

การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบไม่อิงพารามิเตอร์
สำหรับการทดสอบภาวะความเท่ากันของความแปรปรวน
A Comparison of the Efficiency of Nonparametric Test Statistics
For Testing Homogeneity of Variances

วารางคณา เรียนสุทธิ์
Warangkhan Riansut

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ อ.ป่าพะยอม จ.พัทลุง 93210
Department of Mathematics and Statistics, Faculty of Science, Thaksin University, Phatthalung, 93210
E-mail: warang27@gmail.com.com

Received: Dec 14, 2017

Revised: Jan 31, 2018

Accepted: Apr 13, 2019

บทคัดย่อ

วัตถุประสงค์ของการศึกษานี้ คือ เพื่อคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 กำลังการทดสอบ และเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบทั้งหมด 2 วิธี ได้แก่ ตัวสถิติทดสอบของคล็อทซ์และตัวสถิติทดสอบของมู้ด รวมถึงเพื่อเสนอแนะตัวสถิติทดสอบที่เหมาะสม จำแนกสถานการณ์โดยกำหนดประชากรที่มีการแจกแจง 5 รูปแบบ กำหนดจำนวนประชากร 3 และ 4 กลุ่ม และศึกษาทั้งกรณีที่ตัวอย่างมีขนาดเท่ากันและไม่เท่ากัน โดยแบ่งเป็นตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ กำหนดความแตกต่างของความแปรปรวนประชากร 3 ระดับ คือ น้อย ปานกลาง และมาก ระดับนัยสำคัญที่ศึกษา คือ 0.01 และ 0.05 จำลองข้อมูลด้วยการเขียนโปรแกรมสำเร็จรูปทางสถิติ กำหนดจำนวนรอบของการทำซ้ำ คือ 1,000 รอบ ผลการศึกษาพบว่า กรณีตัวอย่างขนาดเล็ก ทุกการแจกแจง และตัวอย่างขนาดกลาง การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของมู้ดมีความเหมาะสมมากกว่าตัวสถิติทดสอบของคล็อทซ์ แต่กรณีความแตกต่างของความแปรปรวนประชากรน้อย (ค่าพารามิเตอร์ไม่ศูนย์กลาง ϕ อยู่ในช่วง 0 ถึง 1.5) หรือตัวอย่างขนาดกลางที่มีขนาดเท่ากัน การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ และตัวอย่างขนาดใหญ่ การแจกแจงปกติ การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของคล็อทซ์มีความเหมาะสมมากกว่าตัวสถิติทดสอบของมู้ด สำหรับกรณีอื่นๆ ตัวสถิติทดสอบทั้ง 2 ตัวมีความเหมาะสมไม่แตกต่างกัน

คำสำคัญ: การทดสอบภาวะความเท่ากันของความแปรปรวน ค่าพารามิเตอร์ไม่ศูนย์กลาง ตัวสถิติทดสอบของคล็อทซ์ ตัวสถิติทดสอบของมู้ด

Abstract

The objectives of this study were to calculate the probability of type I error, power of a test, and to compare the efficiency of two statistics tests: Klotz's test and Mood's test, including the recommending the appropriate test statistic. Classification of the population with the distribution 5 format. Study three and four populations, including the sample sizes were equal and unequal. Sample sizes were divided into small, medium, and large size. Defined the difference of the population variances into three levels: low, moderate, and high difference. The significance levels to study were 0.01 and 0.05. The data were simulated by using Statistical programming with 1,000 replicates. Results of the research were as follows. Small sample size case for all distributions and medium sample size for left- or right-skewed with leptokurtic kurtosis distribution, Mood's test was more appropriate than Klotz's test. But, when the difference of the population variances was low (noncentrality parameter

ϕ in the range 0 to 1.5) or the medium and equal sample size case for left- or right-skewed with platykurtic kurtosis distribution and large sample size case for normal distribution, left- or right-skewed with platykurtic kurtosis distribution, Klotz's test was more appropriate than Mood's test. For other cases, two statistic tests were not significantly different.

