ประสิทธิภาพของการทดสอบสูงสุดไม่อิงพารามิเตอร์เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่มเมื่อมีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน

มนตรี สังข์ทอง*

มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ต.หันตรา อ.พระนครศรีอยุธยา จ.พระนครศรีอยุธยา 13000

** Corresponding Aurthor: montri.so@rmutsb.ac.th*

ผู้ช่วยศาสตราจารย์ สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี

ข้อมูลบทความ

ประวัติบทความ :
รับเพื่อพิจารณา : 11 มีนาคม 2563
แก้ไข : 29 ธันวาคม 2563
ตอบรับ : 22 มีนาคม 2564

คำสำคัญ :
การทดสอบสูงสุด /
การทดสอบตำแหน่ง /
สถิติไม่อิงพารามิเตอร์

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของการทดสอบสูงสุด (Maximum test) แบบไม่อิงพารามิเตอร์เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่มเมื่อมีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน ซึ่งเป็นวิธีการทางสถิติใหม่ที่นำเสนอเปรียบเทียบกับ Brunner-Munzel test, Welch Based on Rank test และ Yuen-Welch test ภายใต้การแจกแจงเบ้ขวา ทั้งกรณีประชากรทั้งสองกลุ่มมีการแจกแจงเดียวกันและมีการแจกแจงต่างกัน กำหนดความแปรปรวน จำนวน 6 เงื่อนไข ผลการวิจัย พบว่า เมื่อประชากรมีการแจกแจงแบบล็อกนอร์มัลและการแจกแจงแบบเลขชี้กำลังที่ระดับนัยสำคัญ 0.05 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Brunner-Munzel test และ Yuen-Welch test แต่สูงกว่า Welch Based on Rank test เมื่อพิจารณากำลังการทดสอบสูงสุดของสถิติทดสอบกลุ่มดังกล่าว พบว่า Welch Based on Rank test มีความสามารถสูงสุด รองลงมา คือ Brunner-Munzel test และ Maximum test โดยที่ระดับนัยสำคัญ 0.01 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Yuen-Welch test แต่สูงกว่า Brunner-Munzel test และ Welch Based on Rank test เมื่อพิจารณากำลังการทดสอบสูงสุดของสถิติทดสอบกลุ่มดังกล่าวสำหรับกรณีประชากรที่ 1 มีการแจกแจงแบบเลขชี้กำลัง และประชากรที่ 2 มีการแจกแจงล็อกนอร์มัล พบว่า ที่ระดับนัยสำคัญ 0.05 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Brunner-Munzel test และ Yuen-Welch test แต่สูงกว่า Welch Based on Rank test เมื่อพิจารณากำลังการทดสอบสูงสุดของสถิติทดสอบกลุ่มดังกล่าว พบว่า Welch Based on Rank test มีความสามารถสูงสุด รองลงมา คือ Maximum test, Yuen-Welch test และ Brunner-Munzel test สำหรับที่ระดับนัยสำคัญ 0.01 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Yuen-Welch test แต่สูงกว่า Brunner-Munzel test และ Welch Based on Rank test เมื่อพิจารณาวิธีทดสอบที่มีกำลังการทดสอบสูงสุดและสามารถควบคุมความผิดพลาดแบบที่ 1 พบว่า Welch Based on Rank test มีความสามารถสูงสุด รองลงมาคือ Yuen-Welch test และ Maximum test

Performance of Non-Parametric Maximum Test for Location Testing between Two Populations with Skewed Distributions and Unequal Variances

Montri Sangthong*

Rajamangala University of Technology Suvarnabhumi, Huntra, Phra Nakhon Si Ayutthaya 13000

* *Corresponding Aurthor: montri.so@rmutsb.ac.th*

Assistant Professor, Division of Mathematics, Faculty of Science and Technology.

Article Info

Article History:
Received: March 11, 2020
Revised: December 29, 2020
Accepted: March 22, 2021

Keywords:
Maximum Test /
Location Testing /
Nonparametric Statistics

Abstract

This research aimed to study the efficiency of the non-parametric Maximum test for location testing between two populations with skewed distributions and unequal variances; the Maximum test is a new statistical technique and was evaluated in comparison with the Brunner-Munzel test, Welch Based on Rank test and Yuen-Welch test. All the tests were performed with positive skewed distributions in both populations, which were identically distributed and non-identically distributed; the variances were set under six different conditions. The results illustrated that under the conditions of log-normal distribution and exponential distribution at a significant level of 0.05, the ability of the Maximum test to control type I error was lower than those of the Brunner-Munzel test and the Yuen-Welch test but higher than tha of the Welch Based on Rank test. When considering the highest test power of all the techniques, Welch Based on Rank test exhibited the highest capacity; this was followed in descending order by the Brunner-Munzel test and Maximum test. At a significant level of 0.01, the ability to control type I error of the Maximum test was lower than that of the Yuen-Welch test but higher than those of the Brunner-Munzel test and Welch based on rank test. When considering the highest test power at this significant level, the highest powers were achieved by the Yuen-Welch test and Welch Based on Rank test. In the case that the first population exhibited exponential distribution and the second population exhibited log-normal distribution, the ability of the Maximum test, at 0.05 significant level, to control type I error was lower than those of the Brunner-Munzel test and the Yuen-Welch test but higher than that of the Welch Based on Rank test. When considering the highest test power at this condition, Welch Based on Rank test had the highest capacity; this was followed in descending oredr by the Maximum test, Yuen-Welch test and Brunner-Munzel test. Furthermore, at a significant level of 0.01, the ability of the Maximum test to control type I error was lower than that of Yuen-Welch test but higher than those of Brunner-Munzel test and Welch Based on Rank test. When considering the test method with the highest test power that can control type I error, Welch based on rank test had the highest capacity; this was followed by the Yuen-Welch test and the Maximum test.

1. บทนำ

การวิจัยเชิงปริมาณถูกนำไปใช้ในการสร้างองค์ความรู้ในศาสตร์สาขาต่างๆ อย่างแพร่หลาย ทั้งวิทยาศาสตร์ธรรมชาติ วิทยาศาสตร์กายภาพ และมนุษยศาสตร์และสังคมศาสตร์ เครื่องมือที่สำคัญยิ่งของการวิจัยเชิงปริมาณในการสรุปอ้างอิง (Inferential) ผลการวิจัยสู่พารามิเตอร์ของประชากร คือ การทดสอบสมมติฐานทางสถิติ สำหรับการทดสอบสมมติฐานเกี่ยวกับตำแหน่งระหว่างประชากรสองกลุ่มมีการนำไปประยุกต์ใช้อย่างกว้างขวาง ซึ่งถ้าประชากรมีการแจกแจงปรกติและความแปรปรวนเท่ากันจะทดสอบสมมติฐานด้วย Independent t test [1-3] ถ้าความแปรปรวนต่างกันจะทดสอบสมมติฐานด้วย Welch t test โดยสถิติดังกล่าวเป็นสถิติอิงพารามิเตอร์ทั้งนี้หลายครั้งข้อมูลที่เก็บรวบรวมหรือจากการทดลองไม่ได้เป็นไปตามข้อตกลงเบื้องต้น ซึ่งสถิติไม่อิงพารามิเตอร์จะถูกเลือกใช้ โดยที่ได้รับความนิยมอย่างแพร่หลาย คือ Wilcoxon-Mann-Whitney test [3-5] นำเสนอโดย Wilcoxon [6] และ Mann-Whitney [7] ทั้งนี้จากการศึกษาของ Stonehouse และ Forrester [8] พบว่า สถิติทดสอบดังกล่าวไม่มีความแกร่งในหลายเงื่อนไข ต่อมา Brunner และ Munzel [9] ได้มีการพัฒนาสถิติทดสอบที่มีความแกร่งเพื่อแก้ปัญหาของเบห์เรนส์-ฟิชเชอร์ (Behrens-Fisher Problem) หรือปัญหาความแปรปรวนไม่เท่ากัน เรียกว่า Brunner-Munzel test จากผลการวิจัย พบว่าสถิติทดสอบดังกล่าวค่อนข้างมีความแกร่ง ทั้งนี้ขึ้นอยู่กับอัตราส่วนของตัวอย่างและความแปรปรวนของการแจกแจง สำหรับขนาดตัวอย่างที่เหมาะสม คือ $n_1, n_2 \geq 10$ ซึ่งเหมาะสมกับการนำไปประยุกต์ใช้กับการวิจัยเชิงทดลองทางคลินิก (Clinical trial) แต่ก็มีผลการวิจัยกล่าวถึงความไม่เหมาะสมในการใช้ Brunner-Munzel test ในการวิเคราะห์ Voxel-based lesion-symptom (VLSM) [10] นอกจากนี้ยังมีสถิติไม่อิงพารามิเตอร์บางสถิติทดสอบที่มีประสิทธิภาพในบางเงื่อนไข คือ Yuen-welch test เหมาะสมสำหรับการทดสอบผลต่างของตำแหน่งที่เป็นมัธยฐานของประชากรที่มีการแจกแจงแบบเบ้ขวาและมีความแปรปรวนต่างกัน [3]

จากที่กล่าวมาข้างต้น พบว่า ไม่มีสถิติทดสอบใดที่เหมาะสมทุกสถานการณ์หรือทุกเงื่อนไข ซึ่งสอดคล้องกับการศึกษาของ Fagerland และ Sandvik [3] ได้ศึกษาประสิทธิภาพของ 5 สถิติทดสอบ โดยจำลองข้อมูลให้มีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน พบว่า ไม่มีสถิติทดสอบใดที่เหมาะสมหรือมีประสิทธิภาพในทุกเงื่อนไข ต่อมา Neuhauser [11] ได้นำเสนอแนวคิดการทดสอบสมมติฐานโดยการใช้สถิติทดสอบร่วมกัน เรียกว่า Maximum test โดยศึกษาสถิติทดสอบร่วมกัน คือ t test และ Wilcoxon's rank-sum test พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ดี และมีกำลังการทดสอบที่สูงกว่า t test หรือ Wilcoxon's rank-sum test ภายหลัง Welz และคณะ [2] ได้มีการพัฒนา Maximum test เพื่อแก้ปัญหาเบห์เรนส์-ฟิชเชอร์ (Behrens-Fisher problem) โดยมีการศึกษาร่วมกันของ Welch Based on Rank test (WBR) และ Brunner-Munzel test (BM) โดยจำลองข้อมูลให้มีการแจกแจงเดียวกันทั้ง 2 ประชากร (Identically distributed) ประกอบด้วยการแจกแจงปรกติ การแจกแจงยูนิฟอร์ม การแจกแจงปัวซง การแจกแจงล็อกนอร์มัล การแจกแจงแบบเลขชี้กำลัง และประชากรทั้งสองมีการแจกแจงต่างกัน (Non-identically distributed) ประกอบด้วย การแจกแจงปรกติกับการแจกแจงไคสแควร์ และการแจกแจงทวิฐานนิยม กำหนดอัตราส่วนความแปรปรวน 3 ขนาด คือ 1:1, 1:2 และ 1:4 โดยกำหนดขนาดตัวอย่าง คือ (10, 10), (10, 20) และ (20, 10) ผลการวิจัย พบว่า การทดสอบสูงสุด (Maximum test) สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบที่สูงหลายการแจกแจงดังเช่นการแจกแจงล็อกนอร์มัล การแจกแจงแบบเลขชี้กำลัง และการแจกแจงปัวซง แม้ว่าขนาดตัวอย่างไม่เท่ากัน

