P-ISSN 0858-4435
E-ISSN 2651-1231

วารสารวิทยาศาสตร์และเทคโนโลยี Homepage : *http://li01.tci-thaijo.org/index.php/tstj*
ปีที่ 31 ฉบับที่ 2 (มีนาคม-เมษายน 2566) หน้า [1-13]

การเปรียบเทียบสถิติทดสอบอิงพารามิเตอร์และสถิติทดสอบไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่าเฉลี่ยสองประชากรที่อิสระกัน

A Comparison of Parametric and Nonparametric Test Statistics for Testing the Mean Difference of Two Independent Populations

ดลลดา วงค์ไชย*, มีนา เลา, อำไพ ทองธีรภาพ

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพมหานคร 10900

Dollada Wongchai*, Mena Lao, Ampai Thongteeraparp

Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900

Received 14 November 2022; Received in revised 8 March 2023; Accepted 16 March 2023

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบสถิติทดสอบอิงพารามิเตอร์ ได้แก่ สถิติทดสอบที และสถิติทดสอบไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ สถิติทดสอบมัธยฐาน สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สถิติทดสอบยูนเวลช์ และสถิติทดสอบบูตสแทร็ปยูนส์ ในการทดสอบความแตกต่างของค่าเฉลี่ยสองประชากรที่อิสระกัน ที่ระดับนัยสำคัญ 0.05 เมื่อประชากร มีการแจกแจงปรกติและไม่ได้มีการแจกแจงปรกติในแต่ละลักษณะความเบ้และความโด่งทั้งหมด 9 ลักษณะ โดยใช้เทคนิค มอนติคาร์โลและกำหนดขนาดตัวอย่างที่สุ่มจากประชากร 2 กลุ่ม (n_1, n_2) คือ (5,10) (10,10) (20,25) (25,25) (40,50) และ (50,50) และทำซ้ำในแต่ละสถานการณ์จำนวน 15,000 รอบ ผลการศึกษาพบว่า สถิติทดสอบทั้ง 5 วิธี สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เกือบทุกสถานการณ์ เมื่อประชากรมีการแจกแจงเบ้ซ้าย การแจกแจงสมมาตรและโด่งมาก การแจกแจงเบ้ขวาและโด่งมาก พบว่าสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน มีค่ากำลังการทดสอบสูงสุด เมื่อประชากรมีการแจกแจงสมมาตรและโด่งน้อย พบว่าสถิติทดสอบยูนเวลช์ มีค่ากำลังการทดสอบสูงสุดเกือบทุกสถานการณ์ เมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งน้อย การแจกแจงเบ้ขวาและโด่งปรกติ พบว่าสถิติทดสอบบูตสแทร็ปยูนส์และสถิติทดสอบผลบวกลำดับที่ของวิลคอกซันมีค่ากำลังการทดสอบสูงสุด ตามลำดับ ในกรณีตัวอย่างขนาดเล็ก

***ผู้รับผิดชอบบทความ:** dollada.woaku.th

doi: 10.14456/tstj.2023.11

คำสำคัญ: สถิติทดสอบไม่อิงพารามิเตอร์; สถิติทดสอบมัธยฐาน; สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน; สถิติทดสอบยูนเวลช์;สถิติทดสอบบูตสแทร็ปยูนส์; กำลังการทดสอบ

Abstract

The purpose of this study is to compare parametric test statistics, t-test, and four non-parametric test statistics: median test, Wilcoxon rank sum test, Yuen Welch test, and bootstrap Yuen test for testing the mean difference of two independent populations at the significant level of 0.05 when the population has a normal distribution and non-normal distribution for nine combinations of population skewness and kurtosis. The Monte Carlo technique is used to simulate data with 15,000 iterations by determining the sample sizes from 2 populations (n_1, n_2): (5,10) (10,10) (20,25) (25,25) (40,50) and (50,50). The study found that all five test statistics have the ability to control the probability of Type 1 error for most of the cases. When the population is negatively skewed distribution, symmetrical and leptokurtic distribution, positively skewed and leptokurtic distribution, the Wilcoxon rank sum test statistic had the highest power of the test. When the population is symmetrical and platykurtic distribution, the Yuen Welch test statistic had the highest power of the test for almost all the cases. When the population is positively skewed and leptokurtic distribution, positively skewed and mesokurtic distribution, the bootstrap Yuen test statistic and the Wilcoxon rank sum test statistic had the highest power of the test, respectively, for small sample sizes.

Keywords: Nonparametric; Median test; Wilcoxon rank sum test; Yuen Welch test; bootstrap Yuen test; power of the test

1. บทนำ

การทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยสองประชากรที่อิสระกัน มีสถิติทดสอบให้เลือกใช้มากมาย ถ้าตัวอย่างสองกลุ่มถูกเลือกมาจากประชากรที่มีการแจกแจงปรกติ สถิติที่มีความเหมาะสมสำหรับการทดสอบความแตกต่างของค่าเฉลี่ยสองประชากรที่อิสระกัน คือ สถิติทดสอบ z ในกรณีที่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร และสถิติทดสอบที ในกรณีที่ไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร ซึ่งทั้งสถิติทดสอบ z และสถิติทดสอบที เป็นสถิติทดสอบอิงพารามิเตอร์ (parametric statistic) แต่ในทางปฏิบัติอาจพบว่าข้อมูลที่นำมาใช้ไม่เป็นไปตามข้อสมมติเบื้องต้น (assumption) ของสถิติทดสอบ กล่าวคือมีการแจกแจงของประชากรเบี่ยงเบนไปจากการแจกแจงปรกติ ถ้าผู้วิจัยยังคงใช้สถิติทดสอบที ผลสรุปที่ได้อาจผิดพลาดได้

จากปัญหาดังกล่าวสถิติทดสอบไม่อิงพารามิเตอร์ (nonparametric statistic) จึงเป็นทางเลือกหนึ่งที่ผู้วิจัยให้ความสนใจในการหาสถิติทดสอบที่เหมาะสมแทนสถิติทดสอบที เมื่อข้อมูลไม่เป็นไปตามข้อสมมติเบื้องต้น ซึ่งได้มีนักสถิติหลายท่านทำการศึกษาพัฒนาสถิติทดสอบขึ้นมาหลายวิธีและเป็นสถิติทดสอบที่นิยมใช้ เช่น สถิติทดสอบมัธยฐาน (median) สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน (Wilcoxon rank sum) นอกจากสถิติทดสอบที่นิยมใช้ ยังมีผู้วิจัยอื่น ๆ ทำการศึกษา ดังนี้

