอที่จะนำไปใช้เป็นเครื่องมือ โดยปรับปรุงข้อคำถามและตัวเลือกของข้อสอบบางข้อที่ไม่ผ่านเกณฑ์ให้เหมาะสมยิ่งขึ้น แล้วให้ผู้เชี่ยวชาญด้านเนื้อหาวิชา จำนวน 2 ท่าน ซึ่งเป็นอาจารย์ที่สอนวิชาคณิตศาสตร์ในระดับมัธยมศึกษาไม่ต่ำกว่า 5 ปี และจบการศึกษาขั้นต้นระดับปริญญาตรี (ดังมีรายชื่อในภาคผนวก) ตรวจสอบความตรงตามเนื้อหา ปรับปรุงภาษาของคำถาม และตัวเลือกตัวลวงก่อนนำไปใช้กับกลุ่มตัวอย่าง

2. แบบสอบร่วมซึ่งเป็นแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 ภาคต้น (ค203) ซึ่งผู้วิจัยเป็นผู้สร้าง ตามหลักสูตรระดับมัธยมศึกษาตอนต้น พุทธศักราช 2521 ของกรมวิชาการ กระทรวงศึกษาธิการจำนวน 40 ข้อ เป็นแบบสอบเลือกตอบชนิด 4 ตัวเลือก สำหรับเป็นแบบสอบร่วม ในแบบสอบเทียบมาตรฐานทั้งสองฉบับ ดังกล่าวข้างต้น

ขั้นตอนการสร้างแบบสอบร่วม

1. ศึกษาหลักสูตร คู่มือครู และแบบเรียนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 2 ภาคต้น เพื่อศึกษาจุดประสงค์ทั่วไปและจุดประสงค์การเรียนรู้ของวิชาตลอดจนเนื้อหาในรายวิชาดังกล่าว ซึ่งประกอบด้วยเนื้อหาวิชา 7 บท ใน 6 เรื่อง ดังนี้

- 1.1 จำนวนและตัวเลข
- 1.2 ระบบจำนวนเต็ม
- 1.3 พหุคูณ
- 1.4 ความเท่ากันทุกประการ
- 1.5 เศษส่วนและทศนิยม
- 1.6 คุณสมบัติของรูปสามเหลี่ยมมุมฉาก

2. ศึกษาตารางวิเคราะห์เนื้อหา และจุดประสงค์ของกลุ่มโรงเรียนมัธยมศึกษาตอนต้น จังหวัดขอนแก่น กลุ่มที่ 2 (ดังรายละเอียดในภาคผนวก)

3. สร้างตารางวิเคราะห์เนื้อหา โดยมีขั้นตอนดังนี้

3.1 วิเคราะห์เนื้อหาและกำหนดจุดประสงค์ เพื่อเป็นแนวทางในการเขียนข้อสอบ โดยนำเนื้อหาวิชาทั้ง 7 บท ไปให้ผู้เชี่ยวชาญจำนวน 5 ท่าน ซึ่งเป็นอาจารย์ที่สอนวิชาคณิตศาสตร์ในระดับมัธยมศึกษาไม่น้อยกว่า 5 ปี และจบการศึกษาขั้นต้นระดับปริญญาตรี (ดังมีรายชื่อในภาคผนวก) เรียงลำดับความสำคัญของเนื้อหาวิชา จากสำคัญและเน้นมากที่สุดเป็นอันดับที่ 1 และรองลงไปตามลำดับ และพิจารณาจุดประสงค์สำคัญเพียง 1 จุดประสงค์ที่ครอบคลุมเนื้อหาในแต่ละบท และนำผลมาพิจารณาพร้อมกับเอกสารหลักสูตร เช่น เวลาที่ใช้ในการสอนแต่ละเนื้อหา ผลการพิจารณาการเรียงลำดับความสำคัญของเนื้อหาวิชาทั้ง 7 บท มีดังนี้

บทที่ 2: ระบบจำนวนเต็ม

บทที่ 4: ระบบจำนวนเต็ม

บทที่ 6: เศษส่วนและทศนิยม

บทที่ 3: พหุนาม

บทที่ 7: คุณสมบัติของรูปสามเหลี่ยมมุมฉาก

บทที่ 5: ความเท่ากันทุกประการ

บทที่ 1: จำนวนและตัวเลข

3.2 สร้างข้อสอบจากจุดประสงค์การเรียนรู้สำคัญ ที่ครอบคลุมเนื้อหาในแต่ละบท บทละ 1 จุดประสงค์ โดยกำหนดจำนวนข้อในแต่ละจุดประสงค์ตามน้ำหนักความสำคัญของเนื้อหาแต่ละบทได้ข้อสอบแบบเลือกตอบชนิด 4 ตัวเลือก จำนวน 40 ข้อ ซึ่งมากกว่าจำนวนที่ต้องการใช้จริง เพื่อคัดเลือกข้อสอบที่ดีไว้เป็นแบบสอบรวม จำนวน 25 ข้อ ซึ่งมีรายละเอียดจำนวนข้อสอบรวมในแต่ละจุดประสงค์ดังตารางที่ 2

