

บทที่ 5

ผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในการศึกษาเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนระหว่างวิธีการสุ่มแบบง่าย กับสุ่มแบบเมตริกซ์เชิงพหุคูณ กับประชากร รวมทั้งการศึกษาความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ยระหว่างวิธีการสุ่มทั้งสองวิธี ซึ่งการวิจัยครั้งนี้ เป็นการวิจัยเชิงทดลองด้วยวิธีมอนติคาร์โล (Monte Carlo Method) โดยมีวัตถุประสงค์ของการวิจัย ดังนี้ (1) เพื่อศึกษาความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ย ระหว่างตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ (2) เพื่อเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่าย กับค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากร (3) เพื่อเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณกับค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากร (4) เพื่อเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่าย กับตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณ ใช้ภาษาคอมพิวเตอร์เขียนโปรแกรมให้เครื่องคอมพิวเตอร์จำลองประชากรผู้สอบและประชากรข้อสอบ และทำการตอบข้อสอบตามสูตรของราชต์ชนิดหนึ่ง-พารามิเตอร์ โลจิสติก โมเดล หาผลการตอบข้อสอบ และคำนวณค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากรเก็บค่าไว้ จากนั้นทำการสุ่มแบบง่ายและสุ่มแบบเมตริกซ์เชิงพหุคูณ นำผลการตอบข้อสอบ มาหาค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างเก็บค่าไว้ โดยมีขั้นตอนในการทดลองดังนี้

(1) ประชากรที่ใช้ในการศึกษา ได้แก่ประชากรผู้สอบจำนวน 1,200 คน โดยกำหนดค่าความสามารถของผู้สอบให้มีค่าอยู่ในช่วง -3 ถึง $+3$ มีลักษณะการแจกแจงแบบปกติ กำหนดค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 1 สัมประสิทธิ์ความเบ้เท่ากับ 0 และสัมประสิทธิ์ความโด่งเท่ากับ 3

(2) ประชากรข้อสอบจำนวน 480 ข้อ โดยกำหนดให้มีลักษณะการแจกแจงค่าความยากของข้อสอบอยู่ในช่วง -2 ถึง $+2$ และมีการแจกแจง 3 ลักษณะคือ

2.1. การแจกแจงค่าความยากของข้อสอบแบบปกติ โดยมีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 1 สัมประสิทธิ์ความเบ้เท่ากับ 0 และสัมประสิทธิ์ความโด่งเท่ากับ 3

2.2 การแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้ายหรือข้อสอบชุดง่าย โดยมีค่าเฉลี่ยเท่ากับ +2 ความแปรปรวนเท่ากับ 1 สัมประสิทธิ์ความเบ้เท่ากับ -1 และสัมประสิทธิ์ความโด่งเท่ากับ 3

2.3 การแจกแจงค่าความยากของข้อสอบแบบเบ้ขวาหรือข้อสอบชุดยาก โดยมีค่าเฉลี่ยเท่ากับ -2 ความแปรปรวนเท่ากับ 1 ค่าสัมประสิทธิ์ความเบ้เท่ากับ +1 และค่าสัมประสิทธิ์ความโด่งเท่ากับ 3

(3) ทาผลการตอบข้อสอบของประชากรผู้สอบและประชากรข้อสอบ โดยตอบถูกเท่ากับ 1 ตอบผิด เท่ากับ 0 ตามสูตรของราชค์ ชนิด หนึ่ง-พารามิเตอร์ โลจิสติก โมเดล คำนวณค่าคะแนนเฉลี่ย และความแปรปรวนคะแนนของประชากรเก็บค่า ไว้

(4) ทำการสุ่มแบบง่าย สามแผนของการทดลองที่กำหนดให้กลุ่มย่อยของการทดลองเป็น สอง สี และหก กลุ่ม และเป็นการสุ่มแบบไม่ใส่คืน มีจำนวนตัวอย่างผู้สอบ เท่ากับ 600 300 และ 200 คน ตามลำดับ ให้ตอบประชากรข้อสอบ 480 ข้อ จัดกระทำซ้ำ 1,000 ครั้ง ในแต่ละแผนของการทดลอง คำนวณค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่าง เก็บค่า ไว้