Keywords: Testing Homogeneity of Variances, Noncentrality Parameter, Klotz's Test, Mood's Test

1. 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 [1] ได้แสดงความคิดเห็นว่า ตามข้อสมมุติเกี่ยวกับการแจกแจงของประชากรซึ่งต้องเป็นการแจกแจงปกตินั้น หากลักษณะของการแจกแจงของประชากรไม่เป็นไปตามข้อสมมุติจะส่งผลให้การประมาณค่าหรือการทดสอบสมมุติฐานเกี่ยวกับค่าความแปรปรวนของประชากรเกิดความผิดพลาดได้ รวมถึงการฝ่าฝืนข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวนจะมีผลอย่างมากต่อการสรุปผลที่ได้จากการใช้ตัวสถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวน โดยเฉพาะเมื่อขนาดตัวอย่างของแต่ละกลุ่มไม่เท่ากัน ดังนั้นก่อนการวิเคราะห์ข้อมูลเกี่ยวกับรวบรวมมาไม่ว่าจะเป็นงานวิจัยในสาขาใดจึงควรจะต้องมีการตรวจสอบข้อสมมุติก่อน และอนันต์ชัย เชื้อนธรรม [2] สรุปไว้ว่า ก่อนที่จะวิเคราะห์ความแปรปรวน ควรต้องมีการทดสอบว่าแต่ละประชากรมีความแปรปรวนเท่ากันหรือไม่ ทั้งนี้เพื่อให้ผลการทดสอบที่ได้เป็นความแตกต่างของค่าเฉลี่ยประชากรจริงๆ หากไม่มีข้อสมมุติข้อนี้ ผลการทดสอบอาจเป็นผลจากความแตกต่างของความแปรปรวนของข้อมูลก็ได้ ซึ่งถ้าหากไม่เป็นไปตามข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวนในกรณีนี้ที่ขนาดตัวอย่างของแต่ละกลุ่มเท่ากันและความแปรปรวนแตกต่างกันเพียงเล็กน้อย ตัวสถิติทดสอบเอฟของการวิเคราะห์ความแปรปรวนยังคงมีความแกร่ง (Robust) ต่อลักษณะที่ไม่เป็นไปตามข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวน แต่ถ้าความแปรปรวนมีความแตกต่างกันค่อนข้างมาก จะมีผลกระทบต่อนัยสำคัญของ การวิเคราะห์ความแปรปรวน ในขณะเดียวกันความคลาดเคลื่อนของการทดลองที่ไม่มีการแจกแจงปกติ เช่น เบ้ไปข้างใดข้างหนึ่ง จะส่งผลทำให้ความแปรปรวนไม่เท่ากัน ซึ่งส่งผลกระทบต่อการมีนัยสำคัญทางสถิติ

ในงานวิจัยทางการแพทย์ การศึกษา การตลาด หรืองานทางด้านจิตวิทยา มักไม่มีการตรวจสอบข้อสมมุติเกี่ยวกับภาวะความเท่ากันของความแปรปรวน จึงอาจจะทำให้ผลสรุปที่ได้จากการวิเคราะห์ความแปรปรวนขัดแย้งกับความจริง รวมถึงค่าระดับนัยสำคัญ (Significance Level) หรือความน่าจะเป็นของความผิดพลาดแบบที่ 1 (Probability of Type I Error) ไม่เป็นไปตามที่กำหนดไว้ และอาจมีกำลังการทดสอบต่ำ อย่างไรก็ตาม การทดสอบภาวะความเท่ากันของความแปรปรวนนั้นมีหลายวิธี ทั้งการทดสอบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ โดยการศึกษาครั้งนี้ต้องการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบไม่อิงพารามิเตอร์ทั้งหมด 2 วิธี ได้แก่ การทดสอบของคล็อทซ์ (Klotz's Test) และการทดสอบของมูด (Mood's Test) ซึ่งเป็นการทดสอบที่มีข้อดี คือ ไม่จำเป็นต้องทราบการแจกแจงของประชากร และเป็นการทดสอบที่ไม่มีปัญหาในกรณีที่ข้อมูลมีความเบ้หรือมีค่าผิดปกติ เช่น มากผิดปกติหรือน้อยผิดปกติ เนื่องจากการทดสอบไม่อิงพารามิเตอร์มักใช้เพียงความถี่ หรือลำดับที่ หรือเครื่องหมายในการวิเคราะห์ข้อมูล ไม่ได้ใช้ค่าของข้อมูลจริง [3] โดยผู้อ่านสามารถศึกษาผลการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบอิงพารามิเตอร์สำหรับการทดสอบภาวะความเท่ากันของความแปรปรวนได้จากอีกบทความหนึ่งของผู้วิจัย [4] การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบในการศึกษาครั้งนี้จะกระทำภายใต้สถานการณ์ต่างๆ ทั้งกรณีการแจกแจงจำนวนประชากร ขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญที่แตกต่างกัน ผู้วิจัยคาดว่าผลการศึกษาที่ได้จะทำให้ทราบความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบภาวะความเท่ากันของ

