

การเปรียบเทียบประสิทธิภาพของการทดสอบซี การทดสอบที การทดสอบ
แบบสุ่ม และการทดสอบแมนท์-วิทนี ยู สำหรับทดสอบค่าเฉลี่ย
หรือค่ากลางของประชากร 2 กลุ่มที่เป็นอิสระกัน
An Efficiency Comparison of Z-Test, T-Test, the Randomization
Test and Mann-Whitney U Test for Testing Two Independent
Population Means or Medians

อัชมา อระวีพร โชติรส ชื่นอารมณ วาสิณี พัททรัพย์ สุปวีณ์ เจตจิตกุลศล อนพ เมืองมีทรัพย์
Autcha Araveeporn, Chotirot Chuenarrom, Wasinee Fuksub, Supawee Choedchitkusol,
and Anop Muangmeesup
ภาควิชาสถิติ คณะวิทยาศาสตร์สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

การวิจัยนี้เป็นการศึกษาเชิงจำลองมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของ
การทดสอบซี การทดสอบที การทดสอบแบบสุ่ม และการทดสอบแมนท์-วิทนี ยู สำหรับทดสอบค่ากลาง
หรือค่าเฉลี่ยของประชากร 2 กลุ่มที่เป็นอิสระกัน ในงานวิจัยนี้ศึกษาค่าเฉลี่ยและความแปรปรวนของทั้ง 2
ประชากร ใน 4 กรณีคือ 1) ค่าเฉลี่ยและความแปรปรวนเท่ากัน 2) ค่าเฉลี่ยเท่ากันและความแปรปรวนไม่
เท่ากัน 3) ค่าเฉลี่ยไม่เท่ากันและความแปรปรวนเท่ากัน 4) ค่าเฉลี่ยและความแปรปรวนไม่เท่ากัน โดยศึกษา
จากข้อมูลที่สุ่มมาจากประชากรที่มีการแจกแจงปกติ กำหนดขนาดตัวอย่าง (n_1, n_2) เท่ากัน คือ (5,5)
(15,15) และ (50,50) ขนาดตัวอย่าง (n_1, n_2) ไม่เท่ากัน (5,15) (15,25) และ (35,50) ในการคำนวณ
ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 กำหนดค่าเฉลี่ยของประชากรเท่ากัน (μ_1, μ_2) เท่ากับ
(9,9) และการคำนวณกำลังการทดสอบ กำหนดค่าเฉลี่ยของประชากรไม่เท่ากัน (μ_1, μ_2) เท่ากับ (9,12)
โดยกรณีที่ความแปรปรวนเท่ากันจะกำหนดค่าความแปรปรวน (σ^2) ของแต่ละประชากรคือ 3 9 27 และ
36 ตามลำดับ ส่วนกรณีที่ความแปรปรวนไม่เท่ากัน จะกำหนดค่าความแปรปรวนตามเกณฑ์ของค่านอนเซน
ทรัลลิตีพารามิเตอร์ ซึ่งได้ค่านอนเซนทรัลลิตีพารามิเตอร์ (ϕ) คือ 0.69 2.49 และ 6.93 ตามลำดับ
กำหนดระดับนัยสำคัญ 2 ระดับคือ 0.01 และ 0.05 โดยใช้โปรแกรมอาร์ในการจำลองและวิเคราะห์ข้อมูล
ทำการจำลองข้อมูลซ้ำ 1,000 รอบในแต่ละสถานการณ์ ผลการวิจัยพิจารณาจากกำลังการทดสอบสูงที่สุด
จากตัวสถิติสามารถควบคุมค่าความผิดพลาดประเภทที่ 1 เมื่อความแปรปรวนเท่ากัน ส่วนใหญ่การทดสอบ
ซี มีกำลังการทดสอบสูงที่สุดในกรณีที่ความแปรปรวนไม่เกิน 27 และการทดสอบแมนท์-วิทนี ยู มีกำลังการ
ทดสอบสูงที่สุดในกรณีที่ความแปรปรวนเป็น 36 และเมื่อความแปรปรวนไม่เท่ากันพบว่า ส่วนใหญ่การ
ทดสอบซี มีกำลังการทดสอบสูงที่สุดในทุกความแปรปรวน

คำสำคัญ : การทดสอบที การทดสอบซี การทดสอบแบบสุ่ม การทดสอบแมนท์-วิทนี ยู ค่าเฉลี่ย ค่ากลาง

Abstract

This research is a simulating research that aimed to study and to compare the efficiency of Z-test, t-test, the randomization test and Mann-Whitney U test for testing means or medians of two independent populations. This research is considered mean and variance of 2 populations in 4 cases such as 1) equal mean and variance 2) equal mean and unequal variance 3) unequal mean and equal variance, 4) unequal mean and variance. The data is generated in term of two populations and normal distribution. The sample sizes are set to equal sample size (n_1, n_2) as (5,5), (15,15), and (50,50), unequal sample size (n_1, n_2) as (5,15), (15,25), and (35,50). The population mean (μ_1, μ_2) are equal as (9,9) for calculating probability of type I error, and unequal as (μ_1, μ_2) to (9,12) for calculating power of a test. The population variances (σ^2) of each population are set to 3, 9, 27 and 36 for equal variance. The unequal variances based on noncentrality parameter (ϕ) that set as 0.69, 2.49, , and 6.93. The significant levels are considered on two levels at 0.01 and 0.05. R program is used for simulation and data analysis with 1,000 times for each situation. The results are considered the maximum power of a test that these tests can control the probability of type I error. For equal variance, Z-test shows the highest power of a test when variance does not exceed 27. Mann-Whitney U test shows the highest power of a test when variance is set as 36. When variance is not equal, Z-test shows the highest power of a test in all situations.