แต่การศึกษาข้างต้น พบว่า การกำหนดเงื่อนไขของความแปรปรวนค่อนข้างน้อย เมื่อทดสอบนัยสำคัญเกี่ยวกับความเท่ากันของความแปรปรวนอัตราส่วน 1:2 และ 1:4 บางเงื่อนไขไม่มีนัยสำคัญทางสถิติ เมื่อพิจารณาจากค่านอนเซนทรัลลิตี้พารามิเตอร์ (noncentrality parameter) พบว่า ขนาดความแตกต่างของความแปรปรวนค่อนข้างต่ำ [12] ทั้งนี้จากผลการวิจัยต่างๆ เป็นไปในทิศทางเดียวกันว่าอิทธิพลของความแปรปรวนหรือปัญหาของเบห์เรนส์ฟิชเชอร์ ส่งผลต่อความแกร่งและกำลังการทดสอบมากกว่าการละเลยข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจง รวมถึงการจำลองแบบปัญหากำหนดขนาดตัวอย่างเพียง 3 ระดับ ทั้งนี้เพื่อให้มีความชัดเจนในประเด็นดังกล่าว จึงได้ศึกษาความแกร่งและกำลังการทดสอบของการทดสอบสูงสุดแบบไม่อิงพารามิเตอร์ในการทดสอบตำแหน่งระหว่าง

ประชากรสองกลุ่ม และเปรียบเทียบประสิทธิภาพกับ Brunner-Munzel test, Welch Based on Rank test (WBR) และ Yuen-Welch test โดยเกณฑ์การพิจารณาประสิทธิภาพ คือ ความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley [13] โดยที่ระดับนัยสำคัญ 0.05 มีค่า [0.025-0.075] และระดับนัยสำคัญ 0.01 มีค่า [0.005-0.015] และกำลังการทดสอบ

2. วิธีการทางสถิติ

การศึกษาประสิทธิภาพของการทดสอบสูงสุดไม่อิงพารามิเตอร์เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน สถิติทดสอบที่นำมาเปรียบเทียบกับ Maximum test ได้แก่ Brunner-Munzel test, Welch Based on Rank test และYuen-Welch test มีรายละเอียด ดังนี้

1. การทดสอบสูงสุด (Max test)

การทดสอบสูงสุด(Maximum test) เป็นสถิติไม่อิงพารามิเตอร์นำเสนอโดย Neuhauser [11] และมีการพัฒนาเพิ่มเติมโดย Welz และคณะ [2] โดยอาศัยสถิติทดสอบที่มีประสิทธิภาพสองตัวสถิติมาพิจารณาร่วมกัน และประยุกต์ร่วมกระบวนการทำซ้ำโดยการสุ่มแบบไม่คืนที่ (permutation) สำหรับการคำนวณความน่าจะเป็น (P-value) เพื่อใช้ในการทดสอบสมมติฐาน มีรายละเอียด ดังนี้

1) สุ่มตัวอย่างขนาด n_1 และ n_2 คำนวณค่า t_{MAX}

$$t_{MAX} = \mathrm{Max}\left[\left|\frac{t_{WBR}}{\sqrt{(df_{WBR}/(df_{WBR}-2))}}\right|, \left|\frac{t_{BM}}{\sqrt{(df_{BM}/(df_{BM}-2))}}\right|\right] \quad (1)$$

เมื่อ t_{WBR} คือ Welch Based on Rank test

t_{BM} คือ Brunner-Munzel test

2) นำตัวอย่างขนาด n_1 และ n_1 มารวมกัน จะได้ $N = n_1 + n_2$

3) สุ่มตัวอย่างแบบไม่คืนที่ (without replacement) ขนาด n_1 และ n_2 จาก N

4) คำนวณค่า t^*_{MAX}

$$t^*_{MAX} = \mathrm{Max}\left[\left|\frac{t^*_{WBR}}{\sqrt{(df_{WBR}/(df_{WBR}-2))}}\right|, \left|\frac{t^*_{BM}}{\sqrt{(df_{BM}/(df_{BM}-2))}}\right|\right] \quad (2)$$

5) ทำซ้ำขั้นที่ 3 และขั้นที่ 4 จำนวน 1,000 รอบ

6) คำนวณค่า P-value = $\dfrac{\text{number of times}\left(t^*_{MAX} \geq t_{MAX}\right)}{1{,}000}$

2. Brunner-Munzel test (BM test)

Brunner and Munzel [9] ได้นำเสนอสถิติทดสอบเพื่อแก้ปัญหาของเบห์เรนส์-ฟิชเชอร์ โดยอาศัยอันดับของค่าสังเกตเพื่อคำนวณค่าสถิติทดสอบ กำหนดให้ r_{1i} คือ อันดับของค่าสังเกตจากตัวอย่างที่ 1 และ r_{2j} คือ อันดับของค่าสังเกตจากตัวอย่างที่ 2 ถ้ามีค่าสังเกตซ้ำกันจะหาค่าเฉลี่ยอันดับร่วมกัน ทั้งนี้เรียงอันดับค่าสังเกตร่วมกันทั้งสองกลุ่มจากน้อยไปมาก และกำหนดให้ v_{1i} อันดับของค่าสังเกตจากตัวอย่างที่ 1 และ v_{2j} คือ อันดับของค่าสังเกตจากตัวอย่างที่ 2 ถ้ามีค่าสังเกตซ้ำกันจะหาค่าเฉลี่ยอันดับร่วมกัน ทั้งนี้เรียงอันดับค่าสังเกตภายในแต่ละตัวอย่าง โดยสามารถคำนวณสถิติทดสอบ ดังนี้

$$t_{BM} = \frac{\bar{R}_1 - \bar{R}_2}{(n_1+n_2)\sqrt{\frac{SB_1^2}{n_1 n_2^2} + \frac{SB_2^2}{n_1^2 n_2}}} \quad (3)$$

เมื่อ $\bar{R}_1 = \frac{\sum_{i=1}^{n_1} r_{1i}}{n_1}$, $\bar{R}_2 = \frac{\sum_{j=1}^{n_2} r_{2j}}{n_2}$

$$SB_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(r_{1i} - v_{1i} - \bar{R}_1 + \frac{n_1+1}{2}\right)^2,$$

$$SB_2^2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}\left(r_{2j} - v_{2j} - \bar{R}_2 + \frac{n_2+1}{2}\right)^2$$

จะปฏิเสธ H_0 เมื่อ $t_{BM} \le -t_{\alpha/2,df}$ หรือ $t_{BM} \ge t_{\alpha/2,df}$ โดย $df = \frac{\left(\frac{SB_1^2}{n_1} + \frac{SB_2^2}{n_2}\right)^2}{\frac{SB_1^4}{n_1^2(n_2-1)} + \frac{SB_2^4}{n_2^2(n_1-1)}}$

3. Welch Based on Rank test (WBR test)

สถิติทดสอบ Welch Based on Rank อาศัยอันดับของค่าสังเกตเพื่อคำนวณค่าสถิติทดสอบ โดยต้องเรียงอันดับของค่าสังเกตร่วมกันทั้งสองกลุ่มจากน้อยไปมาก กำหนดให้ r_{1i} คือ อันดับของค่าสังเกตจากตัวอย่างที่ 1 และ r_{2j} คือ อันดับของค่าสังเกตจากตัวอย่างที่ 2 ถ้ามีค่าสังเกตซ้ำกันจะหาค่าเฉลี่ยอันดับร่วมกัน โดยสามารถคำนวณสถิติทดสอบ ดังนี้ [14 -16]

$$t_{WBR} = \frac{\bar{R}_1 - \bar{R}_2}{\sqrt{\frac{S_{R_1}^2}{n_1} + \frac{S_{R_2}^2}{n_2}}} \quad (4)$$

เมื่อ $\bar{R}_1 = \frac{\sum_{i=1}^{n_1} r_{1i}}{n_1}$, $\bar{R}_2 = \frac{\sum_{j=1}^{n_2} r_{2j}}{n_2}$

$$S_{R_1}^2 = \frac{\sum_{i=1}^{n_1}\left(r_{1i} - \bar{R}_1\right)}{n_1 - 1}\ ,\ S_{R_2}^2 = \frac{\sum_{j=1}^{n_2}\left(r_{2j} - \bar{R}_2\right)}{n_2 - 1}$$

จะปฏิเสธ H_0 เมื่อ $t_{WBR} \le -t_{\alpha/2,df}$ หรือ $t_{WBR} \ge t_{\alpha/2,df}$ โดย $df = \frac{\left(\frac{S_{R_1}^2}{n_1} + \frac{S_{R_2}^2}{n_2}\right)^2}{\frac{\left(S_{R_1}^2/n_1\right)^2}{n_1-1} + \frac{\left(S_{R_2}^2/n_2\right)^2}{n_2-1}}$

4. Yuen-Welch test (YW test)

สถิติทดสอบ Yuen-Welch นำเสนอโดย Yuen [17] โดยอาศัยค่าเฉลี่ยแบบตัดปลาย (trimmed means) ทั้งนี้ การตัดค่าสังเกต 20% หรือ $\gamma = 0.20$ มีประสิทธิภาพดีที่สุด [18 - 19] โดยกำหนดจำนวนค่าสังเกตที่ตัดออก แทนด้วย $h_1 = n_1 - 2\gamma n_1$ และ $h_2 = n_2 - 2\gamma n_2$ สำหรับการคำนวณค่าความคลาดเคลื่อนมาตรฐานบนรากฐานของความแปรปรวนวินเซอร์ไรซ์ (winsorized variance) กำหนดให้ $W_1 = w_{11}, w_{12},..., w_{1n_1}$ และ $W_2 = w_{21}, w_{22},..., w_{2n_2}$ โดยสามารถคำนวณสถิติทดสอบ ดังนี้

$$t_{YW} = \frac{\bar{X}_{\gamma 1} - \bar{X}_{\gamma 2}}{\sqrt{d_1 + d_2}} \quad (5)$$

