

## บทที่ 3

### วิธีการวิจัย

การศึกษานี้เป็นการศึกษาเชิงประจักษ์ ( Empirical Study ) มีวัตถุประสงค์เพื่อเปรียบเทียบความคลาดเคลื่อนและตรวจสอบความเพียงพอของการเทียบมาตรารูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับสามพารามิเตอร์ ในแบบสอบวัดผลสัมฤทธิ์ วิชาคณิตศาสตร์ มีรายละเอียดดังนี้

### ประชากร

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ปีการศึกษา 2539 ภาคเรียนที่ 2 ของโรงเรียนสังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ จังหวัดเพชรบุรี จำนวน 4,819 คน

### กลุ่มตัวอย่าง

กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่ม ดังนี้

1. กลุ่มเทียบมาตรา เบิร์นแนนและโคเลน (Bernnan and Kolen, 1987 : 285) แนะนำว่ากลุ่มเทียบมาตราไม่ควรน้อยกว่า 400 คน ดังนั้นผู้วิจัยจึงใช้กลุ่มตัวอย่าง 2 กลุ่ม กลุ่มที่สอบแบบสอบฉบับ X จำนวน 416 คน และกลุ่มที่สอบแบบสอบฉบับ Y จำนวน 418 คน ได้จากการสุ่มอย่างง่าย (Simple Random Sampling) จากจำนวนประชากรทั้งหมด
2. กลุ่มสอบทานผล โคเลนและวิทนี (Kolen and Whitney, 1982) เสนอว่า กลุ่มสอบทานผลควรมีขนาดไม่น้อยกว่า 10 % ของกลุ่มเทียบมาตรา โดยไม่มีหน่วยตัวอย่างที่ซ้ำกับกลุ่มเทียบมาตรา ในที่นี้ใช้กลุ่มทานผลจำนวน 91 คน ได้จากการสุ่มอย่างง่าย (Simple Random Sampling) การสุ่มจำนวนกลุ่มตัวอย่างจากจำนวนประชากร 22 โรงเรียน จำนวนนักเรียนทั้งหมด 4,819 คน ทำการสุ่มอย่างง่าย (Simple Random Sampling) ทั้งกลุ่มเทียบมาตราและกลุ่มทานผล

โดยมีโรงเรียนเป็นหน่วยการสุ่ม ได้ทั้งหมด 9 โรงเรียน ได้กลุ่มตัวอย่างดังนี้

ตารางที่ 5 จำนวนกลุ่มเทียบมาตราและกลุ่มทานผล

โรงเรียน	กลุ่มเทียบมาตรา	กลุ่มทานผล
โรงเรียนพรหมานุสรณ์	84	10
โรงเรียนเบญจมเทพอุทิศ	81	10
โรงเรียนบางจามวิทยา	103	10
โรงเรียนบ้านแหลมวิทยา	87	10
โรงเรียนเขาย้อยพิทยาคม	106	10
โรงเรียนหนองหญ้าปล้อง	91	10
โรงเรียนบ้านลาดวิทยา	122	11
โรงเรียนหนองจอกวิทยา	95	10
โรงเรียนชะอำคุณหญิงเนื่องบุรี	65	10
รวม	834	91

#### แบบแผนการวิจัย

การเทียบมาตรานี้ออกแบบรวบรวมข้อมูลโดยแบ่งกลุ่มตัวอย่างเป็น 2 กลุ่มดังนี้

1. กลุ่มเทียบมาตรา แบ่งโดยการสุ่มอย่างง่าย 2 กลุ่ม คือ
  - 1.1 กลุ่ม  $\alpha$  จำนวน 416 คน ให้สอบแบบสอบฉบับ X
  - 1.2 กลุ่ม  $\beta$  จำนวน 418 คน ให้สอบแบบสอบฉบับ Y
2. กลุ่มสอบทานผล จำนวน 91 คน จะต้องสอบแบบสอบทั้ง 2 ฉบับ คือ ฉบับ X และฉบับ Y

