

การเปรียบเทียบประสิทธิภาพสำหรับการทดสอบความเท่ากันของความแปรปรวน ภายใต้การแจกแจงปกติแบบผสม

A Comparison of Efficiency for Homogeneity of Variance Tests under Mixture Normal Distribution

วนิดา พงษ์ศักดิ์ชาติ* และ ณัฐวัตร วันทอง

Vanida Pongsakchat* and Natthawat Wanthong

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ประเทศไทย

Department of Mathematics, Faculty of Science, Burapha University, Thailand

Received: 25 February 2026, Received in revised form: 7 May 2026, Accepted: 8 May 2026

Available online: 15 May 2026

บทคัดย่อ

วัตถุประสงค์และที่มา: ในการวิเคราะห์ข้อมูลเชิงสถิติ ข้อสรุปที่ได้จากการทดสอบสมมติฐานจะมีความเที่ยงตรงเชื่อถือได้ ขึ้นอยู่กับข้อสมมุติที่สำคัญหลายประการ หนึ่งในข้อสมมุติที่สำคัญที่สุดคือ ความเท่ากันของความแปรปรวน ซึ่งกำหนดให้ความแปรปรวนของข้อมูลในทุกประชากรที่นำมาศึกษาต้องมีค่าเท่ากัน ข้อสมมุตินี้เป็นเงื่อนไขที่จำเป็นสำหรับการทดสอบทางสถิติที่ใช้กันอย่างแพร่หลาย เช่น การทดสอบที สำหรับการเปรียบเทียบค่าเฉลี่ยของประชากร 2 กลุ่ม และการทดสอบเอฟ สำหรับการเปรียบเทียบค่าเฉลี่ยของประชากรตั้งแต่ 3 กลุ่มขึ้นไป นักสถิติได้พัฒนาวิธีการทดสอบเพื่อตรวจสอบข้อสมมุติเหล่านี้หลายวิธี วิธีการดั้งเดิมอย่างการทดสอบบาร์ตเลตต์ เป็นที่ทราบกันดีว่าการทดสอบนี้มีประสิทธิภาพสูงเมื่อข้อมูลมีการแจกแจงปกติ แต่จะไวต่อการละเมิดการแจกแจงปกติอย่างมาก ด้วยเหตุนี้จึงมีการนำเสนอการทดสอบอื่น ๆ ที่มีความแข็งแกร่งมากกว่า เช่น การทดสอบเลวิน ทั้งในรูปแบบที่อิงตามค่าเฉลี่ยและค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen ซึ่งแต่ละวิธีมีคุณลักษณะและประสิทธิภาพที่แตกต่างกันไปตามลักษณะของข้อมูล เช่น ความเบ้ ความโด่ง และขนาดตัวอย่าง ในปัจจุบัน ข้อมูลจากสถานการณ์จริงมักมีความซับซ้อนมากกว่าการแจกแจงความน่าจะเป็นแบบเดี่ยว โดยมักปรากฏในรูปแบบการแจกแจงแบบผสม การแจกแจงปกติแบบผสมเป็นการแจกแจงความน่าจะเป็นที่พบได้บ่อยมากในงานวิจัยเชิงประจักษ์ เช่น ในตลาดการเงิน การแพทย์ และ ระบาดวิทยา การใช้แบบจำลองปกติมาตรฐานจึงไม่เพียงพอในการอธิบายข้อมูลเหล่านี้ การศึกษาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนภายใต้การแจกแจงแบบผสมจึงมีความจำเป็นอย่างยิ่ง งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวน 6 วิธี ได้แก่ การทดสอบบาร์ตเลตต์ การทดสอบเลวินที่อิงตามค่าเฉลี่ย การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien การทดสอบ Fligner-Killeen โดยมุ่งเน้นศึกษาในประชากร 3 กลุ่ม ภายใต้การแจกแจงปกติแบบผสม 2 องค์ประกอบ และพิจารณาจากเกณฑ์ความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบผ่านสถานการณ์จำลองที่หลากหลาย

วิธีดำเนินการวิจัย: ข้อมูล 3 กลุ่ม ถูกสร้างขึ้นจากการแจกแจงปกติแบบผสม 2 องค์ประกอบ โดยมีสัดส่วนการผสม (p) คือ $p=0.5$ และ $p=0.8$ โดยกำหนดให้ค่าเฉลี่ยของประชากรทั้ง 3 กลุ่มเท่ากับ 8.8 และกำหนดความแปรปรวนเริ่มต้นที่ 5 10 และ 20 กำหนดให้มีขนาดตัวอย่างเท่ากันในทุกกลุ่ม ได้แก่ (10,10,10) (20,20,20) (40,40,40) (60,60,60) (80,80,80) และ (100,100,100) และกำหนดอัตราส่วนของความแปรปรวนเป็น 1:1:1 สำหรับการประเมินความผิดพลาดแบบที่ 1 และ 1:1:2 1:1:4 และ 1:2:4 สำหรับการประเมินกำลังการทดสอบ โดยในการพิจารณาความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ใช้เกณฑ์ความคลาดเคลื่อนของ Bradley ที่ระดับนัยสำคัญ 0.05 โดยการทดสอบที่ถือว่าสามารถควบคุมความผิดพลาดแบบที่ 1 ได้จะต้องมีค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ในช่วง $[0.025, 0.075]$ ในการพิจารณากำลังการทดสอบจะพิจารณาเฉพาะการทดสอบที่ผ่านเกณฑ์ของ Bradley เท่านั้น โดยการทดสอบที่มีประสิทธิภาพต้องมีค่าประมาณกำลังการทดสอบไม่น้อยกว่า 0.80 และวิธีที่มีค่านี้สูงสุดจะถือว่าเป็นวิธีที่มีประสิทธิภาพมากที่สุด

ผลการวิจัย: จากการศึกษาแสดงให้เห็นว่าประสิทธิภาพของการทดสอบทั้ง 6 วิธี มีความแตกต่างกันขึ้นอยู่กับการสัดส่วนการผสมและขนาดตัวอย่าง เมื่อพิจารณาความสามารถในการควบคุมความผิดพลาดแบบที่ 1 กรณีที่ $p=0.8$ การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen มีความแกร่งโดยสามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกขนาดตัวอย่าง ขณะที่การทดสอบบาร์ตเลตต์ต้องใช้ขนาดตัวอย่างตั้งแต่ 60 ขึ้นไปจึงจะควบคุมได้ ส่วนการทดสอบเลวินที่อิงตามค่าเฉลี่ยไม่สามารถควบคุมความผิดพลาดในทุกสถานการณ์ ในกรณี $p=0.5$ การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ และการทดสอบ O'Brien สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกขนาดตัวอย่าง การทดสอบเลวินที่อิงตามค่าเฉลี่ยเริ่มควบคุมได้เมื่อ $n \geq 20$ อย่างไรก็ตาม การทดสอบบาร์ตเลตต์และการทดสอบ Fligner-Killeen ไม่สามารถควบคุมความผิดพลาดได้ภายใต้สัดส่วนการผสมนี้ ในการพิจารณาค่าประมาณกำลังการทดสอบ จะพิจารณาเฉพาะการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ พบว่าค่าประมาณกำลังการทดสอบจะเพิ่มขึ้นตามขนาดตัวอย่างและความแตกต่างของความแปรปรวน โดยเมื่อ $p=0.8$: การทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ทั้ง 4 วิธี ต้องการขนาดตัวอย่าง $n \geq 60$ สำหรับอัตราส่วน 1:1:2 และ $n \geq 20$ สำหรับอัตราส่วน 1:1:4 เพื่อให้ได้กำลังการทดสอบสูงกว่า 0.8 โดยการทดสอบบาร์ตเลตต์จะให้กำลังการทดสอบสูงสุดในเงื่อนไขที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ และเมื่อ $p=0.5$: การทดสอบ O'Brien โดดเด่นที่สุด โดยให้กำลังการทดสอบสูงกว่า 0.8 เมื่อ $n \geq 60$ ในอัตราส่วน 1:1:2 และ $n \geq 20$ ในอัตราส่วน 1:1:4 ซึ่งมีประสิทธิภาพเหนือกว่าการทดสอบเลวินในทุกรูปแบบ

