การพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนแบบใหม่

อรัญ ซุยกระเดื่อง
สำราญ มีแจ้ง รัตนะ บัวสนธ์ และ อรุณี อ่อนสวัสดิ์
มหาวิทยาลัยนเรศวร

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนของประชากร และเปรียบเทียบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้นคือ SA1 และ SA2 กับสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน โดยพิจารณาจากความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I และอำนาจการทดสอบของสถิติทดสอบ เมื่อลักษณะการแจกแจงประชากรเป็นแบบปกติและแบบไม่ปกติ ขนาดกลุ่มตัวอย่างเท่ากันและไม่เท่ากัน และเมื่ออัตราส่วนของความแปรปรวนแตกต่างกัน งานวิจัยนี้จำลองการทดลองด้วยเครื่องคอมพิวเตอร์ โดยใช้เทคนิคมอนติ คาร์โล ซิมูเลชัน กระทำซ้ำ 10,000 ครั้งสำหรับแต่ละสถานการณ์ที่กำหนด

ผลการวิจัยสรุปได้ดังนี้

ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I ปรากฏว่า เมื่อระดับนัยสำคัญ .05 สถิติทดสอบ SA2 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ดีที่สุด สถิติทดสอบ SA1 ไม่มีประสิทธิภาพเท่า SA1 แต่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้มากกว่าสถิติทดสอบเลอวีนและสถิติทดสอบคอครัน ในขณะที่เมื่อระดับนัยสำคัญ .01 สถิติทดสอบ SA1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้มากกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครัน

สำหรับประชากรที่มีการแจกแจงปกติ สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบทุกตัว สถิติทดสอบ SA1 และสถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงใกล้เคียงกันและสูงกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครัน สำหรับประชากรที่มีการแจกแจงแบบเบ้ สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงที่สุด รองลงมาคือสถิติทดสอบ SA1 สถิติทดสอบ SA2 และสถิติทดสอบคอครัน สำหรับประชากรที่มีการแจกแจงโด่งแบน สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบเลอวีน และสถิติทดสอบ SA2 และสำหรับประชากรที่มีการแจกแจงโด่งสูง สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA1 และสถิติทดสอบ SA2

ส่วนหนึ่งของวิทยานิพนธ์ปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิชาวิจัยและประเมินผลการศึกษา มหาวิทยาลัยนเรศวร ปีการศึกษา2550 อาจารย์ที่ปรึกษาหลัก รศ.ดร.สำราญ มีแจ้งและ อาจารย์ที่ปรึกษาร่วม รศ. ดร.รัตนะ บัวสนธ์ รศ.ดร.อรุณี อ่อนสวัสดิ์

The Development of New Statistic Testing for Homogeneity of Variance

Arun Suikraduang

Samran Mejang, Rattana Buosonte, and Arunee On-Sawad

Naresuan University, Thailand

ABSTRACT

This study investigated the efficiency of two new proposed homogeneity of variance test statistics, "SA1" and "SA2", using Monte Carlo methods to compare results with those from Bartlett's, Levene's, and Cochran's tests. Experiments were conducted over normal and non-normal samples, with different variance ratios, using equal and unequal Ns. Each experimental condition was replicated 10,000 times.

With regard to Type I error rates, it was found that SA2 provided the best control at the .05 level. The SA1 test was not as efficient at .05, but nonetheless showed better control than Levene and Cochran. On the other hand, at the .01 level, SA1 was found to be superior to SA2 and Cochran .

For normal distributions, Bartlett's test had the most power. SA1 and Levene were second best, with SA2 and Cochran coming last. With skewed distributions, Bartlett again had the most power, followed by SA1, then SA2, with Cochran having the least power. The SA1 test was found to have the highest power for platykurtic distributions, then Levene, then SA2. With leptokurtic distributions, Levene's test showed the most power, with SA1 second, followed by SA2.

Based on a doctoral dissertation in Educational Research and Evaluation, Naresuan University, under the supervision of Assoc. Prof. Samran Mejang, Ph.D., Assoc. Prof. Rattana Buosonte Ed.D., and Assoc. Prof. Arunee On-Sawad, Ph.D.

ความนำ

การเปรียบเทียบความแตกต่างของค่าเฉลี่ยประชากรที่มากกว่าสองกลุ่มขึ้นไป จะทำภายใต้ข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน 2 ชุดคือชุดของข้อตกลงเบื้องต้นของสถิติทดสอบเอฟ (F - assumption) และชุดที่เป็นข้อตกลงเบื้องต้นของโมเดล (Model assumption) ในทางปฏิบัติจริงเรามักไม่แน่ใจว่าข้อมูลที่ได้มาจะเป็นไปตามข้อตกลงเบื้องต้นทั้งหมด และถ้าไม่เป็นไปตามข้อตกลงเบื้องต้นแล้วการใช้สถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวนจะเป็นอย่างไร สำหรับกรณีที่ไม่เป็นไปตามข้อตกลงเบื้องต้นด้านความเป็นเอกพันธ์ของความแปรปรวน เมื่อขนาดของกลุ่มตัวอย่างแต่ละกลุ่มเท่ากันสามารถที่จะใช้สถิติทดสอบเอฟได้ ทั้งนี้เพราะว่าสถิติทดสอบเอฟมีความแกร่ง (Robustness) ต่อการฝ่าฝืนข้อตกลงข้อนี้ แต่ถ้าขนาดของกลุ่มตัวอย่างในแต่ละกลุ่มไม่เท่ากัน จะเกิดผลกระทบทั้งระดับความมีนัยสำคัญและความไวของการทดสอบ (Sensitivity) ดังนั้นในงานวิจัยเชิงทดลองจึงมีการวางแผนการวิจัยให้ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มเท่ากันเพราะสามารถที่จะใช้สถิติทดสอบเอฟได้โดยไม่จำเป็นต้องทดสอบการเท่ากันของความแปรปรวน แต่ในงานการวิจัยเชิงทดลองบางครั้งอาจจะมีการสูญหายของข้อมูลทำให้ขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากันและในปัจจุบันได้มีการนำการวิเคราะห์ความแปรปรวนมาใช้ในงานวิจัยเชิงบรรยายเป็นจำนวนมาก ซึ่งการวิจัยเชิงบรรยายขนาดของกลุ่มตัวอย่างแต่ละกลุ่มมักไม่เท่ากัน ดังนั้นก่อนที่จะใช้วิธีการวิเคราะห์ความแปรปรวนโดยเฉพาะเมื่อขนาดของกลุ่มตัวอย่างแต่ละกลุ่มไม่เท่ากัน ควรจะตรวจสอบการเท่ากันของความแปรปรวนก่อนเพื่อที่จะไม่เกิดข้อสรุปที่ผิด