ตารางที่ 2 วิเคราะห์เนื้อหา จุดประสงค์ และจำนวนข้อของแบบสอบรวม

เนื้อหา	จุดประสงค์	จำนวนข้อ ที่สร้าง	จำนวนข้อ ที่ใช้จริง
บทที่ 2: ระบบจำนวนเต็ม	นักเรียนสามารถหาผลบวกและผลลบของจำนวนเต็มใด ๆ ได้	6	4
บทที่ 4: ระบบจำนวนเต็ม	นักเรียนสามารถหาผลคูณ ผลหาร ของจำนวนเต็มใด ๆ ได้	6	4
บทที่ 6: เศษส่วนและทศนิยม	นักเรียนสามารถนำความรู้เรื่องบวก ลบ คูณ หาร เศษส่วน และทศนิยม ไปใช้หาคำตอบจากโจทย์ที่กำหนดให้ได้	6	4
บทที่ 3: พื้นที่	นักเรียนสามารถหาพื้นที่ของรูปสามเหลี่ยมและสี่เหลี่ยมต่าง ๆ ได้	6	4
บทที่ 7: คุณสมบัติของรูปสามเหลี่ยมมุมฉาก	นักเรียนสามารถนำความสัมพันธ์ระหว่างความยาวของด้านของรูปสามเหลี่ยมมุมฉากไปใช้ได้	6	3
บทที่ 5: ความเท่ากันทุกประการ	นักเรียนสามารถบอกเงื่อนไขที่ทำให้รูปสามเหลี่ยมสองรูปเท่ากันทุกประการได้	5	3
บทที่ 1: จำนวนและตัวเลข	นักเรียนสามารถเปรียบเทียบค่าของตัวเลขฐานต่าง ๆ ได้	5	3
รวม		40	25

คุณภาพของแบบสอบและการคัดเลือกข้อสอบร่วม

1. ตรวจสอบความตรงตามเนื้อหา

1.1 นำข้อสอบที่สร้างขึ้นไปตรวจสอบความตรงตามเนื้อหา โดยให้ผู้เชี่ยวชาญ ซึ่งมีประสบการณ์ในการสอนวิชาคณิตศาสตร์ในระดับมัธยมศึกษาไม่น้อยกว่า 5 ปี และจบการศึกษาระดับปริญญาตรี จำนวน 6 ท่าน (ดังมีรายชื่อในภาคผนวก) ตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ตามวิธีของโรวีเนลลี และแฮมเบิลตัน (Rovinelli and Hambleton 1977 อ้างถึงใน บุญเชิด วิทยุอนันตพงษ์ 2527: 67-70) มีสูตรดังนี้

$$IOC = R/N$$

เมื่อ IOC คือ ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์

R คือ ผลรวมคะแนนความคิดเห็นของผู้เชี่ยวชาญ

N คือ จำนวนผู้เชี่ยวชาญ

โดยพิจารณาคะแนนการตัดสินเฉลี่ยมีค่ามากกว่า 0.5 ถือว่าข้อสอบนั้นวัดได้สอดคล้องกับจุดประสงค์นั้นจริง โดยให้น้ำหนักคะแนนดังนี้

- +1 หมายถึง แน่ใจว่าข้อสอบวัดได้ตรงกับจุดประสงค์การเรียนรู้
- 0 หมายถึง ไม่แน่ใจว่าข้อสอบวัดได้ตรงกับจุดประสงค์การเรียนรู้
- 1 หมายถึง แน่ใจว่าข้อสอบวัดได้ไม่ตรงตามจุดประสงค์การเรียนรู้

จากการนำข้อสอบไปให้ผู้เชี่ยวชาญตรวจสอบ แล้วนำคะแนนความคิดเห็นมาคำนวณค่า IOC จะได้ค่าคะแนนเฉลี่ยจากผลการตัดสินของผู้เชี่ยวชาญอยู่ในช่วง 0.67-1.00 โดยที่ข้อสอบแต่ละข้อในแต่ละจุดประสงค์ มีค่าเฉลี่ยเกินเกณฑ์ 0.5 แสดงว่าข้อสอบที่สร้างขึ้นทุกข้อมีความตรงตามเนื้อหา ดังผลการตรวจสอบที่ปรากฏในตารางที่ 3 พร้อมกันนี้ผู้วิจัยได้ปรับปรุงข้อสอบบางข้อตามข้อเสนอแนะของผู้เชี่ยวชาญ เช่น การสร้างรูป ตัวลงบางข้อที่ไม่เหมาะสม