สรุปผลการวิจัย: ผลจากการศึกษานี้สรุปได้ว่า 1. การทดสอบบาร์ตเลตต์: แม้จะมีประสิทธิภาพสูงในบางเงื่อนไข ($p=0.8$ และขนาดตัวอย่างใหญ่) แต่ไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ เมื่อข้อมูลมีการแจกแจงแบบผสมที่มีสัดส่วนการผสมเท่ากัน ทำให้มีความเสี่ยงสูงในการนำไปใช้กับข้อมูลที่ซับซ้อน 2. การทดสอบเลวินที่อิงตามค่าเฉลี่ย: มีความอ่อนไหวต่อสัดส่วนการผสมอย่างมาก และจะใช้งานได้เฉพาะเมื่อการแจกแจงแบบผสมมีความสมดุลและขนาดตัวอย่างปานกลางขึ้นไปเท่านั้น 3. การทดสอบ Fligner-Killeen: มีประสิทธิภาพดีเฉพาะในกรณีที่สัดส่วนการผสม

$p = 0.8$ 4.การทดสอบ O'Brien : เป็นวิธีที่มีประสิทธิภาพที่สุด โดยไม่เพียงแต่ควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกเงื่อนไข แต่ยังให้กำลังการทดสอบสูงสุดในเกือบทุกสถานการณ์ โดยเฉพาะในกรณีที่มีขนาดตัวอย่างน้อยถึงปานกลาง

คำสำคัญ: การทดสอบความเท่ากันของความแปรปรวน; การแจกแจงปกติแบบผสม; ความผิดพลาดแบบที่ 1 ; กำลังการทดสอบ

Abstract

Background and Objectives: In statistical data analysis, the validity and reliability of conclusions derived from hypothesis testing depend on several crucial assumptions. One of the most important assumptions is the Homogeneity of Variance, which requires that the variance of the data in every studied population must be equal. This assumption is a necessary condition for widely used statistical tests, such as the t-test for comparing the means of two populations and F-test for comparing the means of three or more populations. Statisticians have developed several testing methods to verify this assumption. Traditional methods like Bartlett's test are well-known for having high efficiency when the data follow a normal distribution but are extremely sensitive to violations of normality. Consequently, other tests with greater robustness have been proposed, such as Levene's test (in both mean-based and median-based forms), the Non-parametric Levene's test, O'Brien's test, and the Fligner-Killeen test, each of which has different characteristics and efficiency depending on the nature of the data, such as skewness, kurtosis, and sample size. Currently, real-world data are often more complex than a single probability distribution, frequently appearing in the form of a mixture distribution. The mixture normal distribution is a very common probability distribution found in empirical research, such as in financial markets, medicine, and epidemiology. Using standard normal models is therefore insufficient for explaining these data. Studying the efficiency of homogeneity of variance tests under mixture distributions is critically necessary. This research, therefore, aims to study and compare the efficiency of six homogeneity of variance tests: Bartlett's test, Mean-based Levene's test, Median-based Levene's test, Non-parametric Levene's test, O'Brien's test, and the Fligner-Killeen test, focusing on three population groups under a two-component mixture normal distribution, and considering the Type I error probability and statistical power criteria across various simulated situations.

Methodology: Data for the three groups were generated from a two-component mixture normal distribution with mixing proportions (p) were $p = 0.5$ and $p = 0.8$. The mean for all three groups was set to 8.8 with initial variances specified at 5, 10, and 20. Equal sample sizes were assigned to every group as (10,10,10), (20,20,20), (40,40,40), (60,60,60), (80,80,80), and (100,100,100). The variance ratios were set at 1:1:1, 1:1:2, 1:1:4 and 1:2:4. To consider the ability to control Type I error, Bradley's liberal criterion was used at a significance level of 0.05. A test considered capable of controlling type I error must have an estimated type I

error probability within the range of [0.025, 0.075]. In considering statistical power, only tests that passed Bradley's criterion were considered; an efficient test must have an estimated statistical power of not less than 0.80, and the method with the highest value is considered the most efficient.

Main Results: The study demonstrates that the efficiency of all six testing methods differs depending on the mixing proportion and sample size. When considering the ability to control type I error: in the case where $p = 0.8$, the Median-based Levene's test, Non-parametric Levene's test, O'Brien's test, and Fligner-Killeen test were robust, being able to control Type I error across all sample sizes, while Bartlett's test required a sample size of 60 or more to maintain control. The Mean-based Levene's test could not control the error in any situation. In the case where $p = 0.5$, the Median-based Levene's test, Non-parametric Levene's test, and O'Brien's test could control Type I error across all sample sizes. The Mean-based Levene's test began to maintain control when $n \geq 20$. However, Bartlett's test and the Fligner-Killeen test could not control the error under this mixing proportion. In considering the estimated statistical power, only tests that could control Type I error were examined. It was found that the estimated statistical power increases according to sample size and the difference in variances. At $p = 0.8$: All four methods capable of controlling Type I error required a sample size of $n \geq 60$ for the 1:1:2 ratio and $n \geq 20$ for the 1:1:4 ratio to achieve statistical power higher than 0.8. Bartlett's test provided the highest statistical power in conditions where it could control Type I error. At $p = 0.5$, O'Brien's test was the most prominent, providing statistical power higher than 0.8 when $n \geq 60$ at a 1:1:2 ratio and $n \geq 20$ at a 1:1:4 ratio, which outperformed all forms of Levene's test.

Conclusions : The results of this study: 1. Bartlett's Test: Although highly efficient in some conditions ($p = 0.8$ and large sample size), it cannot control Type I error when the data have a mixture distribution with equal mixing proportions, making it a high risk for application with complex data. 2. Mean-based Levene's Test: Highly sensitive to mixing proportions and only applicable when the mixture distribution is balanced and the sample size is moderate or larger. 3. Fligner-Killeen Test: Performs well only in cases where the mixing proportion $p = 0.8$. 4. O'Brien's Test: This method is identified as the most suitable and efficient overall. It not only successfully controlled the Type I error under all studied conditions but also provided the highest statistical power in nearly every scenario, particularly with small to moderate sample sizes.

Keywords: Homogeneity of variance tests; Mixture normal distribution; Type I error; Power of the test

