



การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบอิงพารามิเตอร์ สำหรับการทดสอบภาวะความเท่ากันของความแปรปรวน วารางคณา เรียนสุทธิ์

A Comparison of the Efficiency of the Parametric Test Statistics for Testing Homogeneity of Variances

Warangkhan Riansut

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ จังหวัดพัทลุง 93210

Department of Mathematics and Statistics, Faculty of Science, Thaksin University, Phatthalung, 93210

Corresponding author. E-mail address: warang27@gmail.com

บทคัดย่อ

วัตถุประสงค์ของการศึกษาค้นคว้าครั้งนี้ คือ เพื่อเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบของบาร์ตเลตต์ เลวิน บราวน์-ฟอร์ลิตี และ โอไบรน์ รวมถึงเพื่อเสนอแนะตัวสถิติทดสอบที่เหมาะสม โดยการคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ กำหนดประชากรที่ศึกษา คือ การแจกแจงปกติ การแจกแจงเบ้ซ้ายและเบ้ขวาที่มีความโด่งสูงและต่ำกว่าปกติ กำหนดจำนวนประชากร 3 และ 4 กลุ่ม และศึกษาทั้งกรณีที่มีขนาดเท่ากันและไม่เท่ากัน กำหนดความแตกต่างของความแปรปรวนประชากร 3 ระดับ คือ น้อย ปานกลาง และมาก ระดับนัยสำคัญที่ศึกษา คือ 0.01 และ 0.05 จำลองข้อมูลด้วยการเขียนโปรแกรมทางสถิติ กำหนดจำนวนรอบของการทำซ้ำ คือ 1,000 รอบ ผลการศึกษาพบว่า เมื่อประชากรมีการแจกแจงปกติ ตัวสถิติทดสอบของบาร์ตเลตต์มีความเหมาะสมมากที่สุด แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตีมีความเหมาะสมมากที่สุด และเมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์ เลวิน และโอไบรน์มีความเหมาะสมขึ้นอยู่กับขนาดตัวอย่างและระดับนัยสำคัญ

คำสำคัญ: การทดสอบภาวะความเท่ากันของความแปรปรวน บาร์ตเลตต์ เลวิน บราวน์-ฟอร์ลิตี โอไบรน์

Abstract

The objectives of this study were to compare the efficiency of the test statistics: Bartlett, Levene, Brown-Forsythe, and O'Brien, including the recommending the appropriate test statistic by calculate the probability of type I error and power of a test. Defined the study populations were normal distribution, left- and right-skewed with leptokurtic and platykurtic kurtosis distributions. Study three and four populations, including the sample sizes were equal and unequal. Defined the difference of the population variances into three levels: low, moderate, and high difference. The levels of significance to study were 0.01 and 0.05. The Statistical programming was used to simulate 1,000 replicates data. The results found that: when the population was normal distribution, Bartlett's test was the most appropriate. When the population was left- or right-skewed with leptokurtic kurtosis distributions, Brown-Forsythe's test was the most appropriate. When the population was left- or right-skewed with platykurtic kurtosis distributions, Bartlett's test, Levene's test, and O'Brien's test were appropriate depending on sample size and significance level.

Keywords: Testing Homogeneity of Variances, Bartlett, Levene, Brown-Forsythe, O'Brien.

บทนำ

การศึกษาค่าเฉลี่ยของประชากรมากกว่า 2 กลุ่ม โดยใช้ตัวสถิติทดสอบเอฟ (F-Test) ของการวิเคราะห์ความแปรปรวน (Analysis of Variance) จะมีข้อสมมุติ (Assumption) คือ ค่าสังเกตต้องมาจากตัวอย่างที่สุ่ม (Random Sample) มาจากประชากรที่มีการแจกแจงปกติ (Normal Distribution) และมีภาวะความแปรปรวนเท่ากัน



(Homoscedasticity) ดังนั้นก่อนการศึกษาเปรียบเทียบความแตกต่างของค่าเฉลี่ยของประชากรมากกว่า 2 กลุ่ม จึงจำเป็นต้องตรวจสอบข้อสมมุติก่อน ซึ่งปัญหาที่ผู้วิจัยมักพบ คือ ลักษณะการแจกแจงของประชากรไม่เป็นไปตามข้อสมมุติ ถ้ายังคงใช้ตัวสถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวน จะทำให้ผลการทดสอบสมมุติฐานเกิดความผิดพลาดได้ โดย Scheffe (1959) ได้แสดงความคิดเห็นว่า ตามข้อสมมุติเกี่ยวกับการแจกแจงของประชากรซึ่งต้องเป็นการแจกแจงปรกติ นั้น หากลักษณะของการแจกแจงของประชากรไม่เป็นไปตามข้อสมมุติจะส่งผลให้การประมาณค่าหรือการทดสอบสมมุติฐานเกี่ยวกับค่าความแปรปรวนของประชากรเกิดความผิดพลาดได้ รวมถึงการฝ่าฝืนข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวนจะมีผลอย่างมากต่อการสรุปผลที่ได้จากการใช้ตัวสถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวน โดยเฉพาะเมื่อขนาดตัวอย่างของแต่ละกลุ่มไม่เท่ากัน ดังนั้นก่อนการวิเคราะห์ข้อมูลที่เก็บรวบรวมมาไม่ว่าจะเป็นงานวิจัยในสาขาใดจึงควรจะต้องมีการตรวจสอบข้อสมมุติก่อน และ Keanthum (1996) สรุปไว้ว่า ก่อนที่จะวิเคราะห์ความแปรปรวน เราจำเป็นต้องทดสอบว่าแต่ละประชากรมีความแปรปรวนเท่ากันหรือไม่ ทั้งนี้เพื่อให้ผลการทดสอบที่ได้เป็นความแตกต่างของค่าเฉลี่ยประชากรจริง ๆ หากไม่มีข้อสมมุติข้อนี้ ผลการทดสอบอาจเป็นผลจากความแตกต่างของความแปรปรวนของข้อมูลก็ได้ ซึ่งถ้าหากไม่เป็นไปตามข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวนในกรณีที่ขนาดตัวอย่างของแต่ละกลุ่มเท่ากันและความแปรปรวนแตกต่างกันเพียงเล็กน้อย ตัวสถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวนยังคงมีความแกร่ง (Robust) ต่อลักษณะที่ไม่เป็นไปตามข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวน แต่ถ้าความแปรปรวนมีความแตกต่างกันค่อนข้างมาก จะมีผลกระทบต่อระดับนัยสำคัญของการทดสอบ ในขณะที่เดียวกันความคลาดเคลื่อนของการทดลองที่ไม่มีการแจกแจงปรกติ เช่น เบ้ไปข้างใดข้างหนึ่ง จะมีผลทำให้ความแปรปรวนไม่เท่ากัน ซึ่งส่งผลกระทบต่อการมีนัยสำคัญทางสถิติ