Yuen [1] ได้นำเสนอสถิติทดสอบยูนเวลช์ (Yuen Welch) โดยอาศัยค่าเฉลี่ยแบบตัดปลาย

(trimmed mean) เพื่อขจัดข้อมูลผิดปกติ ซึ่งจากการศึกษาของ Fagerland และ Sandvik [2] พบว่าสถิติทดสอบยูนเวลช์เหมาะสำหรับข้อมูลที่มีการแจกแจงเบ้และความแปรปรวนไม่เท่ากันและสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และมีกำลังการทดสอบสูงสุด ต่อมา Dwivedi และคณะ [3] พบว่าสถิติทดสอบไม่อิงพารามิเตอร์ ได้แก่ สถิติทดสอบบูตสแทร็ปที (bootstrap t) โดยการนำตัวอย่างมารวมกัน (pooled sample) เป็นวิธีการที่มีกำลังการทดสอบสูงกว่าหรือเท่ากับสถิติทดสอบที สถิติทดสอบเวลช์ที (Welch t) สถิติทดสอบผลบวกลำดับที่ของวิลคอกซันและสถิติทดสอบการเรียงสับเปลี่ยน (permutation) และสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้เกือบทุกเงื่อนไข และนอกจากนี้ยังมีผู้วิจัยทำการศึกษาสถิติทดสอบที่ใช้วิธีบูตสแทร็ปกับสถิติทดสอบไม่อิงพารามิเตอร์ ได้แก่ มนตรี [4] ได้ศึกษาการเปรียบเทียบวิธีบูตสแทร็ปไม่อิงพารามิเตอร์ เพื่อทดสอบตำแหน่งระหว่างประชากรสองกลุ่ม เมื่อไม่เป็นไปตามข้อสมมติเบื้องต้น ประกอบด้วยสถิติทดสอบไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ สถิติทดสอบบูตสแทร็ปที สถิติทดสอบบูตสแทร็ปเวลช์ที สถิติทดสอบลำดับที่ของบูตสแทร็ปเวลช์ และสถิติทดสอบบูตสแทร็ปยูนส์ โดยจำลองข้อมูลให้มีการแจกแจงล็อกนอร์มัล (log-normal distribution) การแจกแจงแบบเลขชี้กำลัง (exponential distribution) และการแจกแจงแบบแกมมา (gamma distribution) ทั้งกรณีความแปรปรวนและขนาดตัวอย่างเท่ากันและไม่เท่ากัน ซึ่งพบว่าเมื่อประชากรมีการแจกแจงล็อกนอร์มัล การแจกแจงแบบเลขชี้กำลัง และการแจกแจงแกมมา กรณีความแปรปรวนไม่เท่ากัน สถิติทดสอบลำดับที่ของบูตสแทร็ปเวลช์ มีประสิทธิภาพสูงสุดสำหรับกรณีขนาดตัวอย่าง n_1 และ $n_2 > 30$ และความแปรปรวนไม่เท่ากัน พบว่าสถิติทดสอบบูตสแทร็ปยูนส์มีประสิทธิภาพสูงสุด เป็นต้น

ดังนั้นผู้วิจัยจึงสนใจศึกษาสถิติทดสอบไม่อิงพารามิเตอร์สำหรับการทดสอบความแตกต่างของค่าเฉลี่ยสองประชากร ที่อิสระกัน กรณีที่ความแปรปรวนไม่เท่ากัน ได้แก่ สถิติทดสอบมัธยฐาน สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สถิติทดสอบ ยูนเวลช์ และสถิติทดสอบบูตสแทร็ปยูนส์ เมื่อประชากรมีการแจกแจงปรกติและประชากรที่ไม่ได้มีการแจกแจงปรกติ โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และเปรียบเทียบกำลังการทดสอบ

2. วัตถุประสงค์การวิจัย

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำลังการทดสอบของสถิติทดสอบอิงพารามิเตอร์ ได้แก่ สถิติทดสอบที และสถิติทดสอบไม่อิงพารามิเตอร์ 4 วิธี ได้แก่ สถิติทดสอบมัธยฐาน สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สถิติทดสอบยูนเวลช์ และสถิติทดสอบบูตสแทร็ปยูนส์ เมื่อประชากรมีการแจกแจงปรกติและไม่ได้มีการแจกแจงปรกติ ในแต่ละลักษณะการแจกแจงของประชากร

3. เกณฑ์ที่ใช้ในการศึกษา

3.1 การเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยใช้การทดสอบทวินาม (binomial test) [5] โดยกำหนดสมมุติฐาน $H_0 : \alpha \le \alpha_0$ และ $H_1 : \alpha > \alpha_0$ ซึ่งสถิติทดสอบ คือ $Z = \dfrac{\hat{\alpha} - \alpha_0}{\sqrt{\dfrac{\alpha_0(1-\alpha_0)}{r}}}$ โดย $\hat{\alpha}$ แทนค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง, α_0 แทนระดับนัยสำคัญ ที่กำหนดใน H_0 และ H_1 และ r แทนจำนวนครั้งที่ทำการทดลอง

กำหนดระดับนัยสำคัญของการทดสอบ H_0 เทียบกับ H_1 เท่ากับ 0.05 และจำนวนครั้งที่ทำการทดลอง r เท่ากับ 15,000 ครั้ง ดังนั้นช่วงการยอมรับความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 คือ

$$\left(0,\ \alpha_0 + Z_{\alpha^*}\sqrt{\frac{\alpha_0(1-\alpha_0)}{r}}\right)$$

$$\left(0,\ 0.05 + Z_{0.05}\sqrt{\frac{0.05(1-0.05)}{15000}}\right)$$

ดังนั้น สถิติทดสอบใดที่มีค่า $\hat{\alpha}$ อยู่ในช่วง (0.0000, 0.0529) แสดงว่าสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