ตารางที่ 3 จำนวนผู้เชี่ยวชาญที่ตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ และค่าดัชนีความสอดคล้อง

ลำดับที่	จุดประสงค์การเรียนรู้	ข้อสอบที่	ผลการตัดสิน			รวม	ดัชนีความ สอดคล้อง
			+1	0	-1		
1.	นักเรียนสามารถเปรียบเทียบค่า ของตัวเลขฐานต่าง ๆ ได้	1	6			6	1.00
		2	6			6	1.00
		3	6			6	1.00
		4	6			6	1.00
		5	6			6	1.00
2.	นักเรียนสามารถหาผลบวกและผล ลบของจำนวนเต็มใด ๆ ได้	6	6			6	1.00
		7	6			6	1.00
		8	6			6	1.00
		9	6			6	1.00
		10	6			6	1.00
		11	5	1		5	.83
3.	นักเรียนสามารถหาพื้นที่ของรูป สามเหลี่ยมและสี่เหลี่ยมต่าง ๆ ได้	12	6			6	1.00
		13	6			6	1.00
		14	6			6	1.00
		15	5	1		5	.83
		16	6			6	1.00
		17	6			6	1.00
4.	นักเรียนสามารถหาผลคูณ ผลหาร ของจำนวนเต็มใด ๆ ได้	18	6			6	1.00
		19	6			6	1.00
		20	6			6	1.00
		21	6			6	1.00
		22	6			6	1.00
		23	6			6	1.00

ตารางที่ 3 จำนวนผู้เชี่ยวชาญที่ตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ และค่าดัชนีความสอดคล้อง (ต่อ)

ลำดับที่	จุดประสงค์การเรียนรู้	ข้อสอบที่	ผลการตัดสิน			รวม	ดัชนีความ สอดคล้อง
			+1	0	-1		
5.	นักเรียนสามารถบอกเงื่อนไขที่ทำให้รูปสามเหลี่ยมสองรูปเท่ากันทุกประการได้	24	6			6	1.00
		25	6			6	1.00
		26	5		1	4	.67
		27	5		1	4	.67
		28	5	1		5	.83
6.	นักเรียนสามารถนำความรู้เรื่อง บวก ลบ คูณ หาร เศษส่วน และ ทศนิยมไปใช้หาค่าตอบจากโจทย์ที่กำหนดให้ได้	29	6			6	1.00
		30	6			6	1.00
		31	6			6	1.00
		32	6			6	1.00
		33	5	1		5	.83
		34	6			6	1.00
7.	นักเรียนสามารถนำความสัมพันธ์ระหว่างความยาวของด้านของรูปสามเหลี่ยมมุมฉากไปใช้ได้	35	6			6	1.00
		36	6			6	1.00
		37	6			6	1.00
		38	6			6	1.00
		39	5		1	4	.67
		40	6			6	1.00

2. ทดลองใช้แบบสอบร่วม

2.1 ทดลองใช้แบบสอบ (Try Out) เพื่อหาคุณภาพและความเหมาะสมที่จะนำไปใช้กับกลุ่มตัวอย่างจริง โดยนำไปสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2534 ภาคปลาย จำนวน 77 คน เป็นนักเรียนโรงเรียนสวนกุหลาบวิทยาลัยนนทบุรี 40 คน นักเรียนโรงเรียนนนทบุรีวิทยา 37 คน วิเคราะห์หาคุณภาพของข้อสอบ ตามแนวทฤษฎีดั้งเดิม (Classical Theory) เพื่อหาค่าความยาก (p) และค่าอำนาจจำแนก (r) โดยใช้โปรแกรมวิเคราะห์ข้อสอบ IAP (ITEM ANALYSIS PROGRAM) ของมหาวิทยาลัยขอนแก่นได้ค่าความเที่ยงของแบบสอบทั้งฉบับโดยใช้สูตรของ Kuder-Richardson สูตรที่ 20 เท่ากับ .87

2.2 คัดเลือกข้อสอบที่มีคุณภาพ โดยใช้เกณฑ์การคัดเลือก ดังนี้

1. ค่าความยากของข้อสอบอยู่ระหว่าง 0.2-0.8
2. ค่าอำนาจจำแนกตั้งแต่ 0.2 ขึ้นไป ค่าอำนาจจำแนกของตัวลวง