สถิติทดสอบที่ใช้ในการทดสอบการเท่ากันของความแปรปรวนของประชากรที่มากกว่าสองกลุ่ม ที่รู้จักกันมากและนิยมใช้ในการทดสอบสามารถแบ่งออกเป็น 3 กลุ่มดังนี้ กลุ่มที่ 1 พัฒนามาจากตัวสถิติทดสอบของนีแมน-เพียร์สัน ซึ่งสร้างตัวสถิติทดสอบโดยใช้อัตราส่วนภาวะน่าจะเป็นสูงสุด สถิติทดสอบบาร์ตเลตเป็นสถิติทดสอบที่นิยมใช้เนื่องจากมีอำนาจการทดสอบสูงเมื่อตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ใช้ได้ทั้งกรณีที่กลุ่มตัวอย่างมีขนาดเท่ากันและไม่เท่ากัน แต่เมื่อประชากรมีการแจกแจงไม่ปกติ สถิติทดสอบบาร์ตเลตไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I และยังมีความยุ่งยากในการคำนวณ กลุ่มที่ 2 พัฒนาโดยใช้อัตราส่วนของความแปรปรวน สถิติทดสอบคอครันเป็นตัวสถิติทดสอบที่นิยมใช้เนื่องจากใช้ง่ายและรวดเร็ว และยังมีอำนาจการทดสอบสูงเมื่อความแปรปรวนหนึ่งค่ามีขนาดใหญ่กว่าค่าอื่นๆ แต่มีข้อจำกัดคือใช้ในกรณีที่ขนาดกลุ่มตัวอย่างเท่ากัน ถ้าขนาดกลุ่มตัวอย่างไม่เท่ากันสถิติทดสอบคอครันไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I และกลุ่มที่ 3 พัฒนาโดยการแปลงข้อมูลแล้วใช้การวิเคราะห์ความแปรปรวน สถิติทดสอบเลอวีนเป็นสถิติทดสอบที่นิยมใช้เนื่องจากมีความแกร่งเมื่อประชากรมีการแจกแจงไม่ปกติ แต่ต้องมีขนาดกลุ่มตัวอย่างไม่น้อยกว่า 10 ถ้ากลุ่มตัวอย่างมีขนาดเล็กจะเกิดความสัมพันธ์กันสูงระหว่างส่วนเบี่ยงเบน

ในกลุ่มเดียวกันซึ่งทำให้สถิติทดสอบเลอวีนขาดความเที่ยงตรงและการใช้การวิเคราะห์ความแปรปรวนมาวิเคราะห์ จึงต้องอยู่บนข้อตกลงเดียวกันซึ่งทำให้ขาดความสมเหตุสมผล

จากความเป็นมาดังกล่าวแม้ว่าจะมีสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนหลายตัว ซึ่งสถิติทดสอบแต่ละตัวต่างก็มีข้อบกพร่องแตกต่างกันไป และยังไม่มีสถิติทดสอบตัวใดที่เหมาะสมในการทดสอบในทุกสถานการณ์ ผู้วิจัยจึงสนใจที่จะพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนของประชากรที่มากกว่า 2 กลุ่ม โดยใช้แนวคิดการทดสอบความเท่ากันของความแปรปรวนของค่าสังเกตแบบคู่ (Paired Observations) ซึ่งเปรียบเทียบผลต่างของความแปรปรวนระหว่างความแปรปรวนสูงสุดกับความแปรปรวนต่ำสุดเพื่อแก้ไขข้อจำกัดของสถิติทดสอบที่มีอยู่

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน โดยใช้การเปรียบเทียบผลต่างของความแปรปรวนสูงสุดและความแปรปรวนต่ำสุด

2. เพื่อตรวจสอบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้น โดยพิจารณาดังนี้

2.1 เปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ I กับระดับนัยสำคัญที่กำหนดและอำนาจการทดสอบ

2.2 เปรียบเทียบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้นกับสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

สมมติฐานการวิจัย

1. สถิติทดสอบที่พัฒนาขึ้น น่าจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ (มีอัตราความคลาดเคลื่อนประเภทที่ 1 เท่ากับระดับนัยสำคัญที่กำหนด)

2. สถิติทดสอบที่พัฒนาขึ้น น่าจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้มากกว่าสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

3. สถิติทดสอบที่พัฒนาขึ้น น่าจะมีอำนาจการทดสอบสูงกว่า สถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

ขอบเขตของการวิจัย

1. เงื่อนไขที่นำมาศึกษามีดังต่อไปนี้

1.1 ศึกษาเฉพาะกรณีการวิเคราะห์ความแปรปรวนแบบทางเดียวที่มีระดับของปัจจัยเป็น 3, 4 และ 5 ระดับ

1.2 กำหนดขนาดของกลุ่มตัวอย่างที่ศึกษาให้มีค่าตั้งแต่ 10 ถึง 50

1.3 กำหนดค่าเฉลี่ยของประชากรเท่ากับ 50 และความแปรปรวนของประชากรสำหรับการพิจารณาค่าความคลาดเคลื่อนประเภทที่ I ทุกกลุ่มมีค่าเท่ากับ 10

1.4 ความแปรปรวนของประชากรสำหรับการพิจารณาอำนาจการทดสอบ จะกำหนดค่าความแปรปรวนที่แตกต่างกัน โดยอาศัยวิธีของ Game, Winkler and Probert (1972)

1.5 ใช้เกณฑ์ของคอครัน (Cochran) ในการตัดสินความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I กับระดับนัยสำคัญที่กำหนดดังนี้ ที่ระดับนัยสำคัญ .05 อัตราความคลาดเคลื่อนประเภทที่ I มีค่าตั้งแต่ 0.0400 ถึง 0.0600 และที่ระดับนัยสำคัญ .01 อัตราความคลาดเคลื่อนประเภทที่ I มีค่าตั้งแต่ 0.0070 ถึง 0.0150

1.6 การแจกแจงของประชากร ศึกษาลักษณะการแจกแจงของประชากรที่มีการแจกแจงปกติและการแจกแจงไม่ปกติประกอบด้วยการแจกแจงเบ้ซ้าย เบ้ขวา โด่งแบนและโด่งสูง