*Corresponding author. E-mail : vanida@buu.ac.th

Introduction

ในการวิเคราะห์ข้อมูลด้วยวิธีเชิงสถิติ ข้อสรุปที่ได้จากการทดสอบสมมติฐานจะมีความเที่ยงตรงเชื่อถือได้ ขึ้นอยู่กับข้อสมมุติที่สำคัญหลายประการ หนึ่งในข้อสมมุติที่สำคัญที่สุดคือ ความเท่ากันของความแปรปรวน (Homogeneity of Variance หรือ Homoscedasticity) ซึ่งกำหนดให้ความแปรปรวนของข้อมูลในทุกประชากรที่นำมาศึกษาต้องมีค่าเท่ากัน ข้อสมมุตินี้เป็นเงื่อนไขที่จำเป็นสำหรับการทดสอบทางสถิติที่ใช้กันอย่างแพร่หลาย เช่น การทดสอบที (t-test) สำหรับการเปรียบเทียบค่าเฉลี่ยของประชากร 2 กลุ่ม และ การวิเคราะห์ความแปรปรวน (ANOVA) หรือการทดสอบเอฟ (F-test) สำหรับการเปรียบเทียบค่าเฉลี่ยของประชากรตั้งแต่ 3 กลุ่มขึ้นไป เมื่อข้อมูลละเมิดข้อสมมุติเรื่องความเท่ากันของความแปรปรวน ค่าพี (p-value) ที่คำนวณได้อาจคลาดเคลื่อน ซึ่งนำไปสู่การเพิ่มขึ้นของความน่าจะเป็นของความผิดพลาดแบบที่ 1 และอาจส่งผลให้กำลังการทดสอบลดลงอย่างมีนัยสำคัญ ซึ่งทำให้ไม่สามารถตรวจพบความแตกต่างที่แท้จริงของประชากรได้ ซึ่งวิธีเชิงสถิติที่ใช้ในการตรวจสอบข้อสมมุตินี้มีอยู่ด้วยกันหลายวิธี วิธีการดั้งเดิมอย่างการทดสอบบาร์ตเลตต์ (Bartlett's test) ซึ่งเป็นที่ทราบกันดีว่าการทดสอบนี้มีประสิทธิภาพสูงเมื่อข้อมูลมีการแจกแจงปกติ แต่จะไวต่อการละเมิดการแจกแจงปกติอย่างมาก ด้วยเหตุนี้จึงมีการนำเสนอการทดสอบวิธีอื่น ๆ ที่มีความแข็งแกร่งมากกว่า เช่น การทดสอบเลวินที่อิงตามค่าเฉลี่ย (Mean-based Levene's test) การทดสอบเลวินที่อิงตามค่ามัธยฐาน (Median-based Levene's test) การทดสอบเลวินแบบไม่อิงพารามิเตอร์ (Non-parametric Levene's test) การทดสอบ O'Brien (O'Brien test) และการทดสอบ Fligner-Killeen (Fligner-Killeen test) อย่างไรก็ตาม การทดสอบเหล่านี้มีข้อจำกัดที่แตกต่างกัน เช่น การทดสอบบางชนิดจะมีประสิทธิภาพเมื่อข้อมูลที่มีการแจกแจงปกติ (Normal distribution) และการทดสอบบางชนิดมีประสิทธิภาพเมื่อข้อมูลมีการแจกแจงไม่สมมาตร (Asymmetric distribution) จากการศึกษาที่ผ่านมาพบว่าการทดสอบบาร์ตเลตต์เหมาะกับข้อมูลที่มีการแจกแจงปกติ (Levene, 1960; Brown & Forsythe, 1974; Arsham & Lovric, 2011; Sharma & Kibria, 2013; Parra-Frutos, 2013; Rayat, 2018) ในขณะที่การทดสอบเลวินที่อิงตามค่ามัธยฐานมีความเหมาะสมเมื่อข้อมูลไม่มีการแจกแจงปกติ (Levene, 1960; Brown & Forsythe, 1974; Vorapongsathorn *et al.*, 2004; Nordstokke & Zumbo, 2010; Sharma & Kibria, 2013; Parra-Frutos, 2013; Chang *et al.*, 2017) สำหรับการทดสอบ O'Brien เป็นวิธีที่เหมาะสมกับข้อมูลที่ไม่มีการแจกแจงปกติ การแจกแจงแบบพลาติเคอร์ติก (Platykurtic distribution) และข้อมูลมีขนาดไม่เท่ากันในแต่ละกลุ่ม (Marozzi, 2011; Bongbeebina & Rahman, 2025; Wang *et al.*, 2017; Soikliew & Araveeporn, 2018; Hatchavanich, 2018) ส่วนการทดสอบ Fligner-Killeen จะเหมาะกับข้อมูลที่ไม่มีการแจกแจงปกติ แต่ไม่เหมาะกับข้อมูลที่มีขนาดเล็ก และมีกำลังการทดสอบต่ำกว่าการทดสอบเลวินที่อิงตามค่ามัธยฐาน และการทดสอบ O'Brien ในบางกรณี (Marozzi, 2011; Bongbeebina & Rahman, 2025; Katsileros *et al.*, 2024;) สำหรับการทดสอบเลวินแบบไม่อิงพารามิเตอร์เหมาะกับข้อมูลที่มีการแจกแจงแบบเบ้ และมีความแกร่งเมื่อมีค่านอกเกณฑ์ในชุดข้อมูล และมีกำลังการทดสอบสูงกว่าการทดสอบเลวินที่อิงตามค่ามัธยฐานในบางกรณี (Nordstokke & Zumbo, 2010; Yonar *et al.*, 2024) ดังนั้นจากการศึกษาที่ผ่านมาจะเห็นได้ว่าไม่มีการทดสอบวิธีใดที่เป็นวิธีที่ดีที่สุดในทุกสถานการณ์ แต่ละวิธีมีข้อดีและข้อจำกัดขึ้นอยู่กับลักษณะ และขนาดของข้อมูล

ในปัจจุบัน ข้อมูลจากสถานการณ์จริงมักมีความซับซ้อนมากขึ้นกว่าที่จะมีการแจกแจงความน่าจะเป็นแบบเดียว แต่มีการแจกแจงที่เรียกว่าเป็นการแจกแจงแบบผสม (Mixture distribution) ซึ่งประชากรรวมประกอบด้วยประชากร

ย่อยที่มีค่าเฉลี่ยและความแปรปรวนแตกต่างกัน การแจกแจงแบบผสมที่เป็นที่รู้จักกันดี ได้แก่ การแจกแจงปกติแบบผสม (Mixture normal distribution) ซึ่งเป็นการแจกแจงที่เกิดจากการรวมกันของการแจกแจงปกติหลายกลุ่ม (Component) โดยแต่ละกลุ่มจะแทนประชากรย่อย (Subpopulation) ที่มีลักษณะเฉพาะในข้อมูลจริง การแจกแจงชนิดนี้ถูกพัฒนาขึ้นมาเพื่ออธิบายลักษณะของข้อมูลที่มีความซับซ้อน เช่น มีหลายกลุ่มย่อย มีความเบ้ (Skewness) หรือมีหางหนา (Heavy tails) ซึ่งการแจกแจงปกติไม่สามารถอธิบายได้ดีพอ (Wang & Taaffe, 2015) การแจกแจงปกติแบบผสมถูกนำไปใช้ในงานหลากหลายสาขาวิชา เช่น ทางด้านการเงินที่ใช้ในการอธิบายผลตอบแทนของหุ้นในช่วงตลาดปกติเทียบกับช่วงวิกฤต (Yan & Han, 2019) และทางการแพทย์ที่ใช้ในการศึกษาเกี่ยวกับตัวบ่งชี้ทางชีวภาพในผู้ป่วยที่มีฟีโนไทป์ต่างกัน (Biological marker) (Gibbons *et al.*, 1984) และทางระบาดวิทยาที่ใช้อธิบายอัตราการแพร่ระบาดของ COVID-19 ในกลุ่มประชากรที่ต่างกัน (Wang, 2023) ดังนั้น การแจกแจงปกติแบบผสมจึงเป็นการแจกแจงของข้อมูลจริงที่มักพบได้เสมอในการศึกษาต่าง ๆ

จากที่กล่าวมาในข้างต้น จะเห็นได้ว่าการทดสอบความเท่ากันของความแปรปรวนที่มีอยู่ด้วยกันหลายวิธี แต่ละวิธีมีข้อดีและข้อจำกัดแตกต่างกันโดยขึ้นอยู่กับลักษณะการแจกแจงของข้อมูลเป็นสำคัญ อย่างไรก็ตามในงานวิจัยที่ผ่านมา ยังไม่มีการศึกษาในกรณีที่ข้อมูลมีการแจกแจงแบบผสม ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบความน่าจะเป็นของความผิดพลาดแบบที่ 1 (Probability of type I error) และกำลังการทดสอบ (Power of the test) ของการทดสอบความเท่ากันของความแปรปรวนของข้อมูล 3 กลุ่ม ที่มีการแจกแจงปกติแบบผสม 2 องค์ประกอบ โดยการทดสอบที่นำมาศึกษา คือ การทดสอบบาร์ตเลตต์ การทดสอบเลวินที่อิงตามค่าเฉลี่ย การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบ เลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และ การทดสอบ Fligner-Killeen

Methodology

1. การทดสอบความเท่ากันของความแปรปรวน

การทดสอบความเท่ากันของความแปรปรวนของข้อมูล k กลุ่ม มีสมมุติฐานที่ทดสอบคือ

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

$$H_a : \text{มี } \sigma_j^2 \text{ แตกต่างกันอย่างน้อย 2 ค่า } (j = 1, 2, \dots, k)$$

การทดสอบที่นำมาศึกษาได้แก่ การทดสอบบาร์ตเลตต์ การทดสอบเลวินที่อิงตามค่าเฉลี่ย การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen

1.1 การทดสอบบาร์ตเลตต์

Bartlett (1937) ได้นำเสนอการทดสอบบาร์ตเลตต์ ที่สถิติทดสอบมีการแจกแจงไคกำลังสอง (Chi-square distribution) โดย

$$B = \frac{(N-k) \log \left[\frac{\sum_{j=1}^k (n_j - 1) S_j^2}{N-k} \right] - \sum_{j=1}^k (n_j - 1) \log(S_j^2)}{1 + \frac{\sum_{j=1}^k \frac{1}{n_j - 1} - \frac{1}{N-k}}{3(k-1)}} \quad (1)$$

เมื่อ N คือจำนวนข้อมูลทั้งหมด

n_j คือจำนวนข้อมูลในกลุ่มที่ j

k คือจำนวนกลุ่ม

S_j^2 คือความแปรปรวนของข้อมูลในกลุ่มที่ j

ปฏิเสธสมมติฐานหลักเมื่อ $B > \chi_{1-\alpha, k-1}^2$ โดยที่ $\chi_{1-\alpha, k-1}^2$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงไคกำลังสองที่องศาเสรี $k-1$