3.2 การเปรียบเทียบค่ากำลังการทดสอบ นำสถิติทดสอบที่มีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ มาเปรียบเทียบเพื่อหาสถิติทดสอบที่มีค่ากำลังการทดสอบสูงสุด

4. วิธีการวิจัย

4.1 กำหนดขอบเขตการศึกษา

4.1.1 สร้างข้อมูลประชากร 2 กลุ่มที่มีค่าเฉลี่ยประชากรเท่ากันและมีการแจกแจงตามสถานการณ์ที่กำหนด โดยค่าความเบ้และความโด่งต่างกันจำนวน 9 ลักษณะ [5] (Talbe 1)

4.1.2 สมมติฐานในการทดสอบ คือ

H_0: ค่าเฉลี่ยของประชากรทั้งสองกลุ่มไม่แตกต่างกัน

H_1: ค่าเฉลี่ยของประชากรทั้งสองกลุ่มแตกต่างกัน

4.1.3 กำหนดให้ข้อมูลประชากรเป็นไปตาม H_0 นั่นคือกำหนดค่าเฉลี่ยของทั้ง 2 กลุ่มไม่แตกต่างกัน สำหรับการคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำหนดให้ข้อมูลประชากรเป็นไปตาม H_1 นั่นคือกำหนดให้ค่าเฉลี่ยของทั้ง 2 กลุ่มแตกต่างกัน โดยกำหนดอัตราส่วนค่าเฉลี่ยของกลุ่มที่ 1 ต่อกลุ่มที่ 2 เท่ากับ 1.5 : 1 และ 2 : 1 สำหรับการคำนวณค่ากำลังการทดสอบ ซึ่งขนาดประชากรที่สร้าง 2 กลุ่มมีขนาดเท่ากับ 100,000 หน่วย

4.1.4 กำหนดขนาดตัวอย่างที่สุ่มจากประชากร 2 กลุ่ม (n_1, n_2) ดังนี้ (5,10) (10,10) (20,25) (25,25) (40,50) และ (50,50) โดยกำหนดให้ตัวอย่างขนาดเล็กได้แก่ (5,10) และ (10,10) ตัวอย่างขนาดปานกลาง ได้แก่ (20,25) และ (25,25) และตัวอย่างขนาดใหญ่ ได้แก่ (40,50) และ (50,50)

4.1.5 กำหนดระดับนัยสำคัญในการทดสอบเท่ากับ 0.05

4.1.6 กำหนดจำนวนรอบในการทำซ้ำของการเลือกตัวอย่างและคำนวณค่าสถิติทดสอบ เท่ากับ 15,000 ครั้ง

4.2 การจำลองข้อมูล

4.2.1 สร้างข้อมูลประชากรให้ตรงตามสถานการณ์ที่กำหนดไว้ในข้อ 4.1.1

4.2.2 ทำการสุ่มตัวอย่างข้อมูล 2 กลุ่มตามขนาดตัวอย่างที่กำหนดไว้ในข้อ 4.1.4 จากประชากรที่

Table 1 Skewness and Kurtosis of the population defined this study.

Skewness (μ_3^*)	Kurtosis (μ_4^*)		
	Platykurtic ($\mu_4^* = 2.5$)	Mesokurtic ($\mu_4^* = 3$)	Lleptokurtic ($\mu_4^* = 4$)
Negatively skewed ($\mu_3^* = -1$)	(-1, 2.5)	(-1, 3)	(-1, 4)
Symmetrical ($\mu_3^* = 0$)	(0, 2.5)	(0, 3)	(0, 4)
Positively skewed ($\mu_3^* = 1$)	(1, 2.5)	(1, 3)	(1, 4)

สร้างไว้ในข้อ 4.2.1 เพื่อนำมาคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่ากำลังการทดสอบ

4.2.3 คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 กำหนดให้ประชากรเป็นไปตาม H_0 นั่นคือกำหนดค่าเฉลี่ยของทั้ง 2 กลุ่มไม่แตกต่างกัน นำข้อมูลตัวอย่างที่สุ่มได้จากข้อ 4.2.2 ไปทดสอบสมมติฐาน โดยใช้สถิติทดสอบทั้ง 5 วิธี ได้แก่ สถิติทดสอบที สถิติทดสอบมัธยฐาน สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สถิติทดสอบยูนเวลช์ และสถิติทดสอบบูตสแทร็ปยูนส์ ซึ่งสถิติทดสอบทั้ง 5 วิธี ได้แก่

(1) สถิติทดสอบที เป็นการทดสอบความแตกต่างระหว่างค่าเฉลี่ยสองประชากรในกรณีที่ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2) และขนาดตัวอย่าง (n_1, n_2) มีขนาดเล็ก กรณีที่ $\sigma_1^2 \neq \sigma_2^2$ ตัวสถิติที่ใช้ในการทดสอบผลต่างของค่าเฉลี่ยประชากรภายใต้ H_0 เมื่อไม่ทราบค่า σ_1^2 และ σ_2^2 และมีขนาดตัวอย่างมีขนาดเล็ก คือ $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$ เมื่อ d_0 เท่ากับ ผลต่างของค่าเฉลี่ยโดย t มีองศาเสรีเท่ากับ υ ซึ่ง $\upsilon = \dfrac{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}{\dfrac{\left(S_1^2 / n_1\right)^2}{n_1 - 1} + \dfrac{\left(S_2^2 / n_2\right)^2}{n_2 - 1}}$ และเขตวิกฤต คือ $|t| > t_{\frac{\alpha}{2},\upsilon}$ [6] แล้วนำค่าสถิติทดสอบที ที่คำนวณได้ไปเปรียบเทียบกับค่าวิกฤต ถ้าค่าสถิติทดสอบที่คำนวณได้ตกอยู่ในเขตวิกฤต ณ ระดับนัยสำคัญที่กำหนด จะปฏิเสธ H_0

(2) สถิติทดสอบมัธยฐาน (M) ใช้ทดสอบว่าตัวอย่างที่อิสระกัน 2 กลุ่มมีค่ามัธยฐานเท่ากันหรือไม่ ซึ่งข้อมูลประกอบด้วยตัวอย่าง 2 กลุ่มที่เป็นอิสระกัน โดยการทดสอบมีวิธีการดังนี้ [7]

ขั้นที่ 1 นำข้อมูลทั้ง 2 กลุ่มมารวมกันแล้วเรียงลำดับข้อมูลจากน้อยที่สุดไปมากที่สุด เพื่อหาค่ามัธยฐานตัวอย่างรวม (*m*)

ขั้นที่ 2 แบ่งกลุ่มข้อมูลโดยนับจำนวนความถี่ของข้อมูลแต่ละกลุ่ม ว่ามากกว่า m หรือน้อยกว่า m แสดงดังตารางที่ 2

Table 2 Classifying data for the median test statistic calculation.