2. อัตราการควบคุมความคลาดเคลื่อนประเภทที่ I และอำนาจการทดสอบของสถิติทดสอบเปรียบเทียบเฉพาะที่ระดับนัยสำคัญ (α) มีค่าเป็น .01 และ .05 เท่านั้น

3. การวิจัยนี้จำลองการทดลองขึ้นในคอมพิวเตอร์ โดยอาศัยเทคนิคมอนติ คาร์โล ซิมูเลชัน (Monte Carlo Simulation Technique)

4. การจำลองการทดลองนี้ทำการทดลองซ้ำ 10,000 ครั้งในแต่ละสถานการณ์ของการทดลอง

วิธีดำเนินการวิจัย

การวิจัยนี้แบ่งการดำเนินงานเป็น 2 ขั้นตอนคือ ขั้นตอนที่ 1 เป็นการพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน โดยศึกษาจากทฤษฎีและงานวิจัยที่เกี่ยวข้อง และขั้นตอนที่ 2 เป็นการตรวจสอบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้นและเปรียบเทียบประสิทธิภาพกับสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน ใช้การจำลองสถานการณ์ด้วยการทำงานของเครื่องคอมพิวเตอร์โดยใช้เทคนิคมอนติ คาร์โล รายละเอียดมีดังนี้

ขั้นตอนที่ 1 การพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน

1. ศึกษาข้อดีและข้อจำกัดของสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนแต่ละตัว จากเอกสาร แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2. วิเคราะห์แนวคิดของสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนแต่ละตัวโดยพิจารณาจากข้อดีและข้อจำกัดของสถิติทดสอบเหล่านั้นเพื่อหาแนวคิดที่จะนำเสนอสถิติทดสอบตัวใหม่

3. นำเสนอสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน

4. ศึกษานำร่อง (Pilot study) เพื่อตรวจสอบอัตราความคลาดเคลื่อนประเภทที่ I และการแจกแจง ของสถิติทดสอบที่พัฒนาขึ้น

ขั้นตอนที่ 2 การตรวจสอบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้นและเปรียบเทียบประสิทธิภาพกับสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

การทดลองครั้งนี้ ใช้การจำลองสถานการณ์ด้วยการทำงานของเครื่องคอมพิวเตอร์โดยใช้โปรแกรมภาษา C++ โดยการทดลองมีขั้นตอนดังนี้

ขั้นที่ 1 สร้างประชากรจากตัวเลขสุ่มให้มีการแจกแจงปกติ เบ้ซ้าย เบ้ขวา โด่งแบน และโด่งสูงที่มีค่าเฉลี่ยเท่ากับ 50 และค่าความแปรปรวนตามแผนการทดลอง

ขั้นที่ 2 สุ่มตัวอย่างจากประชากรขนาดต่างๆ ตามแผนการทดลอง

ขั้นที่ 3 คำนวณค่าสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน สถิติทดสอบคอครันและสถิติทดสอบที่พัฒนาขึ้น

ขั้นที่ 4 กำหนดค่าวิกฤติของสถิติทดสอบแต่ละตัว แล้วทดสอบนัยสำคัญและนับจำนวนครั้งของการเกิดนัยสำคัญของสถิติทดสอบแต่ละตัว

ขั้นที่ 5 ทำการทดลองตามขั้นตอนที่ 2 – 4 ซ้ำจำนวน 10,000 ครั้ง

ขั้นที่ 6 คำนวณอัตราความคลาดเคลื่อนประเภทที่ I โดยหาค่าสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก (H_0) ที่เป็นจริงในจำนวน 10,000 ครั้งของการทดลองแล้วเปรียบเทียบกับเกณฑ์ของคอครัน ถ้าสถิติทดสอบใดมีอัตราความคลาดเคลื่อนประเภทที่ I อยู่ในเกณฑ์ของคอครัน แสดงว่าสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ก็จะทำการพิจารณาอำนาจการทดสอบต่อไปและคำนวณอำนาจ การทดสอบของสถิติทดสอบ โดยหาค่าสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก (H_0) ที่เป็นเท็จในจำนวน 10,000 ครั้งของการทดลอง

ขั้นที่ 7 ทดสอบความแตกต่างของความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบ ด้วยสถิติทดสอบไค-สแควร์ (χ^2- test) ถ้าพบว่ามีนัยสำคัญทางสถิติจะดำเนินการทดสอบภายหลังโดยการเปรียบเทียบเป็นรายคู่ด้วยสถิติทดสอบซี (Z – test)

ผลการวิจัย

ตอนที่ 1 ผลการพัฒนาสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน

สถิติทดสอบที่พัฒนาขึ้น มีดังนี้

1. $$SA1 = \frac{s^2_{max} - s^2_{min}}{2cs_{max}.s_{min}\sqrt{\frac{1}{n-1}}} \sim t : n-1$$

เมื่อ $c = \left(1 + \frac{1}{4(n-1)}\right)^2 + 0.06 + 0.09(k-2) + \frac{(\hat{\alpha}_4)_{max} - 3}{k}$ โดยที่ k คือจำนวนกลุ่ม

ตัวอย่าง, n คือขนาดกลุ่มตัวอย่าง (กรณีขนาดตัวอย่างไม่เท่ากันใช้ค่าเฉลี่ยฮาร์โมนิคของขนาดกลุ่มตัวอย่างแทนขนาดกลุ่มตัวอย่าง) และ $(\hat{\alpha}_4)_{\max}$ คือค่าความโด่งของกลุ่มตัวอย่างที่มีค่ามากที่สุด

$$2.\ SA\ 2 = \frac{s^2_{\max} - s^2_{\min}}{cs^2_p \sqrt{\frac{2}{n_{s^2_{\max}} - 1} + \frac{2}{n_{s^2_{\min}} - 1}}} \sim t : n_1 + n_2 + \ldots n_k - k$$

เมื่อ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2}{n_1 + n_2 + \ldots + n_k - k}$ โดยที่ k คือจำนวนกลุ่มตัวอย่าง,

$c = \sqrt{1 - \frac{2}{n+1}} + 0.06 + 0.09(k-2) + \frac{(\hat{\alpha}_4)_{\max} - 3}{k}$ โดยที่ n คือขนาดกลุ่มตัวอย่าง (กรณีขนาดตัวอย่างไม่เท่ากันใช้ค่าเฉลี่ยฮาร์โมนิคของขนาดกลุ่มตัวอย่างแทนขนาดกลุ่มตัวอย่าง) และ $(\hat{\alpha}_4)_{\max}$ คือค่าความโด่งของกลุ่มตัวอย่างที่มีค่ามากที่สุด