มีค่าติดลบ และมีค่าใกล้เคียงกัน แต่มีค่าน้อยกว่าค่าอำนาจจำแนกของตัวเลือกถูก

ได้จำนวนข้อคำถามที่อยู่ในเกณฑ์ทั้งค่าความยากและอำนาจจำแนก 33

ข้อ มีค่าความยากของแบบสอบอยู่ระหว่าง .23-.68 และค่าอำนาจจำแนกของแบบสอบระหว่าง .25-.90 ได้จำนวนข้อสอบที่อยู่ในเกณฑ์ในแต่ละจุดประสงค์ครบตามจำนวนที่ต้องการใช้จริงดังนี้

จุดประสงค์ในบทที่ 1	จำนวนข้อที่ได้อยู่ในเกณฑ์ 3 ข้อ	จำนวนข้อที่ต้องการใช้ 3 ข้อ
จุดประสงค์ในบทที่ 2	จำนวนข้อที่ได้อยู่ในเกณฑ์ 4 ข้อ	จำนวนข้อที่ต้องการใช้ 4 ข้อ
จุดประสงค์ในบทที่ 3	จำนวนข้อที่ได้อยู่ในเกณฑ์ 6 ข้อ	จำนวนข้อที่ต้องการใช้ 4 ข้อ
จุดประสงค์ในบทที่ 4	จำนวนข้อที่ได้อยู่ในเกณฑ์ 5 ข้อ	จำนวนข้อที่ต้องการใช้ 4 ข้อ
จุดประสงค์ในบทที่ 5	จำนวนข้อที่ได้อยู่ในเกณฑ์ 4 ข้อ	จำนวนข้อที่ต้องการใช้ 3 ข้อ
จุดประสงค์ในบทที่ 6	จำนวนข้อที่ได้อยู่ในเกณฑ์ 6 ข้อ	จำนวนข้อที่ต้องการใช้ 4 ข้อ
จุดประสงค์ในบทที่ 7	จำนวนข้อที่ได้อยู่ในเกณฑ์ 5 ข้อ	จำนวนข้อที่ต้องการใช้ 3 ข้อ

การจัดเตรียมแบบสอบร่วมเพื่อทำการเทียบมาตรฐาน

1. จัดเตรียมแบบสอบร่วม โดยนำข้อสอบที่ผ่านการวิเคราะห์หาคัดเลือกในแต่ละจุดประสงค์ให้ได้จำนวนตามน้ำหนักความสำคัญของแต่ละจุดประสงค์จากตารางโครงสร้างเนื้อหา โดยคัดเลือกข้อสอบที่มีความยากปานกลาง และค่าอำนาจจำแนกอยู่ในเกณฑ์ดี ในกรณีที่ข้อสอบมีค่าความยากและค่าอำนาจจำแนกใกล้เคียงกัน จะใช้วิธีสัมพันธ์ตามจำนวนข้อในแต่ละจุดประสงค์ ข้อสอบที่ได้รับการคัดเลือกเป็นแบบสอบร่วมที่ใช้ในการวิจัยครั้งนี้ มีค่าความตรงตามเนื้อหา (IOC) ค่าความยาก อำนาจจำแนก ตามตารางที่ 4 ดังนี้

ตารางที่ 4 ค่าความตรงตามเนื้อหา (IOC) ค่าความยาก ค่าอำนาจจำแนก ของแบบสอบรวมที่ใช้ในการวิจัย

ข้อที่ของ แบบสอบ รวม	จุดประสงค์ ในบทที่	ค่า (IOC)	ค่าความยาก	ค่าอำนาจจำแนก
2	1	1	0.42	0.55
3	1	1	0.47	0.85
5	1	1	0.33	0.65
6	2	1	0.42	0.55
7	2	1	0.40	0.50
8	2	1	0.35	0.60
9	2	1	0.47	0.25
12	3	1	0.47	0.65
13	3	1	0.50	0.90
14	3	1	0.45	0.80
17	3	1	0.68	0.45
19	4	1	0.50	0.30
20	4	1	0.45	0.40
21	4	1	0.62	0.35
23	4	1	0.23	0.45
24	5	1	0.45	0.60
25	5	1	0.62	0.65
27	5	0.67	0.42	0.45
29	6	1	0.62	0.45
30	6	1	0.55	0.80
32	6	1	0.33	0.55
34	6	1	0.62	0.65
36	7	1	0.35	0.50
37	7	1	0.47	0.45
40	7	1	0.28	0.55