เมื่อ $d_1 = \frac{SW_1^2 (n_1 - 1)}{h_1 (h_1 - 1)}$

$$d_2 = \frac{SW_2^2 (n_2 - 1)}{h_2 (h_2 - 1)}$$

$$SW_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (w_{1i} - \bar{W}_1)^2$$

$$SW_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (w_{2j} - \bar{W}_2)^2$$

จะปฏิเสธ H_0 เมื่อ $t_{YW} \le -t_{\alpha/2,df}$ หรือ $t_{YW} \ge t_{\alpha/2,df}$ โดย $df = \frac{(d_1 + d_2)^2}{\left(\frac{d_1^2}{h_1 - 1} + \frac{d_2^2}{h_2 - 1} \right)}$

3. วิธีการดำเนินการวิจัย

การศึกษาประสิทธิภาพของการทดสอบสูงสุดไม่อิงพารามิเตอร์เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่มเมื่อมีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน จำลองแบบปัญหาด้วยเทคนิคมอนติคาร์โล โดยโปรแกรม R เวอร์ชั่น 3.5.3 มีรายละเอียด ดังนี้

1. จำลองค่าสังเกตแต่ละประชากร จำนวน 100,000 หน่วย กำหนดให้ประชากรทั้งสองกลุ่มมีการแจกแจงเดียวกัน คือ การแจกแจงล็อกนอร์มัล และการแจกแจงแบบเลขชี้กำลัง และประชากรทั้งสองมีการแจกแจงต่างกัน คือ ประชากรที่ 1 มีการแจกแจงแบบเลขชี้กำลัง และประชาการที่ 2 มีการแจกแจงล็อกนอร์มัล โดยมีรายละเอียด ดังนี้

1) ประชากรที่ 1 และ 2 มีการแจกแจงล็อกนอร์มัล

$$X_1 \sim \log \text{Norm}(\mu = 0, \sigma^2 = 1) - 1 \quad (6)$$

$$X_2 \sim \sigma_2 (\log \text{Norm}(\mu = 0, \sigma^2 = 1) - 1 - \text{shift}) \quad (7)$$

2) ประชากรที่ 1 และ 2 มีการแจกแจงแบบเลขชี้กำลัง

$$X_1 \sim \exp(\lambda = 2) - \frac{\ln(2)}{2} \quad (8)$$

$$X_2 \sim \sigma_Y (\exp(\lambda = 2) - \frac{\ln(2)}{2} - \text{shift}) \quad (9)$$

3) ประชากรที่ 1 มีการแจกแจงแบบเลขชี้กำลัง และประชากรที่ 2 มีการแจกแจงล็อกนอร์มัล

$$X_1 \sim \exp(\lambda = 1) - \frac{\ln(2)}{1} \quad (10)$$

$$X_2 \sim \sigma_2 (\log \text{Norm}(\mu = 0, \sigma^2 = 0.695) - 1 - \text{shift}) \quad (11)$$

ถ้ากำหนดให้ shift = 0 แสดงว่า มัธยฐานทั้ง 2 ประชากรเท่ากัน คือ การคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ถ้า shift = 1, ln(2)/2 และ 1 ตามลำดับ แสดงว่ามัธยฐานทั้ง 2 ประชากรต่างกัน คือ การคำนวณกำลังการทดสอบ

2. กำหนดอัตราส่วนความแปรปรวน 6 ขนาด ประกอบด้วย 1:1, 1:2, 1:4, 1:9, 1:12 และ 1:16

3. กำหนดขนาดตัวอย่าง 7 ขนาด คือ (10, 10), (10, 20), (20, 10), (20, 20), (10, 30), (30, 10), และ (30, 30)

4. คำนวณตัวสถิติทดสอบสมมติฐานความแตกต่างระหว่างค่ากลางของประชากร ประกอบด้วย Maximum test, Brunner-Munzel test, Welch based on rank test และ Yuen-Welch test

5. คำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยพิจารณาสัดส่วนจำนวนการปฏิเสธ H_0 เมื่อ H_0 เป็นจริง สำหรับการพิจารณาความแกร่งใช้เกณฑ์ของ Bradley [13] คือ ที่ระดับนัยสำคัญ 0.05 ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ระหว่าง 0.025-0.075 และที่ระดับนัยสำคัญ 0.01 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ระหว่าง 0.005-0.015 ถือว่าสามารถควบคุมความผิดพลาดแบบที่ 1 ได้ หรือตัวสถิติทดสอบมีความแกร่ง และการคำนวณกำลังการทดสอบโดยพิจารณาสัดส่วนจำนวนการปฏิเสธ H_0 เมื่อ H_0 ไม่จริง ทั้งนี้จำลองข้อมูลแต่ละสถานการณ์จำนวน 10,000 รอบ

6. นำเสนอตัวสถิติทดสอบที่มีประสิทธิภาพ โดยพิจารณาจากความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ตามเกณฑ์ของ Bradley [13] และมีกำลังการทดสอบสูงสุด ภายใต้เงื่อนไขการจำลองแบบปัญหาแต่ละสถานการณ์

4. ผลการวิจัย

4.1 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ เมื่อประชากรมีการแจกแจงล็อกนอร์มัล รายละเอียดดังตารางที่ 1 และตารางที่ 2 ตามลำดับจากตารางที่ 1 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 29 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 69.5 เมื่อเปรียบเทียบกับ Brunner-Munzel test และ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Brunner-Munzel test และ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ทุกเงื่อนไข โดย Welch based on rank test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้น้อยที่สุด จำนวน 26 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 61.90ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 22 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 52.38 เมื่อเปรียบเทียบกับ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 31 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 73.81 โดย Welch based on rank test และ Brunner-Munzel test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เพียง 11 เงื่อนไข และ 3 เงื่อนไข คิดเป็นร้อยละ 26.19 และ 7.14 ตามลำดับ ทั้งนี้เมื่อพิจารณาอิทธิพลของอัตราส่วนของความแปรปรวนที่ต่างกันที่ระดับนัยสำคัญ 0.05 พบว่า ไม่มีอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Brunner-Munzel test และ Yuen-Welch test สำหรับ Maximum test และ Welch based on rank test ได้รับอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 เล็กน้อย สำหรับที่ระดับนัยสำคัญ 0.01 พบว่า Yuen-Welch test และ Maximum test มีความแกร่งกว่า Brunner-Munzel test และ Welch based on rank test โดย Yuen-Welch test ค่อนข้างมีความแกร่งสูงกว่าเมื่อความแปรปรวนมีอัตราส่วนความแตกต่างกันค่อนข้างสูง

ตารางที่ 1 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 10	1:1	**0.0516**	**0.0110**	**0.0566**	0.0192	**0.0542**	**0.0138**	**0.0327**	**0.0054**
	1:2	**0.0562**	**0.0124**	**0.0584**	0.0190	**0.0578**	**0.0126**	**0.0345**	**0.0061**
	1:4	**0.0631**	**0.0136**	**0.0636**	0.0197	**0.0684**	0.0157	**0.0413**	**0.0069**
	1:9	**0.0721**	0.0180	**0.0632**	0.0227	0.0790	0.0229	**0.0544**	**0.0142**
	1:12	**0.0737**	0.0184	**0.0641**	0.0231	0.0824	0.0240	**0.0589**	0.0181
	1:16	0.0776	0.0156	**0.0624**	0.0206	0.0857	0.0220	**0.0623**	0.0182
10, 20	1:1	**0.0502**	**0.0108**	**0.0558**	0.0155	**0.0517**	**0.0117**	**0.0432**	**0.0082**
	1:2	**0.0555**	**0.0127**	**0.0596**	0.0188	**0.0593**	0.0159	**0.0401**	**0.0053**
	1:4	**0.0624**	**0.0126**	**0.0627**	0.0176	**0.0682**	0.0175	**0.0336**	**0.0052**
	1:9	**0.0734**	**0.0131**	**0.0619**	0.0171	0.0835	0.0200	**0.0373**	**0.0064**
	1:12	0.0774	**0.0142**	**0.0621**	0.0170	0.0869	0.0229	**0.0397**	**0.0087**
	1:16	0.0848	0.0163	**0.0660**	0.0180	0.0935	0.0267	**0.0434**	**0.0074**
20, 10	1:1	**0.0547**	**0.0106**	**0.0588**	0.0161	**0.0555**	**0.0128**	**0.0437**	**0.0087**
	1:2	**0.0576**	**0.0147**	**0.0614**	0.0186	**0.0602**	0.0162	**0.0526**	**0.0120**
	1:4	**0.0599**	**0.0142**	**0.0593**	0.0184	**0.0645**	0.0167	**0.0616**	0.0168
	1:9	**0.0661**	0.0168	**0.0605**	0.0183	**0.0702**	0.0212	**0.0658**	0.0206
	1:12	**0.0688**	0.0197	**0.0619**	0.0224	**0.0749**	0.0254	**0.0699**	0.0272
	1:16	**0.0676**	0.0194	**0.0585**	0.0205	**0.0720**	0.0231	**0.0670**	0.0251
20, 20	1:1	**0.0519**	**0.0121**	**0.0537**	**0.0131**	**0.0500**	**0.0111**	**0.0371**	**0.0051**
	1:2	**0.0595**	**0.0124**	**0.0605**	**0.0139**	**0.0591**	**0.0121**	**0.0377**	**0.0061**
	1:4	**0.0716**	0.0179	**0.0672**	0.0198	**0.0740**	0.0198	**0.0439**	**0.0081**
	1:9	0.0799	0.0200	**0.0640**	0.0175	0.0814	0.0227	**0.0432**	**0.0124**
	1:12	0.0761	0.0181	**0.0609**	0.0154	0.0803	0.0204	**0.0451**	**0.0108**
	1:16	0.0872	0.0227	**0.0647**	0.0176	0.0929	0.0258	**0.0503**	**0.0123**
10, 30	1:1	**0.0525**	**0.0102**	**0.0574**	0.0163	**0.0551**	**0.0129**	**0.0566**	**0.0141**
	1:2	**0.0563**	**0.0107**	**0.0585**	0.0167	**0.0581**	**0.0146**	**0.0426**	**0.0068**
	1:4	**0.0646**	**0.0104**	**0.0661**	0.0160	**0.0732**	0.0160	**0.0415**	**0.0055**
	1:9	0.0824	**0.0142**	**0.0686**	0.0179	0.0925	0.0255	**0.0391**	**0.0054**
	1:12	0.0848	**0.0136**	**0.0661**	0.0175	0.0963	0.0250	**0.0401**	**0.0062**
	1:16	0.0922	0.0170	**0.0693**	0.0183	0.1021	0.0306	**0.0437**	**0.0071**
30, 10	1:1	**0.0518**	**0.0118**	**0.0583**	0.0184	**0.0548**	**0.0150**	**0.0566**	**0.0123**
	1:2	**0.0555**	0.0156	**0.0594**	0.0209	**0.0594**	0.0179	**0.0671**	0.0224
	1:4	**0.0587**	**0.0141**	**0.0598**	0.0171	**0.0635**	0.0165	**0.0632**	0.0236
	1:9	**0.0609**	0.0162	**0.0587**	0.0176	**0.0666**	0.0202	**0.0656**	0.0268
	1:12	**0.0642**	0.0180	**0.0599**	0.0205	**0.0697**	0.0228	**0.0653**	0.0253
	1:16	**0.0624**	0.0171	**0.0571**	0.0200	**0.0678**	0.0227	**0.0659**	0.0233