ตอนที่ 2 ผลการตรวจสอบประสิทธิภาพของสถิติทดสอบที่พัฒนาขึ้นและเปรียบเทียบประสิทธิภาพกับสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

1. ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I ของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน สถิติทดสอบคอครันและสถิติทดสอบที่พัฒนาขึ้น พบว่า

1.1 เมื่อระดับนัยสำคัญ (α) .05 สถิติทดสอบบาร์ตเลตสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ทุกกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูง โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 81 กรณีคิดเป็นสัดส่วน 0.6000 สถิติทดสอบเลอวีนสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงโด่งแบนและแจกแจงโด่งสูง ยกเว้นกรณีขนาดกลุ่มตัวอย่างขนาดเล็ก แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 70 กรณีคิดเป็นสัดส่วน 0.5185 สถิติทดสอบคอครันสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้บางกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูง โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 47 กรณีคิดเป็นสัดส่วน 0.3481 สถิติทดสอบ SA1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 96 กรณีคิดเป็นสัดส่วน 0.7111 และสถิติทดสอบ SA2 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง

ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 106 กรณีคิดเป็นสัดส่วน 0.7852

เมื่อเปรียบเทียบสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมด สถิติทดสอบ SA2 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สถิติทดสอบ SA1 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบเลอวีนและสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สถิติทดสอบบาร์ตเลตและสถิติทดสอบเลอวีนมีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

1.2 เมื่อระดับนัยสำคัญ (α) .01 สถิติทดสอบบาร์ตเลตสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ทุกกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูง โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 81 กรณีคิดเป็นสัดส่วน 0.6000 สถิติทดสอบเลอวีนสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงโด่งแบนและแจกแจงโด่งสูง ยกเว้นกรณีขนาดกลุ่มตัวอย่างขนาดเล็ก แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 78 กรณีคิดเป็นสัดส่วน 0.5778 สถิติทดสอบคอครันสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้บางกรณีเมื่อประชากรมีการแจกแจงปกติ แจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูง โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 53 กรณีคิดเป็นสัดส่วน 0.3926 สถิติทดสอบ SA1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 97 กรณีคิดเป็นสัดส่วน 0.7185 และสถิติทดสอบ SA2 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจงเมื่อจำนวนกลุ่มตัวอย่างมี 3 – 4 กลุ่ม แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงเบ้ซ้าย แจกแจงเบ้ขวาและแจกแจงโด่งแบน กรณีจำนวนกลุ่มตัวอย่างมี 5 กลุ่ม โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 62 กรณีคิดเป็นสัดส่วน 0.4953

เมื่อเปรียบเทียบสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมด สถิติทดสอบ SA1 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครันอย่าง

มีนัยสำคัญทางสถิติที่ระดับ .05 สถิติทดสอบบาร์ตเลตและสถิติทดสอบเลอวีนมีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. อำนาจการทดสอบของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน สถิติทดสอบคอครันและสถิติทดสอบที่พัฒนาขึ้น สูงขึ้นเมื่อดัชนีความแตกต่างของความแปรปรวนของประชากร (ϕ) เพิ่มขึ้น ยกเว้นสถิติทดสอบคอครันมีอำนาจการทดสอบลดลงเล็กน้อยเมื่อดัชนีความแตกต่างของความแปรปรวนของประชากรเปลี่ยนจากระดับน้อยเป็นระดับปานกลางและเปลี่ยนจากระดับปานกลางเป็นระดับมาก เมื่อพิจารณาเฉพาะกรณีที่สถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ I พบว่า

2.1 สำหรับประชากรที่มีการแจกแจงปกติ สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบทุกตัว สถิติทดสอบ SA1 และสถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงใกล้เคียงกันและสูงกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครัน

2.2 สำหรับประชากรที่มีการแจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA1 สถิติทดสอบ SA2 และสถิติทดสอบคอครัน สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครัน

2.3 สำหรับประชากรที่มีการแจกแจงโด่งแบน สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบเลอวีนและสถิติทดสอบ SA2 สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2

2.4 สำหรับประชากรมีการแจกแจงโด่งสูง สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA1 และสถิติทดสอบ SA2 สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2

การอภิปรายผลการวิจัย

การเลือกใช้สถิติทดสอบในการวิเคราะห์ข้อมูล นอกจากจะต้องพิจารณาถึงลักษณะของข้อมูลตามข้อตกลงเบื้องต้นของการใช้สถิติวิเคราะห์แต่ละแบบแล้ว ยังจำเป็นที่จะต้องพิจารณาถึงความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I (ความผิดพลาดสำคัญมากที่จะต้องหลีกเลี่ยง) โดยยึดหลักว่าความน่าจะเป็นที่ยอมให้เกิดความคลาดเคลื่อนประเภทที่ I ไม่เกินระดับนัยสำคัญ (α) ที่กำหนดไว้ แล้วจึงพิจารณาอำนาจการทดสอบ (Power of the Test) ของสถิติทดสอบนั้น จากการวิจัยครั้งนี้มีข้อค้นพบดังที่ผู้วิจัยจะได้นำมาอภิปรายดังต่อไปนี้

1. จากสมมติฐานการวิจัยที่กล่าวว่าสถิติทดสอบที่พัฒนาขึ้น น่าจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ (มีอัตราความคลาดเคลื่อนประเภทที่ I เท่ากับระดับนัยสำคัญที่กำหนด)