Keywords : t-test, Z-test, The Randomization test, Mann-Whitney U test, Mean, Median

1. บทนำ

ในการทดสอบสมมติฐานเพื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยหรือค่ากลางของประชากร 2 กลุ่มที่เป็นอิสระกัน โดยการทดสอบที่ใช้นิยมในงานในการวิเคราะห์ข้อมูล ได้แก่ สถิติอิงพารามิเตอร์ คือ การทดสอบซี (Z-test) การทดสอบที (t-test) โดยพิจารณาจากค่าความแปรปรวนของประชากรทั้ง 2 กลุ่มว่าเท่ากันหรือไม่เท่ากัน และเมื่อข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของสถิติอิงพารามิเตอร์จึงสามารถเลือกใช้สถิติไม่อิงพารามิเตอร์ คือการทดสอบแบบสุ่ม (The Randomization Test) และ การทดสอบแมนท์-วิทนี ยู (Mann-Whitney U Test)

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า มนตรี [1] ได้ศึกษาการทดสอบเพื่อเปรียบเทียบความแตกต่างของค่ากลางของประชากร 2 กลุ่ม พบว่า การทดสอบที่ สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้และมีกำลังการทดสอบสูงสุดสำหรับการแจกแจงเบ้ขวาและโค้งสูงกว่าปกติ และการทดสอบแมนท์-วิทนี สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้และมีกำลังการทดสอบสูงสุดสำหรับการแจกแจงเบ้ซ้ายและโค้งสูงกว่าปกติ มนตรี [2] ได้ศึกษาความแกร่งและอำนาจการทดสอบของการทดสอบอิงพารามิเตอร์และไม่

อิงพารามิเตอร์ พบว่าการทดสอบที่ โมดิฟายด์ ยู และการทดสอบแมนท์-วิทนี มีกำลังการทดสอบสูงสุดเมื่อตัวอย่างมีขนาดเล็ก กลาง และใหญ่ ส่วนการทดสอบที่ แลชี มีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างมีขนาดกลางและใหญ่ และความแปรปรวนไม่เท่ากัน Proshan และคณะ [3] ได้ศึกษาความสัมพันธ์ระหว่างการทดสอบแบบสุ่มและการทดสอบที่ พบว่า การทดสอบแบบสุ่มจะให้ผลการทดสอบใกล้เคียงกับการทดสอบที่เมื่อการแจกแจงแบบสุ่มเข้าใกล้การแจกแจงปกติ

จากที่ได้กล่าวมาข้างต้นผู้วิจัยจึงสนใจทำการตรวจสอบประสิทธิภาพของการทดสอบที่ การทดสอบซี การทดสอบแบบสุ่ม และการทดสอบแมนท์-วิทนี ยู สำหรับทดสอบสมมติฐานความแตกต่างระหว่างค่าเฉลี่ยหรือค่ากลางของประชากร 2 กลุ่มที่เป็นอิสระกัน จากการแจกแจงปกติ โดยใช้โปรแกรมอาร์ (R) [4] ในการจำลองและวิเคราะห์ข้อมูล

2. วิธีการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงทดลอง เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของการทดสอบที่ การทดสอบซี การทดสอบแบบสุ่ม และการทดสอบแมนท์-วิทนี ยู สำหรับทดสอบค่าเฉลี่ยหรือค่ากลางของประชากร 2 กลุ่มที่เป็นอิสระกัน

2.1 การวางแผนการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ในการศึกษาเปรียบเทียบดังนี้

2.1.1 กำหนดจำนวนประชากรเท่ากับ 2 ประชากร

2.1.2 กำหนดขนาดตัวอย่างสุ่ม (n_1, n_2) จากแต่ละประชากรเท่ากันและไม่เท่ากันดังนี้

ลักษณะขนาดตัวอย่าง	จำนวนกลุ่ม	ขนาดตัวอย่าง (n_1, n_2)
ขนาดตัวอย่างเท่ากัน	3	(5,5) (15,15) (50,50)
ขนาดตัวอย่างไม่เท่ากัน	3	(5,15) (15,25) (35,50)

2.1.3 ในกรณีที่ค่าความแปรปรวนทั้งสองกลุ่มเท่ากัน

การแจกแจงปกติ มีพารามิเตอร์ (μ, σ^2) คือ (9,3) (9,9) (9,27) และ (9,36) ซึ่งพิจารณาค่าความแปรปรวนคือ σ^2 คือ 3 9 27 และ 36

2.1.4 กำหนดระดับนัยสำคัญ 2 ระดับ คือ 0.01 และ 0.05

2.1.5 ตรวจสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ของการทดสอบจากความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลัก เมื่อสมมติฐานหลักเป็นจริง

2.1.6 ในกรณีที่ค่าความแปรปรวนทั้งสองกลุ่มไม่เท่ากัน

กำหนดความแตกต่างของความแปรปรวน สำหรับการเปรียบเทียบกำลังการทดสอบ โดยใช้ค่านอนเซนทรอลลิตี้พารามิเตอร์ ϕ (Noncentrality Parameter) [5] เป็นเกณฑ์ ดังนี้

ระดับความแตกต่างของความแปรปรวน	ความแปรปรวนแต่ละประชากร	ϕ
มีความแตกต่างกันน้อย ($0 < \phi < 1.5$)	3 : 5.4	0.69
มีความแตกต่างกันปานกลาง ($1.5 \leq \phi < 3$)	3 : 11.65	2.49
มีความแตกต่างกันมาก ($\phi \geq 3$)	3 : 27	6.93

โดยที่ Games และคณะ [6] ได้แนะนำค่าอนเซนทรลิตี้พารามิเตอร์ ϕ เพื่อเป็นเกณฑ์วัดความแตกต่างของความแปรปรวน ดังนี้

$$\phi = \sqrt{\frac{\sum_{i=1}^2 (\sigma_i^2 - \bar{\sigma}^2)^2 / 2}{\sigma_1^2}}$$

โดยที่

σ_1^2 คือ ค่าความแปรปรวนของประชากรที่มีค่าต่ำที่สุด

σ_i^2 คือ ค่าความแปรปรวนของประชากรที่ i โดย $i = 1, 2$

$\bar{\sigma}^2$ คือ ค่าเฉลี่ยความแปรปรวนของประชากรทั้ง 2 กลุ่ม

การแจกแจงปกติ มีค่าพารามิเตอร์ (μ, σ^2) ของประชากรกลุ่มที่ 1 คือ (9,3) ทั้ง 6 สถานการณ์ และประชากรกลุ่มที่ 2 คือ (9,5.4) (9,11.65) และ (9,27) ซึ่งจะได้ค่าความแปรปรวนของทั้งสองประชากรคือ (σ_1^2, σ_2^2) คือ (3,5.4) (3,11.65) และ (3,27) ตามลำดับ โดยจะได้ค่าอนเซนทรลิตี้พารามิเตอร์ คือ 0.69 2.49 และ 6.93 ตามลำดับ