## เครื่องมือในการวิจัย

การวิจัยนี้ผู้วิจัยได้สร้างแบบสอบวัดผลสัมฤทธิ์วิชาคณิตศาสตร์ (ค 204) เรื่องสมการ และอสมการ เป็นแบบสอบเทียบมาตรา 2 ฉบับๆ ละ 25 ข้อ เป็นข้อสอบชนิดเลือกตอบ 4 ตัวเลือก มีรายละเอียดของการสร้างแบบสอบดังต่อไปนี้

### ขั้นตอนการสร้างแบบสอบเทียบมาตรา

1. ศึกษาหลักสูตร คู่มือครู และแบบเรียน วิชาคณิตศาสตร์(ค 204) ตามหลักสูตรของ กระทรวงศึกษาธิการ พ.ศ. 2521 ฉบับปรับปรุง พ.ศ. 2533
2. สร้างตารางวิเคราะห์เนื้อหา จุดประสงค์ และน้ำหนักของข้อสอบ แล้วผู้วิจัยสร้างแบบสอบ ขึ้น 2 ฉบับ ๆ ละ 50 ข้อ โดยสร้างมากกว่าจำนวนที่ต้องการใช้เพื่อประโยชน์ในการคัดเลือกข้อสอบ แล้วนำไปตรวจสอบความตรงตามเนื้อหา โดยให้ผู้เชี่ยวชาญซึ่งมีประสบการณ์ในการสอนวิชาคณิตศาสตร์ (ค 204) ไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน ตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ ตามวิธีของ โรวินเนลลีและแฮมเบิลตัน (Rovinelli and Hambleton, 1937 อ้างจาก บุญเชิด ภิญโญอนันตพงษ์, 2527 : 67 - 70) มีสูตรดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ IOC คือ ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์

$\sum R$  คือ ผลรวมคะแนนความคิดเห็นของผู้เชี่ยวชาญ

N คือ จำนวนผู้เชี่ยวชาญ

โดยพิจารณาคะแนนการตัดสินใจมีค่ามากกว่า 0.5 ถือว่าข้อสอบนั้นวัดได้สอดคล้องกับจุดประสงค์นั้นจริง มีน้ำหนักคะแนนดังนี้

- +1 หมายถึง แน่ใจว่าข้อสอบวัดได้ตรงกับจุดประสงค์การเรียนรู้
- 0 หมายถึง ไม่แน่ใจว่าข้อสอบวัดได้ตรงกับจุดประสงค์การเรียนรู้
- 1 หมายถึง แน่ใจว่าข้อสอบวัดได้ไม่ตรงกับจุดประสงค์การเรียนรู้

3. คัดเลือกข้อสอบที่วัดได้ตรงกับจุดประสงค์การเรียนรู้ไว้เพื่อสร้างแบบสอบจำนวน 2 ฉบับๆ ละ 50 ข้อ นำไปทดลองใช้ เพื่อหาคุณภาพข้อสอบแต่ละข้อ ก่อนนำไปใช้กับกลุ่มตัวอย่างจริง

โดยนำไปสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนคงคาราม จำนวน 81 คน และโรงเรียนท่าขางวิทยา จำนวน 91 คน รวม 172 คน แล้ววิเคราะห์หาคุณภาพของข้อสอบแบบทฤษฎีการทดสอบดั้งเดิม (Classical Test Theory) โดยใช้โปรแกรม ITEM ของอาจารย์ทวี ทองคำ คณะศึกษาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี เพื่อวิเคราะห์ ค่าอำนาจจำแนก (D) และค่าความยาก (p) โดยใช้เกณฑ์ดังนี้ (บุญธรรม กิจปริคาบวิสุทธิ, 2535 : 193-141)

- 1) ค่าอำนาจจำแนกมากกว่า + 0.2
- 2) ค่าความยากอยู่ระหว่าง 0.2 - 0.8

ซึ่งแบบสอบฉบับ X ได้จำนวนข้อสอบ 39 ข้อ และแบบสอบฉบับ Y ได้จำนวนข้อสอบ 36 ข้อ

4. นำแบบสอบที่วิเคราะห์จากทฤษฎีการทดสอบดั้งเดิมแล้วไปทดลองใช้ (Try Out) อีกครั้งหนึ่ง กับนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของโรงเรียนวัดจันทราวาส จำนวน 95 คน และโรงเรียนเขาชัยวิทยา จำนวน 114 คน รวม 209 คน