วิธีการทางสถิติที่นิยมใช้ในการทดสอบภาวะความเท่ากันของความแปรปรวนมีมากมาย ซึ่งแต่ละตัวสถิติทดสอบจะมีข้อจำกัดในการใช้ที่แตกต่างกัน เช่น ตัวสถิติทดสอบของบาร์ตเลตต์ (Bartlett's Test) ที่มีการแจกแจงไคกำลังสอง (Chi-Square Distribution) ได้รับการพัฒนามาจากตัวสถิติทดสอบของนิวแมนและเพียร์สัน สำหรับใช้ทดสอบภาวะความเท่ากันของความแปรปรวน ทั้งกรณีที่มีขนาดเท่ากันและไม่เท่ากัน เป็นตัวสถิติทดสอบที่มีความยุ่งยากในการคำนวณและมีความไวต่อข้อมูลที่ไม่มีการแจกแจงปรกติ โดยเฉพาะอย่างยิ่งเมื่อข้อมูลมีความโด่งผิดปรกติ กล่าวคือ สำหรับข้อมูลที่มีความโด่งสูงกว่าปรกติ (Leptokurtic) ตัวสถิติทดสอบนี้จะมีโอกาสปฏิเสธสมมุติฐานว่างมากกว่าที่ควรจะเป็น และเมื่อข้อมูลมีความโด่งต่ำกว่าปรกติ (Platykurtic) ตัวสถิติทดสอบนี้จะมีโอกาสปฏิเสธสมมุติฐานว่างน้อยกว่าที่ควรจะเป็น ตัวสถิติทดสอบของเลวิน (Levene's Test) มีการแจกแจงโดยประมาณแบบเอฟ (F Distribution) มีวิธีการคำนวณเช่นเดียวกับการวิเคราะห์ความแปรปรวนทางเดียว โดยการแทนค่าสังเกตในการวิเคราะห์ความแปรปรวนทางเดียวแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่าเฉลี่ยของตัวอย่าง ตัวสถิติทดสอบนี้มีความแกร่งเมื่อประชากรไม่มีการแจกแจงปรกติ (Gene & Hopkins, 1996) ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตี (Brown-Forsythe's Test) ได้รับการพัฒนามาจากตัวสถิติทดสอบของเลวิน โดยการใช้ค่ามัธยฐานแทนค่าเฉลี่ยของตัวอย่าง เพราะฉะนั้นจะแทนค่าสังเกตแต่ละค่าด้วยค่าสัมบูรณ์ของส่วนเบี่ยงเบนระหว่างค่าสังเกตแต่ละค่ากับค่ามัธยฐานของตัวอย่าง ตัวสถิติทดสอบนี้ยังคงมีความแกร่งเมื่อประชากรไม่มีการแจกแจงปรกติ เช่นเดียวกับตัวสถิติทดสอบของเลวิน (Conover, Johnson, & Johnson, 1981) สำหรับตัวสถิติทดสอบของโอ'Brien (O'Brien's Test) จะเป็นตัวสถิติทดสอบที่มีกำลังการทดสอบ (Power of a Test) สูงกว่าตัวสถิติทดสอบของเลวิน และตัวสถิติทดสอบของบราวน์-ฟอร์ลิตีเมื่อประชากรไม่มีการแจกแจงปรกติ (Chuenteerawong, 2001)

การศึกษาวิจัยในหลายๆ ด้าน มักไม่มีการตรวจสอบข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวน หรือถ้ามีการตรวจสอบก็มักใช้การทดสอบของเลวิน ทั้ง ๆ ที่มีวิธีการทดสอบที่เหมาะสมมากมายถึง 4 วิธีเป็นอย่างน้อยดังที่ได้กล่าวไปแล้ว อีกทั้งการทดสอบของเลวินไม่ใช้การทดสอบที่ดีที่สุด (Hatchavanich, 2014) ผลสรุปที่ได้จากการทดสอบนี้อาจไม่ตรงกับความเป็นจริง และอาจมีระดับนัยสำคัญ (Significance Level) หรือความน่าจะเป็นของความผิดพลาดแบบที่ 1 (Probability of Type I Error) มากหรือน้อยกว่าที่กำหนดไว้ รวมถึงอาจมีกำลังการทดสอบต่ำ

อย่างไรก็ตาม คำกล่าวที่ว่า ผลสรุปไม่ตรงกับความจริง ระดับนัยสำคัญมากหรือน้อยกว่าที่กำหนด และกำลังการทดสอบต่ำนั้นยังไม่มีที่ยืนยันในเชิงทฤษฎี ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีแนวคิดที่จะศึกษาการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 4 วิธี ได้แก่ ตัวสถิติทดสอบของบาร์ตเลตต์ ตัวสถิติทดสอบของเลวิน ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตี และตัวสถิติทดสอบของโอเบิร์น การเปรียบเทียบประสิทธิภาพจะกระทำภายใต้สถานการณ์ต่าง ๆ ทั้งกรณีการแจกแจง จำนวนประชากร ขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญที่แตกต่างกัน ผู้วิจัยคาดว่าผลการศึกษาที่ได้จะทำให้ทราบความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 4 วิธี ได้ทราบว่าวิธีใดมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่าง ๆ เพื่อเป็นแนวทางให้กับผู้ที่สนใจ สำหรับใช้ประโยชน์ในการตัดสินใจเลือกใช้ตัวสถิติทดสอบที่เหมาะสมกับการทดสอบภาวะความเท่ากันของความแปรปรวนต่อไป