2.1.7 เปรียบเทียบกำลังการทดสอบของการทดสอบจากความน่าจะเป็นที่ตัดสินใจปฏิเสธสมมติฐานหลัก เมื่อสมมติฐานหลักเป็นเท็จ โดยพิจารณาการทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ตามเกณฑ์ Bradley เพื่อหาการทดสอบที่ดีที่สุดในแต่ละสถานการณ์

2.2 การทดสอบความเท่ากันของค่าเฉลี่ยหรือค่ากลางของประชากร

การทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่เป็นอิสระ สมมติฐานที่ต้องการทดสอบ คือ

$$\text{สมมติฐานหลัก} \quad H_0 : \mu_1 = \mu_2$$

$$\text{สมมติฐานรอง} \quad H_1 : \mu_1 \neq \mu_2$$

โดยที่ μ_1 และ μ_2 แทน ค่าเฉลี่ยของประชากรที่ 1 และ 2 ตามลำดับ

2.2.1 การทดสอบซี (Z-test) [7]

เป็นการทดสอบอิงพารามิเตอร์ที่ใช้ทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่เป็นอิสระกัน จะสามารถใช้การทดสอบซีได้ก็ต่อเมื่อข้อมูลเป็นไปตามข้อกำหนดเบื้องต้น คือ ประชากร 2 กลุ่มอิสระต่อกัน ประชากร 2 กลุ่มมีการแจกแจงปกติ ตัวอย่างทั้ง 2 กลุ่มมีขนาดใหญ่ ($n_1, n_2 \geq 30$)

$$\text{การทดสอบ คือ } Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

เมื่อ \bar{x}_1 และ \bar{x}_2 คือ ค่าเฉลี่ยของตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

S_1^2 และ S_2^2 คือ ค่าความแปรปรวนของตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

n_1 และ n_2 คือ จำนวนตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

โดยจะปฏิเสธสมมติฐานหลักเมื่อ $Z_{cal} > Z_{\frac{\alpha}{2}}$ หรือ $Z_{cal} < -Z_{\frac{\alpha}{2}}$ โดยที่ $Z_{\frac{\alpha}{2}}$ และ

$-Z_{\frac{\alpha}{2}}$ เป็นค่าที่ได้จากการเปิดตารางซี ที่ค่าความน่าจะเป็น $\frac{\alpha}{2}$

2.2.2 การทดสอบที (t-test) [7]

เป็นการทดสอบอิงพารามิเตอร์ที่ใช้ทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่เป็นอิสระกัน จะสามารถใช้การทดสอบซีได้ก็ต่อเมื่อข้อมูลเป็นไปตามข้อกำหนดเบื้องต้น คือ ประชากร 2 กลุ่มอิสระต่อกัน ประชากร 2 กลุ่มมีการแจกแจงปกติ ตัวอย่างทั้ง 2 กลุ่มมีขนาดเล็ก ($n_1, n_2 < 30$)

กรณีที่ 1 เมื่อไม่ทราบค่า σ_1^2 และ σ_2^2 แต่ทราบว่า $\sigma_1^2 = \sigma_2^2$

$$\text{การทดสอบ } t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\text{โดยที่ } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

เมื่อ \bar{x}_1 และ \bar{x}_2 คือ ค่าเฉลี่ยของตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

S_p^2 คือ ความแปรปรวนร่วม

n_1 และ n_2 คือ จำนวนตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

โดยจะปฏิเสธสมมติฐานหลักเมื่อ $t_{cal} > t_{\frac{\alpha}{2}; n_1 + n_2 - 2}$ หรือ $t_{cal} < -t_{\frac{\alpha}{2}; n_1 + n_2 - 2}$ โดยที่

$t_{\frac{\alpha}{2}; n_1 + n_2 - 2}$ และ $-t_{\frac{\alpha}{2}; n_1 + n_2 - 2}$ เป็นค่าที่ได้จากการเปิดตารางที่ ที่ระดับนัยสำคัญ $\frac{\alpha}{2}$ และค่าองศาเสรี

(degrees of freedom) ที่ $n_1 + n_2 - 2$

กรณีที่ 2 เมื่อไม่ทราบค่า σ_1^2 และ σ_2^2 แต่ทราบว่า $\sigma_1^2 \neq \sigma_2^2$

$$\text{การทดสอบ } t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$\text{โดยที่ } \nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

เมื่อ \bar{x}_1 และ \bar{x}_2 คือ ค่าเฉลี่ยของตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

S_1^2 และ S_2^2 คือ ค่าความแปรปรวนของตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

n_1 และ n_2 คือ จำนวนตัวอย่างของกลุ่มที่ 1 และ 2 ตามลำดับ

ν คือ ค่าองศาเสรี (degrees of freedom)

โดยจะปฏิเสธสมมติฐานหลักเมื่อ $t_{cal} > t_{\frac{\alpha}{2}; \nu}$ หรือ $t_{cal} < -t_{\frac{\alpha}{2}; \nu}$ โดยที่ $t_{\frac{\alpha}{2}; \nu}$ และ

$-t_{\frac{\alpha}{2}; \nu}$ เป็นค่าที่ได้จากการเปิดตารางที่ ที่ระดับนัยสำคัญ $\frac{\alpha}{2}$ และค่าองศาเสรี (degrees of freedom)

ที่ ν

การทดสอบสมมติฐานเพื่อเปรียบเทียบค่ากลางระหว่างประชากร 2 กลุ่มที่เป็นอิสระ สมมติฐานที่ต้องการทดสอบ คือ

สมมติฐานหลัก $H_0: M_1 = M_2$

สมมติฐานรอง $H_1: M_1 \neq M_2$

โดยที่ M_1 และ M_2 แทน ค่ามัธยฐานของประชากรที่ 1 และ 2 ตามลำดับ

2.2.3 การทดสอบแบบสุ่ม (The Randomization Test) [8]

การทดสอบแบบสุ่มเป็นการทดสอบไม่อิงพารามิเตอร์ ใช้สำหรับกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระต่อกันและขนาดตัวอย่างทั้งสองมีขนาดเล็ก ข้อมูลที่ใช้อย่างน้อยต้องมีมาตรวัดแบบอันดับ ซึ่งมีการกำหนดเบื้องต้น มีวิธีการคำนวณดังนี้

- กำหนดข้อมูลแต่ละค่าให้อยู่ในกลุ่มที่ 1 คือ X_{1n} และกลุ่มที่ 2 คือ X_{2n} กลุ่มใดก็ได้