5. หาจำนวนองค์ประกอบของแบบสอบเพื่อทดสอบความเป็นเอกพันธ์ของแบบสอบ (Unidimensionality Assumption) โดยการวิเคราะห์องค์ประกอบ (Factor Analysis) ใช้โปรแกรม Statistica การวิเคราะห์ใช้วิธีการวิเคราะห์ตัวประกอบสำคัญ (Principal Component) แล้วหมุนแกนด้วยวิธีแวนริแมกซ์ (Varimax) สำหรับการตรวจสอบพิจารณา ค่าไอเกน (Eigen Value) ของตัวประกอบที่ 1 ของแบบทดสอบมีค่าสูงกว่าค่าไอเกนของตัวประกอบที่ 2 อย่างมาก และค่าไอเกนของตัวประกอบที่ 2 สูงกว่าค่าไอเกนของตัวประกอบที่ถัดๆ ไปเพียงเล็กน้อย จึงจะกล่าวได้ว่าแบบสอบฉบับนี้วัดเพียงคุณลักษณะเดียว (Lord, 1980 : 21) ผลการทดสอบความเป็นเอกพันธ์ของแบบสอบปรากฏว่า

5.1 แบบสอบฉบับ X ได้ค่าไอเกนของตัวประกอบที่หนึ่งเท่ากับ 8.10 ค่าไอเกนของตัวประกอบที่สองเท่ากับ 3.37 และค่าไอเกนของตัวประกอบที่สามเท่ากับ 2.23 จึงอนุมานได้ว่าแบบสอบฉบับนี้วัดเพียงคุณลักษณะเดียว

5.2 แบบสอบฉบับ Y ได้ค่าไอเกนของตัวประกอบที่หนึ่งเท่ากับ 7.23 ค่าไอเกนของตัวประกอบที่สองเท่ากับ 3.26 และค่าไอเกนของตัวประกอบที่สามเท่ากับ 2.19 จึงอนุมานได้ว่าแบบสอบฉบับนี้วัดเพียงคุณลักษณะเดียว

6. วิเคราะห์หาคุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) โดยใช้โปรแกรม BILOG 3 เพื่อหา ค่าอำนาจจำแนก (a) ค่าความยาก (b) ค่าสัมประสิทธิ์การเดา (c) และโค้งคุณลักษณะข้อสอบ (Item Characteristic Curve) ที่เหมาะสมกับกลุ่มตัวอย่าง โดยวิเคราะห์ 2 ครั้ง คือ วิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์ และวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ แล้วคัดเลือกข้อสอบที่มีคุณภาพ ดังนี้

6.1 รูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์ โดยพิจารณาข้อสอบที่มีค่าความยากของข้อสอบอยู่ระหว่าง - 2.0 ถึง + 2.0 (สุพัฒน์ สุกลมณัตต์, 2538 : 72)

6.2 รูปแบบทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ โดยใช้เกณฑ์ดังนี้ (สุพัฒน์ สุกลมณัตต์, 2538 : 71-73)

6.2.1 ค่าอำนาจจำแนกอยู่ระหว่าง 0.5 ถึง 2.5

6.2.2 ค่าความยากของข้อสอบอยู่ระหว่าง - 2.0 ถึง + 2.0

6.2.3 ค่าสัมประสิทธิ์การเดาต่ำกว่า 0.3

จากการคัดเลือกข้อสอบตามเกณฑ์ที่กำหนดไว้ ดังนั้นแบบสอบฉบับ X ได้ข้อสอบจำนวน 29 ข้อ และแบบสอบฉบับ Y ได้ข้อสอบจำนวน 26 ข้อ