Classifying data	Samples		Total
	Group 1	Group 2	
Greater than median (>*m*)	A	B	A + B
Less than median (<*m*)	C	D	C + D
Total	n_1 = A + C	n_2 = B + D	$n = n_1 + n_2$

กรณี $n_1 + n_2 \leq 20$ คำนวณความน่าจะเป็นของ Fisher test ดังนี้

$$P(A,B) = \frac{\binom{A+C}{A}\binom{B+D}{B}}{\binom{n_1+n_2}{A+B}} = \frac{(A+B)!(C+D)!n_1!n_2!}{n!A!B!C!D!}$$

เกณฑ์การปฏิเสธ H_0 คือ $P(A,B)<a$

กรณีที่ตัวอย่างมีขนาดใหญ่ สามารถทดสอบด้วยสถิติทดสอบไคกำลังสอง โดยสูตรของสถิติทดสอบไคกำลังสองคือ

$$\chi^2 = \frac{n\left(|AD-BC|-\frac{n}{2}\right)^2}{(A+B)(C+D)n_1n_2}$$

เกณฑ์การปฏิเสธ H_0 คือ $\chi^2 > \chi^2_{\alpha;1}$

(3) สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน (WRS) เป็นสถิติทดสอบไม่อิงพารามิเตอร์ที่มีคุณสมบัติในการทดสอบใกล้เคียงกับสถิติทดสอบทีและสถิติทดสอบ z ใช้สำหรับทดสอบความแตกต่างระหว่างประชากร 2 กลุ่มที่อิสระกัน [8] โดยการคำนวณสถิติทดสอบมีดังนี้

กรณี $n_1 \le 20$ และ $n_2 \le 20$ คำนวณค่า U โดยใช้สูตร

$$U_1 = n_1n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1$$

และ $$U_2 = n_1n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2$$

เมื่อ n_1 คือข้อมูลตัวอย่างสุ่มจากประชากรกลุ่มที่ 1, n_2 คือข้อมูลตัวอย่างสุ่มจากประชากรกลุ่มที่ 2, $\sum R_1$ คือผลรวมของลำดับของข้อมูลตัวอย่างกลุ่มที่ 1, $\sum R_2$ คือผลรวมของลำดับของข้อมูลตัวอย่างกลุ่มที่ 2 โดยที่ $U_1+U_2=n_1n_2$ โดยสถิติทดสอบ คือ $U=\min(U_1,U_2)$ และเกณฑ์การปฏิเสธ H_0 คือ $U \le$ ค่าวิกฤต $U_{\frac{\alpha}{2}}$

กรณี $n_1 > 20$ หรือ $n_2 > 20$ คำนวณค่า z โดยมีขั้นตอนดังนี้

ขั้นที่ 1 นำข้อมูลตัวอย่างสุ่มจากทั้ง 2 กลุ่มประชากร ซึ่งมีขนาด n_1 และ n_2 ตามลำดับมาเรียงข้อมูลจากน้อยไปมาก

ขั้นที่ 2 กำหนดให้ข้อมูลที่น้อยที่สุดเป็นลำดับที่ 1 เรียงจนถึงลำดับที่ n โดยที่ $n=n_1+n_2$

ขั้นที่ 3 นำ U แปลงเป็น z score โดยใช้สูตรดังนี้

$Z=\frac{U-\mu_u}{\sigma_u}$ เมื่อ μ_u แทน ค่าเฉลี่ยของ U

$\mu_u=\frac{n_1n_2}{2}$, σ_u แทน ค่าความเบี่ยงเบนมาตรฐาน U

และ $$\sigma_u=\sqrt{\frac{n_1n_2(n+1)}{12}}$$

เขตวิกฤต คือ $Z<-Z_{\frac{\alpha}{2}}$ และ $Z>Z_{\frac{\alpha}{2}}$ แล้วนำค่าสถิติทดสอบ Z ที่คำนวณได้ไปเปรียบเทียบกับค่าวิกฤต ถ้าค่าสถิติทดสอบที่คำนวณได้ตกอยู่ในเขตวิกฤต ณ ระดับนัยสำคัญที่กำหนด จะปฏิเสธ H_0

(4) สถิติทดสอบยูนเวลช์ (YW) เป็นสถิติทดสอบรูปแบบหนึ่งของสถิติทดสอบที ซึ่งมีหลักการคำนวณ คือ นำเทคนิคค่าเฉลี่ยที่ปรับแล้วมาใช้ในการปรับค่าของข้อมูลและกำหนดจำนวนการตัดออก (trimming: γ) ซึ่งมีค่าอยู่ระหว่าง $0 \le \gamma \le 0.5$ [9] และเลือกใช้จำนวนการตัดออก γ=0.2 หรือ 20% [10] ซึ่งการทดสอบมีวิธีการดังนี้

ขั้นที่ 1 กำหนดให้ $X_1 = x_{11}, x_{12}, \ldots, x_{1n_1}$ คือ ข้อมูลจากตัวอย่างที่ 1 ขนาด n_1 และ $X_2 = x_{21}, x_{22}, \ldots, x_{2n_2}$ คือ ข้อมูลจากตัวอย่างที่ 2 ขนาด n_2