- วิธีที่จะจัดข้อมูลในกลุ่มที่ 1 คือ $X_{1n} = \binom{n_1 + n_2}{n_1}$ วิธี โดยแต่ละวิธีมีโอกาสเกิดขึ้น

เท่าๆกัน และกลุ่มที่ 2 คือ X_{2n} คือข้อมูลที่เหลือจากกลุ่มที่ 1

- ให้ ΣX_{1n} = ผลรวมของข้อมูลจากกลุ่มของ X_{1n} และ ΣX_{2n} = ผลรวมของข้อมูลจากกลุ่มของ X_{2n}

- ทำการสุ่มข้อมูลของแต่ละกลุ่ม และหาผลต่างของค่าเฉลี่ยกลุ่มที่ 1 และกลุ่มที่ 2 คือ \bar{X}_{1n} และ \bar{X}_{2n} เพื่อเลือกวิธีที่ตกอยู่ในเขตวิกฤตเมื่อ $\bar{X}_{1n} - \bar{X}_{2n} > \bar{x}_1 - \bar{x}_2$

โดยจะปฏิเสธสมมติฐานหลักเมื่อ α ที่แท้จริง = $\frac{\text{จำนวนวิธีที่ตกอยู่ในเขตวิกฤต}}{\binom{n_1 + n_2}{n_1}} < \alpha$

2.2.4 การทดสอบแมนท์-วิทนี ยู (Mann – Whitney U Test) [8]

การทดสอบแมนท์-วิทนี ยูเป็นการทดสอบไม่อิงพารามิเตอร์ มักนิยมใช้เพื่อเลี่ยงการทดสอบที่ หรือ เมื่อข้อมูลมีมาตรวัดต่ำกว่าแบบอันดับ ซึ่งมีการกำหนดเบื้องต้นดังต่อไปนี้

- ข้อมูลประกอบด้วยตัวอย่างสุ่ม ด้วยค่า $X_{11}, X_{12}, \dots, X_{1n}$ จากประชากรที่ 1 และตัวอย่างสุ่มอีก 1 ชุด ด้วยค่าสังเกต $X_{21}, X_{22}, \dots, X_{2n}$ จากประชากรที่ 2 ซึ่งเป็นอิสระกัน

- ตัวอย่าง 2 ชุดนี้เป็นอิสระกัน

- ตัวแปรสุ่มมีค่าต่อเนื่อง (continuous)

- มาตรวัดอย่างน้อยเป็นแบบเรียงลำดับ (ordinal scale)

- ฟังก์ชันการแจกแจงของ 2 ประชากร ต่างกันเฉพาะค่ากลาง นั่นคือประชากรทั้ง 2 ต้องมีการแจกแจงที่เหมือนกัน ต่างกันเฉพาะค่ากลางเท่านั้น

กรณีที่ 1 ขนาดตัวอย่างเล็ก (n_1 และ $n_2 \leq 20$)

$$\text{การทดสอบ } T_{cal} = S - \frac{n_1(n_1 + 1)}{2}$$

เมื่อ n_1 คือ จำนวนตัวอย่างของกลุ่มที่ 1

S คือ ผลรวมลำดับที่ของตัวอย่างขนาด n_1 ในข้อมูลรวมทั้งหมดที่เรียงลำดับแล้ว

โดยจะปฏิเสธสมมติฐานหลักเมื่อ $T_{cal} < W_{\frac{\alpha}{2}; n_1, n_2}$ หรือ $T_{cal} > W_{1-\frac{\alpha}{2}}$ เมื่อ $W_{\frac{\alpha}{2}}$ และ $W_{1-\frac{\alpha}{2}}$ เป็นค่าที่ได้จากการเปิดตารางแมนท์-วิทนี ที่ระดับนัยสำคัญ $\frac{\alpha}{2}$ และค่าองศาเสรี (degrees of freedom) ที่ n_1, n_2 และ $W_{1-\frac{\alpha}{2}} = n_1 n_2 - W_{\frac{\alpha}{2}}$

กรณีที่ 2 ขนาดตัวอย่างใหญ่ (n_1 และ $n_2 > 20$)

$$\text{การทดสอบ } Z_{cal} = \frac{T_{cal} - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}}$$

เมื่อ T_{cal} S n_1 และ n_2 มีความหมายเช่นเดียวกับกรณีที่ 1

โดยจะปฏิเสธสมมติฐานหลักเมื่อ $Z_{cal} > Z_{\frac{\alpha}{2}}$ หรือ $Z_{cal} < -Z_{\frac{\alpha}{2}}$ โดยที่ $Z_{\frac{\alpha}{2}}$ และ $-Z_{\frac{\alpha}{2}}$

เป็นค่าที่ได้จากการเปิดตารางซี ที่ค่าความน่าจะเป็น $\frac{\alpha}{2}$

2.3 วิธีการดำเนินการวิจัย

ในการวิจัยนี้มีวิธีดำเนินการวิจัยดังนี้

2.3.1 จำลองข้อมูลและสุ่มตัวอย่างที่ใช้ในการวิจัยด้วยโปรแกรมอาร์ (R) กำหนดค่าเฉลี่ยหรือค่ากลางเท่ากันและค่าความแปรปรวนเท่ากันและไม่เท่ากันระหว่างประชากร 2 กลุ่ม โดยมีการแจกแจง ความแปรปรวน และขนาดตัวอย่าง เป็นไปตามขอบเขตของการวิจัย

2.3.2 ทำการทดสอบการเปรียบเทียบค่าเฉลี่ยหรือค่ากลาง โดยคำนวณค่าการทดสอบซี (Z-test) การทดสอบที (t-test) การทดสอบสุ่ม (The Randomization Test) และการทดสอบแมนท์-วิทนี ยู (Mann – Whitney U Test) นำค่าการทดสอบที่คำนวณได้เทียบกับค่าวิกฤตเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.01 และ 0.05 บันทึกจำนวนครั้งที่ปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักเป็นจริง ทำซ้ำจนครบ 1,000 ครั้ง

2.3.3 หาค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 แล้วคำนวณค่าสัดส่วนที่มีการปฏิเสธสมมติฐานหลักแสดงโดยใช้ค่า $\hat{\alpha} = \frac{2\hat{\alpha}}{1000}$ จากการนำจำนวนครั้งที่ปฏิเสธสมมติฐานหลักหารด้วย 1,000

2.3.4 ทำการเปรียบเทียบความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ของการทดสอบกับเกณฑ์ของ Bradley [9] ถ้าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ที่ระดับนัยสำคัญ 0.01 ตกอยู่ในช่วง [0.005, 0.015] และที่ระดับนัยสำคัญ 0.05 ตกอยู่ในช่วง [0.025, 0.075] จะสรุปได้ว่าการทดสอบนั้นสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้

2.3.5 หากกำลังการทดสอบ เฉพาะการทดสอบที่สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ โดยจำลองข้อมูลให้มีค่าเฉลี่ยหรือค่ากลางไม่เท่ากันและค่าความแปรปรวนเท่ากันและไม่เท่ากันระหว่าง 2 ประชากร แล้วสุ่มตัวอย่างตามขอบเขตของการวิจัย คำนวณค่าการทดสอบเทียบกับค่าวิกฤต และสรุปผล

ทำซ้ำจนครบ 1,000 ครั้ง แล้วคำนวณค่าสัดส่วนที่มีการปฏิเสธสมมติฐานหลัก โดยนำจำนวนครั้งที่ปฏิเสธสมมติฐานหลักหารด้วย 1,000

2.3.6 เปรียบเทียบกำลังการทดสอบของการทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้โดยการทดสอบที่มีกำลังการทดสอบสูงที่สุดเป็นการทดสอบที่ดีที่สุด

3. ผลการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงจำลอง เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของการทดสอบ 4 ตัว ได้แก่ การทดสอบที่กรณีความแปรปรวนเท่ากัน (TEQ) การทดสอบที่กรณีความแปรปรวนไม่เท่ากัน (TNE) การทดสอบซี (Z) การทดสอบสุม (RND) การทดสอบแมนท์-วิทนี ยู (MWU) สำหรับ 2 ประชากรที่เป็นอิสระกัน ซึ่งผลที่ได้จากการวิจัยสามารถสรุปได้เป็น 2 ส่วนใหญ่ๆ คือ

3.1 ความน่าจะเป็นของความผิดพลาดประเภทที่ 1

3.1.1 ประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนเท่ากัน

จากการพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตามเกณฑ์ของ Bradley สรุปผลได้ดังตารางที่ 1 และ 2

ตารางที่ 1. ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 กรณีข้อมูลสุมจากประชากรที่มีการแจกแจงปกติที่ระดับนัยสำคัญ 0.01

ความแปรปรวน (σ^2)	การทดสอบ	ขนาดตัวอย่าง (n_1, n_2)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5)	(15,15)	(50,50)	(5,15)	(15,25)	(35,50)
3	Z	0.011*	0.01*	0.011*	0.01*	0.008*	0.012*
	TEQ	0.012*	0.008*	0.015*	0.008*	0.012*	0.01*
	TNE	0.011*	0.008*	0.015*	0.013*	0.013*	0.014*
	RND	0.008*	0	0	0.003	0	0
	MWU	0.008*	0.008*	0.013*	0.01*	0.009*	0.01*
9	Z	0.011*	0.008*	0.012*	0.011*	0.009*	0.007*
	TEQ	0.012*	0.008*	0.012*	0.009*	0.008*	0.006*
	TNE	0.008*	0.008*	0.012*	0.01*	0.009*	0.005*
	RND	0.008*	0	0	0	0	0
	MWU	0.008*	0.007*	0.014*	0.009*	0.007*	0.009*
27	Z	0.013*	0.008*	0.008*	0.009*	0.012*	0.01*
	TEQ	0.009*	0.011*	0.013*	0.008*	0.009*	0.011*
	TNE	0.006*	0.011*	0.013*	0.011*	0.011*	0.011*
	RND	0.007*	0	0	0	0	0
	MWU	0.006*	0.008*	0.009*	0.01*	0.008*	0.006*
36	Z	0.011*	0.011*	0.017	0.01*	0.014*	0.008*
	TEQ	0.012*	0.011*	0.012*	0.013*	0.013*	0.009*
	TNE	0.01*	0.011*	0.012*	0.009*	0.015*	0.008*
	RND	0.006*	0	0	0	0	0
	MWU	0.01*	0.012*	0.014*	0.004	0.013*	0.007*

* หมายถึงความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 ได้ ตามเกณฑ์ของ Bradley (ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตกอยู่ในช่วง [0.005, 0.015])

จากตารางที่ 1 พบว่า การทดสอบซี (Z) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้เกือบทุกกรณีการศึกษา ยกเว้นในกรณีที่ขนาดตัวอย่าง (50,50) เมื่อความแปรปรวนเท่ากับ 36 การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) และไม่เท่ากัน (TNE) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ทุกกรณีการศึกษา การทดสอบแบบสุ่ม (RND) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ในกรณีที่ขนาดตัวอย่าง (5,5) เท่านั้น และการทดสอบแมนท์-วิทนียู (MWU) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้เกือบทุกกรณี ยกเว้นในกรณีที่ขนาดตัวอย่าง (5,15) เมื่อความแปรปรวนเท่ากับ 27

ตารางที่ 2. ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 กรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปกติที่ระดับนัยสำคัญ 0.05

ความแปรปรวน (σ^2)	การทดสอบ	ขนาดตัวอย่าง (n_1, n_2)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5)	(15,15)	(50,50)	(5,15)	(15,25)	(35,50)
3	Z	0.045*	0.05*	0.06*	0.051*	0.062*	0.058*
	TEQ	0.042*	0.056*	0.062*	0.044*	0.056*	0.053*
	TNE	0.04*	0.056*	0.062*	0.051*	0.06*	0.056*
	RND	0.027*	0.001	0	0.004	0	0
	MWU	0.026*	0.053*	0.057*	0.032*	0.061*	0.055*
9	Z	0.047*	0.051*	0.051*	0.052*	0.055*	0.049*
	TEQ	0.045*	0.057*	0.055*	0.044*	0.053*	0.049*
	TNE	0.044*	0.056*	0.055*	0.055*	0.05*	0.05*
	RND	0.023	0	0	0.007	0	0
	MWU	0.034*	0.045*	0.056*	0.035*	0.054*	0.054*
27	Z	0.037*	0.054*	0.053*	0.044*	0.043*	0.057*
	TEQ	0.051*	0.049*	0.051*	0.053*	0.038*	0.057*
	TNE	0.045*	0.049*	0.051*	0.046*	0.038*	0.054*
	RND	0.03*	0	0	0.003	0	0
	MWU	0.034*	0.043*	0.051*	0.04*	0.042*	0.049*
36	Z	0.058*	0.048*	0.063*	0.044*	0.049*	0.048*
	TEQ	0.054*	0.056*	0.064*	0.047*	0.053*	0.051*
	TNE	0.047*	0.056*	0.064*	0.049*	0.055*	0.051*
	RND	0.027*	0	0	0.006	0	0
	MWU	0.029*	0.056*	0.057*	0.042*	0.058*	0.049*

* หมายถึงความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 ได้ ตามเกณฑ์ของ Bradley (ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตกอยู่ในช่วง [0.025, 0.075])

จากตารางที่ 2 พบว่า การทดสอบซี (Z) การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) และไม่เท่ากัน (TNE) และการทดสอบแมนท์-วิทนียู (MWU) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ทุกกรณีการศึกษา ส่วนการทดสอบแบบสุ่ม (RND) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ในกรณีที่ขนาดตัวอย่าง (5,5) เมื่อความแปรปรวนเท่ากับ 3 27 และ 36

3.1.2 ประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนไม่เท่ากัน

จากการพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตามเกณฑ์ของ Bradley สรุปผลได้ดังตารางที่ 3 และ 4

ตารางที่ 3. ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 กรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปกติที่ระดับนัยสำคัญ 0.01

ความแปรปรวน (σ_1^2, σ_2^2)	ϕ	การ ทดสอบ	ขนาดตัวอย่าง (n_1, n_2)					
			เท่ากัน			ไม่เท่ากัน		
			(5,5)	(15,15)	(50,50)	(5,15)	(15,25)	(35,50)
(3, 5.4)	0.69	Z	0.011*	0.011*	0.011*	0.013*	0.008*	0.011*
		TEQ	0.016	0.01*	0.014*	0.007*	0.006*	0.004
		TNE	0.012*	0.01*	0.013*	0.014*	0.012*	0.008*
		RND	0.007*	0	0	0.001	0	0
		MWU	0.009*	0.011*	0.014*	0.008*	0.008*	0.007*
(3, 11.65)	2.49	Z	0.011*	0.012*	0.015*	0.01*	0.015*	0.01*
		TEQ	0.012*	0.014*	0.012*	0.001	0.007*	0.005*
		TNE	0.01*	0.01*	0.012*	0.013*	0.015*	0.009*
		RND	0.011*	0	0	0	0	0
		MWU	0.014*	0.016	0.014*	0.003	0.012*	0.007*
(3, 27)	6.93	Z	0.01*	0.011*	0.012*	0.01*	0.007*	0.007*
		TEQ	0.021	0.014*	0.014*	0.001	0.003	0.003
		TNE	0.014*	0.011*	0.012*	0.011*	0.011*	0.004
		RND	0.018	0	0	0	0	0
		MWU	0.02	0.017	0.015*	0.002	0.007*	0.009*

* หมายถึงความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 ได้ ตามเกณฑ์ของ Bradley (ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตกอยู่ในช่วง [0.005, 0.015])

จากตารางที่ 3 ที่ระดับนัยสำคัญ 0.01 พบว่า การทดสอบซี (Z) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ทุกกรณีการศึกษา การทดสอบทีในกรณีที่มีความแปรปรวนเท่ากัน (TEQ) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้บางกรณีการศึกษา ยกเว้นในกรณีที่ขนาดตัวอย่าง (5,5) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 กับ 6.93 กรณีที่ขนาดตัวอย่าง (5,15) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 2.49 และ 6.93 กรณีที่ขนาดตัวอย่าง (15,25) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93 และกรณีที่ขนาดตัวอย่าง (35,50) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 และ 6.93 การทดสอบทีในกรณีที่มีความแปรปรวนไม่เท่ากัน (TNE) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้เกือบทุกกรณีการศึกษา ยกเว้นในกรณีที่ขนาดตัวอย่าง (35,50) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93 การทดสอบแบบสุ่ม (RND) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ในกรณีที่ขนาดตัวอย่าง (5,5) ยกเว้นค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93 การทดสอบแมนท์-วิทนียู (MWU) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้บางกรณีการศึกษา ยกเว้นในกรณีที่ขนาดตัวอย่าง (5,5) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93 กรณีที่ขนาดตัวอย่าง (15,15) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 2.49 และ 6.93 กรณีที่ขนาดตัวอย่าง (5,15) และค่าอนเซนทริลลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93

ตารางที่ 4. ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 กรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปกติที่ระดับนัยสำคัญ 0.05

ความแปรปรวน (σ_1^2, σ_2^2)	ϕ	การทดสอบ	ขนาดตัวอย่าง (n_1, n_2)					
			เท่ากัน			ไม่เท่ากัน		
			(5,5)	(15,15)	(50,50)	(5,15)	(15,25)	(35,50)
(3, 5.4)	0.69	Z	0.051*	0.047*	0.059*	0.05*	0.057*	0.055*
		TEQ	0.053*	0.054*	0.063*	0.024	0.037*	0.046*
		TNE	0.045*	0.054*	0.063*	0.046*	0.057*	0.058*
		RND	0.034*	0	0	0.004	0	0
		MWU	0.024	0.056*	0.058*	0.027*	0.046*	0.047*
(3, 11.65)	2.49	Z	0.057*	0.044*	0.062*	0.047*	0.049*	0.044*
		TEQ	0.057*	0.055*	0.061*	0.015	0.027*	0.028*
		TNE	0.053*	0.052*	0.06*	0.039*	0.054*	0.044*
		RND	0.028*	0	0	0	0	0
		MWU	0.039*	0.063*	0.067*	0.015	0.047*	0.044*
(3, 27)	6.93	Z	0.053*	0.049*	0.055*	0.054*	0.062*	0.053*
		TEQ	0.084	0.056*	0.056*	0.004	0.02	0.027*
		TNE	0.068*	0.049*	0.055*	0.055*	0.058*	0.062*
		RND	0.047*	0	0	0.001	0	0
		MWU	0.052*	0.058*	0.077	0.011	0.036*	0.05*

* หมายถึงความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 ได้ ตามเกณฑ์ของ Bradley (ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ตกอยู่ในช่วง [0.025, 0.075])