7. นำข้อสอบที่คัดเลือกไว้มาจัดเป็นกลุ่มข้อสอบที่วัตถุประสงค์เดียวกัน

8. จัดข้อสอบในแต่ละกลุ่มออกเป็นคู่ๆ โดยพิจารณาค่าความยากของข้อสอบ (b) ที่เท่าเทียมกัน หรือใกล้เคียงกัน

9. คู่ข้อสอบในแต่ละคู่เพื่อนำไปสร้างเป็นแบบสอบ 2 ฉบับ จะได้แบบสอบจำนวน 2 ฉบับ ๆ ละ 25 ข้อ

10. เมื่อนำแบบสอบทั้งฉบับ X และฉบับ Y ไปสอบกับกลุ่มตัวอย่างทั้งกลุ่มเทียบมาตรฐานและกลุ่มสอบทานผล

11. นำผลการสอบไปวิเคราะห์ตามทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งและสามพารามิเตอร์แล้วจะได้จำนวนข้อสอบตรงตามเกณฑ์ที่กำหนดไว้ดังนี้

1.1.1 แบบสอบฉบับ X จำนวน 18 ข้อ

1.1.2 แบบสอบฉบับ Y จำนวน 18 ข้อ

### การรวบรวมข้อมูล

1. ติดต่อขอความร่วมมือในการนำแบบสอบไปสอบกับนักเรียนทั้งกลุ่มเทียบมาตรฐานและกลุ่มสอบทานผล กำหนดวันและเวลาในการสอบ
2. นำแบบสอบที่เตรียมไว้ทั้ง 2 ฉบับ ไปสอบกับกลุ่มเทียบมาตรฐานโดยนักเรียนแต่ละคนจะตอบแบบสอบเพียงคนละ 1 ฉบับเท่านั้น ทั้งนี้โดยการแจกข้อสอบแบบสุ่ม เพื่อให้การสอบแบบสอบเทียบมาตรฐานทั้ง 2 ฉบับ มีลักษณะเหมือนกันแต่ละกลุ่มและความสามารถของนักเรียนแต่ละกลุ่มใกล้เคียงกัน
3. นำคะแนนดิบที่ได้จากกลุ่มเทียบมาตรฐาน ไปหาค่าพารามิเตอร์ของคนและข้อสอบรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ จากนั้นจึงคำนวณคะแนนความสามารถ ( $\theta$ ) และค่าคะแนนจริงของนักเรียนเพื่อไปสร้างตารางเทียบมาตรฐาน จะได้ตารางเทียบมาตรฐาน 2 ตาราง คือ ตารางเทียบมาตรฐานรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับตารางเทียบมาตรฐานรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์
4. นำแบบสอบที่เตรียมไว้สำหรับกลุ่มสอบทานผล ไปสอบกับกลุ่มสอบทานผล โดยให้นักเรียนแต่ละคนสอบแบบสอบทั้ง 2 ฉบับ โดยสุ่มนักเรียนครั้งหนึ่งให้สอบแบบสอบฉบับ X ก่อนแล้วให้สอบแบบสอบฉบับ Y นักเรียนอีกครั้งหนึ่งให้สอบแบบสอบฉบับ Y ก่อนแล้วตามด้วยแบบสอบฉบับ X
5. นำคะแนนดิบที่ได้จากกลุ่มสอบทานผล ไปหาค่าพารามิเตอร์ของคนและข้อสอบรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์ จากนั้นจึงคำนวณคะแนนความสามารถ ( $\theta$ ) และค่าคะแนนจริง ของนักเรียนแต่ละคน

6. นำคะแนนจริง (จากแบบสอบฉบับ X) ที่ได้จากกลุ่มสอบทานผล ไปเปรียบเทียบกับคะแนนจริง ( $X^*$ ) จากตารางเทียบมาตรฐานเพื่อคำนวณค่าความคลาดเคลื่อนของการเทียบมาตรฐาน ทั้งการเทียบมาตรฐานรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และการเทียบมาตรฐานรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์
7. เปรียบเทียบค่าความคลาดเคลื่อนของการเทียบมาตรฐานระหว่างรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์กับทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์
8. นำคะแนนจริงที่ได้จากกลุ่มสอบทานผล ไปตรวจสอบความเพียงพอของการเทียบมาตรฐานรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์และทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์