ขั้นที่ 2 เรียงลำดับข้อมูลของตัวอย่างที่ 1 และเรียงลำดับข้อมูลของตัวอย่างที่ 2 จากน้อยไปมาก ตัดข้อมูลปลายหางด้านน้อยและด้านมากออกด้านละ 20% เนื่องจากมีประสิทธิภาพดีที่สุด กำหนดจำนวนข้อมูลที่ตัดออก แทนด้วย $g_1=\gamma n_1$ และ $g_2=\gamma n_2$ กำหนดให้ $W_1 = w_{11}, w_{12}, \ldots, w_{1n_1}$ และ $W_2 = w_{21}, w_{22}, \ldots, w_{2n_2}$ ทั้งนี้ W_1, W_2 คือ ข้อมูลที่เหลือหลังจากการตัดข้อมูลปลายหางด้านน้อยและด้านมากออกด้านละ 20% และแทนข้อมูลคืนเท่ากับจำนวนที่ตัดออกไปด้วยข้อมูลที่น้อยที่สุดหรือมากที่สุดที่ไม่ถูกตัดออกไปของตัวอย่างที่ 1 และตัวอย่างที่ 2 ตามลำดับ

ขั้นที่ 3 คำนวณค่า $YW = \dfrac{\overline{X}_{\gamma_1} - \overline{X}_{\gamma_2}}{\sqrt{d_1 + d_2}}$ โดยที่ $\overline{X}_{\gamma_1} = \dfrac{\sum_{i=1}^{n_1-h_1} X_{\gamma_{1i}}}{n_1 - h_1}$ คือค่าเฉลี่ยของข้อมูลตัวอย่างกลุ่มที่ 1 ที่ตัดข้อมูลออกแล้ว g_1 ตัว เมื่อ $X_{\gamma_{1i}} = X_{\gamma_{11}}, X_{\gamma_{12}}, \ldots, X_{\gamma_{1,n_1-h_1}}$ เป็นข้อมูลตัวอย่างกลุ่มที่ 1 ที่ตัดข้อมูลออกแล้ว g_1 ตัว, $h_1 = n_1 - 2g_1$ คือจำนวนของข้อมูลตัวอย่างกลุ่มที่ 1 ที่ตัดข้อมูลออกแล้ว g_1 ตัว, $d_1 = \dfrac{SW_1^2(n_1 - 1)}{h_1(h_1 - 1)}$ คือความคลาดเคลื่อนมาตรฐานกำลังสอง (squared standard error) ของข้อมูลตัวอย่างกลุ่มที่ 1, $SW_1^2 = \dfrac{\left(\sum_{i=1}^{n_1} w_{1i} - \overline{W}_1\right)^2}{n_1 - 1}$ คือความแปรปรวนวินเซอร์ไรซ์ (winsorized variance) ของข้อมูลตัวอย่างกลุ่มที่ 1 และ $\overline{W}_1 = \dfrac{\sum_{i=1}^{n_1} w_{1i}}{n_1}$ คือค่าเฉลี่ยวินเซอร์ไรซ์ (winsorized mean) ของข้อมูลตัวอย่างกลุ่มที่ 1

$\overline{X}_{\gamma_2} = \dfrac{\sum_{j=1}^{n_2 \; h_2} X_{\gamma_{2j}}}{n_2 - h_2}$ คือค่าเฉลี่ยของข้อมูลตัวอย่างกลุ่มที่ 2 ที่ตัดข้อมูลออกแล้ว g_2 ตัว เมื่อ $X_{\gamma_{2j}} = X_{\gamma_{21}}, X_{\gamma_{22}}, \ldots, X_{\gamma_{2,n_2-h_2}}$ เป็นข้อมูลตัวอย่างกลุ่มที่ 2 ที่ตัดข้อมูลออกแล้ว g_2 ตัว, $h_2 = n_2 - 2g_2$ คือจำนวนของข้อมูลตัวอย่างกลุ่มที่ 2 ที่ตัดข้อมูลออกแล้ว g_2 ตัว, $d_2 = \dfrac{SW_2^2(n_2 - 1)}{h_2(h_2 - 1)}$ คือความคลาดเคลื่อนมาตรฐานกำลังสองของข้อมูลตัวอย่างกลุ่มที่ 2, $SW_2^2 = \dfrac{\left(\sum_{j=1}^{n_2} w_{2j} - \overline{W}_2\right)^2}{n_2 - 1}$ คือความแปรปรวนวินเซอร์ไรซ์ของข้อมูลตัวอย่างกลุ่มที่ 2 และ $\overline{W}_2 = \dfrac{\sum_{j=1}^{n_2} w_{2j}}{n_2}$ คือค่าเฉลี่ยวินเซอร์ไรซ์ของข้อมูลตัวอย่างกลุ่มที่ 2

ขั้นที่ 4 คำนวณ df (เป็นเลขจำนวนเต็ม) ดังนี้ $df_{YW} = \dfrac{(d_1 + d_2)^2}{\left(\dfrac{d_1^2}{h_1 - 1} + \dfrac{d_2^2}{h_2 - 1}\right)}$ และเขตวิกฤต คือ $t < -t_{\frac{\alpha}{2}, df_{YW}}$ และ $t > t_{\frac{\alpha}{2}, df_{YW}}$ แล้วนำค่าสถิติทดสอบที่คำนวณได้ไปเปรียบเทียบกับค่าวิกฤต ถ้าค่าสถิติทดสอบที่คำนวณได้ตกอยู่ในเขตวิกฤต ณ ระดับนัยสำคัญที่กำหนด จะปฏิเสธ H_0

(5) สถิติทดสอบบูตสแทร็ปยูนส์ (BY) เป็นการประยุกต์ใช้เทคนิควิธีบูตสแทร็ปกับสถิติทดสอบยูนเวลช์ [10] ซึ่งการทดสอบมีวิธีการดังนี้

ขั้นที่ 1 ถึง **ขั้นที่ 3** วิธีทำขั้นตอนเหมือนสถิติทดสอบยูนเวลช์

ขั้นที่ 4 ทำการเลือกตัวอย่างขนาด n_1 จากตัวอย่างกลุ่มที่ 1 ด้วยวิธีการเลือกตัวอย่างสุ่มแบบง่ายแบบคืนที่ (simple random sampling with replacement: SRSWR) และในทำนองเดียวกันทำการเลือกตัวอย่างขนาด n_2 จากตัวอย่างกลุ่มที่ 2 ดังกล่าวด้วยวิธี SRSWR และนำตัวอย่างที่สุ่มได้จากกลุ่มที่ 1 และ 2 มาคำนวณหาค่า YW^* ตามขั้นตอนที่ 2 และ 3