จากตารางที่ 4 ที่ระดับนัยสำคัญ 0.05 พบว่า การทดสอบซี (Z) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ทุกกรณีการศึกษา การทดสอบทีในกรณีที่มีความแปรปรวนเท่ากัน (TEQ) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้เกือบทุกกรณีการศึกษา กรณีที่ขนาดตัวอย่าง (5,5) และ (15,25) ที่ค่าอนเซนทรลิตีพารามิเตอร์ (ϕ) เท่ากับ 6.93 กรณีที่ขนาดตัวอย่าง (5,15) และทุกค่าอนเซนทรลิตีพารามิเตอร์ (ϕ) การทดสอบทีในกรณีที่มีความแปรปรวนไม่เท่ากัน (TNE) และการทดสอบแบบสุ่ม (RND) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ในกรณีที่มีขนาดตัวอย่าง (5,5) เท่านั้น ส่วนการทดสอบแมนท์-วิทนี ยู (MWU) สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้บางกรณีการศึกษา ยกเว้นในกรณีที่มีขนาดตัวอย่าง (5,5) ที่ค่าอนเซนทรลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 และกรณีที่มีขนาดตัวอย่าง (5,15) และค่าอนเซนทรลิตีพารามิเตอร์ (ϕ) เท่ากับ 2.49 และ 6.93

3.2 กำลังการทดสอบ

การสรุปผลกำลังการทดสอบของการทดสอบจะแสดงในกรณีที่การทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้เท่านั้น

3.2.1 ประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนเท่ากัน

พิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ในตารางที่ 1 และ 2 จึงนำมาพิจารณากำลังการทดสอบ กรณีที่ระดับนัยสำคัญ 0.01 และ 0.05 สรุปผลได้ดังตารางที่ 5

ตารางที่ 5. การทดสอบที่มีกำลังการทดสอบสูงที่สุดกรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนเท่ากัน

ระดับนัยสำคัญ	ความแปรปรวน	ขนาดตัวอย่าง (n_1, n_2)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5)	(15,15)	(50,50)	(5,15)	(15,25)	(35,50)
0.01	3	Z	Z	Z,TEQ,TNE,MWU	Z	Z	Z,TEQ, TNE,MWU
	9	Z	Z	MWU	Z	Z	Z
	27	Z	Z	TEQ	TEQ	TEQ	TEQ
	36	Z	MWU	MWU	TNE	MWU	MWU
0.05	3	Z	Z	Z,TEQ, TNE,MWU	Z	Z	Z,TEQ,TNE, MWU
	9	Z	Z	Z	Z	Z	Z,TEQ
	27	Z	Z	TEQ,TNE	TEQ	TEQ	TEQ
	36	TEQ	MWU	MWU	TNE	MWU	MWU

Z หมายถึง การทดสอบซี (Z-test)

TEQ หมายถึง การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (t-equal test)

TNE หมายถึง การทดสอบทีในกรณีที่ความแปรปรวนไม่เท่ากัน (t-non equal test)

MWU หมายถึง การทดสอบแมนท์-วิทนี ยู (Mann – Whitney U Test)

จากตารางที่ 5 ที่ระดับนัยสำคัญ 0.01 พบว่าการทดสอบซี (Z) เกือบทุกขนาดตัวอย่างมีกำลังการทดสอบสูงที่สุด เมื่อความแปรปรวนเท่ากับ 3 และ 9 ที่ขนาดตัวอย่าง (5,5) การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) มีกำลังการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50) (5,15) (15,25) และ (35,50) ที่ความแปรปรวน 27 และการทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) และไม่เท่ากัน (TNE) มีกำลังการทดสอบสูงที่สุดและเท่ากันที่ขนาดตัวอย่าง (50,50) และ (35,50) ที่ความแปรปรวน 3 การทดสอบแมนท์-วิทนี ยู (MWU) มีกำลังการทดสอบสูงที่สุดเมื่อความแปรปรวนเท่ากับ 36 ที่ขนาดตัวอย่าง (15,15) (50,50) (15,25) และ (35,50)

ที่ระดับนัยสำคัญ 0.05 พบว่าการทดสอบซี (Z) เกือบทุกขนาดตัวอย่างมีกำลังการทดสอบสูงที่สุด เมื่อความแปรปรวนเท่ากับ 3 และ 9 ส่วนที่ระดับความแปรปรวน 27 การทดสอบซี (Z) มีกำลังการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (5,5) และ(15,15) การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) มีกำลังการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50) (5,15) (15,25) และ (35,50) ที่ความแปรปรวน 27 และการทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน (TEQ) และไม่เท่ากัน (TNE) มีกำลังการทดสอบสูงที่สุดและเท่ากันที่ขนาดตัวอย่าง (50,50) และ (35,50) ที่ความแปรปรวน 3 การทดสอบแมนท์-วิทนี ยู (MWU) มีกำลังการทดสอบสูงที่สุดเมื่อความแปรปรวนเท่ากับ 36 ที่ขนาดตัวอย่าง (15,15) (50,50) (15,25) และ (35,50)

3.2.2 ประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนไม่เท่ากัน

จากการพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้จากตารางที่ 3 และ 4 นำมาพิจารณากำลัการทดสอบ กรณีที่ระดับนัยสำคัญ 0.01 และ 0.05 สรุปผลได้ดังตารางที่ 6

ตารางที่ 6. กำลัการทดสอบ กรณีข้อมูลสุ่มจากประชากรที่มีการแจกแจงปกติ เมื่อความแปรปรวนไม่เท่ากัน

ระดับนัยสำคัญ	ความแปรปรวน (σ_1^2, σ_2^2)	ϕ	ขนาดตัวอย่าง (n_1, n_2)					
			เท่ากัน			ไม่เท่ากัน		
			(5,5)	(15,15)	(50,50)	(5, 15)	(15,25)	(35,50)
0.01	(3, 5.4)	0.69	Z	Z	Z, TEQ, TNE, MWU	Z	Z	Z, TNE, MWU
	(3, 11.65)	2.49	Z	Z	TEQ	Z	Z	Z
	(3, 27)	6.93	Z	Z	Z	Z	Z	Z
0.05	(3, 5.4)	0.69	Z	Z	Z, TEQ, TNE, MWU	Z	Z	Z, TEQ, TNE, MWU
	(3, 11.65)	2.49	Z	Z	Z, TEQ, TNE, MWU	Z	Z	TNE
	(3, 27)	6.93	Z	Z	Z	Z	Z	Z

Z หมายถึง การทดสอบซี (Z-test)

TEQ หมายถึง การทดสอบทีในกรณีที่มีความแปรปรวนเท่ากัน (t-equal test)