ตารางที่ 1 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
30, 30	1:1	**0.0527**	**0.0120**	**0.0526**	**0.0126**	**0.0499**	**0.0106**	**0.0448**	**0.0062**
	1:2	**0.0504**	**0.0129**	**0.0552**	0.0163	**0.0534**	**0.0145**	**0.0480**	**0.0113**
	1:4	0.0759	0.0196	**0.0661**	0.0194	0.0760	0.0212	**0.0427**	**0.0081**
	1:9	0.0913	0.0261	**0.0726**	0.0183	0.0933	0.0276	**0.0472**	**0.0100**
	1:12	0.0931	0.0253	**0.0705**	0.0177	0.0953	0.0264	**0.0481**	**0.0100**
	1:16	0.0968	0.0284	**0.0718**	0.0182	0.0989	0.0292	**0.0489**	**0.0098**

ตัวเข้ม คือ สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

จากตารางที่ 2 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 3 เงื่อนไข จาก 42 เงื่อนไข เมื่อเปรียบเทียบกับ Welch based on rank test และ Brunner-Munzel test พบว่า มีประสิทธิภาพต่ำกว่า โดย Welch based on rank test และ Brunner-Munzel test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 26 เงื่อนไข และ 12 เงื่อนไข คิดเป็นร้อยละ 61.9 และ 28.57 ตามลำดับ

ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 9 เงื่อนไข จาก 42 เงื่อนไข เมื่อเปรียบเทียบกับ Yuen-Welch test และWelch Based on Rank test พบว่า มีประสิทธิภาพต่ำกว่า โดย Yuen-Welch test และ Welch Based on Rank test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 13 เงื่อนไข และ 10 เงื่อนไข คิดเป็นร้อยละ 30.95 และ 23.81 ตามลำดับ

ทั้งนี้เมื่อพิจารณาอัตราส่วนความแตกต่างของความแปรปรวนระดับเดียวกันในแต่ละขนาดตัวอย่าง พบว่า เมื่อขนาดตัวอย่างสูงขึ้นมีอิทธิพลต่อกำลังการทดสอบทำให้สูงขึ้น ทั้งนี้กรณีขนาดตัวอย่างไม่เท่ากัน กลุ่มที่มีขนาดตัวอย่างเล็กกว่าและความแปรปรวนสูงกว่าอีกกลุ่มจะมีกำลังการทดสอบที่ต่ำกว่ากลุ่มที่มีขนาดตัวอย่างใหญ่กว่าและมีความแปรปรวนน้อยกว่าอีกกลุ่ม

ตารางที่ 2 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 10	1:1	0.5227	0.2713	0.5115	0.3145	**0.5469**	**0.3065**	0.3969	0.1565
	1:2	0.5620	0.2991	0.5246	0.3280	**0.5847**	**0.3436**	0.4958	0.2571
	1:4	0.5509	0.2833	0.4853	0.2964	**0.5752**	0.3312	0.5178	**0.3193**
	1:9	**0.5432**	0.2781	0.4469	0.2765	0.5629	0.3233	0.5253	**0.3387**
	1:12	**0.5447**	0.2700	0.4324	0.2613	0.5622	0.3111	0.5311	0.3464
	1:16	0.5486	0.2677	0.4247	0.2554	0.5621	0.3082	**0.5295**	0.3436

ตารางที่ 2 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 20	1:1	0.7719	0.4919	0.7689	0.5618	**0.7901**	**0.5653**	0.4814	0.1897
	1:2	0.8191	**0.5464**	0.7972	0.5909	**0.8358**	0.6346	0.6218	0.3417
	1:4	0.8391	**0.5900**	0.7978	0.5991	**0.8540**	0.6797	0.7166	0.4840
	1:9	**0.8318**	**0.5905**	0.7746	0.5646	0.8483	0.6775	0.7374	0.5525
	1:12	0.8359	**0.5975**	**0.7714**	0.5627	0.8528	0.6850	0.7451	0.5696
	1:16	0.8393	0.6007	**0.7660**	0.5605	0.8555	0.6842	0.7568	0.5853
20, 10	1:1	0.5544	0.3053	0.5486	0.3266	**0.5697**	**0.3375**	0.5251	0.3280
	1:2	0.5468	0.3051	0.5246	0.3122	**0.5666**	0.3424	0.5392	**0.3624**
	1:4	0.5444	**0.2938**	0.5032	0.2925	**0.5610**	0.3316	0.5323	0.3594
	1:9	0.5306	0.2863	0.4608	0.2659	**0.5442**	0.3175	0.5261	0.3394
	1:12	0.5287	0.2838	0.4474	0.2572	**0.5437**	0.3120	0.5228	0.3418
	1:16	0.5533	0.2895	0.4635	0.2610	**0.5664**	0.3175	0.5454	0.3444
20, 20	1:1	0.8356	0.6238	0.8209	0.6249	**0.8406**	**0.6488**	0.6869	0.4377
	1:2	0.8521	0.6573	0.8261	0.6350	**0.8591**	**0.6870**	0.7368	0.5416
	1:4	0.8520	0.6624	0.8126	0.6143	**0.8582**	0.6882	0.7507	**0.5799**
	1:9	0.8479	0.6583	**0.7990**	0.5922	0.8553	0.6845	0.7644	**0.5994**
	1:12	0.8361	0.6467	**0.7796**	0.5705	0.8416	0.6712	0.7543	**0.5877**
	1:16	0.8371	0.6416	**0.7758**	0.5557	0.8419	0.6663	0.7565	**0.5867**
10, 30	1:1	0.8874	0.6546	0.8845	0.7256	**0.8994**	**0.7347**	0.5489	0.2249
	1:2	0.9194	0.7275	0.9038	0.7695	**0.9299**	**0.8077**	0.7208	0.4222
	1:4	0.9409	**0.7724**	0.9182	0.7840	**0.9507**	0.8518	0.8190	0.5933
	1:9	0.9430	**0.7899**	**0.9087**	0.7682	0.9505	0.8626	0.8614	0.6964
	1:12	0.9502	**0.8004**	**0.9120**	0.7724	0.9565	0.8753	0.8736	0.7232
	1:16	0.9445	0.7976	**0.9018**	0.7629	0.9525	0.8693	0.8743	**0.7351**
30, 10	1:1	0.5485	0.3058	0.5473	0.3220	**0.5692**	0.3385	0.5475	**0.3706**
	1:2	0.5522	0.3026	0.5366	0.3107	**0.5704**	0.3391	0.5494	0.3747
	1:4	0.5467	**0.3051**	0.5166	0.2973	**0.5661**	0.3366	0.5454	0.3731
	1:9	0.5292	0.2849	0.4774	0.2679	**0.5435**	0.3126	0.5273	0.3433
	1:12	0.5344	0.2794	0.4733	0.2556	**0.5509**	0.3049	0.5386	0.3480
	1:16	0.5261	0.2732	0.4587	0.2527	**0.5389**	0.2963	0.5222	0.3268
30, 30	1:1	0.9551	0.8534	0.9472	0.8442	**0.9571**	**0.8614**	0.8497	0.6783
	1:2	0.9539	0.8566	0.9394	0.8263	**0.9571**	**0.8672**	0.8622	0.7179
	1:4	0.9548	0.8570	**0.9326**	0.8166	0.9564	0.8688	0.8708	**0.7361**
	1:9	0.9558	0.8664	**0.9325**	0.8046	0.9575	0.8752	0.8898	**0.7609**
	1:12	0.9492	0.8442	**0.9177**	0.7748	0.9495	0.8529	0.8747	**0.7462**
	1:16	0.9439	0.8369	**0.9087**	0.7686	0.9463	0.8485	0.8766	**0.7483**

ตัวเข้ม คือ กำลังการทดสอบสูงสุดภายใต้สถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

4.2 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ เมื่อประชากรมีการแจกแจงแบบเลขชี้กำลัง รายละเอียดดังตารางที่ 3 และตารางที่ 4 ตามลำดับ

จากตารางที่ 3 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 32 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 76.19 เมื่อเปรียบเทียบกับ Brunner-Munzel test และ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Brunner-Munzel test และ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ทุกเงื่อนไข โดย Welch based on rank test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้น้อยที่สุด จำนวน 25 เงื่อนไข จาก 42 เงื่อนไข คิดเป็นร้อยละ 59.52

ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 23 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 54.76 เมื่อเปรียบเทียบกับ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 28 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 66.67 โดย Welch Based on Rank test และ Brunner-Munzel test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้เพียง 11 เงื่อนไข และ 9 เงื่อนไข คิดเป็นร้อยละ 26.19 และ 21.43 ตามลำดับ