ความแปรปรวนทั้งหมด 2 วิธี ได้ทราบว่าวิธีใดมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่างๆ เพื่อเป็นแนวทางให้กับนิสิต คณาจารย์ และบุคคลทั่วไป สำหรับใช้ประโยชน์ในการตัดสินใจเลือกใช้ตัวสถิติทดสอบที่เหมาะสมกับการทดสอบภาวะความเท่ากันของความแปรปรวน ซึ่งจะส่งผลต่อการวิจัย การผลิต การพัฒนาครู อาจารย์ และบุคลากรทางการศึกษาให้มีคุณภาพสูง สอดคล้อง และรองรับกับแนวทางการพัฒนาประเทศ ต่อไป

2. วิธีดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัย แสดงดังนี้

1. จำลองข้อมูลของประชากร 3 และ 4 กลุ่ม ด้วยโปรแกรมสำเร็จรูปทางสถิติ กำหนดการแจกแจงความน่าจะเป็นของประชากรที่ศึกษา 5 รูปแบบ ได้แก่ การแจกแจงปกติ โดยใช้คำสั่ง $\text{rand}('normal', 0, 1)$ 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 $\text{rand}('uniform')$ แล้วแปลงข้อมูลตามวิธีการของ Ramberg และคณะ [5] 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 และคณะ [5] โดยค่าของพารามิเตอร์ทั้ง 4 ตัวจะให้ข้อมูลที่มีการแจกแจงเบ้ซ้าย 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)

$$\lambda_1 = 1.354, \lambda_2 = 0.2582, \lambda_3 = 0.5683$$

$$\text{และ } \lambda_4 = 0.0129$$

- การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ (ความเบ้ = 2, ความโด่ง = 12)

$$\lambda_1 = -0.579, \lambda_2 = -0.1423, \lambda_3 = -0.0273$$

$$\text{และ } \lambda_4 = -0.0995$$

- การแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = 0.4, ความโด่ง = 2.2)

$$\lambda_1 = -1.354, \lambda_2 = 0.2582, \lambda_3 = 0.0129$$

$$\text{และ } \lambda_4 = 0.5683$$

กำหนดขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทั้งตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ดังตารางที่ 1

Table 1 Samples Sizes for Three and Four Populations.

Population	Characteristics of the Sample	Samples Sizes		
		Small	Medium	Large
Three Populations	Equal	(10, 10, 10)	(30, 30, 30)	(100, 100, 100)
	Unequal	(4, 7, 10)	(20, 25, 30)	(80, 90, 100)
Four Populations	Equal	(10, 10, 10, 10)	(30, 30, 30, 30)	(100, 100, 100, 100)
	Unequal	(4, 7, 10, 13)	(20, 25, 30, 35)	(80, 90, 100, 110)

กำหนดค่าเฉลี่ยประชากรทุกกลุ่มเท่ากัน คือ เท่ากับ 10 และกำหนดความแตกต่างของความแปรปรวนประชากร โดยใช้ค่าพารามิเตอร์ไม่ศูนย์กลาง (Noncentrality Parameter: ϕ) ซึ่งเสนอโดย Games, Winkle และ Probert [6] ดังตารางที่ 2 ทำการแปลงข้อมูลแต่ละประชากรให้มีความแปรปรวนตามที่กำหนด โดยใช้สูตร

$$Y = \mu + \sigma X$$

(2)

เมื่อ X แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

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

Table 2 Difference of Population Variances.