TNE หมายถึง การทดสอบทีในกรณีที่มีความแปรปรวนไม่เท่ากัน (t-non equal test)

MWU หมายถึง การทดสอบแมนท์-วิทนี ยู (Mann - Whitney U Test)

จากตารางที่ 6 ที่ระดับนัยสำคัญ 0.01 พบว่า การทดสอบซี (Z) เกือบทุกขนาดตัวอย่างและทุกค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) มีกำลัการทดสอบสูงที่สุด ส่วนการทดสอบทีในกรณีที่มีความแปรปรวนไม่เท่ากัน (TNE) และการทดสอบแมนท์-วิทนียู (MWU) ค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 มีกำลัการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50) และ (35,50) และ การทดสอบทีในกรณีที่มีความแปรปรวนเท่ากัน (TEQ) ค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 และ 2.49 มีกำลัการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50)

ที่ระดับนัยสำคัญ 0.05 พบว่า การทดสอบซี (Z) เกือบทุกขนาดตัวอย่างและทุกค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) มีกำลัการทดสอบสูงที่สุด ส่วนการทดสอบทีในกรณีที่มีความแปรปรวนเท่ากัน (TEQ) และไม่เท่ากัน (TNE) ค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) เท่ากับ 0.69 มีกำลัการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50) และ (35,50) และ การทดสอบทีในกรณีที่มีความแปรปรวนไม่เท่ากัน (TNE) ค่าอนเซนทรีลิตีพารามิเตอร์ (ϕ) เท่ากับ 2.49 มีกำลัการทดสอบสูงที่สุด ที่ขนาดตัวอย่าง (50,50) และ (35,50)

4. สรุปผลการวิจัย

จากผลการวิเคราะห์ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และกำลัการทดสอบ ของการทดสอบ 4 ตัว ได้แก่ การทดสอบซี (Z-test) การทดสอบที (t-test) การทดสอบสุ่ม (The Randomization Test) การทดสอบแมนท์-วิทนี ยู (Mann - Whitney U Test) จาก

การจำลองข้อมูลตามลักษณะต่างๆ ที่ได้กำหนดไว้ในขอบเขตการวิจัย เมื่อขนาดตัวอย่างมีขนาดน้อย (5,5) ทุกการทดสอบสามารถควบคุมค่าความผิดพลาดประเภทที่ 1 ได้ แต่เมื่อขนาดตัวอย่างใหญ่ขึ้น การทดสอบแบบสุ่มจะไม่สามารถควบคุมค่าความผิดพลาดประเภทที่ 1 ได้ เมื่อความแปรปรวนไม่เท่ากัน เมื่อขนาดตัวอย่างมีขนาดน้อย (5,15) การทดสอบซี การทดสอบทีในกรณีที่ความแปรปรวนไม่เท่ากันสามารถควบคุมค่าความผิดพลาดประเภทที่ 1 ได้ แต่เมื่อขนาดตัวอย่างใหญ่ขึ้น ทุกการทดสอบจะสามารถควบคุมค่าความผิดพลาดประเภทที่ 1 ยกเว้นการทดสอบแบบสุ่ม

การทดสอบซีส่วนใหญ่มีกำลังการทดสอบสูงสุด การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน มีกำลังการทดสอบสูงสุดเมื่อค่าความแปรปรวนเท่ากับ 27 แต่เมื่อค่าความแปรปรวนเพิ่มขึ้นการทดสอบแมนท์-วิทนี ยู จะมีกำลังการทดสอบสูงสุด เมื่อความแปรปรวนไม่เท่ากันส่วนใหญ่การทดสอบซี มีกำลังการทดสอบสูงสุด การทดสอบทีในกรณีที่ความแปรปรวนเท่ากัน มีกำลังการทดสอบสูงสุดเมื่อค่าความแปรปรวนเท่ากับ 27 แต่เมื่อค่าความแปรปรวนเพิ่มขึ้นการทดสอบแมนท์-วิทนี ยู จะมีกำลังการทดสอบสูงสุด

เอกสารอ้างอิง (References)

- [1] มนตรี สังข์ทอง. 2556. การศึกษาประสิทธิภาพของสถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม. *วารสารวิทยาศาสตร์ มข*, 41(1), 226-238. [Montri Songthong. 2013. A study of efficiency of parametric and nonparametric statistics in testing of central difference between two populations. *KKU Science Journal*, 41(1), 226-238. (in Thai)]
- [2] มนตรี สังข์ทอง. 2557. ความแกร่งและอำนาจการทดสอบของสถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากรสองกลุ่ม สำหรับข้อมูลแบบลิเคิร์ต 5 ระดับ. *วารสารวิทยาศาสตร์และเทคโนโลยี*, 22(5), 605-609. [Montri Songthong. 2014. Robustness and power of the test of parametric and nonparametric statistics in testing of central difference between two population for Likert-type data 5 point, *Thai Science and Technology Journal*, 22(5), 605-609. (in Thai)]
- [3] Proschan, M., Glimm, E. and Posch, M. 2014. Connection between permutation and t-tests : relevance to adaptive methods, *statistics in medicine*, 33(27), 4734-4742.
- [4] R Development Core Team. 2004 . R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- [5] วราฤทธิ์ พานิชกิจโกศลกุล. 2548. การเปรียบเทียบประสิทธิภาพของการทดสอบของเลวีเน่สำหรับทดสอบความเท่ากันของความแปรปรวน. *วารสารพระจอมเกล้าลาดกระบัง*, 13(1), 29-35. [Wararit Panichkitkosolkul. 2005. A comparative study of the efficiency of Levene's test for testing homogeneity of variances. *Journal of Science Ladkrabang*, 13(1), 29-35. (in Thai)]
- [6] Games, P.A., Winkler, H.B. and Probert, D.A. 1972. Robust tests for homogeneity of variance. *Educational and Psychological Measurement*, 32, 887-909.

- [7] ชูใจ คูหารัตนไชย. 2556. สถิติเบื้องต้น ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง [Choojai Kuharatanachai. 2013. Introduction to Statistics, Department of statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang. (in Thai)]
- [8] อุมภาพร จันทสร. 2541. สถิติที่ไม่ใช้พารามิเตอร์ ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง [Umaporn Janthasorn. 1998. Nonparametrics, Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang. (in Thai)]
- [9] James, V.B. 1978. Robustness?. *The British Journal of Mathematical and Statistical Psychology*, 31(2), 144–152.