2. จัดข้อสอบร่วมตามขนาดความยาว ซึ่งเป็นตัวแปรที่ต้องการศึกษาตามวิธีการสุ่มของฮิลล์ และคณะ (Hills et all 1988: 225) โดยให้แบบสอบร่วมที่มีขนาดสั้นกว่าเป็นชั้นเขตของแบบสอบร่วมที่ยาวกว่า คือ

2.1 จากแบบสอบร่วมจำนวน 25 ข้อ สุ่มข้อสอบในแต่ละจุดประสงค์ขึ้นมาครั้งละ 1 ข้อ เรียงตามลำดับความสำคัญของจุดประสงค์ในแต่ละเนื้อหาจนครบ 25 ข้อ ดังตารางที่ 5

ตารางที่ 5 ตารางการจัดเรียงข้อสอบร่วมในแบบสอบทั้งฉบับ

ข้อที่ของ แบบสอบ ร่วม	จุดประสงค์ ในบทที่	ข้อที่ของ แบบสอบ ทั้งฉบับ	ข้อที่ของ แบบสอบ ร่วม	จุดประสงค์ ในบทที่	ข้อที่ของ แบบสอบ ทั้งฉบับ
1	2	26	14	1	14
2	4	43	15	2	28
3	6	65	16	4	45
4	3	34	17	6	66
5	7	81	18	3	36
6	5	55	19	7	83
7	1	13	20	5	57
8	2	27	21	1	15
9	4	44	22	2	29
10	6	75	23	4	46
11	3	35	24	6	76
12	7	82	25	3	37
13	5	56			

2.2 จากแบบสอบร่วมที่เตรียมไว้ จะได้ว่า เมื่อต้องการศึกษาความยาวของแบบสอบร่วมขนาด 10 15 20 25 ข้อ นั้น แบบสอบร่วมที่มีขนาดสั้นกว่าจะเป็นชั้นเขตของแบบสอบร่วมที่มีขนาดยาวกว่า โดยทุกความยาวของแบบสอบร่วมที่ตรวจสอบจะครอบคลุมทุกจุดประสงค์เหมือนกัน ให้ความยาวของแบบสอบร่วมที่มีขนาดดังกล่าวมีสัญลักษณ์เป็น U10 U15 U20 U25

3. นำแบบสอบรวมจำนวน 25 ข้อมารวมกับแบบสอบของกลุ่มโรงเรียน ปีการศึกษา 2532 (ฉบับ X) และฉบับปีการศึกษา 2533 (ฉบับ Y) โดยเรียงข้อสอบรวมตามเนื้อหาในแบบสอบทั้ง 2 ฉบับ ให้เลขข้อของแบบสอบรวมตรงกันทั้ง 2 ฉบับ จะได้แบบสอบที่ใช้ในการเทียบมาตรฐานเป็นเครื่องมือในการวิจัย 2 ฉบับ ๆ ละ 85 ข้อ