ผลการวิจัย พบว่าเมื่อระดับนัยสำคัญ (α) .05 สถิติทดสอบ SA1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 96 กรณีคิดเป็นสัดส่วน 0.7111 และสถิติทดสอบ SA2 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 106 กรณีคิดเป็นสัดส่วน 0.7852 และเมื่อระดับนัยสำคัญ (α) .01 สถิติทดสอบ SA1 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง ยกเว้นกรณีขนาดตัวอย่างขนาดเล็ก โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 97 กรณีคิดเป็นสัดส่วน 0.7185 และสถิติทดสอบ SA2 สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้เกือบทุกกรณีในแต่ละการแจกแจง เมื่อจำนวนกลุ่มตัวอย่างมี 3 – 4 กลุ่ม แต่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ เมื่อประชากรมีการแจกแจงเบ้ซ้าย แจกแจงเบ้ขวาและแจกแจงโด่งแบน กรณีจำนวนกลุ่มตัวอย่างมี 5 กลุ่ม โดยมีจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ 62 กรณีคิดเป็นสัดส่วน 0.4953 ทั้งนี้อาจเป็นเพราะสถิติทดสอบที่พัฒนาขึ้นทั้ง 2 ตัวนั้น ใช้ตัวประมาณค่าที่ไม่ลำเอียงและมีความแปรปรวนต่ำสุดคือใช้ s^2 ประมาณค่า σ^2 ซึ่งถือว่าเป็นตัวประมาณค่าที่ดีที่สุด โดยทำการเปรียบเทียบผลต่างของความแปรปรวนเพียง 1 คู่ คือความแปรปรวนสูงสุดและความแปรปรวนต่ำสุด ($s^2_{max} - s^2_{min}$) หากผลการเปรียบเทียบคู่นี้ไม่แตกต่างกันแล้ว คู่อื่นๆ ก็ไม่แตกต่างกันด้วยและถ้าผลการเปรียบเทียบคู่นี้แตกต่างกันก็ไม่จำเป็นต้องเปรียบเทียบคู่อื่นๆ อีก จึงสามารถตอบคำถามว่าความแปรปรวนของกลุ่มประชากรเท่ากันได้เช่นเดียวกับการเปรียบเทียบความแปรปรวนรวมทุกกลุ่ม โดยไม่ทำให้อัตราความคลาดเคลื่อนประเภทที่ I เพิ่มขึ้น นอกจากนี้ยังเป็นการเปรียบเทียบทั้งผลต่างและอัตราส่วนของความแปรปรวน โดยสถิติทดสอบ SA1 เปรียบเทียบผลต่างความแปรปรวนกับผลคูณของส่วนเบี่ยงเบนมาตรฐานสูงสุดและต่ำสุด และสถิติทดสอบ SA2 เปรียบเทียบผลต่างความแปรปรวนกับความแปรปรวนรวมทุกกลุ่ม จึงทำให้สถิติทดสอบที่พัฒนาขึ้นมีการแจกแจงสอดคล้องกับการแจกแจงประมาณแบบที (Approximate t – Distribution) ยิ่งถ้าองศาความเป็นอิสระ (Degrees of Freedom) มีค่ามากขึ้นก็จะทำให้สถิติทดสอบนี้มีการแจกแจงเข้าใกล้การแจกแจงปกติ และมีข้อมูลหลายอย่างที่ได้จากการทดลองหรือจากการสังเกตโดยเฉพาะข้อมูลทางสังคมศาสตร์ส่วนมากมักจะสอดคล้องกับการแจกแจงปกติหรือใกล้เคียงปกติ (Approximate Normal Distribution) ซึ่งเป็นการแจกแจงที่สำคัญและเป็นพื้นฐานของการแจกแจงไค-สแควร์ การแจกแจงเอฟและการแจกแจงที อันจะส่งผลให้สถิติทดสอบที่พัฒนาขึ้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ (มีอัตราความคลาดเคลื่อนประเภทที่ I เท่ากับระดับนัยสำคัญที่กำหนด)

2. จากสมมติฐานที่กล่าวว่าสถิติทดสอบที่พัฒนาขึ้น น่าจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้มากกว่าสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

ผลการวิจัยพบว่าเมื่อระดับนัยสำคัญ (α) .05 สถิติทดสอบ SA2 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สถิติทดสอบ SA1 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบเลอวีนและสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และเมื่อระดับนัยสำคัญ (α) .01 สถิติทดสอบ SA1 มีสัดส่วนของจำนวนการทดลองที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I กับจำนวนการทดลองทั้งหมดมากกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทั้งนี้อาจเป็นเพราะสถิติทดสอบที่พัฒนาขึ้นทั้ง 2 ตัวนอกจากใช้ตัวประมาณค่าที่ไม่ลำเอียงและมีความแปรปรวนต่ำสุด ทำการเปรียบเทียบผลต่างของความแปรปรวนเพียง 1 คู่ คือความแปรปรวนสูงสุดและความแปรปรวนต่ำสุด และมีการแจกแจงสอดคล้องกับการแจกแจงประมาณแบบทีแล้วยังถ่วงน้ำหนักด้วยค่าคงที่ (c) ที่ประกอบไปด้วยขนาดตัวอย่าง จำนวนกลุ่มตัวอย่างและค่าความโด่งสูงสุดของกลุ่มตัวอย่าง ทำให้สถิติทดสอบนี้มีค่าค่อนข้างเสถียร (Stable) คือแปรเปลี่ยนไปตามขนาดตัวอย่างและจำนวนกลุ่มตัวอย่างแต่ไม่แปรเปลี่ยนไปตามการแจกแจงของประชากร ทำให้สามารถความคุมความคลาดเคลื่อนประเภทที่ I ได้ทุกการแจกแจงของประชากรที่ศึกษา ในขณะที่สถิติทดสอบบาร์ตเลตนั้นไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูง ซึ่งสอดคล้องกับบราวน์และฟอร์สิตี (Brown & Forsythe, 1974) ที่กล่าวว่าวิธีการทดสอบของบาร์ตเลตจะมีผลทำให้ความคลาดเคลื่อนประเภทที่ I สูงกว่าระดับที่กำหนดถ้าประชากรมีการแจกแจงไม่เป็นปกติ นอกจากนี้แกลสส์และฮอบคินส์ (Glass & Hopkins, 1996) ยังกล่าวว่าสถิติทดสอบบาร์ตเลตเป็นสถิติทดสอบที่มีความไวต่อการแจกแจงของประชากรที่ไม่เป็นแบบปกติโดยเฉพาะอย่างยิ่งข้อมูลที่มีความโด่งสูงกว่าปกติ (Leptokurtic) กล่าวคือสำหรับข้อมูลที่มีการแจกแจงแบบไม่ปกติ ตัวสถิติทดสอบนี้จะมีโอกาสที่จะปฏิเสธสมมติฐานหลักมากกว่าที่ควรจะเป็น และเมื่อข้อมูลมีการแจกแจงแบบโด่งต่ำกว่าปกติ (Platykurtic) ตัวสถิติทดสอบนี้จะมีโอกาสที่จะปฏิเสธสมมติฐานหลักน้อยกว่าที่ควรจะเป็น เกมและคณะ (Game et. al., 1972) ได้กล่าวว่าในการทดสอบเพื่อเปรียบเทียบความแปรปรวน ถ้าไม่คำนึงถึงรูปแบบการแจกแจงของประชากร ส่วนเบี่ยงเบนมาตรฐานของความแปรปรวนของตัวอย่างคือ $\sigma_{S^2} = \sigma^2_{X_i}\sqrt{\frac{2}{n-1}+\frac{\gamma_2}{n}}$ โดยที่ γ_2 คือดัชนีความโด่งของประชากร เมื่อประชากรมีการแจกแจงปกติ $\gamma_2 = 0$ แต่ถ้าประชากรมีการแจกแจงโด่งสูงกล่าวคือ $\gamma_2 > 0$ จะทำให้ค่าที่แท้จริงของ $\sigma^2_{S^2}$ มากกว่ากรณีที่ประชากรมีการแจกแจงปกติ ซึ่งจะทำให้เกิดความคลาดเคลื่อนประเภทที่ I ของการทดสอบสูงกว่าระดับที่กำหนด แต่ถ้าประชากรมีการแจกแจงโด่งแบนกล่าวคือ $\gamma_2 < 0$ ค่าที่แท้จริงของ $\sigma^2_{S^2}$ จะน้อยกว่ากรณีที่ประชากรมีการแจกแจงปกติ ซึ่งจะทำให้เกิดความคลาดเคลื่อนประเภทที่ I ของการทดสอบต่ำกว่าระดับที่กำหนด ซึ่งผลการทดลองครั้งนี้ก็สอดคล้องกับ คำกล่าวของแกลสส์-ฮอบคินส์และเกม สำหรับสถิติทดสอบเลอวีนนั้นไม่สามารถควบคุมความคลาดเคลื่อน

