

สรุปผลการวิจัย อภิปรายผลและข้อเสนอนแนะ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อศึกษาเปรียบเทียบอำนาจการทดสอบการเท่ากันของความแปรปรวนของประชากรโดยใช้ตัวสถิติทดสอบ 3 ประเภทคือ ตัวสถิติทดสอบ X^2 หรือตัวสถิติทดสอบของบาร์ตเลต ตัวสถิติทดสอบ C หรือตัวสถิติทดสอบของแคตเวล และตัวสถิติทดสอบ F' หรือตัวสถิติทดสอบของเลอวิน และหาผลสรุปว่าตัวสถิติใดมีความเหมาะสมที่จะใช้ทดสอบการเท่ากันของความแปรปรวนของประชากรในแต่ละสถานการณ์ ซึ่งได้กำหนดขึ้นดังต่อไปนี้

1. การแจกแจงของประชากรเป็นแบบปกติ และแบบไม่ปกติคือ การแจกแจงแบบยูนิฟอร์ม แบบดับเบิลเอกซ์โพเนนเชียล แบบพลาติเคอร์ติค ซึ่งกำหนดความโด่ง 2.8 และความเบ้ 0.25, 0.5, 0.75 แบบเมโซเคอร์ติค ซึ่งกำหนดความโด่ง 3.0 และความเบ้ 0.25, 0.5, 0.75 แบบเลปโตเคอร์ติค ซึ่งกำหนดความโด่ง 3.2 และความเบ้ 0.25, 0.5, 0.75 และการแจกแจงแบบเอกซ์โพเนนเชียล รวมการแจกแจงของประชากรทั้งหมด 13 รูปแบบ
2. ขนาดของกลุ่มตัวอย่างซึ่งสุ่มจาก 3 ประชากรมีขนาดเท่ากันคือ 5, 11, 21, 31, และ 61 ตามลำดับ
3. อัตราส่วนความแตกต่างของความแปรปรวนของประชากรกำหนด 3 ระดับคือ แตกต่างกันน้อย ปานกลาง และแตกต่างกันมาก
4. ระดับนัยสำคัญของการทดสอบที่กำหนดขึ้นเท่ากับ 0.01 และ 0.05

วิธีดำเนินการวิจัยครั้งนี้ใช้วิธีการจำลองการทดลองด้วยเทคนิคมอนติคาร์โลซิมูเลชัน ด้วยการทำงานของเครื่องคอมพิวเตอร์ IBM 370/3010 ใช้โปรแกรมภาษาฟอร์แทรน (FORTRAN) เพื่อสร้างข้อมูลให้มีลักษณะตามแผนการทดลองที่กำหนด และกำหนดให้เครื่องคอมพิวเตอร์จำลองการทดลองทั้งหมด 1,000 ครั้งในแต่ละสถานการณ์



5.1 สรุปผลการวิจัย

การหาผลสรุปว่าตัวสถิติใดมีความเหมาะสมที่จะใช้ทดสอบการเท่ากันของความแปรปรวนของประชากรนอกจากจะเปรียบเทียบค่าอำนาจการทดสอบของตัวสถิติทั้งสามจากผลการทดลองแล้ว จะต้องพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ที่เกิดขึ้นในการทดลองของตัวสถิติทดสอบในแต่ละสถานการณ์ด้วย ดังนั้นการสรุปผลการวิจัยครั้งนี้จะเสนอเป็น 2 ตอนคือ

5.1.1 การเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ของตัวสถิติทดสอบทั้งสามที่ระดับนัยสำคัญ 0.01 และ 0.05 สรุปผลได้ดังนี้คือ

1. สถิติทดสอบของบาร์ตเลตสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีที่สุด และมีความเชื่อถือได้ในการสรุปผล เมื่อมีการนำผืนข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของประชากรที่ต้องเป็นแบบปกติถ้าการแจกแจงของประชากรมีลักษณะที่ไม่เบี่ยงเบนจากการแจกแจงแบบปกติมากนัก