ทั้งนี้เมื่อพิจารณาอิทธิพลของอัตราส่วนของความแปรปรวนที่ต่างกัน ที่ระดับนัยสำคัญ 0.05 พบว่า ไม่มีอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Brunner-Munzel test และ Yuen-Welch test สำหรับ Maximum test และ Welch Based on Rank test ได้รับอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 เล็กน้อย สำหรับที่ระดับนัยสำคัญ 0.01 พบว่า Yuen-Welch test และ Maximum test มีความแกร่งกว่า Brunner-Munzel test และ Welch Based on Rank test โดย Yuen-Welch test ค่อนข้างมีความแกร่งสูงกว่าเมื่อความแปรปรวนมีอัตราส่วนความแตกต่างกันค่อนข้างสูง

ตารางที่ 3 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 10	1:1	**0.0485**	**0.0108**	**0.0541**	0.0163	**0.0510**	**0.0109**	**0.0347**	**0.0067**
	1:2	**0.0570**	**0.0128**	**0.0602**	0.0191	**0.0599**	**0.0142**	**0.0426**	**0.0080**
	1:4	**0.0582**	**0.0140**	**0.0576**	0.0201	**0.0651**	0.0170	**0.0481**	**0.0108**
	1:9	**0.0682**	**0.0148**	**0.0585**	0.0198	0.0754	0.0203	**0.0618**	0.0157
	1:12	**0.0697**	0.0182	**0.0561**	0.0227	0.0781	0.0249	**0.0648**	0.0211
	1:16	**0.0706**	0.0165	**0.0553**	0.0209	0.0798	0.0233	**0.0649**	0.0216
10, 20	1:1	**0.0516**	**0.0117**	**0.0560**	0.0179	**0.0533**	**0.0139**	**0.0488**	**0.0120**
	1:2	**0.0541**	**0.0124**	**0.0582**	0.0173	**0.0584**	**0.0138**	**0.0427**	**0.0070**
	1:4	**0.0638**	**0.0119**	**0.0619**	0.0176	**0.0698**	0.0173	**0.0408**	**0.0079**
	1:9	**0.0742**	**0.0134**	**0.0596**	0.0163	0.0825	0.0210	**0.0424**	**0.0082**
	1:12	**0.0733**	**0.0140**	**0.0549**	0.0161	0.0836	0.0224	**0.0441**	**0.0102**
	1:16	0.0770	**0.0147**	**0.0539**	0.0159	0.0844	0.0247	**0.0465**	**0.0116**
20, 10	1:1	**0.0485**	**0.0113**	**0.0536**	0.0158	**0.0496**	**0.0124**	**0.0478**	**0.0102**
	1:2	**0.0583**	**0.0150**	**0.0631**	0.0218	**0.0609**	0.0176	**0.0607**	**0.0149**
	1:4	**0.0618**	0.0183	**0.0628**	0.0215	**0.0657**	0.0206	**0.0623**	0.0215
	1:9	**0.0634**	0.0176	**0.0574**	0.0210	**0.0689**	0.0224	**0.0666**	0.0231
	1:12	**0.0619**	0.0170	**0.0549**	0.0187	**0.0669**	0.0217	**0.0643**	0.0250
	1:16	**0.0664**	0.0199	**0.0554**	0.0215	**0.0708**	0.0243	**0.0675**	0.0228

ตารางที่ 3 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
20, 20	1:1	**0.0480**	**0.0106**	**0.0487**	**0.0124**	**0.0455**	**0.0102**	**0.0425**	**0.0073**
	1:2	**0.0536**	**0.0129**	**0.0525**	**0.0150**	**0.0541**	**0.0136**	**0.0438**	**0.0073**
	1:4	**0.0675**	0.0166	**0.0599**	0.0164	**0.0701**	0.0178	**0.0519**	**0.0102**
	1:9	**0.0730**	0.0185	**0.0571**	0.0157	0.0752	0.0207	**0.0517**	**0.0130**
	1:12	0.0808	0.0201	**0.0612**	0.0159	0.0853	0.0224	**0.0532**	**0.0135**
	1:16	0.0786	0.0210	**0.0589**	**0.0148**	0.0838	0.0241	**0.0577**	**0.0146**
10, 30	1:1	**0.0462**	**0.0093**	**0.0516**	**0.0144**	**0.0481**	**0.0124**	**0.0575**	0.0159
	1:2	**0.0529**	**0.0088**	**0.0551**	**0.0148**	**0.0558**	**0.0124**	**0.0507**	**0.0098**
	1:4	**0.0669**	**0.0100**	**0.0649**	**0.0146**	0.0754	0.0170	**0.0457**	**0.0092**
	1:9	0.0814	**0.0133**	**0.0644**	0.0173	0.0917	0.0258	**0.0471**	**0.0088**
	1:12	0.0859	**0.0140**	**0.0631**	**0.0148**	0.0984	0.0260	**0.0462**	**0.0076**
	1:16	0.0872	0.0159	**0.0600**	**0.0149**	0.0979	0.0287	**0.0512**	**0.0095**
30, 10	1:1	**0.0520**	**0.0112**	**0.0576**	0.0171	**0.0542**	**0.0136**	**0.0589**	0.0157
	1:2	**0.0561**	**0.0147**	**0.0589**	0.0196	**0.0583**	0.0173	**0.0650**	0.0225
	1:4	**0.0562**	**0.0141**	**0.0561**	0.0169	**0.0605**	0.0173	**0.0694**	0.0267
	1:9	**0.0598**	0.0157	**0.0561**	0.0167	**0.0647**	0.0191	**0.0656**	0.0251
	1:12	**0.0617**	0.0171	**0.0568**	0.0196	**0.0669**	0.0222	**0.0701**	0.0278
	1:16	**0.0668**	0.0168	**0.0583**	0.0188	**0.0728**	0.0224	**0.0710**	0.0264
30, 30	1:1	**0.0495**	**0.0105**	**0.0497**	**0.0106**	**0.0474**	**0.0088**	**0.0423**	**0.0065**
	1:2	**0.0670**	0.0158	**0.0657**	0.0174	**0.0668**	0.0168	**0.0478**	**0.0090**
	1:4	0.0797	0.0205	**0.0712**	0.0186	0.0822	0.0210	**0.0481**	**0.0107**
	1:9	0.0801	0.0215	**0.0611**	0.0166	0.0823	0.0233	**0.0494**	**0.0102**
	1:12	0.0813	0.0219	**0.0602**	0.0156	0.0833	0.0234	**0.0488**	**0.0111**
	1:16	0.0821	0.0228	**0.0563**	0.0153	0.0832	0.0234	**0.0477**	**0.0101**

ตัวเข้ม คือ สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

จากตารางที่ 4 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 7 เงื่อนไข จาก 42 เงื่อนไข เมื่อเปรียบเทียบกับ Welch Based on Rank test และ Brunner-Munzel test พบว่า มีประสิทธิภาพต่ำกว่า โดย Welch Based on Rank test และ Brunner-Munzel test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 25 เงื่อนไข และ 9 เงื่อนไข คิดเป็นร้อยละ 59.52 และ 21.43 ตามลำดับ

ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 9 เงื่อนไข จาก 42 เงื่อนไข เมื่อเปรียบเทียบกับ Yuen-Welch test และWelch Based on Rank test พบว่า มีประสิทธิภาพต่ำกว่า โดย Yuen-Welch test และ Welch Based on Rank test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 11 เงื่อนไขเท่ากัน คิดเป็นร้อยละ 26.19

ทั้งนี้เมื่อพิจารณาอัตราส่วนความแตกต่างของความแปรปรวนระดับเดียวกันในแต่ละขนาดตัวอย่าง พบว่า เมื่อขนาดตัวอย่างสูงขึ้นมีอิทธิพลต่อกำลังการทดสอบทำให้สูงขึ้น ทั้งนี้กรณีขนาดตัวอย่างไม่เท่ากัน กลุ่มที่มีขนาดตัวอย่างเล็กกว่าและความแปรปรวนสูงกว่าอีกกลุ่มจะมีกำลังการทดสอบที่ต่ำกว่ากลุ่มที่มีขนาดตัวอย่างใหญ่กว่าและมีความแปรปรวนน้อยกว่าอีกกลุ่ม

ตารางที่ 4 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	**0.01**	**0.05**	**0.01**	**0.05**	**0.01**	**0.05**	**0.01**
10, 10	1:1	0.4915	0.2515	0.4914	0.2991	**0.5069**	**0.2730**	0.3562	0.1325
	1:2	0.5222	0.2640	0.4976	0.2983	**0.5487**	**0.3055**	0.4477	0.2203
	1:4	0.5228	0.2668	0.4730	0.2849	**0.5470**	0.3124	0.4831	**0.2733**
	1:9	**0.5118**	**0.2591**	0.4286	0.2611	0.5319	0.3013	0.4874	0.2963
	1:12	**0.5235**	0.2652	0.4214	0.2618	0.5422	0.3054	0.5022	0.2984
	1:16	**0.5175**	0.2521	0.3979	0.2422	0.5343	0.2902	0.4954	0.2875
10, 20	1:1	0.7292	0.4541	0.7364	0.5239	**0.7455**	**0.5044**	0.4480	0.1681
	1:2	0.7785	0.5034	0.7641	0.5600	**0.7995**	**0.5837**	0.5989	0.3084
	1:4	0.7974	**0.5265**	0.7594	0.5481	**0.8141**	0.6175	0.6838	0.4359
	1:9	**0.8096**	**0.5505**	0.7492	0.5315	0.8278	0.6443	0.7305	0.5215
	1:12	**0.8006**	**0.5403**	0.7335	0.5118	0.8172	0.6314	0.7244	0.5151
	1:16	0.8065	**0.5544**	**0.7370**	0.5152	0.8250	0.6425	**0.7370**	0.5367
20, 10	1:1	0.5314	0.2875	0.5327	0.3141	**0.5466**	**0.3146**	0.4843	0.2857
	1:2	0.5316	0.2818	0.5177	0.2955	**0.5477**	0.3203	0.4958	**0.3116**
	1:4	0.5161	0.2789	0.4822	0.2782	**0.5335**	0.3164	0.4873	0.3085
	1:9	0.5141	0.2823	0.4559	0.2673	**0.5308**	0.3136	0.4973	0.3031
	1:12	0.5166	0.2770	0.4423	0.2599	**0.5281**	0.3091	0.4965	0.2928
	1:16	0.5117	0.2688	0.4220	0.2421	**0.5252**	0.2972	0.4955	0.3001
20, 20	1:1	0.8084	0.5846	0.8000	0.5995	**0.8115**	**0.5997**	0.6516	0.3883
	1:2	0.8385	0.6279	0.8158	0.6167	**0.8437**	**0.6588**	0.7313	0.5104
	1:4	0.8382	0.6212	0.7997	0.5861	**0.8443**	0.6522	0.7431	**0.5403**
	1:9	**0.8288**	0.6145	0.7723	0.5496	0.8372	0.6439	0.7531	**0.5522**
	1:12	0.8074	0.5952	**0.7444**	0.5263	0.8153	0.6191	0.7352	**0.5390**
	1:16	0.8052	0.5949	0.7391	0.5137	0.8149	0.6226	**0.7426**	**0.5404**
10, 30	1:1	0.8222	0.5550	0.8274	0.6431	**0.8372**	**0.6265**	0.4820	0.1803
	1:2	0.9020	0.6866	0.8930	0.7462	**0.9128**	**0.7726**	0.7111	0.3885
	1:4	**0.9202**	0.7252	0.8986	**0.7506**	0.9321	0.8118	0.8129	0.5718
	1:9	0.9303	**0.7505**	**0.8923**	0.7312	0.9404	0.8357	0.8589	0.6768
	1:12	0.9369	**0.7647**	**0.8934**	0.7400	0.9450	0.8465	0.8769	0.7166
	1:16	0.9231	0.7497	**0.8723**	**0.7139**	0.9334	0.8305	0.8593	0.7079