1.2 การทดสอบเลวินที่อิงค่าเฉลี่ย

Levene (1960) ได้นำเสนอการทดสอบเลวินที่อิงค่าเฉลี่ยโดยใช้ค่าสัมบูรณ์ของส่วนเหลือ (Residual) $Z_{ij} = |y_{ij} - \bar{y}_{.j}|$ ในการวิเคราะห์ความแปรปรวน ซึ่งเป็นการแปลงรูปจากการทดสอบค่าเฉลี่ยเป็นการทดสอบความแปรปรวนทำให้ได้การทดสอบที่มีความแข็งแกร่งต่อข้อสมมุติเกี่ยวกับการแจกแจงปกติของข้อมูล โดยมีสถิติทดสอบคือ

$$L = \frac{(N-k) \sum_{j=1}^k n_j (\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1) \sum_{j=1}^k \sum_{i=1}^{n_j} (Z_{ij} - \bar{Z}_{.j})^2} \quad (2)$$

เมื่อ $Z_{ij} = |y_{ij} - \bar{y}_{.j}|$ คือค่าสัมบูรณ์ของส่วนเหลือ (Residual)

y_{ij} คือค่าสังเกตที่ i ในกลุ่มที่ j

$\bar{y}_{.j}$ คือค่าเฉลี่ยของข้อมูลในกลุ่มที่ j

$\bar{Z}_{.j}$ คือค่าเฉลี่ยของ Z_{ij} ในกลุ่มที่ j

$\bar{Z}_{..}$ คือค่าเฉลี่ยรวมของ Z_{ij} ทั้งหมด

ปฏิเสธสมมติฐานหลักเมื่อ $L > F_{1-\alpha, k-1, N-k}$ โดยที่ $F_{1-\alpha, k-1, N-k}$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

1.3 การทดสอบเลวินที่อิงค่ามัธยฐาน

Brown & Forsythe (1974) ได้นำเสนอการทดสอบเลวินที่อิงค่ามัธยฐานหรือการทดสอบ Brown-Forsythe โดยการทดสอบนี้ใช้หลักการเดียวกันกับการทดสอบเลวินที่อิงค่าเฉลี่ยแต่ใช้ค่ามัธยฐานแทนค่าเฉลี่ยนั่นคือ $Z_{ij} = |y_{ij} - \tilde{y}_{.j}|$ มีสถิติทดสอบคือ

$$BF = \frac{(N-k) \sum_{j=1}^k n_j (\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1) \sum_{j=1}^k \sum_{i=1}^{n_j} (Z_{ij} - \bar{Z}_{.j})^2} \quad (3)$$

เมื่อ y_{ij} คือค่าสังเกตที่ i ในกลุ่มที่ j
 $\tilde{y}_{.j}$ คือค่ามัธยฐานของข้อมูลในกลุ่มที่ j
 $\bar{Z}_{.j}$ คือค่าเฉลี่ยของ Z_{ij} ในกลุ่มที่ j
 $\bar{Z}_{..}$ คือค่าเฉลี่ยรวมของ Z_{ij} ทั้งหมด

ปฏิเสธสมมติฐานหลักเมื่อ $BF > F_{1-\alpha, k-1, N-k}$ โดยที่ $F_{1-\alpha, k-1, N-k}$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

1.4 การทดสอบเลวินแบบไม่อิงพารามิเตอร์

Nordstokke & Zumbo (2010) ได้นำเสนอการทดสอบเลวินแบบไม่อิงพารามิเตอร์ซึ่งปรับปรุงมาจากการทดสอบเลวินที่อิงค่าเฉลี่ย โดยใช้การแปลงรูป y_{ij} ให้เป็นลำดับที่ R_{ij} และ $Z_{ij} = |R_{ij} - \bar{R}_{.j}|$ เมื่อ $\bar{R}_{.j}$ คือค่าเฉลี่ยของลำดับที่ในกลุ่มที่ j และมีสถิติทดสอบรูปแบบเดียวกับการทดสอบเลวินที่อิงค่าเฉลี่ย คือ

$$NL = \frac{(N-k) \sum_{j=1}^k n_j (\bar{Z}_{.j} - \bar{Z}_{..})^2}{(k-1) \sum_{j=1}^k \sum_{i=1}^{n_j} (Z_{ij} - \bar{Z}_{.j})^2} \quad (4)$$

ปฏิเสธสมมติฐานหลักเมื่อ $NL > F_{1-\alpha, k-1, N-k}$ โดยที่ $F_{1-\alpha, k-1, N-k}$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

1.5 การทดสอบ Fligner-Killeen

การทดสอบ Fligner-Killeen ได้รับการนำเสนอโดย Fligner & Killeen (1976) โดยการหาค่าสถิติทดสอบนี้ทำได้โดยนำข้อมูล y_{ij} มาแปลงรูปเป็น $Z_{ij} = |y_{ij} - \bar{y}_{.j}|$ และหาลำดับที่ของ Z_{ij} กำหนดให้เป็น R_{ij} จากนั้นนำมาหาค่าสกออร์

(Score) a_{N,R_j} จาก $a_{N,R_j} = \Phi^{-1}\left(0.5 + \frac{0.5R_j}{N+1}\right)$ ซึ่งคือควอนไทล์ที่ $\left(0.5 + \frac{0.5R_j}{N+1}\right)$ ของการแจกแจงปกติ และได้สถิติทดสอบคือ

$$FK = \frac{\sum_{j=1}^k n_j (\bar{A}_j - \bar{a})^2}{V^2} \quad (5)$$

เมื่อ $\bar{A}_j$ คือค่าเฉลี่ยของสกอาร์ a_{N,R_j} ในกลุ่มที่ j

$\bar{a}$ คือค่าเฉลี่ยรวมของ a_{N,R_j} และ V^2 ความแปรปรวนของ a_{N,R_j}

ปฏิเสธสมมุติฐานหลักเมื่อ $FK > \chi^2_{1-\alpha, k-1}$ โดยที่ $\chi^2_{1-\alpha, k-1}$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงไคกำลังสองที่องศาเสรี $k-1$

1.6 การทดสอบ O'Brien

O'Brien (1981) ได้นำเสนอการทดสอบ O'Brien ที่ใช้การแปลงรูปข้อมูล y_{ij} ให้เป็น V_{ij} ดังนี้

$$V_{ij} = \frac{(n_j - 1.5)n_j(y_{ij} - \bar{y}_j)^2 - 0.5S_j^2(n_j - 1)}{(n_j - 1)(n_j - 2)}$$

และสถิติทดสอบของการทดสอบนี้คือ

$$OB = \frac{(N-k) \sum_{j=1}^k n_j (\bar{V}_j - \bar{V})^2}{(k-1) \sum_{j=1}^k \sum_{i=1}^{n_j} (V_{ij} - \bar{V}_j)^2} \quad (6)$$

เมื่อ $\bar{V}_j$ ค่าเฉลี่ยของ V_{ij} ในกลุ่มที่ j และ $\bar{V}$ คือค่าเฉลี่ยรวมของ V_{ij}

ปฏิเสธสมมุติฐานหลักเมื่อ $OB > F_{1-\alpha, k-1, N-k}$ โดยที่ $F_{1-\alpha, k-1, N-k}$ คือควอนไทล์ที่ $1-\alpha$ ของการแจกแจงเอฟที่องศาเสรี $k-1$ และ $N-k$

2. การแจกแจงปกติแบบผสม

ให้ Y เป็นตัวแปรสุ่มที่มีการแจกแจงปกติแบบผสมสององค์ประกอบ ที่ประกอบด้วยการแจกแจงปกติตัวแปรเดียวสองการแจกแจง โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็นคือ

$$f(y;p,\mu_1,\sigma_1,\mu_2,\sigma_2)=p\times\frac{1}{\sigma_1\sqrt{2\pi}}\exp\left[-\frac{(y-\mu_1)^2}{2\sigma_1^2}\right]+(1-p)\times\frac{1}{\sigma_2\sqrt{2\pi}}\exp\left[-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right] \quad (7)$$

เมื่อ p คือพารามิเตอร์ของการผสมและ $0 < p < 1$

μ_1 และ σ_1 คือค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของการแจกแจงปกติองค์ประกอบที่ 1

μ_2 และ σ_2 คือค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของการแจกแจงปกติองค์ประกอบที่ 2

โดยมีค่าเฉลี่ยคือ $\mu_y = p\mu_1 + (1-p)\mu_2$ และความแปรปรวนคือ

$$Var(y) = p\sigma_1^2 + (1-p)\sigma_2^2 + p(1-p)(\mu_1 - \mu_2)^2$$

(KOCA et al., 2019)

3. ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และค่าประมาณกำลังการทดสอบ

ในการวิจัยนี้ได้ใช้ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และค่าประมาณของกำลังการทดสอบเป็นเกณฑ์ในการพิจารณาประสิทธิภาพของการทดสอบแต่ละชนิด

3.1 ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1

ความน่าจะเป็นของความผิดพลาดแบบที่ 1 (α) คือความน่าจะเป็นของการปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นจริง โดยค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 หาได้จาก

$$\hat{\alpha} = \frac{\sum_{i=1}^M I_i}{M} \quad (8)$$

เมื่อ $I_i = 1$ หากปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นจริงในการทำซ้ำครั้งที่ i และ $I_i = 0$ หากไม่สามารถปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นจริงในการทำซ้ำครั้งที่ i และ M คือจำนวนรอบของการทำซ้ำ

ในการใช้ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในการพิจารณาประสิทธิภาพของการทดสอบจะนำค่าประมาณที่ได้ไปเปรียบเทียบกับเกณฑ์ของ Bradley (Bradley, 1978) ที่ระดับนัยสำคัญ 0.05 โดยการทดสอบใดที่ให้ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ในช่วง [0.025, 0.075] จะถือว่าการทดสอบนั้นสามารถควบคุมความผิดพลาดแบบที่ 1 ได้

3.2 ค่าประมาณกำลังการทดสอบ

กำลังการทดสอบ ($1 - \beta$) คือความน่าจะเป็นของการปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นเท็จ โดยค่าประมาณกำลังการทดสอบหาได้จาก

$$1 - \hat{\beta} = \frac{\sum_{i=1}^M I_i}{M} \quad (9)$$

เมื่อ $I_i = 1$ หากปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นเท็จในการทำซ้ำครั้งที่ i และ $I_i = 0$ หากไม่สามารถปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นเท็จในการทำซ้ำครั้งที่ i และ M คือจำนวนรอบของการทำซ้ำ

การใช้ค่าประมาณกำลังการทดสอบในการพิจารณาประสิทธิภาพของการทดสอบจะถือว่าการทดสอบที่มีค่าประมาณกำลังการทดสอบสูงที่สุดคือการทดสอบที่มีประสิทธิภาพมากที่สุด อย่างไรก็ตามจาก Cohen (1988) การทดสอบที่มีประสิทธิภาพควรมีกำลังการทดสอบไม่น้อยกว่า 0.80

1. ขอบเขตการศึกษา

1. การทดสอบที่ศึกษา คือ การทดสอบบาร์ตเลตต์ (B) การทดสอบเลวินที่อิงตามค่าเฉลี่ย (L-mean) การทดสอบเลวินที่อิงตามค่ามัธยฐาน (L-med) การทดสอบเลวินแบบไม่อิงพารามิเตอร์ (L-non) การทดสอบ O'Brien (OB) และการทดสอบ Fligner-Killeen (FK)
2. จำนวนกลุ่มของข้อมูลเท่ากับ 3 กลุ่ม ($k = 3$)
3. ข้อมูลทั้ง 3 กลุ่ม มีการแจกแจงปกติแบบผสม 2 องค์ประกอบมีสัดส่วนของการผสม (Mixing proportion; p) ที่แสดงถึงน้ำหนักขององค์ประกอบที่ 1 ที่เป็นองค์ประกอบหลัก ซึ่งในการศึกษานี้พิจารณา 2 กรณี:

$p = 0.5$: แทนกรณีที่มีข้อมูลเกิดจากประชากรย่อยสองกลุ่มในสัดส่วนที่เท่ากัน

$p = 0.8$: แทนกรณีที่มีประชากรกลุ่มใหญ่เป็นหลักและมีกลุ่มย่อยที่เป็นข้อมูลผิดปกติ (Outliers) แทรกอยู่ โดยกำหนดลักษณะการแจกแจงของข้อมูลที่กำลังศึกษาดังนี้

- MN1 คือกรณีที่ข้อมูล 3 กลุ่ม มีการแจกแจงปกติแบบผสมที่มีค่าเฉลี่ยเท่ากับ 8.8 และมีอัตราส่วนความแปรปรวน 1:1:1 กำหนดค่าความแปรปรวนเท่ากับ 5 เท่ากันทั้ง 3 กลุ่ม
 - MN2 คือกรณีที่ข้อมูล 3 กลุ่ม มีการแจกแจงปกติแบบผสมที่มีค่าเฉลี่ยเท่ากับ 8.8 และมีอัตราส่วนความแปรปรวน 1:1:2 กำหนดค่าความแปรปรวนของแต่ละกลุ่มเท่ากับ 5 5 และ 10 ตามลำดับ
 - MN3 คือกรณีที่ข้อมูล 3 กลุ่ม มีการแจกแจงปกติแบบผสมที่มีค่าเฉลี่ยเท่ากับ 8.8 และมีอัตราส่วนความแปรปรวน 1:1:4 กำหนดค่าความแปรปรวนของแต่ละกลุ่มเท่ากับ 5 5 และ 20 ตามลำดับ
 - MN4 คือกรณีที่ข้อมูล 3 กลุ่ม มีการแจกแจงปกติแบบผสมที่มีค่าเฉลี่ยเท่ากับ 8.8 และมีอัตราส่วนความแปรปรวน 1:2:4 กำหนดค่าความแปรปรวนของแต่ละกลุ่มเท่ากับ 5 10 และ 20 ตามลำดับ
- และเพื่อให้ได้ข้อมูลที่มีการแจกแจงตามที่กำหนด จึงกำหนดค่าพารามิเตอร์ในองค์ประกอบที่ I และ II ของแต่ละกลุ่มดังแสดงใน Table 1 และ 2

4. จำนวนข้อมูลในแต่ละกลุ่ม คือ n_1 n_2 และ n_3 เท่ากันทุกกลุ่ม (10,10,10) (20,20,20) (40,40,40) (60,60,60) (80,80,80) และ (100,100,100)

5. ระดับนัยสำคัญของการทดสอบเท่ากับ $\alpha = 0.05$

6. ทำซ้ำในแต่ละสถานการณ์ 10,000 ครั้ง

Table 1 Parameters of mixture normal distribution when $p = 0.8$

Distribution	Variance Ratio $\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	Group 1		Group 2		Group 3	
		I	II	I	II	I	II
		μ_1, σ_1^2	μ_2, σ_2^2	μ_1, σ_1^2	μ_2, σ_2^2	μ_1, σ_1^2	μ_2, σ_2^2
MN1	1:1:1	8, 2.44	12, 2.44	8, 2.44	12, 2.44	8, 2.44	12, 2.44
MN2	1:1:2	8, 2.44	12, 2.44	8, 2.44	12, 2.44	8, 7.44	12, 7.44
MN3	1:1:4	8, 2.44	12, 2.44	8, 2.44	12, 2.44	8, 17.44	12, 17.44
MN4	1:2:4	8, 2.44	12, 2.44	8, 7.44	12, 7.44	8, 17.44	12, 17.44

Table 2 Parameters of mixture normal distribution when $p = 0.5$

Distribution	Variance Ratio $\sigma_1^2 : \sigma_2^2 : \sigma_3^2$	Group 1		Group 2		Group 3	
		I	II	I	II	I	II
		μ_1, σ_1^2	μ_2, σ_2^2	μ_1, σ_1^2	μ_2, σ_2^2	μ_1, σ_1^2	μ_2, σ_2^2
MN1	1:1:1	8, 1	12, 1	8, 1	12, 1	8, 1	12, 1
MN2	1:1:2	8, 1	12, 1	8, 1	12, 1	8, 6	12, 6
MN3	1:1:4	8, 1	12, 1	8, 1	12, 1	8, 16	12, 16
MN4	1:2:4	8, 1	12, 1	8, 6	12, 6	8, 16	12, 16

ขั้นตอนการศึกษา

ในการศึกษาและเปรียบเทียบประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนของประชากร 3 กลุ่ม ที่ประชากรแต่ละกลุ่มมีการแจกแจงปกติแบบผสม 2 องค์ประกอบ โดยใช้การจำลองมอนติคาร์โล (Monte Carlo simulation) มีขั้นตอนดังนี้

1. สร้างข้อมูลที่มีการแจกแจงปกติแบบผสมที่มีขนาดตัวอย่างและพารามิเตอร์ตามที่กำหนดใน Table 1 และ Table 2 จำนวน 3 กลุ่ม
2. นำข้อมูลที่ได้มาทดสอบความเท่ากันของความแปรปรวนด้วยการทดสอบทั้ง 6 ชนิด
3. ทำซ้ำข้อ 1 และ 2 จำนวน 10,000 ครั้ง นับจำนวนครั้งของการปฏิเสธสมมติฐานหลัก