### การวิเคราะห์ข้อมูล

#### 1. การวิเคราะห์ข้อมูล

##### 1.1 หากคุณภาพของข้อสอบตามทฤษฎีการทดสอบดั้งเดิม

1.1.1 ค่าอำนาจจำแนก (D) ใช้สูตรดังนี้ (บุญธรรม กิจปรีดาบริสุทธิ์, 2535 : 143)

$$D = \frac{P_H - P_L}{n}$$

เมื่อ  $P_H$  คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มสูง

$P_L$  คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มต่ำ

$n$  คือ จำนวนนักเรียนในกลุ่มสูงหรือต่ำ

1.1.2 ค่าความยาก (p) ใช้สูตรดังนี้ (บุญธรรม กิจปรีดาบริสุทธิ์, 2535 : 143)

$$P = \frac{P_H + P_L}{2n}$$

เมื่อ  $P_H$  คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มสูง

$P_L$  คือ จำนวนนักเรียนที่ตอบถูกในกลุ่มต่ำ

$n$  คือ จำนวนนักเรียนในกลุ่มสูงหรือต่ำ

1.2 หากคุณภาพของข้อสอบตามทฤษฎีการตอบสนองข้อสอบ

1.2.1 รูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์

ค่าความยาก ( $b$ ) และโค้งลักษณะข้อสอบ (Item Characteristic Curve) โดยใช้

โปรแกรม BILOG 3

1.2.2 รูปแบบทฤษฎีการตอบสนองข้อสอบที่มีสามพารามิเตอร์

ค่าอำนาจจำแนก ( $a$ ) ค่าความยาก ( $b$ ) ค่าสัมประสิทธิ์การเดา ( $c$ ) และ

โค้งลักษณะข้อสอบ (Item Characteristic Curve) โดยใช้โปรแกรม BILOG 3

1.3 การเทียบมาตราโดยรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีหนึ่งพารามิเตอร์ ดังนี้

1.3.1 หากค่าความน่าจะเป็นในการตอบถูกของข้อสอบที่ระดับความสามารถต่างๆ โดยการนำค่าความสามารถ ( $\theta$ ) ของผู้สอบแต่ละคนที่ได้จากโปรแกรม BILOG 3 ไปหาค่า  $P_i(\theta)$  จาก

โปรแกรม Microsoft Excel 5.0 จากสูตร

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

เมื่อ  $P_i(\theta)$  คือ ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ  $\theta$  ตอบข้อสอบข้อ  $i$  ถูกต้อง

$\theta$  คือ ระดับความสามารถของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบสอบ

$a_i$  คือ ค่าอำนาจจำแนกของข้อสอบข้อ  $i$

$b_i$  คือ ค่าความยากของข้อสอบข้อ  $i$

$c_i$  คือ ค่าการเดาของข้อสอบข้อ  $i$

$D$  คือ Scaling Factor มีค่าเท่ากับ 1.7

$e$  คือ ค่าคงที่มีค่าเท่ากับ 2.71828...

1.3.2 การประมาณคะแนนจริงของผู้สอบแต่ละคน ซึ่งค่าประมาณที่ใช้ คือ คะแนนจริงตามความน่าจะเป็นในการตอบถูกทั้งหมดของผู้สอบ จากโปรแกรม Statistica คังสูตร (Lord, 1980 : 202)

$$\xi = \xi(\theta) = \sum_{i=1}^m P_i(\theta_i)$$



$$\eta = \eta(\theta) = \sum_{j=1}^n P_j(\theta_y)$$

- เมื่อ  $\xi$  คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ X  
 $\eta$  คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ Y  
 $m$  คือ จำนวนข้อของแบบสอบฉบับ X  
 $n$  คือ จำนวนข้อของแบบสอบฉบับ Y  
 $\theta_x$  คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ X  
 $\theta_y$  คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ Y