วิธีการศึกษาและวัสดุอุปกรณ์

ขั้นตอนการดำเนินการวิจัย แสดงดังนี้

1. จำลองข้อมูลของประชากร 3 และ 4 กลุ่ม ด้วยโปรแกรมทางสถิติ กำหนดการแจกแจงความน่าจะเป็นของประชากรที่ศึกษา 5 การแจกแจง ได้แก่ การแจกแจงปกติ โดยใช้คำสั่ง $\text{rand}(\text{'normal'}, 0, 1)$ การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ (ความเบ้ = -2, ความโด่ง = 12) การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = -0.4, ความโด่ง = 2.2) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ (ความเบ้ = 2, ความโด่ง = 12) และการแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = 0.4, ความโด่ง = 2.2) สำหรับการจำลองข้อมูลที่มีลักษณะไม่สมมาตร (เบ้ซ้าย เบ้ขวา ความโด่งสูงกว่าปกติ และความโด่งต่ำกว่าปกติ) จะเริ่มต้นจากการจำลองข้อมูลที่มีการแจกแจงเอกรูปในช่วง (0, 1) โดยใช้คำสั่ง $\text{rand}(\text{'uniform'})$ แล้วแปลงข้อมูลตามวิธีการในบทความของ Ramberg, Tadikamalla, Dudewicz, and Mykytka (1979) ซึ่งสร้างตัวแปรสุ่ม X ที่ขึ้นอยู่กับความเบ้ (Skewness) และความโด่ง (Kurtosis) ดังนี้

$$X = \lambda_1 + \frac{U^{\lambda_3} - (1-U)^{\lambda_4}}{\lambda_2} \quad (1)$$

เมื่อ U แทนตัวเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0, 1)

λ_1 แทนพารามิเตอร์ตำแหน่ง (Location Parameter)

λ_2 แทนพารามิเตอร์มาตราส่วน (Scale Parameter)

λ_3, λ_4 แทนพารามิเตอร์รูปร่าง (Shape Parameter)

ค่าของพารามิเตอร์ $\lambda_1, \lambda_2, \lambda_3$ และ λ_4 ได้จากตารางที่ 1 หน้า 205 หรือตารางที่ 4 หน้า 210 - 214 ในบทความของ Ramberg et al. (1979) โดยค่าของพารามิเตอร์ทั้ง 4 ตัวจะให้ข้อมูลที่มีการแจกแจงเบ้ขวา สำหรับข้อมูลที่มีการแจกแจงเบ้ซ้ายให้สลับค่าของ λ_3 และ λ_4 และปรับเครื่องหมายของ λ_1 เป็นตรงข้าม ดังนั้นถ้าต้องการข้อมูลที่มีการแจกแจงตามที่กำหนด ต้องมีค่าของพารามิเตอร์ $\lambda_1, \lambda_2, \lambda_3$ และ λ_4 ดังนี้

- การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ (ความเบ้ = -2, ความโด่ง = 12)

$$\lambda_1 = 0.579, \lambda_2 = -0.1423, \lambda_3 = -0.0995 \text{ และ } \lambda_4 = -0.0273$$

- การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = -0.4, ความโด่ง = 2.2)

$$\lambda_1 = 1.354, \lambda_2 = 0.2582, \lambda_3 = 0.5683 \text{ และ } \lambda_4 = 0.0129$$



- การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปรกติ (ความเบ้ = 2, ความโด่ง = 12)
 $\lambda_1 = -0.579, \lambda_2 = -0.1423, \lambda_3 = -0.0273$ และ $\lambda_4 = -0.0995$
- การแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปรกติ (ความเบ้ = 0.4, ความโด่ง = 2.2)
 $\lambda_1 = -1.354, \lambda_2 = 0.2582, \lambda_3 = 0.0129$ และ $\lambda_4 = 0.5683$

กำหนดขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทั้งตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ดังตารางที่ 1

ตารางที่ 1 ขนาดตัวอย่างสำหรับประชากร 3 และ 4 กลุ่ม

ประชากร	ลักษณะขนาดตัวอย่าง	ขนาดตัวอย่างของแต่ละกลุ่ม		
		ขนาดเล็ก	ขนาดกลาง	ขนาดใหญ่
3 กลุ่ม	เท่ากัน	(10, 10, 10)	(30, 30, 30)	(100, 100, 100)
	ไม่เท่ากัน	(4, 7, 10)	(20, 25, 30)	(80, 90, 100)
4 กลุ่ม	เท่ากัน	(10, 10, 10, 10)	(30, 30, 30, 30)	(100, 100, 100, 100)
	ไม่เท่ากัน	(4, 7, 10, 13)	(20, 25, 30, 35)	(80, 90, 100, 110)

กำหนดค่าเฉลี่ยประชากรทุกกลุ่มเท่ากัน คือ เท่ากับ 10 และกำหนดความแตกต่างของความแปรปรวนประชากร โดยใช้ค่าพารามิเตอร์ไม่ศูนย์กลาง (ϕ) ซึ่งเสนอโดย Games, Winkle, and Probert (1972) ดังตารางที่ 2