Population	Variances	Difference of Population Variances	ϕ
Three Populations	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.0, \sigma_3^2 = 1.0$	Not difference	0
	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.2, \sigma_3^2 = 1.4$	Low ($0 < \phi < 1.5$)	0.1633
	$\sigma_1^2 = 1.0, \sigma_2^2 = 2.0, \sigma_3^2 = 3.0$		0.8165
	$\sigma_1^2 = 1.0, \sigma_2^2 = 3.0, \sigma_3^2 = 5.0$	Moderate ($1.5 \leq \phi < 3$)	1.6330
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 7.0$		2.4495
	$\sigma_1^2 = 1.0, \sigma_2^2 = 5.0, \sigma_3^2 = 9.0$	High ($\phi \geq 3$)	3.2660
	$\sigma_1^2 = 1.0, \sigma_2^2 = 6.0, \sigma_3^2 = 11.0$		4.0825
Four Populations	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.0, \sigma_3^2 = 1.0, \sigma_4^2 = 1.0$	Not difference	0
	$\sigma_1^2 = 1.0, \sigma_2^2 = 1.2, \sigma_3^2 = 1.4, \sigma_4^2 = 1.6$	Low ($0 < \phi < 1.5$)	0.2236
	$\sigma_1^2 = 1.0, \sigma_2^2 = 2.0, \sigma_3^2 = 3.0, \sigma_4^2 = 4.0$		1.1180
	$\sigma_1^2 = 1.0, \sigma_2^2 = 3.0, \sigma_3^2 = 5.0, \sigma_4^2 = 7.0$	Moderate ($1.5 \leq \phi < 3$)	2.2361
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 6.0, \sigma_4^2 = 8.0$		2.5860
	$\sigma_1^2 = 1.0, \sigma_2^2 = 4.0, \sigma_3^2 = 7.0, \sigma_4^2 = 10.0$	High ($\phi \geq 3$)	3.3541
	$\sigma_1^2 = 1.0, \sigma_2^2 = 5.0, \sigma_3^2 = 9.0, \sigma_4^2 = 13.0$		4.4721

2. ค่าของค่าของตัวสถิติทดสอบ (Test Statistic) และ p-value ของการทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 2 วิธี เพื่อเปรียบเทียบกับระดับนัยสำคัญที่กำหนด (0.01 และ 0.05) โดยถ้า p-value น้อยกว่าหรือเท่ากับระดับนัยสำคัญ จะตัดสินใจปฏิเสธสมมติฐานว่าง (ปฏิเสธ H_0) และถ้า p-value มากกว่าระดับนัยสำคัญ จะตัดสินใจไม่ปฏิเสธสมมติฐานว่าง รายละเอียดของแต่ละตัวสถิติทดสอบและวิธีการคำนวณ p-value แสดงดังนี้

2.1 ตัวสถิติทดสอบของคล็อทซ์ มีสูตรในการคำนวณดังนี้ [7]

$$\chi_0^2 = \frac{1}{s^2} \sum_{i=1}^a \frac{(A_i - n_i \bar{A})^2}{n_i} \quad (3)$$

เมื่อ A_{ij} แทนคะแนนคล็อทซ์ (Klotz Score) มีค่า

$$\text{เท่ากับ } \left[\Phi^{-1} \left(\frac{R(y_{ij})}{n+1} \right) \right]^2$$

y_{ij} แทนค่าสังเกตตัวที่ j ของตัวอย่างกลุ่มที่ i ,

$i = 1, 2, \dots, a$ และ $j = 1, 2, \dots, n_i$

$R(y_{ij})$ แทนลำดับที่ของ y_{ij} ซึ่งได้จากการเรียงลำดับค่าสังเกตจากตัวอย่างทุกกลุ่มปนกันจากน้อยไปมาก