ประเภทที่ I เมื่อประชากรมีการแจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา ซึ่งแกลสส์และ ฮอบกินส์ได้กล่าวถึงสถิติทดสอบเลอวีน ว่าเนื่องจากใช้ ANOVA วิเคราะห์จึงต้องอยู่บนข้อตกลงเบื้องต้นอันเดียวกันทำให้ขาดความสมเหตุสมผล นอกจากนั้นไม่ใช้ค่าสังเกต X_i แต่เป็น $\left|X_i - \overline{X}\right|$ เมื่อทำการวิเคราะห์แล้วผลที่ได้จะสรุปผลได้ว่าค่าเฉลี่ยของ $\left|X_i - \overline{X}\right|$ ของแต่ละกลุ่มแตกต่างกันหรือไม่ ซึ่งสถิติทดสอบเลอวีนน่าจะถูกต้องเฉพาะกรณีขนาดกลุ่มตัวอย่างเท่ากันเท่านั้น ซึ่งสอดคล้องกับคำกล่าวของเคเยส (Keyes, 1997) ที่ว่าค่าเฉลี่ยของ $\left|X_i - \overline{X}\right|$ แปรเปลี่ยนไปตามขนาดกลุ่มตัวอย่าง ยิ่งไปกว่านั้นการแจกแจงของ $\left|X_i - \overline{X}\right|$ มักจะมีลักษณะเบ้เสมอ ซึ่งขัดกับลักษณะของประชากรและค่าเฉลี่ยของ $\left|X_i - \overline{X}\right|$ ไม่ใช่ตัวประมาณที่ไม่เอนเอียงและมีความแปรปรวนต่ำสุดของความแปรปรวนของประชากร นอกจากนี้มิลเลอร์ (Miller, 1968) ได้พิสูจน์ให้เห็นว่าถ้ากลุ่มตัวอย่างมีขนาดเล็ก การแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของกลุ่มตัวอย่างนั้นจะเกิดความสัมพันธ์กันสูงระหว่างส่วนเบี่ยงเบนในกลุ่มเดียวกัน ถ้าการแจกแจงของประชากรมีลักษณะเบ้ (Skew) ทำให้ตัวสถิติทดสอบเลอวีนขาดความเที่ยงตรง ซึ่งจะทำให้มีผลต่อความคลาดเคลื่อนประเภทที่ I เช่นเดียวกับสถิติทดสอบบาร์ตเลต สำหรับสถิติทดสอบคอครันเป็นตัวสถิติทดสอบที่ใช้ง่ายและรวดเร็ว ใช้เมื่อขนาดของกลุ่มตัวอย่างเท่ากันเท่านั้น ถ้าขนาดกลุ่มตัวอย่างไม่เท่ากันการทดลองนี้ใช้ค่าเฉลี่ย ฮาร์โมนิคของขนาดของทุกกลุ่มตัวอย่างแทนขนาดกลุ่มตัวอย่าง แต่ก็ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I เมื่อประชากรมีการแจกแจงโด่งแบนและแจกแจงโด่งสูงคล้ายกันกับสถิติทดสอบบาร์ตเลต สำหรับการแจกแจงปกติ แจกแจงเบ้ซ้ายและแจกแจงเบ้ขวากรณีที่ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I นั้นส่วนใหญ่อัตราความคลาดเคลื่อนประเภทที่ I จะน้อยกว่าระดับนัยสำคัญที่กำหนดและเกิดขึ้นมากที่สุดเมื่อเปรียบเทียบกับสถิติทดสอบตัวอื่นๆ ซึ่งสอดคล้องกับการศึกษาของโคโนเวอร์และคณะ (Conover et al., 1981) และไรเวสท์ (Rivest, 1986) ที่พบว่าสถิติทดสอบ คอครันไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ I ได้เมื่อประชากรแจกแจงแบบไม่สมมาตรและแบบหางยาว

เมื่อพิจารณาระดับนัยสำคัญ (α) ของการทดสอบ ที่ระดับนัยสำคัญ .05 สถิติทดสอบที่พัฒนาขึ้นทั้ง 2 ตัว สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้มากกว่าสถิติทดสอบตัวอื่นๆ และเนื่องจากในการทดสอบสมมติฐานเกี่ยวกับความเป็นเอกพันธ์ของความแปรปรวนของประชากรนั้นควรจะควบคุมความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ II (β) ให้มีค่าน้อยลงโดยการกำหนดระดับนัยสำคัญ (α) ให้มีค่ามากขึ้น เพราะการทดสอบสมมติฐานกรณีนี้เป็นการตรวจสอบว่าข้อมูลนั้นมีลักษณะเป็นไปตามข้อตกลงเบื้องต้นหรือไม่ ถ้าไม่ควบคุมค่า β จะทำให้มีโอกาสในการยอมรับสมมติฐานหลัก (H_0) ทั้งๆ ที่สมมติฐานหลักนั้น ผิดมีมากขึ้น ถ้าผลจากการทดสอบสมมติฐานปรากฏว่าไม่ปฏิเสธ H_0 ก็ไม่แน่ว่าการที่ยอมรับ $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$ นั้นเป็นเพราะว่าความแปรปรวนของประชากรแต่ละกลุ่มเท่ากันจริงๆ หรือเพราะว่าเปิดโอกาสในการยอมรับ H_0 มากขึ้น แต่ถ้าผู้วิจัยควบคุมค่า β โดยกำหนดค่า α ให้สูงขึ้น เช่นจากระดับนัยสำคัญ .01 เป็น .05 ถ้าผลจากการทดสอบสมมติฐานปรากฏว่าไม่ปฏิเสธ H_0 ก็มั่นใจได้ว่าความแปรปรวนของประชากรแต่ละกลุ่ม