(5) ทำการสุ่มแบบเมตริกซ์เชิงพหุคูณ ตามสามแผนของการทดลองที่กำหนดให้กลุ่มย่อยของการทดลองเป็น สอง สี และหก กลุ่ม ให้มีจำนวนเท่ากันในแต่ละกลุ่มของแต่ละแผนการทดลอง และเป็นการสุ่มแบบไม่ใส่คืน โดยแผนหนึ่งมีจำนวนตัวอย่างผู้สอบ 600 คน จำนวนตัวอย่างข้อสอบ 240 ข้อ และแผนสอง มีจำนวนตัวอย่างผู้สอบ 300 คน จำนวนตัวอย่างข้อสอบ 120 ข้อ และแผนสาม มีจำนวนตัวอย่างผู้สอบ 200 คน จำนวนตัวอย่างข้อสอบ 80 ข้อ จัดกระทำซ้ำ 1,000 ครั้ง ในแต่ละแผนของการทดลอง คำนวณค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่าง

(6) นำค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนที่ได้จากประชากรและจากตัวอย่าง มาหาความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ย และเปรียบเทียบผลการทดลองระหว่างตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ และกับประชากร

สรุปผลการวิจัย

1. ผลการศึกษาความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ย ระหว่างตัวอย่างที่ลุ่มแบบง่ายและลุ่มแบบเมตริกซ์เชิงพหุคูณ ทั้งสามแผน พบว่า วิธีการลุ่มแบบง่ายมีความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ยมีแนวโน้มสูงกว่า วิธีการลุ่มแบบเมตริกซ์เชิงพหุคูณ และมีความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ยเพิ่มมากขึ้น เมื่อจำนวนตัวอย่างลดลง

2. ผลการเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่ลุ่มแบบง่ายกับค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากร

2.1 ผลการเปรียบเทียบค่าคะแนนเฉลี่ยของตัวอย่างที่ลุ่มแบบง่ายทั้งสามแผน กับค่าคะแนนเฉลี่ยของประชากร

(1) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ พบว่า ค่าคะแนนเฉลี่ยของแผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 3 ครั้ง แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 3 และ 7 ครั้ง ตามลำดับ และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 2 และ 19 ครั้ง ตามลำดับ

(2) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้ายพบว่า ค่าคะแนนเฉลี่ย ไม่พบความแตกต่างกัน ที่ระดับ .01 ส่วนที่ระดับ .05 แผนหนึ่ง จำนวนตัวอย่าง 600 คน แผนสอง จำนวนตัวอย่าง 300 คน และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญ จำนวน 1 10 และ 10 ครั้ง ตามลำดับ

(3) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ขวาพบว่า ค่าคะแนนเฉลี่ยของแผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 1 ครั้ง แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 10 ครั้ง และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 1 และ 13 ครั้ง ตามลำดับ

2.2 ผลการเปรียบเทียบความแปรปรวนคะแนนของตัวอย่างที่ลุ่มแบบง่ายกับความแปรปรวนคะแนนของประชากร

(1) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ พบว่าความแปรปรวนคะแนนของตัวอย่าง ทั้งสามแผน ไม่แตกต่างกันที่ระดับ .01 โดยแผนสอง และแผนสาม จำนวนตัวอย่าง 300 คน และ 200 คน ตามลำดับ ความแปรปรวนคะแนนแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวนแผนละ 9 ครั้ง จากการกระทำซ้ำ 1,000 ครั้งในแต่ละแผน

(2) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้ายพบว่าความแปรปรวนคะแนนของแผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 13 ครั้ง แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 2 และ 26 ครั้ง ตามลำดับ และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 3 และ 60 ครั้ง ตามลำดับ

(3) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ขวาพบว่าความแปรปรวนคะแนนของแผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 13 ครั้ง แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 1 และ 48 ครั้ง ตามลำดับ และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 39 ครั้ง

4. ผลการเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ

4.1 ผลการเปรียบเทียบค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ

(1) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ พบว่าค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ แผนหนึ่ง จำนวนตัวอย่าง 600 คน และแผนสาม จำนวนตัวอย่าง 200 คน ไม่พบความแตกต่างกัน จากการกระทำซ้ำ 1,000 ครั้ง ส่วนแผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 จำนวน 1 ครั้ง

(2) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้าย พบว่าค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ แผนหนึ่ง จำนวนตัวอย่าง 600 คน และแผนสาม จำนวนตัวอย่าง 200 คน ไม่พบความแตกต่างกัน จากการกระทำซ้ำ 1,000 ครั้ง และแผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 1 ครั้ง

(3) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ขวา พบว่าค่าคะแนนเฉลี่ยทั้งสามแผน ไม่พบความแตกต่างกัน จากการกระทำซ้ำ 1,000 ครั้ง