ขั้นที่ 5 ทำซ้ำขั้นที่ 4 จำนวน B รอบ (กำหนด B = 1,000 รอบ) จึงได้ YW^* จำนวน 1,000 ค่า

ขั้นที่ 6 คำนวณ $P - value$ =

$$\frac{\text{จำนวนครั้งที่}\left(\left|YW^*\right| \geq \left|YW\right|\right)}{B}$$

ขั้นที่ 7 เกณฑ์การปฏิเสธ H_0 คือ $P - value < \alpha$

ทำข้อ 4.2.2 และ 4.2.3 ซ้ำกัน 15,000 ครั้ง จากนั้นคำนวณค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ดังนี้ $\hat{\alpha}$ =

$$\frac{\text{จำนวนครั้งของการปฏิเสธ } H_0 \text{ เมื่อสมมุติฐาน } H_0 \text{ เป็นจริง}}{\text{จำนวนครั้งในการทำซ้ำ}}$$

จากนั้นนำ $\hat{\alpha}$ ของแต่ละสถิติทดสอบไปเปรียบเทียบกับเกณฑ์ที่กำหนด โดยในที่นี้กำหนดระดับนัยสำคัญ 0.05 ถ้าสถิติทดสอบใดมีค่า $\hat{\alpha}$ อยู่ในช่วง (0.0000, 0.0529) จะถือว่าสถิติทดสอบนั้นสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

4.2.4 คำนวณค่ากำลังการทดสอบ

ถ้าสถิติทดสอบใดสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ จะนำสถิติทดสอบนั้นมาคำนวณค่าประมาณกำลังการทดสอบ และกำหนดให้ประชากรเป็นไปตาม H_1 นั่นคือกำหนดค่าเฉลี่ยของทั้ง 2 กลุ่มไม่แตกต่างกัน นำข้อมูลที่สุ่มได้จากข้อ 4.2.2 ไปทำการทดสอบสมมติฐาน โดยใช้สถิติทดสอบทั้ง 5 วิธี และทำข้อ 4.2.2 และ 4.2.4 ซ้ำกัน 15,000 ครั้ง จากนั้นคำนวณค่าประมาณกำลังการทดสอบ ดังนี้

ค่าประมาณกำลังการทดสอบ =

$$\frac{\text{จำนวนครั้งของการปฏิเสธ } H_0 \text{ เมื่อสมมุติฐาน } H_1 \text{ เป็นจริง}}{\text{จำนวนครั้งในการทำซ้ำ}}$$

4.2.5 เปรียบเทียบค่ากำลังการทดสอบของสถิติทดสอบ โดยที่ค่ากำลังการทดสอบของสถิติทดสอบใดมีค่ามากที่สุด แสดงว่าสถิติทดสอบนั้นมีกำลังการทดสอบสูงสุด

5. ผลการวิจัยและวิจารณ์

5.1 ผลการวิจัย

การศึกษาเปรียบเทียบสถิติทดสอบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ที่ใช้ในการทดสอบความแตกต่างของค่าเฉลี่ยระหว่าง 2 ประชากรที่อิสระกัน โดยประชากรที่ศึกษามีความเบ้และความโด่ง 9 ลักษณะดังตารางที่ 1 โดยสถิติทดสอบที่ใช้ 5 วิธี ได้แก่ สถิติทดสอบที สถิติทดสอบมัธยฐาน สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สถิติทดสอบยูนเวลช์และสถิติทดสอบบูตสแทร็ปยูนส์ โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และเปรียบเทียบค่ากำลังการทดสอบ สรุปได้ดังนี้

5.1.1 ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบทั้ง 5 วิธี ที่ระดับนัยสำคัญ 0.05 (Table 3)

Table 3 สรุปได้ว่าเมื่อประชากรมีการแจกแจงปรกติ พบว่าสถิติทดสอบทั้ง 5 วิธี สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้เกือบทุกขนาดตัวอย่าง ยกเว้นที่ขนาดตัวอย่างเท่ากับ (5,10) พบว่าสถิติทดสอบยูนเวลช์ ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ และที่ขนาดตัวอย่างเท่ากับ (5,10) พบว่าสถิติทดสอบมัธยฐานและสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทุกลักษณะของประชากรที่มีความเบ้และความโด่ง 9 ลักษณะ

เมื่อประชากรมีการแจกแจงเบ้ซ้าย พบว่าสถิติทดสอบทั้ง 5 วิธีสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ที่ขนาดตัวอย่างเท่ากับ (10,10) (20,25) (25,25) และ (50,50) ส่วนกรณีเมื่อประชากรมีการแจกแจงสมมาตร พบว่าสถิติทดสอบทั้ง 5 วิธีสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ที่ขนาดตัวอย่างเท่ากับ (20,25) (25,25) (40,50) และ (50,50) และเมื่อประชากรมีการแจกแจงเบ้ขวา พบว่าสถิติทดสอบทั้ง 5 วิธีสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ที่ขนาดตัวอย่างเท่ากับ (10,10) (20,25) และ (25,25)

5.1.2 ค่ากำลังการทดสอบในแต่ละลักษณะประชากรที่มีความเบ้และความโด่ง 9 ลักษณะ โดย

Table 3 Test statistics which can control probability of Type I error for nine combinations of population skewness and kurtosis at the significant level of 0.05.

Skewness	Kurtosis	(n_1,n_2)					
		(5,10)	**(10,10)**	**(20,25)**	**(25,25)**	**(40,50)**	**(50,50)**
Negatively skewed	platykurtic	M, WRS	*	*	*	M, WRS, YW, BY	*
	mesokurtic	M, WRS, BY	*	*	*	*	*
	leptokurtic	M, WRS	*	*	*	*	*
Symmetrical	platykurtic	M, WRS	t, M, WRS, BY	*	*	*	*
	mesokurtic or normal	t, M, WRS, BY	*	*	*	*	*
	leptokurtic	t, M, WRS, BY	*	*	*	*	*
Positively skewed	platykurtic	M, WRS, BY	*	*	*	t, M, WRS, BY	t, M, WRS, BY
	mesokurtic	t, M, WRS	*	*	*	*	*
	leptokurtic	M, WRS, BY	*	*	*	*	*