การรวบรวมข้อมูล

1. ติดต่อขอความร่วมมือในการนำแบบสอบไปสอบกับนักเรียนในโรงเรียนของกลุ่มตัวอย่างเทียบมาตรฐานและกลุ่มสอบทานผล ตลอดจนกำหนดวัน และเวลาในการสอบ โดยจัดเตรียมเนื้อหาวิชา ค203 ฉบับย่อให้นักเรียนทบทวนล่วงหน้า 1 สัปดาห์
2. นำแบบสอบที่ได้จัดเตรียมไว้ทั้ง 2 ฉบับ แต่ละฉบับมีความยาวเต็มฉบับ 85 ข้อ ไปทดสอบกับกลุ่มตัวอย่างเทียบมาตรฐานโดยนักเรียนแต่ละคนในแต่ละห้อง ของแต่ละโรงเรียน ถูกสุ่มอย่างมีระบบให้สอบแบบสอบ เพียงคนละ 1 ฉบับเท่านั้น เพื่อให้การสอบแบบสอบทั้ง 2 ฉบับ มีลักษณะเหมือนสุ่ม และนักเรียนมีความสามารถทางการเรียนใกล้เคียงกัน
3. นำคะแนนของนักเรียนมาจัดแยกออกตามลักษณะแบบสอบเทียบมาตรฐาน จะได้ผู้สอบแบบสอบฉบับที่ 1 (ฉบับ X) จำนวน 426 คน และฉบับที่ 2 (ฉบับ Y) จำนวน 405 คน สุ่มจำนวนนักเรียนในฉบับที่ 1 ออก 21 คน เพื่อให้ได้จำนวนนักเรียนแต่ละแบบสอบเท่ากัน แล้วนำคะแนนของนักเรียนแต่ละคนที่สอบแบบสอบขนาด 85 ข้อ ในแต่ละฉบับนั้นมาตรวจตามความยาวแบบสอบรวม จะได้คะแนนของนักเรียนเมื่อทำแบบสอบ 70 ข้อ 75 ข้อ 80 ข้อ และ 85 ข้อ ทั้งนักเรียนที่สอบฉบับ X และฉบับ Y นั้น (มีสัญลักษณ์เป็น XU70 XU75 XU80 XU85 และ YU70 YU75 YU80 YU85) คือเมื่อใช้แบบสอบรวมขนาดความยาว 10 15 20 และ 25 ข้อ (U10 U15 U20 U25) ดังนั้น จึงได้คะแนนของแบบสอบตามขนาดความยาวของแบบสอบรวมที่ต้องการเทียบมาตรฐานในแต่ละฉบับ เป็นคนคนเดียวกัน มีจำนวนกลุ่มตัวอย่างขนาด 405 คน ทุกขนาดความยาวของแบบสอบรวมในแบบสอบแต่ละฉบับ
4. นำแบบสอบที่ได้จัดเตรียมไว้ไปสอบกับกลุ่มสอบทานผล จำนวน 136 คน โดยให้นักเรียนแต่ละคน ในแต่ละห้องเรียน แต่ละโรงเรียน ได้สอบแบบสอบทั้ง 2 ฉบับ โดยสุ่มอย่างมีระบบให้แต่ละคนสอบแบบสอบ ฉบับ X ก่อนแล้วตามด้วย ฉบับ Y หรือสอบ ฉบับ Y ก่อนแล้วตามด้วยฉบับ X โดยสอบวันละ 1 ฉบับ เว้น 1 วัน จึงสอบอีก 1 ฉบับ
5. นำคะแนนของกลุ่มสอบทานผลมาดำเนินการเช่นเดียวกับกลุ่มตัวอย่างเทียบมาตรฐานตามข้อ 3 จะได้คะแนนจากผลการสอบแต่ละคู่ในแต่ละความยาวของแบบสอบรวม ตัดนักเรียนที่เข้าสอบไม่ครบทั้ง 2 ครั้ง จำนวน 19 คนออกไป เพื่อให้ได้กลุ่มสอบทานผลที่สอบแบบสอบทั้ง 2 ฉบับมีจำนวน 117 คน

การวิเคราะห์ข้อมูล

1. การวิเคราะห์ข้อมูลกลุ่มตัวอย่างเทียบมาตรฐาน

1.1 วิเคราะห์ข้อสอบตามทฤษฎีดั้งเดิม (Classical Theory)

โดยใช้โปรแกรม ITEMX ที่สถาบันบริการคอมพิวเตอร์จุฬาลงกรณ์มหาวิทยาลัย วิเคราะห์รายข้อของแบบสอบแต่ละฉบับ เพื่อหาค่าความยากง่าย อำนาจจำแนก และค่าความเที่ยงของแบบสอบ

1.2 หาค่าสถิติพื้นฐานของกลุ่มตัวอย่าง

โดยใช้โปรแกรมสำเร็จรูป SPSSX (Statistical Package for Social Science Version X) จากสถาบันบริการคอมพิวเตอร์จุฬาลงกรณ์มหาวิทยาลัย วิเคราะห์หาค่าเฉลี่ย (Mean) ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation)

1.3 เทียบมาตรฐานเชิงเส้นตรง (Linear Equating)

ตามสูตรซึ่งใช้ในแบบแผนการรวบรวมข้อมูลของแองกอฟฟ์ รูปแบบที่ 3 (Angoff 1984: 104-109) เมื่อแบบสอบเทียบมาตรฐานมีความเที่ยงไม่เท่ากัน (3B) มีขั้นตอนในการคำนวณดังนี้

1.3.1 คำนวณค่าความชัน (Slope) และจุดตัดแกน (intercept) ของสมการเส้นตรง เพื่อแปลงคะแนนของสเกล y ให้อยู่ในสเกล x จากสูตร

$$A_{XY} = (b_{YU\beta} \hat{\sigma}_X^2) / (b_{XU\alpha} \hat{\sigma}_Y^2)$$

$$B_{XY} = \hat{\mu}_X - A_{XY} \hat{\mu}_Y$$

$$\hat{\mu}_X = M_{X\alpha} + b_{XU\alpha} (\hat{\mu}_U - M_{U\alpha})$$

$$\hat{\sigma}_X^2 = S_{X\alpha}^2 + b_{XU\alpha}^2 (\hat{\sigma}_U^2 - S_{U\alpha}^2)$$

$$\hat{\mu}_Y = M_{Y\beta} + b_{YU\beta} (\hat{\mu}_U - M_{U\beta})$$

$$\hat{\sigma}_Y^2 = S_{Y\beta}^2 + b_{YU\beta}^2 (\hat{\sigma}_U^2 - S_{U\beta}^2)$$

$$\text{โดยที่ } \hat{\mu}_U = M_{Ut} ; \hat{\sigma}_U^2 = S_{Ut}^2 ; t = \alpha + \beta$$