ตารางที่ 2 ความแตกต่างของความแปรปรวนประชากร

ประชากร	ความแปรปรวน	ความแตกต่างของความแปรปรวน	ϕ
3 กลุ่ม	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.0, \sigma_3^2 = 1.0$	ไม่แตกต่าง	0
	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.2, \sigma_3^2 = 1.4$	น้อย ($0 < \phi < 1.5$)	0.1633
	$\sigma_1^2 = 1.0, \sigma_2^2 = 2.0, \sigma_3^2 = 3.0$		0.8165
	$\sigma_1^2 = 1.0, \sigma_2^2 = 3.0, \sigma_3^2 = 5.0$	ปานกลาง ($1.5 \leq \phi < 3$)	1.6330
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 7.0$		2.4495
	$\sigma_1^2 = 1.0, \sigma_2^2 = 5.0, \sigma_3^2 = 9.0$	มาก ($\phi \geq 3$)	3.2660
$\sigma_1^2 = 1.0, \sigma_2^2 = 6.0, \sigma_3^2 = 11.0$	4.0825		
4 กลุ่ม	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.0, \sigma_3^2 = 1.0, \sigma_4^2 = 1.0$	ไม่แตกต่าง	0
	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.2, \sigma_3^2 = 1.4, \sigma_4^2 = 1.6$	น้อย ($0 < \phi < 1.5$)	0.2236
	$\sigma_1^2 = 1.0, \sigma_2^2 = 2.0, \sigma_3^2 = 3.0, \sigma_4^2 = 4.0$		1.1180
	$\sigma_1^2 = 1.0, \sigma_2^2 = 3.0, \sigma_3^2 = 5.0, \sigma_4^2 = 7.0$	ปานกลาง ($1.5 \leq \phi < 3$)	2.2361
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 6.0, \sigma_4^2 = 8.0$		2.5860
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 7.0, \sigma_4^2 = 10.0$	มาก ($\phi \geq 3$)	3.3541
$\sigma_1^2 = 1.0, \sigma_2^2 = 5.0, \sigma_3^2 = 9.0, \sigma_4^2 = 13.0$	4.4721		

ทำการแปลงข้อมูลแต่ละประชากรให้มีความแปรปรวนตามที่กำหนด โดยใช้สูตร

$$Y = \mu + \sigma X \tag{2}$$

เมื่อ X แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

Y แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ μ (การศึกษาครั้งนี้กำหนด $\mu = 10$) และความแปรปรวนเท่ากับ σ^2

2. คำนวณค่าของตัวสถิติทดสอบ (Test Statistic) และ p-value ของการทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 4 วิธี แล้วเปรียบเทียบกับระดับนัยสำคัญที่กำหนด (0.01 และ 0.05) โดยถ้า p-value น้อยกว่า

หรือเท่ากับระดับนัยสำคัญ จะตัดสินใจปฏิเสธสมมุติฐานว่าง (ปฏิเสธ H_0) และถ้า p-value มากกว่าระดับนัยสำคัญ จะตัดสินใจไม่ปฏิเสธสมมุติฐานว่าง รายละเอียดของแต่ละตัวสถิติทดสอบและวิธีการคำนวณ p-value แสดงดังนี้

2.1 ตัวสถิติทดสอบของบาร์ตเลตต์ (Bartlett's Test: B) (Riansut, 2016)

$$\chi_0^2 = \frac{(n-a)\ln s_p^2 - \sum_{i=1}^a (n_i-1)\ln s_i^2}{1 + \frac{1}{3(a-1)} \left[\sum_{i=1}^a \frac{1}{n_i-1} - \frac{1}{n-a} \right]} \quad (3)$$

เมื่อ $s_i^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$, $s_p^2 = \frac{\sum_{i=1}^a \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{\sum_{i=1}^a (n_i-1)} = \frac{\sum_{i=1}^a (n_i-1)s_i^2}{n-a}$

y_{ij} แทนค่าสังเกตตัวที่ j ของตัวอย่างกลุ่มที่ i , $i = 1, 2, \dots, a$ และ $j = 1, 2, \dots, n_i$, $\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$

a แทนจำนวนประชากร, n_i แทนขนาดตัวอย่างกลุ่มที่ i , $i = 1, 2, \dots, a$, $n = \sum_{i=1}^a n_i$

$$p\text{-value} = P(\chi_{a-1}^2 > \chi_0^2) \quad (4)$$

2.2 ตัวสถิติทดสอบของเลวี (Levene's Test: L) (Riansut, 2016)

$$F_0 = \frac{\left[\sum_{i=1}^a \frac{d_i^2}{n_i} - \frac{d^2}{n} \right] / (a-1)}{\left[\sum_{i=1}^a \sum_{j=1}^{n_i} d_{ij}^2 - \sum_{i=1}^a \frac{d_i^2}{n_i} \right] / (n-a)} \quad (5)$$

เมื่อ $d_{ij} = |y_{ij} - \bar{y}_i|$, $d_i = \sum_{j=1}^{n_i} d_{ij}$ และ $d = \sum_{i=1}^a \sum_{j=1}^{n_i} d_{ij}$

$$p\text{-value} = P(F_{a-1, n-a} > F_0) \quad (6)$$

2.3 ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตี (Brown-Forsythe's Test: BF) (Riansut, 2016) มีสูตรการคำนวณค่าของตัวสถิติทดสอบดังสมการที่ (5) และวิธีการคำนวณ p-value ดังสมการที่ (6) โดยเปลี่ยนแปลงค่า d_{ij} เป็น $d_{ij} = |y_{ij} - \tilde{y}_i|$ เมื่อ \tilde{y}_i แทนค่ามัธยฐานของตัวอย่างกลุ่มที่ i , $i = 1, 2, \dots, a$

2.4 ตัวสถิติทดสอบของโอบรีน (O'Brien's Test: O) (O'Brien, 1981) มีสูตรการคำนวณค่าของตัวสถิติทดสอบดังสมการที่ (5) และวิธีการคำนวณ p-value ดังสมการที่ (6) โดยเปลี่ยนแปลงค่า d_{ij} เป็น

$$d_{ij} = \frac{n_i(n_i-1.5)(y_{ij} - \bar{y}_i)^2 - 0.5s_i^2(n_i-1)}{(n_i-1)(n_i-2)}$$

3. คำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 4 วิธี โดยการนับจำนวนครั้งที่ตัดสินใจปฏิเสธสมมุติฐานว่าง นำมาหารด้วยจำนวนรอบของการทำซ้ำ ซึ่งคือ 1,000 รอบ ดังนี้

- กรณีความแปรปรวนของประชากรไม่แตกต่างกัน ($\phi=0$) จะได้ความน่าจะเป็นของความผิดพลาดแบบที่ 1 นั่นคือ

$$\alpha = \frac{\text{จำนวนครั้งที่ตัดสินใจปฏิเสธ } H_0 \text{ เมื่อ } \sigma_1^2 = \sigma_2^2 = \dots = \sigma_a^2}{1,000} \quad (7)$$