Note: “ * ” = All five test statistics

กำหนดอัตราส่วนค่าเฉลี่ยประชากรกลุ่มที่ 1 ต่อกลุ่มที่ 2 เท่ากับ 1.5 : 1 และ 2 : 1 ที่ระดับนัยสำคัญ 0.05 สามารถสรุปผลดังตารางที่ 4

จากตารางที่ 4 เมื่อกำหนดอัตราส่วนค่าเฉลี่ยประชากรกลุ่มที่ 1 ต่อกลุ่มที่ 2 เท่ากับ 1.5 : 1 และ 2 : 1 พบว่า เมื่อประชากรมีการแจกแจกปรกติ หรือประชากรที่มีความเบ้เท่ากับ 0 และความโด่งเท่ากับ 3 พบว่าสถิติทดสอบที มีค่ากำลังการทดสอบสูงสุดที่ทุกขนาดตัวอย่าง ส่วนเมื่อประชากรมีการแจกแจงเบ้ซ้ายพบว่าสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน มีค่ากำลังการทดสอบสูงสุดที่ทุกขนาดตัวอย่าง

เมื่อประชากรมีการแจกแจงสมมาตรและโด่งต่ำกว่าปรกติ พบว่าสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่างเท่ากับ (5,10) สถิติทดสอบบูตสแทร็ปยูนส์ มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่างเท่ากับ (10,10) และสถิติทดสอบยูนเวลช์ มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่างเท่ากับ (20,25) (25,25) (40,50) และ (50,50) ส่วนกรณีเมื่อประชากรมีการแจกแจงสมมาตรและโด่งสูงกว่าปรกติ พบว่าสถิติทดสอบบูตสแทร็ปยูนส์ มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่าง (5,10) และ (20,25) สถิติทดสอบยูนเวลช์ มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่าง (10,10) และสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่าง (25,25) (40,50) และ (50,50)

เมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งต่ำกว่าปรกติ พบว่าสถิติทดสอบที มีค่ากำลังการทดสอบสูงสุดที่ทุกขนาดตัวอย่าง ยกเว้นสถิติทดสอบ

Table 4 Test statistics having the highest power of the test for nine combinations skewness and kurtosis when the ratio of the population means are 1.5 : 1 and 2 : 1 at the significant level of 0.05.

Distribution	(n_1,n_2)	**Mean ($\mu_1 : \mu_2$) is 1.5 : 1**	**Mean ($\mu_1 : \mu_2$) is 2 : 1**
Negatively skewed and platykurtic ($\mu_3^* = -1, \mu_4^* = 2.5$)	(5,10)	WRS	WRS
	(10,10)	WRS	WRS
	(20,25)	WRS	WRS
	(25,25)	WRS	WRS
	(40,50)	WRS	WRS, YW, BY
	(50,50)	t, WRS	*
Negatively skewed and mesokurtic ($\mu_3^* = -1, \mu_4^* = 3$)	(5,10)	WRS	WRS
	(10,10)	WRS	WRS
	(20,25)	WRS	*
	(25,25)	t, WRS	*
	(40,50)	*	*
	(50,50)	*	*
Negatively skewed and leptokurtic ($\mu_3^* = -1, \mu_4^* = 4$)	(5,10)	WRS	M, WRS
	(10,10)	WRS	*
	(20,25)	WRS	*
	(25,25)	t, WRS, YW	*
	(40,50)	*	*
	(50,50)	*	*
Symmetrical and platykurtic ($\mu_3^* = 0, \mu_4^* = 2.5$)	(5,10)	WRS	WRS
	(10,10)	BY	BY
	(20,25)	YW	YW
	(25,25)	YW	YW
	(40,50)	YW	YW
	(50,50)	YW, BY	YW
Normal ($\mu_3^* = 0, \mu_4^* = 3$)	(5,10)	t	t
	(10,10)	t	t
	(20,25)	t	t
	(25,25)	t	t
	(40,50)	t	t
	(50,50)	t	t

Table 4 Test statistics having the highest power of the test for nine combinations skewness and kurtosis when the ratio of the population means are 1.5 : 1 and 2 : 1 at the significant level of 0.05 (continued).

Distribution	(n_1,n_2)	Mean $(\mu_1 : \mu_2)$ is 1.5 : 1	Mean $(\mu_1 : \mu_2)$ is 2 : 1
Symmetrical and	(5,10)	BY	BY
leptokurtic	(10,10)	YW	YW
$(\mu_3^* = 0, \mu_4^* = 4)$	(20,25)	BY	BY
	(25,25)	WRS	WRS
	(40,50)	WRS	WRS
	(50,50)	WRS	WRS
Positively skewed and	(5,10)	BY	BY
platykurtic	(10,10)	t	t
$(\mu_3^* = 1, \mu_4^* = 2.5)$	(20,25)	t	t
	(25,25)	t	t
	(40,50)	t	t
	(50,50)	t	t
Positively skewed and	(5,10)	WRS	WRS
mesokurtic	(10,10)	t	t
$(\mu_3^* = 1, \mu_4^* = 3)$	(20,25)	t	t
	(25,25)	t	t
	(40,50)	t	t
	(50,50)	t	t
Positively skewed and	(5,10)	WRS	WRS
leptokurtic	(10,10)	WRS	WRS
$(\mu_3^* = 1, \mu_4^* = 4)$	(20,25)	t, WRS	t, M, WRS
	(25,25)	t, WRS	t, M, WRS, YW
	(40,50)	*	*
	(50,50)	*	*

Note: " * " = All five test statistics

บูตสแทร็ปยูนส์ที่ขนาดตัวอย่างเท่ากับ (5,10) ส่วนกรณีเมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งปรกติ พบว่าสถิติทดสอบที มีค่ากำลังการทดสอบสูงสุดที่ทุกขนาดตัวอย่าง ยกเว้นสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน ที่ขนาดตัวอย่างเท่ากับ (5,10) และเมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งสูงกว่าปรกติ พบว่าสถิติทดสอบทั้ง 5 วิธี ที่ขนาดตัวอย่างเป็น (40,50) และ (50,50) และสถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน ที่ขนาดตัวอย่างเท่ากับ (5,10) และ (10,10) มีค่ากำลังการทดสอบสูงสุด