4. ในกรณีที่ประชากร 3 กลุ่ม มีความแปรปรวนเท่ากัน คำนวณค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1
ในกรณีที่ประชากร 3 กลุ่ม มีความแปรปรวนแตกต่างกันอย่างน้อย 2 กลุ่ม คำนวณค่าประมาณกำลังการทดสอบ
5. พิจารณาประสิทธิภาพของการทดสอบ 6 ชนิด โดยพิจารณาค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ก่อน หากการทดสอบชนิดใดมีค่านี้อยู่ในช่วง $[0.025 - 0.075]$ จะถือว่าการทดสอบนี้สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ และจะนำการทดสอบนี้มาพิจารณาค่าประมาณกำลังการทดสอบต่อ โดยการทดสอบที่มีค่าประมาณกำลังการทดสอบมากกว่า 0.80 จะถือว่าเป็นการทดสอบที่มีประสิทธิภาพ และหากการทดสอบใดมีค่าประมาณกำลังการทดสอบมากที่สุดจะถือว่าเป็นการทดสอบที่มีประสิทธิภาพมากที่สุด ส่วนการทดสอบที่มีค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 นอกเกณฑ์ที่กำหนดจะถือว่าเป็นการทดสอบที่ไม่มีประสิทธิภาพ

Results

ผลการศึกษาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวนของการทดสอบ 6 ชนิด มีดังนี้

1. ค่าประมาณความน่าจะเป็นของความผิดพลาดที่ 1

Table 3 และ Table 4 แสดงค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของการทดสอบ 6 ชนิด เมื่อข้อมูลมีการแจกแจงปกติแบบผสม ที่มีการกำหนดสัดส่วนของการผสม 0.8 และ 0.5 โดยการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ ต้องมีค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ในช่วง $[0.025, 0.075]$

จาก Table 3 กรณีที่ $p = 0.8$ จะเห็นได้ว่าการทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ทุกขนาดตัวอย่าง ในขณะที่การทดสอบบาร์ตเลตต์สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ขนาดตัวอย่างเท่ากับ (60,60,60) ขึ้นไป ส่วนการทดสอบเลวินที่อิงตามค่าเฉลี่ยไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกกรณี

จาก Table 4 กรณีที่ $p = 0.5$ จะเห็นได้ว่าการทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ และการทดสอบ O'Brien สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ที่ทุกขนาดตัวอย่าง ในขณะที่การทดสอบเลวินที่อิงตามค่าเฉลี่ยสามารถควบคุมความผิดพลาดแบบที่ 1 ได้เมื่อขนาดตัวอย่างเท่ากับ (20,20,20) ขึ้นไป ส่วนการทดสอบบาร์ตเลตต์และการทดสอบ Fligner-Killeen ไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกกรณี

2. ค่าประมาณกำลังการทดสอบ

Table 5 และ Table 6 แสดงค่าประมาณกำลังการทดสอบของการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ โดยการทดสอบที่มีประสิทธิภาพต้องมีค่าประมาณกำลังการทดสอบไม่น้อยกว่า 0.8

Table 3 Empirical probability of type I error, $p = 0.8$

Sample size	Tests					
	B	L-mean	L-med	L-non	OB	FK
(10,10,10)	0.1389	0.1613	0.0519*	0.0488*	0.0692*	0.0532*
(20,20,20)	0.0973	0.1625	0.0528*	0.0519*	0.0696*	0.0544*
(40,40,40)	0.0767	0.1547	0.0526*	0.0484*	0.0603*	0.0584*
(60,60,60)	0.0710*	0.1576	0.0501*	0.0493*	0.0562*	0.0555*
(80,80,80)	0.0677*	0.1515	0.0477*	0.0483*	0.0518*	0.0529*
(100,100,100)	0.0634*	0.1506	0.0473*	0.0548*	0.0513*	0.0552*

* $\hat{\alpha}$ falls between [0.025, 0.075].

Table 4 Empirical probability of type I error, $p = 0.5$

Sample size	Tests					
	B	L-mean	L-med	L-non	OB	FK
(10,10,10)	0.0074	0.0767	0.0270*	0.0511*	0.0359*	0.0195
(20,20,20)	0.0006	0.0642*	0.0260*	0.0469*	0.0389*	0.0137
(40,40,40)	0.0005	0.0651*	0.0250*	0.0535*	0.0475*	0.0185
(60,60,60)	0.0001	0.0529*	0.0258*	0.0456*	0.0448*	0.0197
(80,80,80)	0.0002	0.0551*	0.0271*	0.0504*	0.0460*	0.0237
(100,100,100)	0.0002	0.0575*	0.0278*	0.0527*	0.0486*	0.0248

* $\hat{\alpha}$ falls between [0.025, 0.075].

จาก Table 5 จะเห็นได้ว่าเมื่อ $p = 0.8$ และอัตราส่วนความแปรปรวนเท่ากับ 1:1:2 การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen มีค่าประมาณกำลังการทดสอบมากกว่า 0.8 เมื่อตัวอย่างมีขนาด (60,60,60) ขึ้นไป ส่วนเมื่ออัตราส่วนความแปรปรวนเท่ากับ 1:1:4 การทดสอบ 4 ชนิดนี้ มีค่าประมาณกำลังการทดสอบมากกว่า 0.8 เมื่อตัวอย่างมีขนาด (20,20,20) ขึ้นไป และเมื่ออัตราส่วนความแปรปรวนเท่ากับ 1:2:4 การทดสอบ 4 ชนิดนี้ มีค่าประมาณกำลังการทดสอบมากกว่า 0.8 เมื่อตัวอย่างมีขนาด (40,40,40) ขึ้นไป สำหรับการทดสอบบาร์ตเลตต์มีค่าประมาณกำลังการทดสอบมากกว่า 0.8 ที่ขนาดตัวอย่าง (60,60,60) ขึ้นไป และที่ทุกอัตราส่วนความแปรปรวน

จาก Table 6 เมื่อ $p = 0.5$ และอัตราส่วนความแปรปรวนเท่ากับ 1:1:2 การทดสอบ O'Brien มีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (60,60,60) ขึ้นไป ส่วนการทดสอบเลวินที่อิงตามค่าเฉลี่ยและที่อิงตามค่ามัธยฐานมีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (100,100,100) เมื่ออัตราส่วนความแปรปรวนเท่ากับ 1:1:4 การทดสอบ O'Brien มีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (20,20,20) ขึ้นไป ส่วนการทดสอบเลวินที่อิงตามค่าเฉลี่ยและที่อิงตามค่ามัธยฐานมีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (40,40,40) ขึ้นไป และการทดสอบเลวินแบบไม่อิงพารามิเตอร์มีค่าประมาณกำลังการ

ทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (60,60,60) ขึ้นไป และเมื่ออัตราส่วนความแปรปรวนเท่ากับ 1:2:4 การทดสอบ O'Brien การทดสอบเลวินที่อิงตามค่าเฉลี่ยและที่อิงตามค่ามัธยฐานมีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (40,40,40) ขึ้นไป และการทดสอบเลวินแบบไม่อิงพารามิเตอร์มีค่าประมาณกำลังการทดสอบสูงกว่า 0.8 เมื่อขนาดตัวอย่างเท่ากับ (60,60,60) ขึ้นไป

นอกจากนั้นจะเห็นได้ว่าค่าประมาณกำลังการทดสอบของการทดสอบมีค่าสูงขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้น และอัตราส่วนความแปรปรวนมีความแตกต่างกันมากขึ้น

Table 5 Empirical power of the test when $p = 0.8$

Variance ratio	Sample size	Tests					
		B	L-mean	L-med	L-non	OB	FK
1 : 1 : 2	(10,10,10)	-	-	0.1810	0.2004	0.1715	0.1768
	(20,20,20)	-	-	0.4089	0.3908	0.4136	0.3970
	(40,40,40)	-	-	0.7533	0.7061	0.7753	0.7522
	(60,60,60)	0.9407*	-	0.9214*	0.8751*	0.9351*	0.9223*
	(80,80,80)	0.9828*	-	0.9767*	0.9510*	0.9814*	0.9767*
	(100,100,100)	0.9953*	-	0.9944*	0.9833*	0.9951*	0.9950*
1 : 1 : 4	(10,10,10)	-	-	0.5150	0.4304	0.4916	0.4663
	(20,20,20)	-	-	0.9034*	0.8752*	0.9131*	0.8738*
	(40,40,40)	-	-	0.9986*	0.9758*	0.9990*	0.9970*
	(60,60,60)	1.0000*	-	1.0000*	0.9978*	1.0000*	1.0000*
	(80,80,80)	1.0000*	-	1.0000*	0.9999*	1.0000*	1.0000*
	(100,100,100)	1.0000*	-	1.0000*	1.0000*	1.0000*	1.0000*
1 : 2 : 4	(10,10,10)	-	-	0.3409	0.3465	0.2868	0.3267
	(20,20,20)	-	-	0.7669	0.6910	0.7222	0.7400
	(40,40,40)	-	-	0.9890*	0.9584*	0.9878*	0.9849*
	(60,60,60)	1.0000*	-	0.9997*	0.9958*	0.9996*	0.9995*
	(80,80,80)	1.0000*	-	1.0000*	0.9998*	1.0000*	1.0000*
	(100,100,100)	1.0000*	-	1.0000*	1.0000*	1.0000*	1.0000*

* $1 - \hat{\beta}$ is greater than 0.8.