1.4 การเทียบมาตรฐานโดยรูปแบบทฤษฎีการตอบสนองข้อสอบที่มีตามพารามิเตอร์ ดังนี้

1.4.1 หากค่าความน่าจะเป็นในการตอบถูกของข้อสอบที่ระดับความสามารถต่างๆ โดยการนำค่าความสามารถ ( $\theta$ ) ของผู้สอบแต่ละคนที่ได้จากโปรแกรมโปรแกรม BILOG 3 ไปหาค่า  $P_i(\theta)$  จากโปรแกรม Microsoft Excel 5.0 จากสูตร

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

เมื่อ  $P_i(\theta)$  คือ ความน่าจะเป็นที่ผู้สอบที่มีระดับความสามารถ  $\theta$  ตอบข้อสอบข้อ  $i$  ถูกต้อง

$\theta$  คือ ระดับความสามารถของผู้สอบที่ประมาณได้จากคะแนนรวมของการตอบแบบสอบ

$a_i$  คือ ค่าอำนาจจำแนกของข้อสอบข้อ  $i$

$b_i$  คือ ค่าความยากของข้อสอบข้อ  $i$

$c_i$  คือ ค่าการเดาของข้อสอบข้อ  $i$

$D$  คือ Scaling Factor มีค่าเท่ากับ 1.7

$e$  คือ ค่าคงที่มีค่าเท่ากับ 2.71828...

1.4.2 การประมาณคะแนนจริงของผู้สอบแต่ละคน ซึ่งค่าประมาณที่ใช้ คือ คะแนนจริงตามความน่าจะเป็นในการตอบถูกทั้งหมดของผู้สอบ จากโปรแกรม Statistica ดังสูตร (Lord, 1980 : 199)

$$\xi = \xi(\theta) = \sum_{i=1}^m P_i(\theta_x)$$

$$\eta = \eta(\theta) = \sum_{j=1}^n P_j(\theta_y)$$

- เมื่อ  $\xi$  คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ X  
 $\eta$  คือ คะแนนจริงของผู้สอบที่ได้จากแบบสอบฉบับ Y  
 $m$  คือ จำนวนข้อของแบบสอบฉบับ X  
 $n$  คือ จำนวนข้อของแบบสอบฉบับ Y  
 $\theta_x$  คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ X  
 $\theta_y$  คือ ค่าความสามารถของผู้สอบที่ได้จากแบบสอบฉบับ Y

1.5 ปรับคะแนนที่อยู่ต่ำกว่าระดับการคาดด้วยวิธีการแปลงเชิงเส้นตรง (Linear Interpolation) ปรับจากคะแนนจริงจากแบบสอบ X และฉบับ Y ที่อยู่ต่ำกว่าระดับการคาด (Lord, 1982 : 211 อ้างจาก ภาวิณี ศรีสุขวัฒนานันท์, 2529 : 95) โดยใช้สูตร

$$x^* = x(y) = \frac{\sum_i c_i (1 - c_i)}{\sqrt{\sum_j c_j (1 - c_j)}} (y - \sum_j c_j) + \sum_j c_j$$

- เมื่อ  $x^*$  คือ คะแนนจริงของแบบสอบฉบับที่ 1 ที่สมมูลกับคะแนน  $y$   
 $c_i$  คือ คะแนนการคาดของแบบสอบฉบับที่ 1 ข้อที่  $i$   
 $c_j$  คือ คะแนนการคาดของแบบสอบฉบับที่ 2 ข้อที่  $j$   
 $y$  คือ คะแนนจริงของแบบสอบฉบับที่ 2

1.6 หาค่าความคลาดเคลื่อนของการเทียบมาตราแต่ละรูปแบบ โดยใช้สูตร (พรทิมล นาคเวช, 2537 : 63-34)

$$E = |X_i - X_i^*|$$

- เมื่อ  $E$  คือ ค่าความคลาดเคลื่อน  
 $X_i$  คือ คะแนนจริงของแบบสอบคนที่  $i$  ที่ได้จากการสอบ  
 $X_i^*$  คือ คะแนนจริงของแบบสอบคนที่  $i$  ที่ได้จากการเทียบมาตรา

## 2. ทดสอบความแตกต่างของความคลาดเคลื่อนของการเทียบมาตรา

ทดสอบความแตกต่างของความคลาดเคลื่อนของการเทียบมาตราทั้งสองรูปแบบโดยใช้สูตร  $t$ -test แบบกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระต่อกัน ( $t$ -Independent) ใช้สูตรดังนี้ (James D. Wynne, 1982 : 169)

$$t = \frac{\sum D}{\sqrt{\frac{N \sum D^2 - (\sum D)^2}{N-1}}}$$

เมื่อ  $\sum D$  คือ ผลรวมของผลต่างระหว่างค่าความคลาดเคลื่อนของการเทียบมาตราแต่ละคู่

$\sum D^2$  คือ ผลรวมกำลังสองของผลต่างระหว่างค่าความคลาดเคลื่อนของการเทียบมาตราแต่ละคู่

$N$  คือ จำนวนคู่ของผลต่าง

### 3. การตรวจสอบความเพียงพอของการเทียบมาตรา

ใช้วิธีการประเมินความคงที่ (Stability) ในการเทียบมาตราโดยโคเลนและวิทนี (Kolen and Whitney, 1982) แนะนำให้ใช้ข้อมูลจากคะแนนของผู้สอบเองเป็นเกณฑ์ในการหาความแตกต่าง โดยข้อมูลเหล่านี้ได้มาจากการออกแบบใช้กลุ่มทานผล ที่ได้มาจากประชากรเดียวกันกับกลุ่มเทียบมาตรา และต้องไม่มีหน่วยตัวอย่างซ้ำกัน เพื่อให้สอบแบบสอบทั้ง 2 ฉบับ นำผลจากแบบสอบฉบับ Y มาแปลงคะแนนจากตารางเทียบมาตราให้เป็นคะแนนในแบบสอบฉบับ X แล้วหาความแตกต่างระหว่างคะแนนที่แปลงกับคะแนนสอบจากฉบับ X ซึ่งเป็นคะแนนเกณฑ์ มีขั้นตอนและเกณฑ์การประเมินดังนี้

3.1 คำนวณค่าดัชนีความแตกต่าง (Discrepancy Index : C) (ภาวิณี ศรีสุขวัฒนานันท์, 2529 : 102)

$$C = \frac{\sum (X_i - X_i^*)^2}{nS_x^2}$$

เมื่อ  $C$  คือ ค่าดัชนีความแตกต่าง

$X_i$  คือ คะแนนจริงจากแบบสอบฉบับ X ของผู้สอบคนที่  $i$

$X_i^*$  คือ คะแนนจริงจากแบบสอบฉบับ X ที่ได้จากการนำคะแนนจากแบบสอบฉบับ Y ไปแปลงจากตารางเทียบมาตราของคนที  $i$

$n$  คือ จำนวนคนของกลุ่มสอบทานผล

$S_x^2$  คือ ค่าความแปรปรวนของคะแนนจริงจากแบบสอบฉบับ X

3.2 ประเมินค่าดัชนีความแตกต่าง (C) โดยอาศัยหลักเกณฑ์การประเมินของปีเตอร์เซนและคณะ (Petersen, et al. 1982 : 93 - 94 อ้างจาก ภาวิณี ศรีสุขวัฒนานันท์, 2529 : 103) ซึ่งกำหนด

คุณภาพของการเทียบมาตรฐานตามระดับการยอมรับ ดังนี้

ระดับน่าพอใจอย่างยิ่ง	เมื่อ	$C \leq (0.05 S.D_{ix})^2$
ระดับน่าพอใจ	เมื่อ	$(.05 S.D_{ix})^2 < C \leq (0.10 S.D_{ix})^2$
ระดับปานกลาง	เมื่อ	$(0.10 S.D_{ix})^2 < C \leq (0.15 S.D_{ix})^2$
ระดับไม่น่าพอใจ	เมื่อ	$(0.15 S.D_{ix})^2 < C \leq (0.20 S.D_{ix})^2$
ระดับไม่น่าพอใจอย่างยิ่ง	เมื่อ	$(0.20 S.D_{ix})^2 < C$