Φ^{-1} แทนฟังก์ชันควอนไทล์ปรกติ (Normal Quantile Function) เช่น $\Phi^{-1}(0.5) = 0$

$$A_i = \sum_{j=1}^{n_i} A_{ij}, \quad \bar{A} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^{n_i} A_{ij}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^a \sum_{j=1}^{n_i} (A_{ij} - \bar{A})^2$$

a แทนจำนวนประชากร

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

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

(4)

2.2 ตัวสถิติทดสอบของมู้ด มีสูตรการคำนวณค่าของตัวสถิติทดสอบดังสมการที่ (3) และวิธีการคำนวณ p-value ดังสมการที่ (4) โดยเปลี่ยนแปลงเป็นคะแนนมู้ด (Mood Score) หรือ A_{ij} ดังนี้ [7]

$$\left[R(y_{ij}) - \frac{n+1}{2} \right]^2$$

3. คำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 2 วิธี โดยการนับจำนวนครั้งที่ตัดสินใจปฏิเสธสมมติฐานว่างนำมาหารด้วยจำนวนรอบของการทำซ้ำ ซึ่งคือ 1,000 รอบ ดังนี้

● กรณีความแปรปรวนของประชากรไม่แตกต่างกัน ($\phi=0$) จะได้ว่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 (α) การศึกษาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ดำเนินการโดยนำค่า α ที่คำนวณได้ไปเปรียบเทียบกับเกณฑ์การทดสอบทวินาม [8] ดังนี้

- กรณีระดับนัยสำคัญเท่ากับ $\alpha_0 =$

0.01 จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ก็ต่อเมื่อค่า α ที่คำนวณได้อยู่ในช่วง

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

(5)

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 ที่คำนวณได้อยู่ในช่วงที่ (5) 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 $(1-\beta)$

4. สรุปผลและเสนอแนะว่าตัวสถิติทดสอบสำหรับการทดสอบภาวะความเท่ากันของความแปรปรวนทั้งหมด 2 วิธี วิธีใดมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุดภายใต้สถานการณ์ต่างๆ ทั้งกรณีการแจกแจง จำนวนประชากร ขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญที่แตกต่างกัน

3. ผลการวิจัย

การศึกษาครั้งนี้จะแบ่งผลการวิจัยออกเป็น 2 ส่วน ดังนี้

1. ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของแต่ละตัวสถิติทดสอบ สรุปได้ดังตารางที่ 3 – 4

Table 3 Ability to Control the Probability of Type I Error from Klotz's Test.

Population	Samples Sizes	Significance Level = 0.01					Significance Level = 0.05				
		Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic	Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic
Three Populations	(10, 10, 10)	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
	(30, 30, 30)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(100, 100, 100)	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
	(4, 7, 10)	✓	✓	✓	✓	✓	✓	×	×	✓	×
	(20, 25, 30)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(80, 90, 100)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Four Populations	(10, 10, 10, 10)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(30, 30, 30, 30)	✓	✓	✓	✓	✓	✓	×	✓	✓	✓
	(100, 100, 100, 100)	✓	✓	✓	✓	✓	✓	✓	✓	✓	×
	(4, 7, 10, 13)	✓	✓	✓	✓	✓	×	✓	✓	×	✓
	(20, 25, 30, 35)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(80, 90, 100, 110)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ Can be controlled.

×

จากตารางที่ 3 พบว่า ที่ระดับนัยสำคัญ 0.01 ทุกการแจกแจง ตัวสถิติทดสอบของคล็อทซ์สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ขนาดตัวอย่างเท่ากันและไม่เท่ากัน แต่ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี

จากตารางที่ 4 พบว่า ที่ระดับนัยสำคัญ 0.01 ทุกการแจกแจง ตัวสถิติทดสอบของมู้ดสามารถ

ควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดทุกกรณี ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ขนาดตัวอย่างเท่ากันและไม่เท่ากัน แต่ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี โดยตัวสถิติทดสอบของมู้ดสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีกว่าตัวสถิติทดสอบของคล็อทซ์

Table 4 Ability to Control the Probability of Type I Error from Mood's Test.