ตารางที่ 4 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	**0.01**	**0.05**	**0.01**	**0.05**	**0.01**	**0.05**	**0.01**
30, 10	1:1	0.5316	0.2894	0.5403	0.3157	**0.5526**	**0.3206**	0.5045	0.3211
	1:2	0.5292	**0.2906**	0.5186	0.3022	**0.5456**	0.3248	0.5111	0.3256
	1:4	0.5120	**0.2806**	0.4875	0.2841	**0.5316**	0.3152	0.4995	0.3125
	1:9	0.5164	0.2788	0.4677	0.2671	**0.5335**	0.3058	0.5052	0.3063
	1:12	0.4898	0.2615	0.4321	0.2436	**0.5031**	0.2866	0.4824	0.2820
	1:16	0.4985	0.2612	0.4284	0.2398	**0.5136**	0.2835	0.4927	0.2864
30, 30	1:1	0.9334	0.8054	0.9287	0.8062	**0.9360**	**0.8156**	0.8211	0.6144
	1:2	0.9478	0.8330	0.9371	0.8170	**0.9498**	0.8475	0.8611	**0.6974**
	1:4	0.9438	0.8394	**0.9221**	0.7974	0.9436	0.8491	0.8752	**0.7212**
	1:9	0.9397	0.8304	**0.9133**	0.7674	0.9416	0.8424	0.8866	**0.7469**
	1:12	0.9423	0.8282	**0.9112**	0.7617	0.9435	0.8405	0.8928	**0.7531**
	1:16	0.9358	0.8201	**0.8998**	0.7456	0.9374	0.8334	0.8886	**0.7426**

ตัวเข้ม คือ กำลังการทดสอบสูงสุดภายใต้สถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

4.3 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ เมื่อประชากรมีการแจกแจงแบบเลขชี้กำลัง และการแจกแจงล็อกนอร์มัล รายละเอียดดังตารางที่ 5 และตารางที่ 6 ตามลำดับ

จากตารางที่ 5 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 35 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 83.33 เมื่อเปรียบเทียบกับ Brunner-Munzel test และ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Brunner-Munzel test และ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ทุกเงื่อนไข โดย Welch Based on Rank test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้น้อยที่สุด จำนวน 28 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 66.67

ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 24 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 57.14 เมื่อเปรียบเทียบกับ Yuen-Welch test พบว่า มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า ซึ่ง Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 31 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 73.81 โดย Welch Based on Rank test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 13 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 30.95 สำหรับ Brunner-Munzel test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ 8 เงื่อนไขจาก 42 เงื่อนไข คิดเป็นร้อยละ 19.05

ทั้งนี้เมื่อพิจารณาอิทธิพลของอัตราส่วนของความแปรปรวนที่ต่างกัน ที่ระดับนัยสำคัญ 0.05 พบว่า ไม่มีอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Brunner-Munzel test และ Yuen-Welch test สำหรับ Maximum test และ Welch Based on Rank test ได้รับอิทธิพลของความแปรปรวนต่อความสามารถในการควบคุมความผิดพลาดแบบที่ 1 เล็กน้อย สำหรับที่ระดับนัยสำคัญ 0.01 พบว่า Yuen-Welch test และ Maximum test มีความแกร่งกว่า Brunner-Munzel test และ Welch Based on Rank test โดย Yuen-Welch test ค่อนข้างมีความแกร่งสูงกว่าเมื่อความแปรปรวนมีอัตราส่วนความแตกต่างกันค่อนข้างสูง

ตารางที่ 5 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง และการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 10	1:1	**0.0496**	**0.0074**	**0.0537**	**0.0139**	**0.0524**	**0.0093**	**0.0396**	**0.0062**
	1:2	**0.0572**	**0.0126**	**0.0618**	0.0197	**0.0598**	**0.0131**	**0.0403**	**0.0087**
	1:4	**0.0573**	**0.0140**	**0.0588**	0.0201	**0.0646**	0.0158	**0.0440**	**0.0110**
	1:9	**0.0620**	**0.0137**	**0.0561**	0.0192	**0.0673**	0.0186	**0.0533**	**0.0123**
	1:12	**0.0683**	0.0164	**0.0596**	0.0223	0.0753	0.0224	**0.0596**	0.0176
	1:16	0.0783	0.0180	**0.0562**	**0.0146**	0.0813	0.0212	**0.0506**	**0.0124**
10, 20	1:1	**0.0523**	**0.0116**	**0.0568**	0.0168	**0.0541**	**0.0129**	**0.0511**	**0.0115**
	1:2	**0.0570**	**0.0115**	**0.0610**	0.0181	**0.0597**	**0.0143**	**0.0420**	**0.0075**
	1:4	**0.0632**	**0.0119**	**0.0654**	0.0171	**0.0697**	0.0163	**0.0428**	**0.0070**
	1:9	**0.0677**	**0.0129**	**0.0591**	0.0160	**0.0743**	0.0194	**0.0431**	**0.0080**
	1:12	**0.0726**	**0.0145**	**0.0591**	0.0166	0.0829	0.0216	**0.0460**	**0.0086**
	1:16	**0.0720**	**0.0123**	**0.0547**	**0.0140**	0.0821	0.0216	**0.0419**	**0.0076**
20, 10	1:1	**0.0537**	**0.0122**	**0.0593**	0.0179	**0.0561**	**0.0141**	**0.0493**	**0.0102**
	1:2	**0.0543**	**0.0133**	**0.0600**	0.0177	**0.0568**	**0.0136**	**0.0488**	**0.0090**
	1:4	**0.0605**	0.0153	**0.062**0	0.0199	**0.0639**	0.0182	**0.0622**	0.0189
	1:9	**0.0641**	0.0173	**0.0599**	0.0197	**0.0700**	0.0212	**0.0656**	0.0219
	1:12	**0.0613**	0.0182	**0.0570**	0.0202	**0.0662**	0.0233	**0.0624**	0.0207
	1:16	**0.0717**	0.0182	**0.0627**	0.0205	0.0778	0.0241	**0.0724**	0.0243
20, 20	1:1	**0.0513**	**0.0105**	**0.0532**	**0.0131**	**0.0493**	**0.0102**	**0.0403**	**0.0063**
	1:2	**0.0629**	**0.0125**	**0.0627**	**0.0149**	**0.0603**	**0.0119**	**0.0421**	**0.0067**
	1:4	**0.0633**	**0.0133**	**0.0583**	0.0153	**0.0650**	0.0151	**0.0419**	**0.0079**
	1:9	**0.0717**	0.0183	**0.0583**	0.0174	**0.0740**	0.0210	**0.0502**	**0.0109**
	1:12	**0.0731**	0.0157	**0.0549**	**0.0136**	0.0760	0.0185	**0.0504**	**0.0116**
	1:16	0.0800	0.0195	**0.0593**	0.0152	0.0825	0.0235	**0.0529**	**0.0134**
10, 30	1:1	**0.0522**	**0.0099**	**0.0559**	0.0153	**0.0531**	**0.0127**	**0.0572**	0.0182
	1:2	**0.0541**	**0.0106**	**0.0585**	0.0163	**0.0577**	**0.0143**	**0.0534**	**0.0108**
	1:4	**0.0641**	**0.0089**	**0.0659**	0.0167	0.0697	0.0178	**0.0479**	**0.0094**
	1:9	**0.0730**	**0.0109**	**0.0619**	0.0161	0.0837	0.022	**0.0440**	**0.0084**
	1:12	0.0798	**0.0135**	**0.0633**	0.0162	0.0904	0.0233	**0.0474**	**0.0081**
	1:16	0.0872	**0.0137**	**0.0642**	0.0167	0.1007	0.0252	**0.0473**	**0.0082**
30, 10	1:1	**0.0550**	**0.0107**	**0.0609**	0.0176	**0.0573**	**0.0146**	**0.0595**	**0.0144**
	1:2	**0.0550**	**0.0147**	**0.0581**	0.0197	**0.0568**	0.0173	**0.0632**	0.0205
	1:4	**0.0582**	0.0156	**0.0604**	0.0198	**0.0630**	0.0201	**0.0639**	0.0206
	1:9	**0.0627**	0.0171	**0.0607**	0.0194	**0.0682**	0.0201	**0.0643**	0.0216
	1:12	**0.0655**	0.0176	**0.0610**	0.0198	**0.0705**	0.0215	**0.0634**	0.0216
	1:16	**0.0613**	0.0166	**0.0560**	0.0187	**0.0684**	0.0217	**0.0621**	0.0205

ตารางที่ 5 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง และการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
30, 30	1:1	**0.0531**	**0.0128**	**0.0542**	**0.0131**	**0.0519**	**0.0115**	**0.0470**	**0.0093**
	1:2	**0.0563**	0.0131	**0.0568**	**0.014**	**0.0565**	**0.0127**	**0.0441**	**0.0067**
	1:4	**0.0701**	0.0182	**0.0643**	0.0178	**0.0711**	0.0186	**0.0473**	**0.0108**
	1:9	0.0797	0.0211	**0.0654**	0.0158	0.0818	0.0234	**0.0457**	**0.0109**
	1:12	0.0877	0.0240	**0.0649**	0.0168	0.0899	0.0253	**0.0500**	**0.0105**
	1:16	0.0833	0.0233	**0.0624**	0.0151	0.085	0.0237	**0.0499**	**0.0109**