4.2 ผลการเปรียบเทียบความแปรปรวนคะแนนของกลุ่มตัวอย่างที่ลุ่มแบบง่าย กับลุ่มแบบเมตริกซ์เชิงพหุคูณ

(1) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ พบว่าความแปรปรวนคะแนนของกลุ่มตัวอย่างที่ลุ่มแบบง่ายกับลุ่มแบบเมตริกซ์เชิงพหุคูณ แผนหนึ่ง จำนวนตัวอย่าง 600 คน และแผนสาม จำนวนตัวอย่าง 200 คน ไม่พบความแตกต่างกัน จากการกระทำซ้ำ 1,000 ครั้ง และแผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 1 ครั้ง จากการกระทำซ้ำ 1,000 ครั้ง

(2) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้าย พบว่าความแปรปรวนคะแนนของกลุ่มตัวอย่างที่ลุ่มแบบง่ายกับลุ่มแบบเมตริกซ์เชิงพหุคูณ แผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 1 และ 9 ครั้ง ตามลำดับ แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 18 ครั้ง และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 2 และ 3 ครั้ง ตามลำดับ

(3) กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ขวา พบว่าความแปรปรวนคะแนนของกลุ่มตัวอย่างที่ลุ่มแบบง่ายกับลุ่มแบบเมตริกซ์เชิงพหุคูณ แผนหนึ่ง จำนวนตัวอย่าง 600 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 1 และ 2 ครั้ง ตามลำดับ แผนสอง จำนวนตัวอย่าง 300 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 และ .05 จำนวน 1 และ 28 ครั้ง ตามลำดับ และแผนสาม จำนวนตัวอย่าง 200 คน แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 จำนวน 1 ครั้ง

อภิปรายผล

1. ผลการศึกษาความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ย ระหว่างตัวอย่างที่ลุ่มแบบง่ายกับลุ่มแบบเมตริกซ์เชิงพหุคูณทั้งสามแผน พบว่าในการทดลองทั้งสามแผนความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ย มีค่าเพิ่มขึ้น เมื่อจำนวนตัวอย่างลดลง และความคลาดเคลื่อน

มาตรฐานของค่าคะแนนเฉลี่ย จากวิธีการสุ่มแบบง่ายมีค่าสูงกว่าวิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณ ซึ่งเป็นไปตามหลักการสุ่มที่ว่า เมื่อจำนวนตัวอย่างเพิ่มขึ้นสี่เท่า ความคลาดเคลื่อน (Error) ของการสุ่มจะลดลงครึ่งหนึ่ง (Yamane, 1973, p.205) โดยค่าความคลาดเคลื่อนมาตรฐานของค่าคะแนนเฉลี่ยจากวิธีการสุ่มแบบง่าย มีค่าสูงกว่าวิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณ สอดคล้องกับคำกล่าวของ ลอยด์ (Loyd, 1988, อ้างใน กัมพล ไชยพันธ์, 2535, หน้า 27) และชาร์ทอค (Chartock, 1980, p.638-A) ที่ว่า วิธีการสุ่มตัวอย่างแบบเมตริกซ์เชิงพหุคูณประมาณค่าคะแนนเฉลี่ยได้แม่นยำมากกว่าวิธีการสุ่มแบบง่าย รวมทั้งไซรอนนิค (Sirotnik, 1974, p.461) ที่กล่าวว่า เมื่อมีจำนวนตัวอย่างเท่ากัน วิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณมีค่าคะแนนเฉลี่ยใกล้เคียงกับค่าคะแนนเฉลี่ยของประชากรมากกว่าวิธีการสุ่มแบบง่าย

2. ผลการเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่ายกับค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากร

2.1 ผลการเปรียบเทียบค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบง่ายกับค่าคะแนนเฉลี่ยของประชากร พบว่า เมื่อจำนวนตัวอย่างเพิ่มขึ้นมีจำนวนครั้งของความแตกต่างอย่างมีนัยสำคัญทางสถิติลดลง สอดคล้องกับ ยามาเน่ (Yamane, 1973, p.205) ที่กล่าวว่า เมื่อขนาดของกลุ่มตัวอย่างเพิ่มขึ้นสี่เท่า ความคลาดเคลื่อนของการสุ่มจะลดลงครึ่งหนึ่ง โดยพิจารณาตามลักษณะการแจกแจงค่าความยากของข้อสอบ ทั้งสามแบบ มีจำนวนครั้งใกล้เคียงกันจากการกระทำซ้ำ 1,000 ครั้ง

2.2 ผลการเปรียบเทียบความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่ายกับความแปรปรวนคะแนนของประชากร พบว่า ในข้อสอบที่มีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้าย ในแผนสอง และแผนสามจำนวนตัวอย่าง 300 และ 200 คน มีจำนวนครั้งของความแตกต่างกันอย่างมีนัยสำคัญมากที่สุดคือ 60 ครั้ง จากการกระทำซ้ำ 1,000 ครั้ง และในกรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ขวา แผนสองมีจำนวนตัวอย่าง 300 คน มีจำนวนครั้งของความแตกต่างกันอย่างมีนัยสำคัญ 48 ครั้ง จากการกระทำซ้ำ 1,000 ครั้ง ขณะที่ในแผนสอง และแผน 3 จำนวนตัวอย่าง 300 และ 200 คน กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ มีจำนวนครั้งของความแตกต่างกันอย่างมีนัยสำคัญเกิดขึ้นแผนละ 9 ครั้ง จากการกระทำซ้ำ 1,000 ครั้ง จากผลการทดลองนี้พบว่ามีความแตกต่างเกิดขึ้นสูงสุดจำนวน 60 ครั้ง

จากการกระทำซ้ำ 1,000 ครั้ง หรือใน 100 ครั้ง มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ เกิดขึ้น 6 ครั้ง นั่นคือความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภท I เท่ากับ 0.6 ซึ่งสูงกว่าเกณฑ์ที่ยอมรับ จึงสรุปได้ว่าความแปรปรวนคะแนนของแผนสอง จำนวนตัวอย่าง 200 คน ในข้อสอบที่มีลักษณะการแจกแจงแบบเบ้ซ้าย แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนแผนอื่นๆ เมื่อพิจารณาจากการกระทำซ้ำ 1,000 ครั้ง พบความแตกต่างของความแปรปรวนคะแนนเกิดขึ้นไม่เกิน 50 ครั้ง คือความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภท I ไม่เกิน .05 ซึ่งอยู่ในเกณฑ์ที่ยอมรับได้ จึงสรุปได้ว่าความแปรปรวนคะแนนของตัวอย่างไม่แตกต่างจากความแปรปรวนคะแนนของประชากร ที่ระดับนัยสำคัญ .05

ในข้อสอบที่มีลักษณะการแจกแจงค่าความยากของข้อสอบแบบเบ้ซ้ายพบจำนวนครั้งของความแตกต่างกันอย่างมีนัยสำคัญมากกว่าแบบเบ้ขวา และมากกว่าแบบปกติตามลำดับนั้น เกิดจากในข้อสอบชุดยากคนอ่อนที่ควรตอบข้อสอบผิดเกิดความผิดพลาดไปตอบข้อสอบข้อนั้นถูก และในข้อสอบชุดง่ายคนเก่งที่ควรตอบข้อสอบข้อนั้นถูกกลับไปตอบผิด จึงเกิดความแปรปรวนคะแนนเพิ่มขึ้นมากกว่าความแปรปรวนคะแนนของข้อสอบชุดปกติ สอดคล้องกับงานของ สมพร บุญอิม (2529, หน้า 5) และไรท์ (Wright, 1949, p.134) ที่พบว่าลักษณะการแจกแจงค่าความยากของข้อสอบทำให้คะแนนที่ได้รับจากการทดสอบมีลักษณะการแจกแจงไม่เป็นปกติตามไปด้วย คือคะแนนที่ได้จากข้อสอบชุดปกติก็จะมีลักษณะเป็นปกติ และคะแนนจากข้อสอบแบบสม่ำเสมอ (Uniform) ก็จะมีลักษณะ เป็นแบบสม่ำเสมอด้วย

3. ผลการเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณกับค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของประชากร

3.1 ผลการเปรียบเทียบค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณกับค่าคะแนนเฉลี่ยของประชากร กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากข้อสอบแบบปกติ แบบเบ้ซ้าย และแบบเบ้ขวา ไม่พบจำนวนครั้งของความแตกต่างกันเกิดขึ้น จากการกระทำซ้ำ 1,000 ครั้ง เพราะการสุ่มแบบเมตริกซ์เชิงพหุคูณมีค่าสถิติใกล้เคียงกับค่าพารามิเตอร์ของประชากร (Sirotnik, 1974, p.461) คือค่าคะแนนเฉลี่ยที่ได้จากการสุ่มมีค่าใกล้เคียงกับค่าคะแนนเฉลี่ยของประชากร ดังนั้นความแปรปรวนคะแนนก็ไม่มีความแตกต่างกันมากด้วย

3.2 ผลการเปรียบเทียบความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณกับความแปรปรวนคะแนนของประชากร กรณีข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ แบบเบ้ซ้าย และแบบเบ้ขวา ไม่พบจำนวนครั้งของความแตกต่างกันเกิดขึ้นจากการกระทำซ้ำ 1,000 ครั้ง นั่นคือ ค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่าง มีค่าใกล้เคียงกับค่าพารามิเตอร์ของประชากร ตามที่ ไชโรตนิค (Sirotnik, 1974, p.461) ได้สรุปไว้ว่า ค่าคะแนนเฉลี่ยจากการสุ่มตัวอย่างแบบเมตริกซ์เชิงพหุคูณมีค่าใกล้เคียงกับค่าคะแนนเฉลี่ยของประชากรมากกว่าวิธีการสุ่มผู้สอบเพียงอย่างเดียวที่นักวิจัยส่วนมากนิยมใช้กัน หรือที่เรียกการสุ่มวิธีนี้ว่า การสุ่มแบบง่าย และผลที่ได้นี้ เป็นไปตามสมมุติฐานที่ตั้งไว้ว่าค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบเมตริกซ์เชิงพหุคูณกับค่าคะแนนเฉลี่ย และความแปรปรวนคะแนนของประชากร ไม่มีความแตกต่างกัน ทั้งนี้อาจเป็นเพราะข้อดีของวิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณ ที่ว่าวิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณให้ค่าประมาณค่าเฉลี่ยคะแนนได้ใกล้เคียงกับค่าคะแนนเฉลี่ยของประชากร มากกว่าวิธีสุ่มแบบง่าย และการจำลองสถานการณ์ด้วยวิธีมอนติคาร์โล ของการทดลองในครั้งนี้ สามารถควบคุมตัวแปรแทรกซ้อนต่าง ๆ ได้ เช่น ตัวแปรในด้านบริบทของข้อสอบ และตัวแปรด้านความสามารถของผู้สอบที่เปลี่ยนแปลงไปตามกาลเวลา และสามารถจัดการกระทำซ้ำได้โดยไม่สิ้นเปลืองเวลามาก ตลอดจนตัวแปรทางด้านจิตวิทยาของผู้สอบ เช่น ความเบื่อหน่าย ความล้าในการสอบ การเดาตอบข้อสอบ เป็นต้น

4. ผลการเปรียบเทียบค่าคะแนนเฉลี่ยและความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ

4.1 ผลการเปรียบเทียบค่าคะแนนเฉลี่ยของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ พบว่า ที่ระดับนัยสำคัญ .01 ของข้อสอบที่มีลักษณะการแจกแจงค่าความยากของข้อสอบแบบปกติ ในแผนสอง จำนวนตัวอย่าง 300 คน พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติเกิดขึ้น 1 ครั้ง จากการกระทำซ้ำ 1,000 ครั้ง ส่วนที่ระดับนัยสำคัญ .05 พบความแตกต่างกันของค่าคะแนนเฉลี่ยเกิดขึ้นเพียงหนึ่งครั้ง จากการกระทำซ้ำ 1,000 ครั้ง ในแผนสอง จำนวนตัวอย่าง 300 คน ของแต่ละลักษณะการแจกแจงค่าความยากของข้อสอบ นั่นคือ ความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภท I เท่ากับ .001 ซึ่งตามทฤษฎีการสุ่มแล้ว เมื่อจำนวนตัวอย่างลดลง ควรเกิดความคลาดเคลื่อนมากขึ้น แต่ในการทดลองกลับพบว่า แผนสาม จำนวนตัวอย่าง

200 คน ไม่พบความแตกต่างของค่าคะแนนเฉลี่ยเกิดขึ้น จากการกระทำซ้ำ 1,000 ครั้ง ส่วนแผนสอง จำนวนตัวอย่าง 300 คน กลับพบความแตกต่างของค่าคะแนนเฉลี่ยเกิดขึ้นหนึ่งครั้ง จากการกระทำซ้ำ 1,000 ครั้ง ทั้งนี้อาจเกิดจากการแกว่ง (Fluctuation) ของการทดลอง ซึ่ง การ์ค และคณะ (Garg et al., 1986, p.123) พบว่าการแกว่งของการทดลองเกิดขึ้นในแต่ละแผนการสุ่มได้