Population	Samples Sizes	Significance Level = 0.01						Significance Level = 0.05					
		Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic	Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic	Normal	Left-Skewed, Leptokurtic
Three Populations	(10, 10, 10)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(30, 30, 30)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(100, 100, 100)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(4, 7, 10)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×
	(20, 25, 30)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(80, 90, 100)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Four Populations	(10, 10, 10, 10)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(30, 30, 30, 30)	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
	(100, 100, 100, 100)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(4, 7, 10, 13)	✓	✓	✓	✓	✓	×	✓	×	×	×	✓	✓
	(20, 25, 30, 35)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	(80, 90, 100, 110)	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓

✓ Can be controlled.

× Cannot be controlled.

2. ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด จึงมีความ

เหมาะสมกับการทดสอบภาวะความเท่ากันของความแปรปรวน สรุปได้ดังตารางที่ 5

Table 5 The Appropriate Test Statistic for Testing Homogeneity of Variances

Sample Sizes	Distributions				
	Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic
Small and Equal	M	M	M	M	M
Medium and Equal	K, M	M	K	M	K
Large and Equal	K ¹	K, M ¹	K ¹	K, M ¹	K ¹
	K, M ²	K, M ²	K, M ²	K, M ²	K, M ²
Small and Unequal	M	M	M	M	M

Table 5 The Appropriate Test Statistic for Testing Homogeneity of Variances (Continue)

Sample Sizes	Distributions				
	Normal	Left-Skewed, Leptokurtic	Left-Skewed, Platykurtic	Right-Skewed, Leptokurtic	Right-skewed, platykurtic
Large and Unequal	K ¹	K, M ¹	K ¹	K, M ¹	K ¹
	K, M ²	K, M ²	K, M ²	K, M ²	K, M ²

K: Klotz's Test

M: Mood's Test

1: Low Difference of the Population Variances

2: Moderate to High Difference of the Population Variances

จากตารางที่ 5 พบว่า กรณีตัวอย่างขนาดเล็ก ทุกการแจกแจง และตัวอย่างขนาดกลาง การแจกแจง เบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของมู้ดมีความเหมาะสมมากกว่าตัวสถิติทดสอบของคล็อทซ์ แต่กรณีตัวอย่างขนาดกลางที่มีขนาดเท่ากัน การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ และตัวอย่างขนาดใหญ่ การแจกแจงปกติ การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของคล็อทซ์มีความเหมาะสมมากกว่าตัวสถิติทดสอบของมู้ด สำหรับกรณีอื่นๆ ตัวสถิติทดสอบทั้ง 2 ตัวมีความเหมาะสมไม่แตกต่างกัน อีกทั้งการศึกษาค้างนี้ยังพบว่า กำลังการทดสอบจะแปรผันตรงกับจำนวนประชากร ขนาดตัวอย่าง ความแตกต่างของความแปรปรวนประชากร และระดับนัยสำคัญ

4. สรุปและข้อเสนอแนะ

จากผลการศึกษาเกี่ยวกับความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และกำลังการทดสอบของแต่ละตัวสถิติทดสอบ พบว่า ตัวสถิติทดสอบของคล็อทซ์และตัวสถิติทดสอบของมู้ดสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดที่ระดับนัยสำคัญ 0.01 ในทุกสถานการณ์ที่ศึกษา แต่ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบทั้ง 2 ตัวนี้จะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี โดยตัวสถิติทดสอบของมู้ดสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีกว่าตัวสถิติทดสอบของคล็อทซ์ กำลังการทดสอบของตัวสถิติทดสอบของคล็อทซ์และตัวสถิติทดสอบของมู้ดจะแปรผันตรงกับ