ตัวเข้ม คือ สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

จากตารางที่ 6 ที่ระดับนัยสำคัญ 0.05 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 8 เงื่อนไข จาก 42 เงื่อนไข คิดเป็นร้อยละ 19.05 เมื่อเปรียบเทียบกับ Brunner-Munzel test และ Yuen-Welch test พบว่า มีประสิทธิภาพสูงกว่า โดย Brunner-Munzel test และ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 2 เงื่อนไข และจำนวน 5 เงื่อนไข คิดเป็นร้อยละ 4.76 และ 11.90 ตามอันดับ ทั้งนี้ Welch Based on Rank test มีประสิทธิภาพสูงสุด คือ สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 28 เงื่อนไข คิดเป็นร้อยละ 66.67

ที่ระดับนัยสำคัญ 0.01 พบว่า Maximum test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 8 เงื่อนไข จาก 42 เงื่อนไข คิดเป็นร้อยละ 19.43 เมื่อเปรียบเทียบกับ Welch Based on Rank test และ Yuen-Welch test พบว่า มีประสิทธิภาพต่ำกว่า โดย Welch Based on Rank test และ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบสูงสุด จำนวน 13 เงื่อนไข และ 12 เงื่อนไข คิดเป็นร้อยละ 30.95 และ 28.57 ตามลำดับ ทั้งนี้เมื่อพิจารณาอัตราส่วนความแตกต่างของความแปรปรวนระดับเดียวกันในแต่ละขนาดตัวอย่าง พบว่า เมื่อขนาดตัวอย่างสูงขึ้นมีอิทธิพลต่อกำลังการทดสอบทำให้สูงขึ้น ทั้งนี้กรณีขนาดตัวอย่างไม่เท่ากัน กลุ่มที่มีขนาดตัวอย่างเล็กกว่าและความแปรปรวนสูงกว่าอีกกลุ่มจะมีกำลังการทดสอบที่ต่ำกว่ากลุ่มที่มีขนาดตัวอย่างใหญ่กว่าและมีความแปรปรวนน้อยกว่าอีกกลุ่ม

ตารางที่ 6 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง และการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 10	1:1	0.7655	0.5115	0.7559	0.5554	**0.7842**	**0.5579**	0.6611	0.3517
	1:2	0.7986	0.5350	0.7578	0.5521	**0.8175**	**0.5955**	0.7425	0.4937
	1:4	0.8012	0.5445	0.7350	0.5409	**0.8197**	0.6020	0.7758	**0.5599**
	1:9	0.8019	0.5356	0.6935	0.5045	**0.8131**	0.5882	0.7719	**0.5624**
	1:12	**0.7930**	0.5228	0.6676	0.4834	0.8030	0.5701	0.7774	0.5700
	1:16	0.7957	0.5308	0.6515	0.4856	0.8063	0.5748	**0.7771**	**0.5609**

ตารางที่ 6 กำลังการทดสอบของสถิติทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อมีการแจกแจงแบบเลขชี้กำลัง และการแจกแจงล็อกนอร์มัล จำแนกตามขนาดตัวอย่าง ความแปรปรวน และระดับนัยสำคัญ (ต่อ)

Sample size	Var. Ratio	Max		BM		WBR		YW	
		0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
10, 20	1:1	0.9563	0.8288	0.9533	0.8686	**0.9610**	**0.8708**	0.8205	0.4860
	1:2	0.9712	0.8645	0.9656	0.8873	**0.9760**	**0.9132**	0.9183	0.7134
	1:4	0.9744	0.8748	0.9632	0.8807	**0.9781**	0.9216	0.9499	**0.8314**
	1:9	0.9763	**0.8838**	0.9593	0.8690	**0.9794**	0.9274	0.9612	0.8792
	1:12	**0.9761**	**0.8867**	0.9587	0.8655	0.9801	0.9302	0.9661	0.8860
	1:16	**0.9792**	0.8958	0.9606	0.8699	0.9818	0.9360	0.9703	**0.8974**
20, 10	1:1	0.8069	0.5797	0.7989	0.5982	**0.8182**	**0.6174**	0.7691	0.5648
	1:2	0.8033	0.5752	0.7814	0.5728	**0.8170**	**0.6209**	0.7701	0.5771
	1:4	0.8018	0.5615	0.7552	0.5346	**0.8151**	0.6049	0.7791	0.5766
	1:9	0.7737	0.5329	0.6936	0.4817	**0.7832**	0.5714	0.7524	0.5386
	1:12	0.7894	0.5511	0.6922	0.4912	**0.7975**	0.5839	0.7757	0.5586
	1:16	**0.7824**	0.5403	0.6761	0.4804	0.7917	0.5735	0.7653	0.5529
20, 20	1:1	0.9770	0.9047	0.9733	0.9046	**0.9780**	**0.9150**	0.9379	0.8076
	1:2	0.9796	0.9163	0.9738	0.9059	**0.9815**	**0.9309**	0.9580	0.8647
	1:4	0.9806	**0.9195**	0.9724	0.8998	**0.9813**	0.9336	0.9648	0.8847
	1:9	0.9801	0.9177	0.9661	0.8811	**0.9823**	0.9283	0.9666	**0.8953**
	1:12	**0.9782**	0.9141	0.9642	0.8735	0.9798	0.9267	0.9686	**0.8964**
	1:16	0.9763	0.9115	0.9585	0.8659	0.9777	0.9243	**0.9655**	**0.8943**
10, 30	1:1	0.9864	0.9262	0.9857	0.9525	**0.9887**	**0.9560**	0.8655	0.5260
	1:2	0.9951	0.9621	0.9938	0.9721	**0.9963**	**0.9809**	0.9668	0.8058
	1:4	**0.9972**	**0.9713**	0.9944	0.9747	0.9982	0.9868	0.9896	0.9294
	1:9	**0.9981**	**0.9809**	0.9959	0.9759	0.9984	0.9912	0.9949	0.9736
	1:12	0.9979	**0.9787**	0.9948	0.9716	0.9981	0.9911	**0.9949**	0.9757
	1:16	0.9983	**0.9789**	0.9936	0.9676	0.9988	0.9903	**0.9950**	0.9757
30, 10	1:1	0.8163	0.5838	0.8091	0.6059	**0.8269**	**0.6277**	0.7858	0.6005
	1:2	0.8174	**0.5808**	0.7975	0.5838	**0.8294**	0.6292	0.7932	0.5987
	1:4	0.7921	0.5565	0.7565	0.5339	**0.8057**	0.5958	0.7780	0.5688
	1:9	0.7749	0.5376	0.7073	0.4888	**0.7854**	0.5699	0.7630	0.5444
	1:12	0.7781	0.5329	0.7011	0.4788	**0.7895**	0.5627	0.7661	0.5504
	1:16	0.7771	0.5292	0.6892	0.4686	**0.7901**	0.5535	0.7690	0.5440
30, 30	1:1	0.9982	0.9877	0.9973	0.9873	**0.9985**	**0.9900**	0.9906	0.9570
	1:2	**0.9989**	0.9918	0.9981	0.9897	**0.9989**	**0.9927**	0.9946	0.9763
	1:4	0.9989	0.9909	0.9983	0.9858	**0.9990**	0.9928	0.9962	**0.9800**
	1:9	0.9987	0.9905	**0.9971**	0.9806	0.9987	0.9915	0.9959	**0.9807**
	1:12	0.9987	0.9886	0.9967	0.9767	0.9987	0.9902	**0.9968**	**0.9792**
	1:16	0.9981	0.9885	**0.9960**	0.9755	0.9982	0.9895	0.9955	**0.9795**

ตัวเข้ม คือ กำลังการทดสอบสูงสุดภายใต้สถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

5. ตัวอย่าง

เพื่อความเข้าใจวิธีการวิเคราะห์ข้อมูลด้วย Maximum test ซึ่งเป็นวิธีการใหม่นำเสนอโดย Welz และคณะ [2] มีรายละเอียดการวิเคราะห์ข้อมูล ดังนี้ จากข้อมูลระยะเวลาการมีชีวิตรอดของผู้ป่วยมะเร็งรังไข่ และมะเร็งเต้านม ของ Hand และคณะ [20] ในตารางที่ 7 พบว่า ระยะเวลามีชีวิตของผู้ป่วยมะเร็งรังไข่ จำนวน 6 ราย มีค่าเฉลี่ย 884.333 วัน มัธยฐาน 406 วัน และความแปรปรวน 1,206,875.4667 มีความเบ้ 1.825 และความโด่ง 6.2018 และระยะเวลามีชีวิตของผู้ป่วยมะเร็งเต้านม จำนวน 11 ราย มีค่าเฉลี่ย 1,395.9091 วัน มัธยฐาน 1,166 วัน และความแปรปรวน 1,535,038.4909 ความเบ้ 1.0912 และความโด่ง 3.4628

ตารางที่ 7 ระยะเวลาการมีชีวิตรอดของผู้ป่วยมะเร็ง

มะเร็งรังไข่	1,234	89	201	356	2,970	456					
มะเร็งเต้านม	1,235	24	1,581	1,166	40	727	3,808	791	1,804	3,460	719

เมื่อพิจารณาระยะเวลาการมีชีวิตรอดระหว่างผู้ป่วยมะเร็งรังไข่ และมะเร็งเต้านม พบว่า ข้อมูลมีการแจกแจงแบบเบ้ขวา และมีความแปรปรวนต่างกันมาก นอกจากนี้มีค่าเฉลี่ยและมัธยฐานค่อนข้างต่างกัน เมื่อทดสอบสมมติฐานด้วย Welch Based on Rank test จะได้ $t_{WBR} = -1.0407$, $df = 11.468$, $P-value = 0.3194$ Brunner-Munzel test จะได้ $t_{BM} = 0.9950$, $df = 12.725$, $P-value = 0.3383$ และ Maximum test จะได้ $t_{MAX} = 0.9456$, $P-value = 0.3040$ ให้ผลไปในทิศทางเดียวกัน คือ ระยะเวลาการมีชีวิตรอดไม่แตกต่างกัน แต่จะเห็นได้ว่า Maximum test ให้ค่า P-value ที่ต่ำกว่าการทดสอบด้วย Brunner-Munzel test และ Welch Based on Rank test และเมื่อทดสอบสมมติฐานด้วย Yuen-Welch test จะได้ $t_{YW} = -1.482$, $df = 6.28$, $P-value\ 0.1867$ ซึ่งจะให้ค่า P-value ที่ต่ำกว่าสถิติทดสอบดังกล่าวข้างต้น สำหรับการวิเคราะห์ข้อมูลครั้งนี้ได้เขียนคำสั่งด้วยโปรแกรม R