$\hat{\mu}_X, \hat{\sigma}_X^2$ คือ ค่าประมาณค่าเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่าง
ที่ทาแบบสอบชุด x ที่ปรับแล้ว

$\hat{\mu}_Y, \hat{\sigma}_Y^2$ คือ ค่าประมาณค่าเฉลี่ย และความแปรปรวนของกลุ่มตัวอย่าง
ที่ทาแบบสอบชุด y ที่ปรับแล้ว

$M_{X\alpha}, S_{X\alpha}^2$ คือ ค่าเฉลี่ย และความแปรปรวนของคะแนนชุด x

$M_{Y\beta}, S_{Y\beta}^2$ คือ ค่าเฉลี่ย และความแปรปรวนของคะแนนชุด y

$M_{U\alpha}, S_{U\alpha}^2$ คือ ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบสอบร่วม U จากกลุ่ม α

$M_{Y\beta}, S_{U\beta}^2$ คือ ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบสอบร่วม U จากกลุ่ม β

M_{Ut}, S_{Ut}^2 คือ ค่าเฉลี่ย และความแปรปรวนของคะแนนแบบสอบร่วม U จากกลุ่ม
รวม t

$b_{XU\alpha}$ คือ สัมประสิทธิ์ถดถอย X จาก U ในกลุ่มตัวอย่าง α

$b_{YU\beta}$ คือ สัมประสิทธิ์ถดถอย Y จาก U ในกลุ่มตัวอย่าง β

1.3.2 นำค่าพารามิเตอร์มาสร้างสมการเส้นตรงเพื่อแปลงคะแนน y
ให้อยู่ในมาตราของ x

$$x^* = e(y) = AY + B \quad (1)$$

1.3.3 สร้างตารางเทียบมาตราคะแนนจากแบบสอบฉบับ Y ไปยังคะแนน
ตามมาตราของแบบสอบฉบับ X ซึ่งได้ตารางของคะแนนแปลงทั้งหมด 4 ตาราง

1.4 การวิเคราะห์ความคลาดเคลื่อนมาตรฐาน

1.4.1 คำนวณค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราเชิง
เส้นตรง โดยใช้สูตรการประมาณความคลาดเคลื่อนที่เสนอโดยลอร์ด (Lord 1950 cited
by Angoff 1984: 106) ดังนี้

$$SEE^2 = 2\hat{\sigma}_X^2 (1 - \hat{r}^2) \frac{(1 + \hat{r}^2)Z_Y^2 + 2}{N_t} \quad (2)$$

เมื่อ SEE^2 คือ ความแปรปรวนของความคลาดเคลื่อนในการเทียบมาตรฐานของคะแนนที่แปลง

$\hat{\sigma}_X^2$ คือ ค่าประมาณความแปรปรวนของคะแนนฉบับ X ที่ปรับแล้ว

\hat{r} คือ ค่าประมาณของความสัมพันธ์ระหว่างแบบสอบ Y กับ U

โดยมี
$$\hat{r} = \frac{b_{YU}\hat{\sigma}_U}{\hat{\sigma}_Y}$$

การคำนวณค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานได้คำนวณทั้งหมด 4 ชุด ตามจำนวนการแปลงที่กระทำในขั้นก่อน แต่เนื่องจากค่าความคลาดเคลื่อนมาตรฐาน จะแตกต่างกันตามระดับของคะแนนและมีลักษณะสมมาตรกัน เมื่อคะแนนอยู่เหนือและต่ำกว่าค่าเฉลี่ยจึงกำหนดการเสนอค่าความคลาดเคลื่อนมาตรฐานที่ค่าคะแนนมาตรฐาน Z_Y เท่ากับ 0.0 0.5 1.0 1.5 2.0 2.5 และ 3.0