นำค่า α ที่คำนวณได้ไปเปรียบเทียบกับเกณฑ์การทดสอบทวินาม (Seksan, 2008) ดังนี้



- กรณีระดับนัยสำคัญเท่ากับ $\alpha_0 = 0.01$ จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ก็ต่อเมื่อค่า α ที่คำนวณได้จากสมการที่ (7) อยู่ในช่วง

$$\alpha_0 - Z_{0.01/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}} < \alpha < \alpha_0 + Z_{0.01/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}} \tag{8}$$

แทนค่า α_0 ที่กำหนด คือ 0.01, $Z_{0.01/2} = Z_{0.005} = 2.576$, n แทนจำนวนรอบของการทำซ้ำ = 1,000 จะได้

$$0.0019 < \alpha < 0.0181$$

- กรณีระดับนัยสำคัญเท่ากับ $\alpha_0 = 0.05$ จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ก็ต่อเมื่อค่า α ที่คำนวณได้จากสมการที่ (7) อยู่ในช่วงที่ (8) เมื่อแทนค่า α_0 ที่กำหนด คือ 0.05, $Z_{0.05/2} = Z_{0.025} = 1.96$, n แทนจำนวนรอบของการทำซ้ำ = 1,000 จะได้

$$0.0365 < \alpha < 0.0635$$

- กรณีความแปรปรวนของประชากรแตกต่างกัน ($\phi > 0$) จะได้กำลังการทดสอบ นั่นคือ

$$1-\beta = \frac{\text{จำนวนครั้งที่ตัดสินใจปฏิเสธ } H_0 \text{ เมื่อ } \sigma_i^2 \neq \sigma_j^2 \text{ อย่างน้อยหนึ่งค่าของ } i \neq j}{1,000} \tag{9}$$

4. สรุปผลและเสนอแนะว่าตัวสถิติทดสอบสำหรับการทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 4 วิธีวิธีใดมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่างๆ ทั้งกรณีการแจกแจง จำนวนประชากร ขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญที่แตกต่างกัน

ผลการศึกษา

การศึกษาครั้งนี้จะแบ่งผลการวิจัยออกเป็น 2 ส่วน ดังนี้

1. ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด แสดงดังตารางที่ 3 - 7

ตารางที่ 3 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงปกติ

ประชากร	n	ระดับนัยสำคัญ 0.01				ระดับนัยสำคัญ 0.05			
		B	L	BF	OB	B	L	BF	OB
3 กลุ่ม	(10, 10, 10)	0.013	0.011	0.005	0.005	0.053	0.063	0.036*	0.035*
	(30, 30, 30)	0.010	0.007	0.005	0.005	0.039	0.055	0.038	0.040
	(100, 100, 100)	0.004	0.007	0.005	0.007	0.042	0.042	0.036*	0.038
	(4, 7, 10)	0.010	0.017	0.006	0.009	0.061	0.073*	0.036*	0.035*
	(20, 25, 30)	0.013	0.013	0.008	0.009	0.045	0.052	0.040	0.037
	(80, 90, 100)	0.016	0.008	0.007	0.011	0.060	0.053	0.044	0.054
	(10, 10, 10, 10)	0.016	0.019*	0.007	0.007	0.061	0.075*	0.033*	0.041
4 กลุ่ม	(30, 30, 30, 30)	0.009	0.010	0.007	0.010	0.052	0.061	0.050	0.052
	(100, 100, 100, 100)	0.006	0.009	0.008	0.008	0.049	0.052	0.050	0.041
	(4, 7, 10, 13)	0.010	0.020*	0.006	0.009	0.048	0.058	0.022*	0.036*
	(20, 25, 30, 35)	0.013	0.015	0.009	0.009	0.055	0.056	0.040	0.049
	(80, 90, 100, 110)	0.016	0.014	0.014	0.017	0.060	0.065*	0.061	0.063

* แทนไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด



ตารางที่ 4 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ

ประชากร	n	ระดับนัยสำคัญ 0.01				ระดับนัยสำคัญ 0.05			
		B	L	BF	OB	B	L	BF	OB
3 กลุ่ม	(10, 10, 10)	0.141*	0.024*	0.001*	0.006	0.276*	0.125*	0.034*	0.037
	(30, 30, 30)	0.239*	0.033*	0.008	0.009	0.404*	0.133*	0.048	0.038
	(100, 100, 100)	0.313*	0.033*	0.013	0.006	0.465*	0.115*	0.047	0.036*
	(4, 7, 10)	0.076*	0.042*	0.015	0.015	0.202*	0.141*	0.053	0.057
	(20, 25, 30)	0.233*	0.042*	0.008	0.010	0.380*	0.130*	0.041	0.049
	(80, 90, 100)	0.328*	0.039*	0.011	0.006	0.497*	0.133*	0.054	0.035*
4 กลุ่ม	(10, 10, 10, 10)	0.191*	0.063*	0.012	0.018	0.346*	0.152*	0.034*	0.055
	(30, 30, 30, 30)	0.348*	0.043*	0.004	0.011	0.496*	0.137*	0.042	0.040
	(100, 100, 100, 100)	0.394*	0.048*	0.002	0.002	0.562*	0.135*	0.046	0.027*
	(4, 7, 10, 13)	0.143*	0.065*	0.009	0.014	0.298*	0.171*	0.041	0.083*
	(20, 25, 30, 35)	0.315*	0.048*	0.011	0.011	0.482*	0.148*	0.041	0.044
	(80, 90, 100, 110)	0.400*	0.040*	0.003	0.005	0.564*	0.135*	0.041	0.035*

* แทนไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 5 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ

ประชากร	n	ระดับนัยสำคัญ 0.01				ระดับนัยสำคัญ 0.05			
		B	L	BF	OB	B	L	BF	OB
3 กลุ่ม	(10, 10, 10)	0.002	0.015	0.006	0.007	0.024*	0.092*	0.040	0.054
	(30, 30, 30)	0.002	0.018	0.009	0.008	0.009*	0.064*	0.039	0.054
	(100, 100, 100)	0.000*	0.012	0.010	0.006	0.007*	0.075*	0.051	0.053
	(4, 7, 10)	0.007	0.022*	0.008	0.008	0.033*	0.095*	0.035*	0.044
	(20, 25, 30)	0.002	0.018	0.007	0.007	0.012*	0.074*	0.036*	0.053
	(80, 90, 100)	0.000*	0.015	0.009	0.010	0.007*	0.071*	0.049	0.059
4 กลุ่ม	(10, 10, 10, 10)	0.002	0.018	0.007	0.010	0.011*	0.080*	0.033*	0.039
	(30, 30, 30, 30)	0.000*	0.017	0.008	0.007	0.003*	0.067*	0.031*	0.046
	(100, 100, 100, 100)	0.001*	0.016	0.009	0.014	0.009*	0.073*	0.051	0.057
	(4, 7, 10, 13)	0.004	0.029*	0.005	0.016	0.023*	0.108*	0.030*	0.050
	(20, 25, 30, 35)	0.001*	0.008	0.002	0.006	0.004*	0.076*	0.032*	0.051
	(80, 90, 100, 110)	0.000*	0.020*	0.010	0.007	0.006*	0.076*	0.051	0.056

* แทนไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 6 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ

ประชากร	n	ระดับนัยสำคัญ 0.01				ระดับนัยสำคัญ 0.05			
		B	L	BF	OB	B	L	BF	OB
3 กลุ่ม	(10, 10, 10)	0.154*	0.042*	0.013	0.008	0.286*	0.130*	0.040	0.043
	(30, 30, 30)	0.267*	0.044*	0.013	0.006	0.413*	0.140*	0.045	0.050
	(100, 100, 100)	0.353*	0.039*	0.011	0.004	0.502*	0.125*	0.055	0.051
	(4, 7, 10)	0.091*	0.053*	0.017	0.015	0.224*	0.153*	0.041	0.059
	(20, 25, 30)	0.226*	0.037*	0.005	0.011	0.373*	0.133*	0.049	0.050
	(80, 90, 100)	0.314*	0.033*	0.009	0.005	0.479*	0.116*	0.045	0.034*
4 กลุ่ม	(10, 10, 10, 10)	0.177*	0.055*	0.012	0.008	0.346*	0.156*	0.027*	0.050
	(30, 30, 30, 30)	0.327*	0.041*	0.008	0.010	0.481*	0.147*	0.039	0.034*
	(100, 100, 100, 100)	0.414*	0.057*	0.013	0.010	0.567*	0.153*	0.049	0.045
	(4, 7, 10, 13)	0.143*	0.055*	0.007	0.015	0.279*	0.152*	0.035*	0.072*
	(20, 25, 30, 35)	0.326*	0.046*	0.008	0.011	0.490*	0.141*	0.040	0.045
	(80, 90, 100, 110)	0.424*	0.045*	0.009	0.007	0.602*	0.146*	0.047	0.042

* แทนไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด



ตารางที่ 7 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ

ประชากร	n	ระดับนัยสำคัญ 0.01				ระดับนัยสำคัญ 0.05			
		B	L	BF	OB	B	L	BF	OB
3 กลุ่ม	(10, 10, 10)	0.006	0.023*	0.014	0.012	0.031*	0.082*	0.039	0.051
	(30, 30, 30)	0.000*	0.018	0.013	0.010	0.012*	0.055	0.040	0.047
	(100, 100, 100)	0.000*	0.021*	0.013	0.014	0.015*	0.076*	0.061	0.056
	(4, 7, 10)	0.007	0.019*	0.003	0.011	0.026*	0.089*	0.029*	0.037
	(20, 25, 30)	0.003	0.015	0.009	0.009	0.012*	0.077*	0.039	0.047
	(80, 90, 100)	0.000*	0.016	0.006	0.007	0.006*	0.077*	0.056	0.058
4 กลุ่ม	(10, 10, 10, 10)	0.001*	0.020*	0.003	0.009	0.015*	0.091*	0.030*	0.046
	(30, 30, 30, 30)	0.002	0.020*	0.005	0.009	0.012*	0.091*	0.046	0.050
	(100, 100, 100, 100)	0.000*	0.010	0.003	0.005	0.002*	0.059	0.038	0.044
	(4, 7, 10, 13)	0.006	0.031*	0.002	0.013	0.025*	0.104*	0.024*	0.048
	(20, 25, 30, 35)	0.000*	0.018	0.004	0.012	0.009*	0.068*	0.026*	0.048
	(80, 90, 100, 110)	0.001*	0.013	0.006	0.006	0.004*	0.067*	0.039	0.040

* แทนไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 3 - 7 พบว่า เมื่อประชากรมีการแจกแจงปกติ ตัวสถิติทดสอบของบาร์ตเลตต์สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี แต่ตัวสถิติทดสอบของเลวิน ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตี และตัวสถิติทดสอบของโอปรีนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี โดยจะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีเมื่อระดับนัยสำคัญ 0.01 แต่ถ้าระดับนัยสำคัญเพิ่มขึ้นเป็น 0.05 จะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ลดลง เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์และตัวสถิติทดสอบของเลวินไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี แต่ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตีและตัวสถิติทดสอบของโอปรีนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี โดยจะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีเมื่อระดับนัยสำคัญ 0.01 แต่ถ้าระดับนัยสำคัญเพิ่มขึ้นเป็น 0.05 จะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ลดลง เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์และตัวสถิติทดสอบของเลวินสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี แต่ตัวสถิติทดสอบของบราวน์-ฟอร์ลิตีและตัวสถิติทดสอบของโอปรีนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี โดยจะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีเมื่อระดับนัยสำคัญ 0.01 แต่ถ้าระดับนัยสำคัญเพิ่มขึ้นเป็น 0.05 จะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ลดลง

2. ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด จึงมีความเหมาะสมกับการทดสอบภาวะความเท่ากันของความแปรปรวน สรุปได้ดังตารางที่ 8 (สามารถศึกษากำลังการทดสอบได้จากงานวิจัยฉบับสมบูรณ์ของ Riansut (2017))