“-“ The test cannot control type I error.

Discussion

จากผลการศึกษาที่ได้ พบว่าเมื่อสัดส่วนของการผสมแตกต่างกันผลการศึกษาที่ได้มีความแตกต่างกันบางส่วน เนื่องจาก การแจกแจงปกติแบบผสมที่มีสัดส่วนของการผสมแตกต่างกันจะมีลักษณะการแจกแจงที่แตกต่างกัน โดยเมื่อพิจารณา ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 จะเห็นได้ว่าการทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบ เลวินแบบไม่อิงพารามิเตอร์ และการทดสอบ O'Brien สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกกรณี เนื่องจากการทดสอบเหล่านี้มีความแข็งแกร่งต่อลักษณะการแจกแจงความน่าจะเป็นของข้อมูล ซึ่งสอดคล้องกับผลการศึกษาของ Soikliew

Table 6 Empirical power of the test when $p = 0.5$

Variance ratio	Sample size	Tests					
		B	L-mean	L-med	L-non	OB	FK
1 : 1 : 2	(10,10,10)	-	-	0.0728	0.0833	0.1250	-
	(20,20,20)	-	0.2416	0.1643	0.1120	0.3611	-
	(40,40,40)	-	0.4344	0.3708	0.1701	0.7487	-
	(60,60,60)	-	0.6157	0.5710	0.2373	0.9157*	-
	(80,80,80)	-	0.7385	0.7074	0.2866	0.9759*	-
	(100,100,100)	-	0.8268*	0.8068*	0.3364	0.9933*	-
1 : 1 : 4	(10,10,10)	-	-	0.2991	0.2266	0.3966	-
	(20,20,20)	-	0.7841	0.7084	0.4388	0.8684*	-
	(40,40,40)	-	0.9764*	0.9685*	0.7350	0.9972*	-
	(60,60,60)	-	0.9988*	0.9980*	0.8816*	0.9999*	-
	(80,80,80)	-	0.9998*	0.9998*	0.9546*	1.0000*	-
	(100,100,100)	-	1.0000*	1.0000*	0.9813*	1.0000*	-
1 : 2 : 4	(10,10,10)	-	-	0.1924	0.1831	0.2293	-
	(20,20,20)	-	0.5916	0.5036	0.3396	0.6386	-
	(40,40,40)	-	0.9097*	0.8910*	0.6302	0.9750*	-
	(60,60,60)	-	0.9861*	0.9827*	0.8196*	0.9992*	-
	(80,80,80)	-	0.9984*	0.9983*	0.9143*	1.0000*	-
	(100,100,100)	-	0.9999*	0.9998*	0.9633*	1.0000*	-

* $1 - \hat{\beta}$ is greater than 0.8.

“-“ The test cannot control type I error.

& Araveeporn (2018) และ Yonar *et al.* (2024) ส่วนการทดสอบ Fligner-Killeen สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เมื่อ $p = 0.8$ เท่านั้น ซึ่งสอดคล้องกับผลการศึกษาของ Katsileros *et al.* (2024) ที่กล่าวว่าการทดสอบชนิดนี้ไม่เสถียรในเรื่องของการควบคุมความผิดพลาดแบบที่ 1 ในขณะที่การทดสอบเลวินที่อิงตามค่าเฉลี่ยสามารถคุมความผิดพลาดแบบที่ 1 ได้เมื่อ $p = 0.5$ และขนาดตัวอย่างเท่ากับ 20 ขึ้นไปเท่านั้น ซึ่งสอดคล้องกับผลการศึกษาของ Soikliew & Araveeporn (2018) และ Yonar *et al.* (2024) ที่กล่าวว่าการทดสอบชนิดนี้ไม่เหมาะกับข้อมูลที่ไม่ได้มีการแจกแจงปกติ สำหรับการทดสอบบาร์ตเลตต์นั้นมีความสามารถคุมความผิดพลาดแบบที่ 1 ได้น้อยที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ เนื่องจากการทดสอบชนิดนี้จะไม่สามารถควบคุมความผิดพลาดแบบที่ 1 เมื่อข้อมูลไม่มีการแจกแจงปกติ และตัวอย่างมีขนาดเล็กซึ่งสอดคล้องกับผลการศึกษาของ Sharma & Kibria (2013)

เมื่อนำการทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้มาศึกษาค่าประมาณกำลังการทดสอบ เพื่อพิจารณาว่าการทดสอบชนิดใดที่มีประสิทธิภาพในการทดสอบความเท่ากันของความแปรปรวนเมื่อข้อมูลมีการแจกแจงปกติแบบผสมทั้ง 3 กลุ่ม จาก Table 7 และ Table 8 จะเห็นได้ว่าการทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ และการทดสอบ O'Brien เป็นการทดสอบที่มีประสิทธิภาพในเกือบทุกสถานการณ์ โดยเฉพาะเมื่ออัตราส่วนความแปรปรวนมีความแตกต่างกันมาก 1:1:4 ซึ่งสอดคล้องกับผลการศึกษาของ Yonar *et al.* (2024) และ

Wang *et al.* (2017) โดยเฉพาะการทดสอบ O'Brien มีประสิทธิภาพดีถึงแม้อัตราส่วนความแปรปรวนจะมีความแตกต่างกันน้อย 1:1:2 และการทดสอบชนิดนี้ยังให้กำลังการทดสอบสูงที่สุดในหลาย ๆ สถานการณ์อีกด้วย ซึ่งสอดคล้องกับผลการศึกษาของ Wang *et al.* (2017) สำหรับการทดสอบบาร์ตเลตต์และการทดสอบ Fligner-Killeen เป็นการทดสอบที่มีประสิทธิภาพเมื่อข้อมูลมีการแจกแจงปกติแบบผสมเมื่อ $p = 0.8$ และตัวอย่างมีขนาดใหญ่

Table 7 The tests performance when $p = 0.8$

Sample size	Variance ratio		
	01:01:02	01:01:04	01:02:04
(10,10,10)	-	-	-
(20,20,20)	-	L-med, L-non, OB*, FK	-
(40,40,40)	-	L-med, L-non, OB*, FK	L-med*, L-non, OB, FK
(60,60,60)	B*, L-med, L-non,	B*, L-med*, L-non, OB*,	B*, L-med, L-non, OB, FK
(80,80,80)	B*, L-med, L-non,	B*, L-med*, L-non, OB*,	B*, L-med*, L-non, OB*,
(100,100,100)	B*, L-med, L-non,	B*, L-med*, L-non*, OB*,	B*, L-med*, L-non*, OB*,

* Test has the greatest estimate of power of the test.

“-“ In cases where no test proves effective.

Table 8 The tests performance when $p = 0.5$

Sample size	Variance ratio		
	1:1:2	1:1:4	1:2:4
10,10,10	-	-	-
20,20,20	-	OB*	-
40,40,40	-	L-mean, L-med, OB*	L-mean, L-med, OB*
60,60,60	OB*	L-mean, L-med, L-non, OB*	L-mean, L-med, L-non, OB*
80,80,80	OB*	L-mean, L-med, L-non, OB*	L-mean, L-med, L-non, OB*
100,100,100	L-mean, L-med, OB*	L-mean*, L-med*, L-non, OB*	L-mean, L-med, L-non, OB*

* Test has the greatest estimate of power of the test.

“-“ In cases where no test proves effective.