1.4.2 คำนวณค่าประสิทธิภาพสัมพัทธ์ (Relative efficiency: RE) เพื่อเปรียบเทียบขนาดของความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน เมื่อใช้แบบสอบร่วมขนาด 10 15 และ 20 ข้อ เทียบกับขนาด 25 ข้อ ซึ่งเป็นขนาดความยาวที่ให้ค่าความคลาดเคลื่อนมาตรฐานน้อยที่สุดตามสมมุติฐาน ดังนี้

$$RE_{U10.U25} = \frac{SEEX^*(U25)}{SEEX^*(U10)} \times 100\%$$

$$RE_{U15.U25} = \frac{SEEX^*(U25)}{SEEX^*(U15)} \times 100\%$$

$$RE_{U20.U25} = \frac{SEEX^*(U25)}{SEEX^*(U20)} \times 100\%$$

2. การวิเคราะห์กลุ่มสอบทานผล

วิธีวิเคราะห์กลุ่มสอบทานผล เป็นวิธีประเมินความคงที่ (Stability) ในการเทียบมาตราโดยโคเลนและวิทนี (Kolen and Whitney 1982) เป็นผู้แนะนำให้ใช้ข้อมูลจากคะแนนของผู้สอบเองเป็นเกณฑ์ในการหาความแตกต่าง ข้อมูลเหล่านี้ได้มาจากการออกแบบใช้กลุ่มสอบทานผล ซึ่งสุ่มมาจากประชากรเดียวกันกับกลุ่มตัวอย่างเทียบมาตรา (ที่ใช้สร้างตารางเทียบคะแนน) และไม่มีหน่วยตัวอย่างที่ซ้ำกัน ได้รับการทดสอบด้วยแบบสอบเทียบมาตราทั้งสองฉบับ โดยนำผลสอบจากแบบสอบเทียบมาตราฉบับ Y (ฉบับปีการศึกษา 2533) มาแปลงคะแนนจากตารางเทียบมาตราให้เป็นคะแนนในแบบสอบฉบับ X (x^*) แล้ววัดความแตกต่างระหว่างคะแนนที่แปลงกับคะแนนสอบจากแบบสอบฉบับ X ซึ่งเป็นคะแนนเกณฑ์ (ฉบับปีการศึกษา 2532) ถ้าความแตกต่างโดยเฉลี่ยทั้งกลุ่มมีค่าน้อย แสดงว่ารูปแบบการเทียบมาตราที่นำมาสร้างตารางคะแนนแปลงนั้น มีความเหมาะสมและเพียงพอที่จะให้ผลการแปลงคะแนนอย่างคงเส้นคงวา มีขั้นตอนดังนี้

2.1 คำนวณค่าดัชนีความแตกต่าง (Discrepancy Index: C) จากสูตรของปีเตอร์เซนและคณะ (Petersen and others 1982:71-135)

$$c = \frac{\sum (x_i - x_i^*)^2}{nS_x^2} \quad (3)$$

เมื่อ c คือ ค่าดัชนีความแตกต่าง

x_i คือ คะแนนจากแบบสอบฉบับ x ของผู้สอบคนที่ i

x_i^* คือ คะแนนจากแบบสอบฉบับ x ที่ได้จากการนำคะแนนจากแบบสอบฉบับ y ไปแปลงจากตารางเทียบมาตราของคนที่ i

n คือ จำนวนคนของกลุ่มสอบทานผล

S_x^2 คือ ค่าความแปรปรวนของการแจกแจงคะแนนจากแบบสอบฉบับ x

2.2 ประเมินค่าดัชนีความแตกต่าง (C) โดยอาศัยหลักเกณฑ์การประเมินของปีเตอร์เซนและคณะ (Petersen and Others 1982: 93-94) ซึ่งกำหนดคุณภาพในการเทียบมาตรฐานตามระดับการยอมรับ ดังนี้

ระดับน่าพอใจอย่างมาก เมื่อ $C \leq (.05 SD_x)^2$

ระดับน่าพอใจ เมื่อ $(.05 SD_x)^2 < C \leq (.10 SD_x)^2$

ระดับปานกลาง เมื่อ $(.10 SD_x)^2 < C \leq (.15 SD_x)^2$

ระดับไม่น่าพอใจ เมื่อ $(.15 SD_x)^2 < C \leq (.20 SD_x)^2$

ระดับไม่น่าพอใจอย่างยิ่ง เมื่อ $(.20 SD_x)^2 \leq C$

2.3 ทดสอบความมีนัยสำคัญของความแตกต่างของค่าดัชนีความแตกต่าง (C) โดยใช้ Friedman test ซึ่งเป็นสถิติทดสอบนัยพาราเมตริก(รายละเอียดในภาคผนวก) โดยการวิเคราะห์ความแปรปรวนสองทางที่ใช้ข้อมูลจัดลำดับที่ทำการวัดซ้ำ และไม่ต้องมีข้อตกลงเบื้องต้นของการแจกแจงของค่าดัชนีความแตกต่าง (Marascuilo and McSweeney 1977: 357-361)