ตารางที่ 8 ตัวสถิติทดสอบที่เหมาะสมกับการทดสอบภาวะความเท่ากันของความแปรปรวน

ขนาดตัวอย่าง	การแจกแจง				
	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
เล็กและเท่ากัน	B	OB	$B, L^3 / OB^4$	BF	OB
กลางและเท่ากัน	B	BF	$B, L^3 / OB^4$	BF	$B, L^3 / OB^4$
ใหญ่และเท่ากัน	B^1	BF	OB^1	BF	OB^1
	* ²	BF	* ²	BF	* ²
เล็กและไม่เท่ากัน	B	BF	B^3 / OB^4	BF	B^3 / OB^4
กลางและไม่เท่ากัน	B	BF	$B, L^3 / OB^4$	BF	$B, L^3 / OB^4$
ใหญ่และไม่เท่ากัน	B^1	BF	OB^1	BF	OB^1
	* ²	BF	* ²	BF	* ²

* แทนตั้งแต่ 2 ตัวสถิติทดสอบขึ้นไปมีความเหมาะสม

- 1 แทนเมื่อพิจารณาที่ความแปรปรวนของประชากรมีความแตกต่างกันน้อย
- 2 แทนเมื่อพิจารณาที่ความแปรปรวนของประชากรมีความแตกต่างกันปานกลางจนถึงมาก
- 3 แทนเมื่อพิจารณาที่ระดับนัยสำคัญ 0.01
- 4 แทนเมื่อพิจารณาที่ระดับนัยสำคัญ 0.05

จากผลการศึกษาเกี่ยวกับกำลังการทดสอบ พบว่า กำลังการทดสอบจะแปรผันตรงกับจำนวนประชากร ขนาดตัวอย่าง ความแตกต่างของความแปรปรวนประชากร และระดับนัยสำคัญ และจากตารางที่ 8 พบว่า เมื่อประชากรมีการแจกแจงปกติ ทุกขนาดตัวอย่าง ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ตัวสถิติทดสอบของบาร์ตเลตต์มีความเหมาะสมมากที่สุด และถ้าตัวอย่างมีขนาดใหญ่ ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากันที่มีความแปรปรวนของประชากรแตกต่างกันปานกลางจนถึงมาก พบว่า ตัวสถิติทดสอบโดยส่วนใหญ่มีความเหมาะสมไม่แตกต่างกัน แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของบราวน์-ฟอรัลตีมีความเหมาะสมมากที่สุด และเมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์ ตัวสถิติทดสอบของเลวิน และตัวสถิติทดสอบของโอปรีนมีความเหมาะสมขึ้นอยู่กับขนาดตัวอย่างและระดับนัยสำคัญ เช่น เมื่อตัวอย่างมีขนาดเล็กและเท่ากัน สำหรับการแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์และตัวสถิติทดสอบของเลวินมีความเหมาะสมที่ระดับนัยสำคัญ 0.01 แต่ตัวสถิติทดสอบของโอปรีนมีความเหมาะสมที่ระดับนัยสำคัญ 0.05

สรุปผล อภิปรายผลการศึกษา และข้อเสนอแนะ

จากผลการศึกษาเกี่ยวกับความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และกำลังการทดสอบของแต่ละตัวสถิติทดสอบ สามารถวิจารณ์ผลได้ดังนี้

ตัวสถิติทดสอบของบาร์ตเลตต์สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด เมื่อประชากรมีการแจกแจงปกติ และมีกำลังการทดสอบสูงที่สุดในทุกสถานการณ์ที่ศึกษา แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบนี้ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี และเมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี สอดคล้องกับที่ Gene and Hopkins (1996) สรุปไว้ว่า ตัวสถิติทดสอบของบาร์ตเลตต์เป็นตัวสถิติทดสอบที่มีความยุ่งยากในการคำนวณและมีความไวต่อข้อมูลที่ไม่มีการแจกแจงปกติ โดยถ้าข้อมูลมีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบนี้จะมีโอกาสปฏิเสธสมมุติฐานว่างมากกว่าที่ควรจะเป็น จึงส่งผลให้ความน่าจะเป็นของความผิดพลาดแบบที่ 1 สูงเกินขอบเขตที่กำหนด และเมื่อข้อมูลมีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบนี้จะมีโอกาสปฏิเสธสมมุติฐานว่างน้อยกว่าที่ควรจะเป็น จึงส่งผลให้ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ต่ำกว่าขอบเขตที่



กำหนด สอดคล้องกับที่ Srisomboon (2010) สรุปไว้ว่า กรณีประชากรมีการแจกแจงปกติ ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ตัวสถิติทดสอบของบาร์ตเลตต์สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุด และมีกำลังการทดสอบสูงที่สุด และสอดคล้องกับที่ Hatchavanich (2014) สรุปไว้ว่า กรณีประชากรมีการแจกแจงปกติ ตัวสถิติทดสอบของบาร์ตเลตต์เป็นตัวสถิติทดสอบที่ดีที่สุด เพราะไม่ว่าตัวอย่างขนาดเล็กหรือตัวอย่างขนาดใหญ่ก็มีกำลังการทดสอบสูง ดังนั้นจากผลการศึกษาคั้งนี้จึงสรุปได้ว่า เมื่อข้อมูลมีการแจกแจงปกติ ตัวสถิติทดสอบของบาร์ตเลตต์จะมีความเหมาะสมมาก และถ้าข้อมูลมีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบนี้อาจมีความเหมาะสมในบางกรณี แต่ถ้าข้อมูลมีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบนี้จะไม่มีที่เหมาะสมเลย

ตัวสถิติทดสอบของเลวินสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อประชากรมีการแจกแจงปกติ แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบนี้ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี และเมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี อีกทั้งกำลังการทดสอบของตัวสถิติทดสอบนี้ก็ยิ่งต่ำกว่าตัวสถิติทดสอบของบาร์ตเลตต์เป็นส่วนใหญ่ สอดคล้องกับที่ Hatchavanich (2014) สรุปไว้ว่า ตัวสถิติทดสอบของเลวินไม่ใช่ตัวสถิติทดสอบที่ดีที่สุด แต่ยังมีตัวสถิติทดสอบอื่นๆ ที่ให้ผลการทดสอบดีกว่า ทั้งนี้ขึ้นอยู่กับ การแจกแจงของข้อมูล