2. สถิติทดสอบของเลอวันไม่สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ถ้าตัวอย่างมีขนาดเล็กและถ้าตัวอย่างมีขนาดใหญ่เพียงพอจะสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้เมื่อประชากรมีการแจกแจงแบบดับเบิลเอกซ์โปเนนเชียลและแบบเอกซ์โปเนนเชียล ในขณะที่สถิติทดสอบอื่น ไม่สามารถควบคุมได้

3. สถิติทดสอบของแคตเวลควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้มีน้อยมากเมื่อขนาดของตัวอย่างใหญ่

4. สถิติทดสอบทั้ง 3 ประเภทสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ที่ระดับนัยสำคัญ 0.01 ถ้าการแจกแจงของประชากรเป็นแบบเมโซเคอร์ติก $(K,S) = (3.0, 0.25)$ เมื่อ $n = 21$ และ 31 $(K,S) = (3.0, 2.5)$ เมื่อ $n = 31$ แบบเลปโตเคอร์ติก $(K,S) = (3.2, 0.25)$ และ $(3.2, 0.5)$ เมื่อ $n = 21$ ส่วนที่ระดับนัยสำคัญ 0.05 สถิติทดสอบทั้ง 3 ประเภทสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ถ้าการแจกแจงของประชากรเป็นแบบเมโซเคอร์ติก $(K,S) = (3.0, 0.25)$ เมื่อ $n = 21$ และ 61

5.1.2 การเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบการเท่ากันของความแปรปรวนของประชากร 3 ประเภทที่ระดับนัยสำคัญ 0.01 และ 0.05 สรุปผลได้ดังนี้คือ

1. ตัวสถิติทดสอบของบาร์ตเลต มีอำนาจการทดสอบสูงที่สุด ไม่ว่าจะกำหนดระดับนัยสำคัญของการทดสอบที่ 0.01 หรือ 0.05 เกือบทุกการแจกแจงของประชากร และขนาดของกลุ่มตัวอย่างที่กำหนด ยกเว้นการแจกแจงแบบยูนิฟอร์ม

2. ตัวสถิติทดสอบของแคตเวลมีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบของเลอวินเมื่อกลุ่มตัวอย่างมีขนาดเล็ก และจะมีอำนาจการทดสอบใกล้เคียงกับอำนาจการทดสอบของตัวสถิติทดสอบของบาร์ตเลต

3. ตัวสถิติทดสอบของเลอวินมีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบของแคตเวลเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ และจะมีอำนาจการทดสอบใกล้เคียงกับอำนาจการทดสอบของตัวสถิติทดสอบของบาร์ตเลต

4. เมื่ออัตราส่วนของความแปรปรวนของประชากรแตกต่างกันมาก อำนาจการทดสอบของตัวสถิติทดสอบทั้งสามสูงใกล้เคียงกัน และเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ อำนาจการทดสอบของตัวสถิติทั้งสามเกือบเท่ากับ 1 สำหรับการทดสอบที่ระดับนัยสำคัญ 0.05

5. เมื่อจำแนกตามลักษณะการแจกแจงของประชากรที่กำหนดส่วนมากแล้ว ถ้าเพิ่มขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญของการทดสอบจะทำให้อำนาจการทดสอบของตัวสถิติทั้งสามเพิ่มขึ้น

6. ตัวสถิติทดสอบของเลอวินซึ่งคำนวณโดยใช้มัธยฐานของตัวอย่างแทนค่าเฉลี่ยเลขคณิตของตัวอย่าง จะมีลักษณะ asymptotically distribution free ซึ่งสอดคล้องตามที่ Miller (1968 : 567-582) ได้กล่าวเสนอแนะไว้

7. ถ้าเปรียบเทียบอำนาจการทดสอบของตัวสถิติในสถานการณ์ ซึ่งสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ทั้ง 3 ประเภทแล้ว ตัวสถิติทดสอบของบาร์ตเลตยังคงมีอำนาจการทดสอบสูงที่สุด และถ้าขนาดของกลุ่มตัวอย่างใหญ่ ตัวสถิติทดสอบของเลอวินจะมีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบของแคตเวล

5.2 ปัญหาที่เกิดขึ้นในการวิจัยครั้งนี้

จากผลการทดลองจะสังเกตได้ว่า กรณีที่กลุ่มตัวอย่างมีขนาดเล็กมากคือ ขนาดเท่ากับ 5 หรือ 11 บางกรณีของการทดสอบที่ระดับนัยสำคัญ 0.01 ถ้ากำหนดระดับความแตกต่างของอัตราส่วนความแปรปรวนของประชากรทั้งสามคือค่า 5 อยู่ในระดับต่ำมาก และความแตกต่างระหว่างค่า 5 น้อย จะทำให้เกิดปัญหากรณีที่ค่าอำนาจการทดสอบของตัวสถิติแทนที่จะเพิ่มขึ้นหรือคงที่เมื่อ

ค่า θ เพิ่มขึ้นแต่ผลจากการทดลองค่าอำนาจการทดสอบกลับลดลง เมื่อ θ มีค่าเพิ่มขึ้นเล็กน้อย แต่ก็ยังคงทำให้ผลลัพธ์ของการเปรียบเทียบอำนาจการทดสอบยังคงเดิม ซึ่งปัญหาเช่นนี้ได้เกิดขึ้นเช่นเดียวกับผลการวิจัยของ Game (1972 : 887-909) ที่ได้ ทำการวิจัยเรื่องความแปรปรวนของตัวสถิติทดสอบการเท่ากันของความแปรปรวนซึ่งเมื่อกำหนดขนาดของกลุ่มตัวอย่างเท่ากับ 18 และลักษณะการแจกแจงของประชากรเป็นแบบยูนิฟอร์ม เปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบของบาร์ตเลต สถิติทดสอบ M สถิติทดสอบ Q สถิติทดสอบ LEV3 และสถิติทดสอบ LEV2 ที่ระดับนัยสำคัญ 0.05 ผลปรากฏว่าเมื่อ θ เพิ่มขึ้น อำนาจการทดสอบของตัวสถิติทดสอบ M จะลดลงเรื่อย ๆ ซึ่ง Game ให้เหตุผลว่าอาจเนื่องมาจากสูตรการคำนวณของตัวสถิติ M หรืออาจเป็นเพราะความคลาดเคลื่อนจากการคำนวณตัวของเครื่องคอมพิวเตอร์ ซึ่งบางฟังก์ชันในสูตรการคำนวณ Game ใช้วิธี 'double precision' และบางฟังก์ชันใช้ 'single precision' ซึ่งให้ค่าตัวเลขเพียง 7 ตำแหน่ง แทนที่จะเป็น 16 ตำแหน่งเหมือนกัน นอกจากนี้จากผลการวิจัยของ Blair และ Higgins (1980) ซึ่งทำการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบวิลคอกซอน และสถิติทดสอบที่ ภายใต้การแจกแจงของประชากรที่เป็นแบบไม่ปกติ เมื่อประชากรมีการแจกแจงแบบ Mixed Normal สำหรับการทดสอบที่ระดับนัยสำคัญ 0.025 และขนาดของกลุ่มตัวอย่างเท่ากับ 6 อำนาจการทดสอบของตัวสถิติทดสอบวิลคอกซอน จะลดลงบางจุดเมื่อ $\mu_1 - \mu_2$ เพิ่มขึ้น

อย่างไรก็ดีกรณีที่เกิดปัญหาเช่นนี้ขึ้นในการทดลองครั้งนี้ ผู้วิจัยคิดว่าอาจเนื่องมาจากสาเหตุต่อไปนี้คือ

1. การวิจัยครั้งนี้กำหนดจำนวนของการทำการทดลองซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ ซึ่งอาจจะน้อยเกินไปสำหรับการทดสอบในบางรูปแบบ ของการแจกแจงของประชากรที่ใช้กลุ่มตัวอย่างขนาดเล็กมาก ซึ่งถ้าทำการทดลองซ้ำเป็น 10,000 ครั้งอาจจะทำให้ไม่เกิดปัญหาดังกล่าวขึ้น ดังที่ผู้วิจัยได้ทดลองบางรูปแบบการแจกแจงของประชากร ปรากฏผลแสดงในภาคผนวก