จำนวนประชากร ขนาดตัวอย่าง ความแตกต่างของความแปรปรวนประชากร และระดับนัยสำคัญ สำหรับกรณีตัวอย่างขนาดเล็ก ทุกการแจกแจง และตัวอย่างขนาดกลาง การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งสูงกว่าปกติ ตัวสถิติทดสอบของมู้ดมีความเหมาะสมมากกว่าตัวสถิติทดสอบของคล็อทซ์ แต่กรณีตัวอย่างขนาดกลางที่มีขนาดเท่ากัน การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ และตัวอย่างขนาดใหญ่ การแจกแจงปกติ การแจกแจงเบ้ซ้ายหรือเบ้ขวาที่มีความโด่งต่ำกว่าปกติ ตัวสถิติทดสอบของคล็อทซ์มีความเหมาะสมมากกว่าตัวสถิติทดสอบของมู้ด ขณะที่กรณีอื่นๆ ตัวสถิติทดสอบทั้ง 2 ตัวมีความเหมาะสมไม่แตกต่างกัน อย่างไรก็ตาม การศึกษาค้างนี้ได้ศึกษาเพียงประสิทธิภาพของการทดสอบไม่อิงพารามิเตอร์เท่านั้น ดังนั้นการศึกษาค้างต่อไป ควรนำการทดสอบอิงพารามิเตอร์มาเปรียบเทียบกับประสิทธิภาพด้วย เช่น การทดสอบทีสาม การทดสอบของเนย์แมน-เพียร์สัน การทดสอบเลয়ারด์ไคกำลังสอง การทดสอบของเลวิน การทดสอบของโอบริน [9] การทดสอบของเลวินแบบถ่วงน้ำหนัก [10] การทดสอบของบาร์ตเลตต์ การทดสอบของบ็อกซ์ การทดสอบของแจ๊คไนฟ์ [11] การทดสอบของเลห์แมน การทดสอบของบราวน์-ฟอร์ลิต [12]

5. กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ กองทุนวิจัยมหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2560

6. Reference

- [1] Scheffe, H. 1959. **The Analysis of zariance.** 6th Edition. New York: Jonh Wiley & Sons.
- [2] Keanthum, A. 1996. **Principles of Experimental Designs.** Bangkok: Kasetsart University.
- [3] National Statistical Office. 2015. **The Use of a Statistical Test.** <http://service.nso.go.th/nso/nsopublish/know/estat04.pdf>. 21 December.
- [4] Riansut, W. 2018. A Comparison of the Efficiency of the Parametric Test Statistics for Testing Homogeneity of Variances. **Naresuan University Journal Science and Technology** 26(3): In Press.
- [5] Ramberg, J.S., Tadikamalla, P.R., Dudewicz, E.J. and Mykytka, E.F. 1979. A Probability Distribution and Its Uses in Fitting Data. **Technometrics** 21(2): 201-214.
- [6] Games, P.A., Winkle, H.J. and Probert, D.A. 1972. Robust Tests for Homogeneity of Variance. **Educational and Psychological Measurement** 32(4): 274-285.
- [7] SAS Help. 2016. **Chapter 47 The NPAR1WAY Procedure.** <http://www.okstate.edu/sas/v8/saspdf/stat/chap47.pdf>. 31 August.
- [8] Seksan, J. 2008. **Effect of Failure to Meet The Assumptions of Homogeneity of Variances on Type I Error in Two Population Tests of Difference.** Master's Thesis: Kasetsart University.
- [9] Thukaeow, P. 2009. **Comparison on Homogeneity Test of Variances.** Master's Thesis: Kasetsart University.
- [10] Pornsawad, T. 2002. **A Comparative Study of the Efficiency Tests Statistic for Test of Homogeneity of Variances.** Master's Thesis: Thammasat University.
- [11] Chandee, P. 2006. **Robust Test Statistics for Homogeneity of Population Variances.** Master's Thesis: Kasetsart University.
- [12] Sittisombat, S. 2007. Comparison of the Power of the Test **for Homogeneity of Variances.** Master's Thesis: Chiang Mai University.