ตัวสถิติทดสอบของบรวาน์-ฟอร์ลิตีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดที่ระดับนัยสำคัญ 0.01 เกือบทุกสถานการณ์ที่ศึกษา แต่ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ลดลง อีกทั้งเมื่อประชากรมีการแจกแจงปกติ พบว่า โดยส่วนใหญ่กำลังการทดสอบของตัวสถิติทดสอบนี้จะต่ำกว่าตัวสถิติทดสอบของบาร์ตเลตต์และตัวสถิติทดสอบของเลวิน แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงหรือต่ำกว่าปกติ ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดมากกว่าตัวสถิติทดสอบของบาร์ตเลตต์และตัวสถิติทดสอบของเลวิน จึงนับเป็นข้อดีของตัวสถิติทดสอบนี้ ซึ่งอาจเนื่องมาจากสูตรการคำนวณนั้นใช้ คำนวณฐาน แทนค่าเฉลี่ยของตัวอย่างในสูตรการคำนวณของตัวสถิติทดสอบของเลวิน สอดคล้องกับที่ Conover, Johnson, and Johnson (1981) สรุปไว้ว่า ตัวสถิติทดสอบของบรวาน์-ฟอร์ลิตียังคงมีความแกร่งเมื่อประชากรไม่มีการแจกแจงปกติ

ตัวสถิติทดสอบของโอเบิร์นสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดที่ระดับนัยสำคัญ 0.01 ในทุกสถานการณ์ที่ศึกษา แต่ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ลดลง อย่างไรก็ตาม ถึงแม้ว่าตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดมากกว่าตัวสถิติทดสอบของบาร์ตเลตต์ ตัวสถิติทดสอบของเลวิน และตัวสถิติทดสอบของบรวาน์-ฟอร์ลิตี แต่กำลังการทดสอบของตัวสถิติทดสอบของนี้ไม่ได้สูงมากนัก ดังนั้นตัวสถิติทดสอบนี้จึงไม่ใช่ตัวสถิติทดสอบที่ดีที่สุด ผลการศึกษาคั้งนี้ขัดแย้งกับ Chuenteerawong (2001) ที่พบว่า ตัวสถิติทดสอบของโอเบิร์นมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบของเลวินและตัวสถิติทดสอบของบรวาน์-ฟอร์ลิตีเมื่อประชากรไม่มีการแจกแจงปกติ อาจเนื่องมาจากสถานการณ์ที่ศึกษาแตกต่างกัน

จากการวิจารณ์ผลข้างต้น สามารถสรุปได้ว่า เมื่อประชากรมีการแจกแจงปกติ ทุกขนาดตัวอย่าง ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ตัวสถิติทดสอบของบาร์ตเลตต์มีความเหมาะสมมากที่สุด และถ้าตัวอย่างมีขนาดใหญ่ ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากันที่มีความแปรปรวนของประชากรแตกต่างกันปานกลางจนถึงมาก พบว่า ตัวสถิติทดสอบโดยส่วนใหญ่มีความเหมาะสมไม่แตกต่างกัน แต่เมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของบรวาน์-ฟอร์ลิตีมีความเหมาะสมมากที่สุด และเมื่อประชากรมีการแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของบาร์ตเลตต์ ตัวสถิติทดสอบของเลวิน และตัวสถิติทดสอบของโอเบิร์นมีความเหมาะสมขึ้นอยู่กับขนาดตัวอย่างและระดับนัยสำคัญ



ในการวิจัยครั้งต่อไป ควรศึกษาประสิทธิภาพของตัวสถิติทดสอบอื่นๆ เพิ่มเติมในการทดสอบภาวะความเท่ากันของความแปรปรวน เพื่อให้ครอบคลุมตัวสถิติทดสอบให้มากที่สุด เช่น ตัวสถิติทดสอบของคัลลอร์ทซ์ ตัวสถิติทดสอบของมูต ตัวสถิติทดสอบของเลวินแบบถ่วงน้ำหนัก ตัวสถิติทดสอบของจีนี ตัวสถิติทดสอบเลয়ারต์ไคกำลังสอง ตัวสถิติทดสอบของบ็อกซ์ ตัวสถิติทดสอบของแจ๊คไนฟ์ ตัวสถิติทดสอบที่สาม ตัวสถิติทดสอบของเนย์แมน-เพียร์สัน ตัวสถิติทดสอบ F-K: med χ^2 และตัวสถิติทดสอบ F-K: med F เป็นต้น

กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ กองทุนวิจัยมหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2560

เอกสารอ้างอิง

- Chuenteerawong, W. (2001). *A comparison of the power of the tests for homogeneity of variances when populations are not normally distributed*. Bangkok: Thammasat University.
- Conover, W. J., Johnson, M. E., & Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, 23(4), 351-361.
- Games, P. A., Winkle, H. J., & Probert, D. A. (1972). Robust tests for homogeneity of variance. *Educational and Psychological Measurement*, 32(4), 274-285.
- Gene, V. G., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). New Jersey: Allyn and Becon.
- Hatchavanich, D. (2014). A comparison of type I error and power of statistics for homogeneity of variance tests. *Journal of Science Ladkrabang*, 23(1), 17-28.
- Keanthum, A. (1996). *Principles of experimental designs*. Bangkok: Kasetsart University.
- O'Brien, R. G. (1981). A simple test for variance effects in experimental design. *Psychological Bulletin*, 89(3), 570-574.
- Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J., & Mykytka, E. F. (1979). A probability distribution and its uses in fitting data. *Technometrics*, 21(2), 201-214.
- Riansut, W. (2016). *Experimental designs*. Songkhla: Thaksin University.
- Riansut, W. (2017). *A comparison of the efficiency of the test statistics for testing homogeneity of variance*. Research Report. Thaksin University.
- Scheffe, H. (1959). *The analysis of variance* (6th ed.). New York: John Wiley & Sons.
- Seksan, J. (2008). *Effect of failure to meet the assumptions of homogeneity of variances on type I error in two population tests of difference*. Bangkok: Kasetsart University.
- Srisomboon, S. (2010). *A comparison of the power of the test statistics for homogeneity testing of variances*. Bangkok: Kasetsart University.