Conclusions

ในการศึกษาประสิทธิภาพของการทดสอบความเท่ากันของความแปรปรวน 6 ชนิด คือ การทดสอบบาร์ตเลตต์ การทดสอบเลวินที่อิงตามค่าเฉลี่ย การทดสอบเลวินที่อิงตามค่ามัธยฐาน การทดสอบเลวินแบบไม่อิงพารามิเตอร์ การทดสอบ O'Brien และการทดสอบ Fligner-Killeen โดยพิจารณาจากค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และค่าประมาณกำลังการทดสอบ เมื่อข้อมูลทั้ง 3 กลุ่ม มีการแจกแจงปกติแบบผสม 2 องค์ประกอบ ที่มีสัดส่วนการผสมเท่ากับ 0.8 และ 0.5 โดยใช้การจำลองมอนติคาร์โล

เมื่อ $p = 0.8$ และตัวอย่างมีขนาดใหญ่ (60,60,60) ขึ้นไป การทดสอบบาร์ตเลตต์เป็นการทดสอบที่มีประสิทธิภาพมากที่สุดในทุกอัตราส่วนความแปรปรวน แต่เมื่อตัวอย่างมีขนาด (20,20,20) และ (40,40,40) การทดสอบ O'Brien เป็นการทดสอบที่มีประสิทธิภาพมากที่สุด ส่วนการทดสอบ Fligner-Killeen และการทดสอบเลวินที่อิงตามค่ามัธยฐานมีประสิทธิภาพมากที่สุดเมื่ออัตราส่วนความแปรปรวนแตกต่างกันมากและตัวอย่างมีขนาดใหญ่

เมื่อ $p = 0.5$ การทดสอบ O'Brien เป็นการทดสอบที่มีประสิทธิภาพมากที่สุดที่ทุกอัตราส่วนความแปรปรวน เมื่อตัวอย่างมีขนาด (60,60,60) ขึ้นไป และเมื่อตัวอย่างมีขนาด (20,20,20) และ (40,40,40) เมื่ออัตราส่วนความแปรปรวนแตกต่างกันมากการทดสอบ O'Brien ยังคงมีประสิทธิภาพดีที่สุด

ดังนั้น สรุปในภาพรวมแล้วการทดสอบ O'Brien ถือเป็นการทดสอบที่มีประสิทธิภาพดีในเกือบทุกสถานการณ์ เนื่องจากสถิติทดสอบของการทดสอบนี้ได้จากการแปลงข้อมูลที่มีการนำขนาดตัวอย่างมาปรับค่าทำให้การทดสอบนี้มีประสิทธิภาพดีในเกือบทุกขนาดตัวอย่าง นอกจากนั้นการทดสอบนี้ยังมีความแข็งแกร่งต่อความเอนกของการแจกแจงของข้อมูล ดังนั้นจึงใช้ได้ดีแม้ข้อมูลที่ศึกษาจะไม่ได้มีการแจกแจงปกติ

References

- Arsham, H., & Lovric, M. (2011). Bartlett's Test. In: Lovric, M. (eds) *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_132.
- Bartlett, M. S. (1937). Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901), 268–282.
<http://www.jstor.org/stable/96803>
- Bongbeebina, C, & Rahman, M. (2025). ASSESSING HOMOGENEITY: A COMPARATIVE STUDY FOR ROBUST STATISTICAL ANALYSIS. *Far East Journal of Mathematical Sciences (FJMS)*, 142(2), 139-171. <https://doi.org/10.17654/0972087125009>
- Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31(2), 144–152.
- Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, 69(346), 364–367. <https://doi.org/10.2307/2285659>
- Chang, C. H., Pal, N., & Lin, J. J. (2017). A revisit to test the equality of variances of several populations. *Communications in Statistics - Simulation and Computation*, 46(8), 6360–6384.
<https://doi.org/10.1080/03610918.2016.1202277>

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Fligner, M. A., & Killeen, T. J. (1976). Distribution-Free Two-Sample Tests for Scale. *Journal of the American Statistical Association*, 71(353), 210–213. <https://doi.org/10.2307/2285771>
- Gibbons, R. D., Dorus, E., Ostrow, D. G., Pandey, G. N., Davis, J. M., & Levy, D. L. (1984). Mixture distributions in psychiatric research. *Biological psychiatry*, 19(7), 935–961.
- Hatchavanich, D. (2018). A Comparison of Type I Error and Power of Parametric Statistics and Nonparametric Statistics for Homogeneity of Variance Test. *Asian Health, Science and Technology Reports*, 26(1), 106–117. <https://ph03.tci-thaijo.org/index.php/ahstr/article/view/1771>
- Katsileros, A., Antonetsis, N., Mouzaidis, P., Tani, E., Bebeli, P. J., & Karagrigoriou, A. (2024). A comparison of tests for homoscedasticity using simulation and empirical data. *Communications for Statistical Applications and Methods*, 31, 1-35. <https://doi.org/10.29220/csam.2024.31.1.001>
- KOCA, M. B, KILIÇ, M. B, & ŞAHİN, Y (2019). Assessing wind energy potential using finite mixture distributions. *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(3), 2276-2294. <https://doi.org/10.3906/elk-1802-109>
- Levene, H. (1960) *Robust Tests for Equality of Variances*. In: Olkin, I., Ed., *Contributions to Probability and Statistics*, Stanford University Press, Palo Alto, 278-292.
- Marozzi, M. (2011). Levene type tests for the ratio of two scales. *Journal of Statistical Computation and Simulation*, 81(7), 815–826. <https://doi.org/10.1080/00949650903499321>
- Nordstokke, D. W., & Zumbo, B. D. (2010). A new nonparametric Levene test for equal variances. *Psicológica*, 31(2), 401–430.
- O'Brien, R. G. (1981). A simple test for variance effects in experimental designs. *Psychological Bulletin*, 89(3), 570–574. <https://doi.org/10.1037/0033-2909.89.3.570>
- Parra-Frutos, I. (2013). Testing homogeneity of variances with unequal sample sizes. *Computational Statistics*, 28, 1269–1297. <https://doi.org/10.1007/s00180-012-0353-x>

- Rayat, C.S. (2018). *Variance-Ratio Test and Analysis of Variance (ANOVA)*. In: Statistical Methods in Medical Research. Springer, Singapore. https://doi.org/10.1007/978-981-13-0827-7_12
- Sharma, D., & Kibria, B. M. G. (2013). On some test statistics for testing homogeneity of variances: a comparative study. *Journal of Statistical Computation and Simulation*, 83(10), 1944–1963. <https://doi.org/10.1080/00949655.2012.675336>
- Soikliew, K., & Araveeporn, A. (2018). Modifications of Levene's and O'Brien's Tests for Testing the Homogeneity of Variance Based on Median and Trimmed Mean. *Thailand Statistician*, 16(2), 106–128. <https://ph02.tci-thaijo.org/index.php/thaistat/article/view/135555>
- Vorapongsathorn, T., Taejaroenkul, S., & Viwatwongkasem, C. (2004). A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions. *Songklanakarin Journal of Science and Technology*, 26(4).
- Wang, H. (2023). Tolerance limits for mixture-of-normal distributions with application to COVID-19 data. *WIREs Computational Statistics*, 15(6), e1611. <https://doi.org/10.1002/wics.1611>
- Wang, J., & Taaffe, M. R. (2015). Multivariate Mixtures of Normal Distributions: Properties, Random Vector Generation, Fitting, and as Models of Market Daily Changes. *INFORMS Journal on Computing*, 27(2), 193-203. <https://doi.org/10.1287/ijoc.2014.0616>
- Wang, Y., Rodríguez de Gil, P., Chen, Y.-H., Kromrey, J. D., Kim, E. S., Pham, T., Nguyen, D., & Romano, J. L. (2017). Comparing the Performance of Approaches for Testing the Homogeneity of Variance Assumption in One-Factor ANOVA Models. *Educational and Psychological Measurement*, 77(2), 305-329. <https://doi.org/10.1177/0013164416645162>
- Yan, H., & Han, L. (2019). Empirical distributions of stock returns: Mixed normal or kernel density?. *Physica A: Statistical Mechanics and its Applications*, 514, 473-486. <https://doi.org/10.1016/j.physa.2018.09.080>
- Yonar, A., Yonar, H., Demirsöz, M., & Tekindal, M. A. (2024). A COMPARATIVE ANALYSIS FOR HOMOGENEITY OF VARIANCE TESTS. *Journal of Science and Arts*, 24(2), 305-328. <https://doi.org/10.46939/j.sci.arts-24.2-a06>