และเมื่อมีการแจกแจงเบ้ซ้ายและการแจกแจงเบ้ขวาและโด่งมาก ที่ตัวอย่างขนาดปานกลางและใหญ่เกือบทุกสถานการณ์ ค่ากำลังการทดสอบของสถิติทดสอบทีและสถิติทดสอบผลบวกลำดับที่ของวิลคอกซันมีค่ากำลังการทดสอบที่สูงสุดและเท่ากัน ซึ่งสอดคล้องกับ Dwivedi และคณะ [3]

6. สรุป

จากผลการศึกษาพบว่าสถิติทดสอบทั้ง 5 วิธี สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้เกือบทุกสถานการณ์

เมื่อพิจารณาค่ากำลังการทดสอบเมื่อประชากรมีการแจกแจงปรกติควรใช้สถิติทดสอบที ส่วนกรณีเมื่อประชากรมีการแจกแจงเบ้ซ้ายควรเลือกใช้สถิติทดสอบผลบวกลำดับที่ของวิลคอกซันที่ทุกขนาดตัวอย่าง

เมื่อประชากรมีการแจกแจงสมมาตรและโด่งน้อยควรเลือกใช้สถิติทดสอบยูนเวลช์ ในกรณีตัวอย่างขนาดปานกลางและใหญ่ และเมื่อประชากรมีการแจกแจงสมมาตรและโด่งมากควรเลือกใช้สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน ในกรณีตัวอย่างขนาดใหญ่

เมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งน้อยควรเลือกใช้สถิติทดสอบบูตสแทร็ปยูนส์ ส่วนเมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งปรกติเลือกใช้สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน ในกรณีตัวอย่างขนาดเล็ก นอกจากนี้เมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งมากเลือกใช้สถิติทดสอบได้ทั้ง 5 วิธี ในกรณีตัวอย่างขนาดใหญ่ และเลือกใช้สถิติทดสอบผลบวกลำดับที่ของวิลคอกซัน ในกรณีตัวอย่างขนาดเล็กและปานกลาง

7. อภิปรายผล

เมื่อพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เมื่อประชากรมีการแจกแจงเบ้และโด่ง 9 ลักษณะ พบว่าสถิติทดสอบทั้ง 5 วิธี สามารถควบคุมความน่าจะเป็นของความเคลื่อนประเภทที่ 1 ได้ ซึ่งสอดคล้องกับมนตรี [5] ยกเว้นกรณีที่ขนาดตัวอย่างเท่ากัน เมื่อประชากรมีการแจกแจงสมมาตรและโด่งน้อยที่มีตัวอย่างขนาดเล็ก และเมื่อประชากรมีการแจกแจงเบ้ขวาและโด่งน้อยที่มีตัวอย่างขนาดใหญ่

เมื่อพิจารณาค่ากำลังการทดสอบ พบว่าการแจกแจงเบ้ซ้าย สถิติทดสอบผลบวกลำดับที่ของวิลคอกซันมีค่ากำลังการทดสอบสูงสุด ที่ขนาดตัวอย่างเล็ก ปานกลางและใหญ่ ส่วนกรณีการแจกแจงสมมาตรและโด่งน้อย สถิติทดสอบยูนเวลช์มีค่ากำลังการทดสอบสูงสุด ที่ขนาดตัวอย่างปานกลางและใหญ่ และกรณีการแจกแจงสมมาตรและโด่งมาก สถิติทดสอบผลบวกลำดับที่ของวิลคอกซันมีค่ากำลังการทดสอบสูงสุด ที่ขนาดตัวอย่างใหญ่

กรณีการแจกแจงเบ้ขวาและโด่งน้อย สถิติทดสอบบูตสแทร็ปยูนส์มีค่ากำลังการทดสอบสูงสุด ที่ขนาดตัวอย่างเล็ก และกรณีการแจกแจงเบ้ขวาและโด่งปรกติและโด่งมาก สถิติทดสอบผลบวกลำดับที่ของวิลคอกซันมีค่ากำลังการทดสอบสูงสุด ที่ขนาดตัวอย่างเล็ก ซึ่งแตกต่างจาก Fagerland และ Sandvik [2] พบว่าการแจกแจงเบ้ขวา สถิติทดสอบยูนเวลช์มีค่ากำลังการทดสอบสูงสุดที่ขนาดตัวอย่างมากกว่า 30

8. References

[1] Yuen, K.K., 1974, The Two-Sample Trimmed T for Unequal Population Variances, Biometrika, 61: 165-170.

[2] Fagerland, M.W. and L., Sandvik., 2009, Performance of Five Two-Sample Location Tests for Skewed Distributions with Unequal Variances, Contemporary Clinical Trial, 30 (5): 490-496.

[3] Dwivedi, A.K., I, Mallawaarachchi. and L.A., Alvarado., 2017, Analysis of Small Sample Size Studies Using Nonparametric Bootstrap Test with Pooled Resampling Method, Statistic in Medicine, 36 (14): 2187-2205.

[4] Montri, S., 2020, Nonparametric Bootstrap Method for Location Testing between Two Populations under Combined Assumption Violations, Science Journal, 25(3): 864-879. (in Thai)

[5] Montri, S., 2004, A Comparative Study of Parametric and Nonparametric Test Statistic for Testing the Difference between Two population, Master Thesis, Kasetsart University, Bangkok, 95 p. (in Thai)

[6] Pitsamai, H., 2017, Principles of Statistics I, (7th ed), Kasetsart University Press, Bangkok, 339 p. (in Thai)

[7] Saichon, S., 2020, Nonparametric, (2nd ed), Chamchuree Products co.,ltd., Bangkok, 650 p. (in Thai)

[8] Arpha, W., 2015, A Comparison of Nonparametric Statistics for Testing the Mean Difference between Two Independent Populations with Small Sample Sizes, Master Thesis, Kasetsart University, Bangkok, 95 p. (in Thai)

[9] Luh, W.M. and J.H., Guo., 2007, Approximate Sample Size Formulas for the Two-Sample Trimmed Mean Test with Unequal Variances, British Journal of Mathematics and Statistical Psychology, 60: 137-146.

[10] Wilcox, R.R., 2005, Introduction to Robust Estimation and Hypothesis Testing, (2nd ed), San Diego, CA: Acedamic Press.