4.2 ผลการเปรียบเทียบความแปรปรวนคะแนนของตัวอย่างที่สุ่มแบบง่ายกับสุ่มแบบเมตริกซ์เชิงพหุคูณ พบว่า ที่ระดับนัยสำคัญ .01 พบความแตกต่างเกิดขึ้น จำนวน 1 และ 2 ครั้ง ในข้อสอบแบบเบื้อซ้าย และจำนวนอย่างละ 1 ครั้ง ในแผนหนึ่งและแผนสอง ของข้อสอบแบบเบื้อขวา ส่วนข้อสอบแบบปกติ ไม่พบจำนวนครั้งของความแตกต่างเกิดขึ้น จากการกระทำซ้ำ 1,000 ครั้ง ของแต่ละแผน ในข้อสอบชุดเบื้อขวา (หรือข้อสอบชุดยาก) แผนหนึ่ง แผนสอง และแผนสาม จำนวนตัวอย่าง 600 300 และ 200 คน พบความแตกต่างเกิดขึ้นที่ระดับ .05 จำนวน 2 28 และ 11 ครั้ง ตามลำดับ จากการจัดการกระทำซ้ำ 1,000 ครั้ง ในแต่ละแผนในข้อสอบแบบเบื้อซ้าย หรือข้อสอบชุดง่าย แผนหนึ่ง แผนสอง และแผนสาม จำนวนตัวอย่าง 600 300 และ 200 คน พบความแตกต่างเกิดขึ้นที่ระดับ .05 จำนวน 9 18 และ 3 ครั้ง ตามลำดับ จากการจัดการกระทำซ้ำ 1,000 ครั้ง ส่วนข้อสอบแบบปกติ พบว่า ในแผนสอง จำนวนตัวอย่าง 300 คน มีจำนวนครั้งของความแตกต่างเกิดขึ้นหนึ่งครั้ง จากการกระทำซ้ำ 1,000 ครั้ง ทั้งนี้เพราะผู้สอบที่มีความสามารถแบบปกติ คือ มีความสามารถปานกลาง ให้ตอบข้อสอบชุดยาก พบว่า คนส่วนใหญ่ได้คะแนนน้อย การกระจายคะแนน (Discrepancy) เกิดขึ้นน้อย เมื่อให้ผู้สอบที่มีความสามารถปานกลาง ไปตอบข้อสอบชุดปานกลาง พบว่า คนส่วนใหญ่ได้คะแนนปานกลาง การกระจายคะแนน (Discrepancy) มีสูง และถ้าให้ผู้สอบที่มีความสามารถปานกลาง ไปตอบข้อสอบชุดง่าย พบว่า คนส่วนใหญ่ได้คะแนนสูง การกระจายคะแนน (Discrepancy) เกิดขึ้นน้อยครั้งเช่นเดียวกันกับข้อสอบชุดยาก ดังนั้นข้อสอบที่ยากมากๆ หรือง่ายมากๆ นั้น ไม่เหมาะสมกับทฤษฎีการตอบข้อสอบของราซด์ เพราะถือว่าไม่มีความหมายในการวัด แสดงว่าถ้าต้องการใช้ข้อสอบกับผู้สอบกลุ่มใด ควรจัดให้ค่าความยากและความสามารถของผู้สอบมีความเหมาะสมกัน ไรท์ (Wright, 1979, pp.128-140 และ สมพร บุญอ้อม, 2529, หน้า 52)

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 ในการจัดการสอบเพื่อวัดความสามารถของผู้สอบ ควรจัดผู้สอบและข้อสอบให้มีความเหมาะสมกัน เพราะข้อสอบที่ยาก ง่าย หรือปานกลาง เมื่อให้เด็กกลุ่มปกติทำ พบว่ามีความคลาดเคลื่อน (Error) น้อยที่สุด

1.2 วิธีการสุ่มแบบเมตริกซ์เชิงพหุคูณ เป็นการเพิ่มทางเลือก (Alternative) ของการใช้วิธีการสุ่มให้เหมาะสมกับงานวิจัย งานสำรวจ หรืองานประเมินโครงการ ที่ต้องการอ้างอิงไปสู่ประชากรที่กำหนด (Defined population)

2. ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

ควรทำการศึกษาผู้สอบที่มีลักษณะการแจกแจงค่าความสามารถของผู้สอบแบบอื่น ๆ คือ แบบเบ้ซ้าย หรือ แบบเบ้ขวา เมื่อข้อสอบมีลักษณะการแจกแจงค่าความยากของข้อสอบทั้งสามแบบ คือ แบบปกติ แบบเบ้ซ้าย และแบบเบ้ขวา