6. วิจารณ์ผล

ผลการศึกษา พบว่า Brunner-Munzel test และ Yuen-Welch test เป็นสถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ทุกเงื่อนไขที่ระดับนัยสำคัญ 0.05 เมื่อประชากรมีการแจกแจงเดียวกันทั้ง 2 ประชากร (Identically distributed) คือ การแจกแจงล็อกนอร์มัล และการแจกแจงแบบเลขชี้กำลัง และประชากรทั้งสองมีการแจกแจงต่างกัน คือ ประชากรที่ 1 มีการแจกแจงแบบเลขชี้กำลัง และประชาการที่ 2 มีการแจกแจงล็อกนอร์มัล ทั้งนี้ที่ระดับนัยสำคัญ 0.01 พบว่า ความสามารถควบคุมความผิดพลาดแบบที่ 1 ลดลง แต่ Yuen-Welch test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ดีกว่าสถิติทดสอบอื่น ทั้งนี้อาจเนื่องมาจากการจำลองข้อมูลดังกล่าวมีการแจกแจงแบบเบ้ขวาในการคำนวณสถิติทดสอบจะมีการตัดข้อมูลส่วนปลายด้านละ 20% ซึ่งจะให้ประสิทธิภาพสูงสุด [18-19] ทำให้ข้อมูลที่มีค่าสุดขีดถูกตัดออกไป และมีการคำนวณค่าความคลาดเคลื่อนมาตรฐานด้วยวิธีวินเซอร์ไรซ์ (winsorized variance) จึงทำให้มีสถิติทดสอบดังกล่าวมีความไว (sensitive) ในการปฏิเสธสมมติฐานว่างค่อนข้างต่ำ ถึงแม้ว่าสถิติทดสอบดังกล่าวจะมีความแกร่งแต่เมื่อพิจารณากำลังการทดสอบจะค่อนข้างต่ำกว่าสถิติทดสอบในกลุ่มเดียวกันที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ สอดคล้องกับผลการวิจัยของ Fagerland และ Sandvik [3] กล่าวว่า Yuen-Welch test เหมาะสมกับการวิเคราะห์ข้อมูลที่มีการแจกแจงแบบเบ้ขวา สำหรับ Brunner-Munzel test เป็นสถิติทดสอบที่มีรากฐานมาจาก Wilcoxon-Mann-Whitney test และมีการปรับแก้เพื่อให้มีความแกร่งเมื่อเกิดปัญหาของเบห์เรนส์-ฟิชเชอร์ (Behrens-Fisher Problem) แต่ที่ระดับนัยสำคัญ 0.01 ค่อนข้างควบคุมความผิดพลาดแบบที่ 1 ได้ค่อนข้างต่ำ นอกจากนี้จากการศึกษา พบว่า เหมาะสมกับขนาดตัวอย่างตั้งแต่ 30 ขึ้นไป [21]

สำหรับ Maximum test อาศัยสถิติทดสอบที่มีประสิทธิภาพสองตัวสถิติมาพิจารณาร่วมกัน ทั้งนี้จะนำสถิติทดสอบดังกล่าวมาแปลงเป็นค่ามาตรฐาน (standardized) และประยุกต์ร่วม

กระบวนการทำซ้ำโดยการสุ่มแบบไม่คืนที่ (permutation) เพื่อคำนวณค่า P-value ซึ่งกระบวนการดังกล่าวเหมาะสมกับการทดสอบสมมติฐานเมื่อตัวอย่างมีขนาดเล็ก จากผลการวิจัยพบว่า ความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ค่อนข้างสูงกว่าสถิติทดสอบอื่นๆ ยกเว้น Yuen-Welch test ทั้งที่ระดับนัยสำคัญ 0.05 และ 0.01 และอัตราส่วนความแปรปรวนระดับต่างๆ กรณีที่ควบคุมไม่ได้ส่วนใหญ่แตกต่างจากเกณฑ์เล็กน้อย และมีกำลังการทดสอบที่ใกล้เคียงกันกับ Brunner-Munzel test และ Welch Based on Rank test สอดคล้องกับผลการวิจัยของ Welz และคณะ [2]

7. สรุป

การศึกษาประสิทธิภาพของการทดสอบสูงสุดไม่อิงพารามิเตอร์เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่มเมื่อมีการแจกแจงเบ้และความแปรปรวนไม่เท่ากัน สามารถสรุปผลการวิจัยได้ดังนี้

1. เมื่อประชากรทั้งสองกลุ่มมีการแจกแจงล็อกนอร์มัล และการแจกแจงแบบเลขชี้กำลัง พบว่า ที่ระดับนัยสำคัญ 0.05 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Brunner-Munzel test และ Yuen-Welch test สำหรับที่ระดับนัยสำคัญ 0.01 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Yuen-Welch test เมื่อพิจารณากำลังการทดสอบในกลุ่มที่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ พบว่า ที่ระดับนัยสำคัญ 0.05 คือ Welch Based on Rank test รองลงมา คือ Brunner-Munzel test และ Maximum test และที่ระดับนัยสำคัญ 0.01 คือ Yuen-Welch test รองลงมา Welch Based on Rank test และ Maximum test

2. เมื่อประชากรมีการแจกแจงแบบเลขชี้กำลังและการแจกแจงล็อกนอร์มัล พบว่า ที่ระดับนัยสำคัญ 0.05 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Brunner-Munzel test และ Yuen-Welch test แต่สูงกว่า Welch Based on Rank test สำหรับที่ระดับนัยสำคัญ 0.01 Maximum test มีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ต่ำกว่า Yuen-Welch test แต่สูงกว่า Brunner-Munzel test และ Welch Based on Rank test เมื่อพิจารณาสถิติทดสอบที่มีกำลังการทดสอบสูงสุดและสามารถควบคุมความผิดพลาดแบบที่ 1 พบว่า ที่ระดับนัยสำคัญ 0.05 คือ Welch Based on Rank test รองลงมา คือ Maximum test, Yuen-Welch test และ Brunner-Munzel test ตามลำดับ และที่ระดับนัยสำคัญ 0.01 คือ Welch Based on Rank test รองลงมา Yuen-Welch test และ Maximum test ตามลำดับ

8. เอกสารอ้างอิง

1. Nguyen, D.T., Kim, E.S., Gil, P.R., Kellermann, A., Chen, YH., Kromrey, J.D. and Bellara, A., 2016, "Parametric Tests for Two Population Means under Normal and Non-Normal Distribution," *Journal of Modern Applied Statistical Methods,* 15 (1), pp. 141-159.

2. Welz, A., Ruxton, G.D. and Neuhauser, M., 2018, "A non-parametric maximum test for the Behrens-Fisher problem," *Journal of Statistical Computation and Simulation*, DOI: 10.1080/00949655.2018.1431236.

3. Fagerland, M.W. and Sandvik, L., 2009, "Performance of five two-sample location tests for skewed distributions with unequal variances," *Contemporary Clinical Trial*, 30, pp. 490-496.

4. Neuhauser, M., 2012, Nonparametric Statistical Tests: A Computational Approach, CRC Press, Florida.

5. Mickelson, W.T., 2013, "A Monte Carlo Simulation of the Robust Rank-order Test under Various Population Symmetry Conditions," *Journal of Modern Applied Statistical Methods,* 12 (1), pp. 21-33.

6. Wilcoxon, F., 1945, "Individual Comparisons by Ranking Methods," *Biometrics*, 1, pp. 80-83.

7. Mann, H.B. and Whitney, D.R., 1947, "On a Test of Whether One of Two Random Variables is Stochastically Larger Than the Other," *Annals of Mathematical Statistics*, 18, pp. 50-60.

8. Stonehouse, J.M. and Forrester, G.J., 1998, "Robustness of the t and U Tests under Combined Assumption Violations," *Journal of Applied Statistics,* 25 (1), pp. 63-74.

9. Brunner, E. and Munzel, U., 2000, "The Non-

parametric Behrens-fisher Problem: Asymptotic Theory and a Small-sample Approximation," *Biometrical Journal*, 42 (1), pp. 17–25.

10. Medina, J.M., Kimberg, D.Y., Chatterjee, A. and Coslett, H.B., 2010, "Inappropriate Usage of the Brunner-Munzel Test in Recent Voxel-based Lesion-symtom Mapping Studies," *Neuropsychologia*, 48, pp. 341-343.

11. Neuhauser, M., 2015, "Combining the t test and Wilcoxon's Rank-sum Test," *Journal of Applied Statistics*, 42 (12), pp. 2769-2775.

12. Game, P.A., Winkler, H.B. and Probert, D.A., 1972, "Robust Tests for Homogeneity of Variance," *Educational and Psychological Measurement*, 32, pp. 877-909.

13. Bradley, J.V., 1978, "Robustness?," *British Journal of Mathematical and Statistical Psychology,* 31, pp. 144-152.

14. Welch, B.L., 1937, "The Significance of the Difference between Two Means when the Population Variances are Unequal," *Biometrika*, 29, pp. 350-362.

15. Reiczigel, J., Zakarias, I. and Rozsa, L., 2005, "A Bootstrap Test of Stochastic Equality of Two Populations," *The American Statistician*, 59 (2), pp. 1-6.

16. Winter, J.C.F., 2013, "Using the Student's t-test with Extremely Small Sample Sizes," *Practical Assessment. Research and Evaluation*, 18 (10), pp. 1-12.

17. Yuen, K.K., 1974, "The Two-sample Trimmed t for Unequal Population Variances," *Biometrika*, 61, pp. 165-170.

18. Wilcox, R.R., 1994, "Some Results on the Tukey-Mclaughlin and Yuen Methods for Trimmed Means when Distribution are Skewed," *Biometrical Journal*, 3, pp. 259-273.

19. Wilcox, R.R., 2005, Introduction to Robust Estimation and Hypothesis Testing, 2nd ed., Academic Press, San Diego, CA.

20. Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. and Ostrowski, E., 1994, A Handbook of Small Data Sets, Chapman and Hall, Boca Raton.

21. Reiczigel, J. Zakarias, I. and Rozsa, L., 2005, "A Bootstrap Test of Stochastic Equality of Two Populations," *The American Statistician*, 59 (2), pp. 1-6.