เท่ากันจริงๆ เพราะว่าถึงแม้ว่าจะกำหนดโอกาสของการยอมรับ H_0 ให้ต่ำลงแต่ก็ยังยอมรับ H_0 อยู่ ก็แสดงว่าความแปรปรวนของประชากรแต่ละกลุ่มเท่ากันจริง

3. จากสมมติฐานที่กล่าวว่าสถิติทดสอบที่พัฒนาขึ้น น่าจะมีอำนาจการทดสอบสูงกว่า สถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีนและสถิติทดสอบคอครัน

ผลการวิจัยพบว่าอำนาจการทดสอบของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลอวีน สถิติทดสอบคอครันและสถิติทดสอบที่พัฒนาขึ้นมีค่าสูงขึ้น เมื่อดัชนีความแตกต่างของความแปรปรวนของประชากร (ϕ) เพิ่มขึ้นทั้งนี้เนื่องมาจากอำนาจการทดสอบจะแปรตามขนาดของความแตกต่างระหว่างค่าความแปรปรวนของประชากรภายใต้สมมติฐานหลักและสมมติฐานทางเลือก กล่าวคือถ้าค่าความแปรปรวนของประชากรภายใต้สมมติฐานทางเลือกยิ่งต่างจากค่าความแปรปรวนของประชากรภายใต้สมมติฐานหลักมากเท่าไรจะยิ่งทำให้มีอำนาจการทดสอบเพิ่มมากขึ้นเท่านั้น ส่วนสถิติทดสอบคอครันมีอำนาจการทดสอบลดลงเล็กน้อยเมื่อดัชนีความแตกต่างของความแปรปรวนของประชากรเปลี่ยนจากระดับน้อยเป็นระดับปานกลาง และเปลี่ยนจากระดับปานกลางเป็นระดับมากนั้นยัง ทำให้ผลสรุปของการเปรียบเทียบอำนาจการทดสอบคงเดิม ซึ่งผลการทดลองเช่นนี้เกิดขึ้นเช่นเดียวกับผลการวิจัยของอารยา กุลานุช (2528) และเกมและคณะ (Game et al., 1972) ที่ค่าดัชนีความแตกต่างของความแปรปรวนของประชากรเพิ่มขึ้น แต่ผลจากการทดลองอำนาจการทดสอบกลับลดลง ซึ่งเกมให้เหตุผลว่าอาจเนื่องมาจากสูตร การคำนวณของตัวสถิติทดสอบหรืออาจเป็นเพราะความคลาดเคลื่อนจากการคำนวณตัวสถิติทดสอบของเครื่องคอมพิวเตอร์ เมื่อพิจารณาเฉพาะกรณีที่สถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ I พบว่าถ้าประชากรมีการแจกแจงปกติ สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบทุกตัว สถิติทดสอบ SA1 และสถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงใกล้เคียงกันและสูงกว่าสถิติทดสอบ SA2 และสถิติทดสอบคอครัน ถ้าประชากรมีการแจกแจงเบ้ซ้ายและแจกแจงเบ้ขวา สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA1 สถิติทดสอบ SA2 และสถิติทดสอบคอครัน สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2 และสถิติทดสอบ คอครัน ถ้าประชากรมีการแจกแจงโด่งแบน สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบเลอวีนและสถิติทดสอบ SA2 สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2 ถ้าประชากรมีการแจกแจงโด่งสูง สถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA1 และสถิติทดสอบ SA2 สถิติทดสอบ SA1 มีอำนาจการทดสอบสูงกว่าสถิติทดสอบ SA2 ทั้งนี้อาจเป็นเพราะสถิติทดสอบบาร์ตเลตพัฒนามาจากตัวสถิติทดสอบของนีแมน-เพียร์สัน โดยการแทนตัวประมาณภาวะน่าจะเป็นสูงสุดที่เอนเอียงด้วยตัวประมาณที่ไม่เอนเอียงของความแปรปรวน เปลี่ยนองศาความเป็นอิสระจาก n_i เป็น $n_i - 1$ และใช้ค่าคงที่ C ที่ประกอบไปด้วยขนาดตัวอย่างและจำนวนกลุ่มตัวอย่างเป็นค่าปรับแก้เพื่อประมาณด้วยการแจกแจงแบบไค-สแควร์ เป็นสถิติทดสอบที่คำนวณจากข้อมูลดิบทุกจำนวนที่เป็นอิสระต่อกัน และใช้การเปรียบเทียบแบบผลต่างของลอการิทึมของความแปรปรวน ซึ่งผลการทดลองนี้สอดคล้องกับเยาวภา ไชยศรี (2542) ที่พบว่าเมื่อลักษณะการแจกแจงประชากรเป็นแบบปกติและแบบไวบูลล์ (เบ้ขวา) สถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูง

กว่าสถิติทดสอบเลอวีน นอกจากนี้ อารยา กุลานุช (2528) สุวรรณ มุ่งฝากกลาง (2535) เกมและคณะ (1972) การ์ทไซด์ (Gartside, 1972) และ เลยาร์ด (Layard, 1973) ยังพบว่าเมื่อประชากรมีการแจกแจงปกติสถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงสุด ส่วนแกสส์และฮอบคินส์ได้กล่าวถึงสถิติทดสอบบาร์ตเลตว่าเป็นสถิติทดสอบที่ให้ผลการทดสอบที่ถูกต้อง ถ้ากลุ่มของประชากรมีการแจกแจงปกติ แต่เป็นสถิติทดสอบที่มีความยุ่งยากซับซ้อนในการคำนวณมากเพราะจะต้องมีการคำนวณจากค่าลอการิทึมของความแปรปรวนของกลุ่มตัวอย่าง สำหรับสถิติทดสอบเลอวีนใช้ ANOVA ในการวิเคราะห์โดยคำนวณจากค่าสัมบูรณ์ของผลต่างของข้อมูลแต่ละจำนวนกับค่าเฉลี่ยของกลุ่มตัวอย่างนั้นๆ อาจทำให้ข้อมูลที่นำมาคำนวณไม่เป็นอิสระต่อกันและข้อมูลที่เปลี่ยนแปลงนี้อาจ ทำให้สารสนเทศของข้อมูลบางส่วนหายไป จึงส่งผลให้อำนาจการทดสอบของสถิติทดสอบเลอวีนน้อยลง ในขณะที่สถิติทดสอบ SA1 และสถิติทดสอบ SA2 ทำการเปรียบเทียบผลต่างของความแปรปรวนสูงสุดและต่ำสุดนั้น มื่อผลต่างของความแปรปรวนมีค่าน้อยอาจทำให้อำนาจการทดสอบของสถิติทดสอบ SA1 และ SA2 น้อยลงได้ แต่ถ้าดัชนีความแตกต่างของความแปรปรวนของประชากรมีค่ามากแล้วสถิติทดสอบบาร์ตเลต สถิติทดสอบ SA1 และสถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงใกล้เคียงกัน ส่วนสถิติทดสอบคอครันทำการเปรียบเทียบอัตราส่วนของความแปรปรวนสูงสุดและผลรวมของความแปรปรวนทุกกลุ่ม ใช้ทดสอบการเท่ากันของความแปรปรวนเมื่อขนาดของกลุ่มตัวอย่างเท่ากัน แต่กรณีที่ขนาดตัวอย่างไม่เท่ากันการใช้ค่าเฉลี่ยฮาร์โมนิคของขนาดของทุกกลุ่มตัวอย่างแทนขนาดกลุ่มตัวอย่าง อาจทำให้การหาค่าองศาความเป็นอิสระคลาดเคลื่อนได้และส่งผลให้อำนาจการทดสอบของสถิติทดสอบคอครันน้อยลง อย่างไรก็ตามกรณีที่สถิติทดสอบ SA1 สถิติทดสอบ SA2 และสถิติทดสอบ คอครันไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I นั้นส่วนใหญ่อัตราความคลาดเคลื่อนประเภทที่ I จะน้อยกว่าระดับนัยสำคัญที่กำหนดและเกิดขึ้นมากที่สุดเมื่อเปรียบเทียบกับสถิติทดสอบตัวอื่นๆ ซึ่งอาจหมายถึงว่าโอกาสที่อำนาจการทดสอบของสถิติทดสอบ SA1 สถิติทดสอบ SA2 และสถิติทดสอบคอครันจะต่ำกว่าที่เป็นจริงและจะเกิดขึ้นมากกว่าเมื่อเปรียบเทียบกับสถิติทดสอบตัวอื่นๆ ในทางตรงกันข้ามสถิติทดสอบเลอวีนไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ I นั้นส่วนใหญ่อัตราความคลาดเคลื่อนประเภทที่ I จะมากกว่าระดับนัยสำคัญที่กำหนดและเกิดขึ้นมากที่สุดเมื่อเปรียบเทียบกับสถิติทดสอบตัวอื่นๆ ซึ่งอาจหมายถึงว่าโอกาสที่อำนาจการทดสอบของสถิติทดสอบเลอวีนจะสูงกว่าที่เป็นจริงและจะเกิดขึ้นมากกว่าเมื่อเปรียบเทียบกับสถิติทดสอบตัวอื่นๆ ก็เป็นได้

ข้อเสนอแนะ

การนำผลการวิจัยไปใช้

เมื่อไม่ทราบการแจกแจงของประชากร ควรเลือกใช้สถิติทดสอบ SA1 เนื่องจากสามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ทุกการแจกแจงของประชากร และเมื่ออัตราส่วนของความแปรปรวนของประชากร

แตกต่างกันมาก สถิติทดสอบบาร์ตเลต สถิติทดสอบ SA1 และสถิติทดสอบเลอวีนมีอำนาจการทดสอบสูงใกล้เคียงกัน

การวิจัยต่อไป

ควรศึกษาถึงลักษณะการแจกแจงประชากรแบบไม่ปกติ โดยกำหนดค่าความเบ้และค่าความโด่งให้ครอบคลุมยิ่งขึ้น

เอกสารอ้างอิง

เยาวภา ไชยศรี. (2542). *การเปรียบเทียบประสิทธิภาพของตัวทดสอบสถิติบาร์ตเลต เลอวีน และบราวน์-ฟอร์สิตีสำหรับทดสอบความเท่ากันของความแปรปรวนของประชากร*. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

สุวรรณ มุ่งฝากกลาง. (2535). *การเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I และอำนาจการทดสอบของสถิติทดสอบบาร์ตเลต เลอวีน และสแควร์แรงค์ ในการทดสอบความเป็นเอกพันธ์ของความแปรปรวน*. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต. ภาควิชาวิจัยการศึกษา. จุฬาลงกรณ์มหาวิทยาลัย.

อารยา กุลานุช. (2528). *การเปรียบเทียบอำนาจการทดสอบการเท่ากันของความแปรปรวนของประชากรโดยใช้สถิติทดสอบ 3 ประเภท*. วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต. ภาควิชาสถิติ. จุฬาลงกรณ์มหาวิทยาลัย.

Brown, M.B.,& Forsythe, A.B. (1974). Robust test for the equality of variances. *Journal of the American Statistical Association,* 69, 364-367.

Conover, W.J., Johnson, M.E., & Johnson, M.M. (1981). A comparative study of tests for homogeneity of variances, with application to the outer continental shelf bidding data. *Technometrics,* 23, 351-361.

Game, P.A. , Winkler, H.R., & Probert, D.A. (1972). Robust test for homogeneity of variance. *Educational and Psychological Measurement,* 32, 887 – 909.

Gartside, P.S. (1972). A study of methods for comparing several variances. *Journal of the American Statistical Association*, 67, 342-346.

Keyes, T.K., & Levy, M.S. (1997). Analysis of Levene's test under design imbalance. *Journal of Educational and Behavioral Statistics*, 22, 227-236.

Glass, V.G., & Hopkins, K.D. (1996). *Statistical Methods in Education and Psychology.*(3 rd.th.). New Jersey: Allyn & Bacon.

Layard, M.W.J. (1973). Robust large-sample tests for homogeneity of variances. *Journal of the American Statistical Association*, 68, 195-198.

Miller, R.G. (1968). Jacknifing Variances. *Annals of Mathematical Statistics*, 39, 567-582.

Rivest, L.P. (1986). Bartlett's, Cochan's, and Hartley's tests of variances are liberal when the underlying distribution is long-tailed. *Journal of the American Statistical Association,* 81, 124-128.