2. ความคลาดเคลื่อนอาจเนื่องมาจากการสุ่มตัวอย่างจากประชากร เนื่องจากอัตราส่วนของประชากรที่กำหนดแตกต่างกันน้อย และระดับของ θ ก็ต่างกันน้อยมาก จึงทำให้ค่าความแปรปรวนของตัวเลขสุ่มที่สร้างขึ้นแล้วสุ่มตัวอย่างมาทดสอบเกิดความคลาดเคลื่อนได้มาก

5.3 ย้อเล่นอแนะ ผลจากการวิจัยครั้งนี้จะเล่นอแนะเป็น 2 ด้านคือ

ด้านการนำไปใช้ประโยชน์

เพื่อเป็นแนวทางในการเลือกใช้ลัทธิทดสอบการเท่ากันของความแปรปรวนของประชากรที่มากกว่า 2 ประชากรขึ้นไป ดังนี้

1. เมื่อทราบลักษณะการแจกแจงของประชากร

1.1 ถ้าการแจกแจงของประชากรมีลักษณะไม่เบี่ยงเบนจากแจกแจงแบบปกติมากนัก ควรใช้ลัทธิทดสอบของบาร์ตเลต เพราะมีอำนาจการทดสอบสูงที่สุด และมีความเชื่อถือได้ แต่ถ้ากลุ่มตัวอย่างมีขนาดเล็กลงและผู้ทดสอบต้องการผลอย่างรวดเร็วแล้ว น่าจะเลือกใช้ลัทธิทดสอบของแคดเวล เพราะมีวิธีการคำนวณที่ง่ายกว่าลัทธิทดสอบของบาร์ตเลตและมีอำนาจการทดสอบใกล้เคียงกับลัทธิทดสอบของบาร์ตเลตมาก

1.2 เมื่อการแจกแจงของประชากรเป็นแบบยูนิฟอร์ม ถึงแม้ว่าจะไม่มีตัวสถิติใดที่สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ควรเลือกใช้ลัทธิทดสอบของบาร์ตเลต เมื่อกลุ่มตัวอย่างมีขนาดเล็ก เพราะมีอำนาจการทดสอบสูงที่สุด และมีความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ใกล้เคียงกับระดับนัยสำคัญที่กำหนดมากที่สุด และเมื่อกลุ่มตัวอย่างมีขนาดใหญ่ควรใช้ลัทธิทดสอบของเลอวิน

1.3 เมื่อลักษณะการแจกแจงของประชากรเป็นแบบดับเบิลเอกซ์โปเนนเชียลและแบบเอกซ์โปเนนเชียล ถ้ากลุ่มตัวอย่างมีขนาดใหญ่พอควรเลือกใช้ลัทธิทดสอบของเลอวิน เพราะควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีที่สุด

2. เมื่อไม่ทราบลักษณะการแจกแจงของประชากรควรเลือกใช้ลัทธิทดสอบของบาร์ตเลตและลัทธิทดสอบของเลอวินเมื่อใช้กลุ่มตัวอย่างขนาดใหญ่ เพราะสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้มากกว่าลัทธิทดสอบของแคดเวลและลัทธิทดสอบของบาร์ตเลตที่มีอำนาจการทดสอบสูงที่สุด

ด้านการศึกษาวิจัย

1. ถ้าเป็นไปได้ควรมีการศึกษาวิจัยตนเองเดียวกันนี้โดยกำหนดจำนวนซ้ำของการทดลองมากกว่า 5,000 ครั้งขึ้นไป
2. การทดลองครั้งนี้กำหนดลักษณะการแจกแจงของประชากรในแบบที่ไม่สมมาตร ซึ่งมีความเบ้แบบบวก (the positive skewness) ทั้งหมด ดังนั้นจึงน่าจะมีการเปรียบเทียบกรณีการแจกแจงของประชากรที่มีความเบ้แบบลบ (the negative skewness) ด้วย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย