การเปรียบเทียบตัวสถิติทดสอบไม่อิงพารามิเตอร์ สำหรับการทดสอบค่าเฉลี่ยระหว่างประชากร 2 กลุ่ม ที่เป็นอิสระกัน

Comparison of Nonparametric Test Statistics for Testing Population Mean between Two Independent Populations

ธิดารัตน์ สมฤทธิ์, ปรียา อยู่เล่ห์, รติวัฒน์ สิทธิบุตร์*,
วีรยุทธ รักษาธรรม และสายชล สินสมบูรณ์ทอง

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ถนนฉลองกรุง เขตลาดกระบัง กรุงเทพมหานคร 10520

Thidarat Somrit, Preeya Yoole, Ratiwat Sitdhibutr*,
Wirayut Raksatham and Saichon Sinsomboonthong

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang,
Chalongkrung Road, Ladkrabang, Bangkok 10520

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบกำลังการทดสอบของตัวสถิติทดสอบไม่อิงพารามิเตอร์ (nonparametric statistic) ระหว่างค่าเฉลี่ยของประชากร 2 กลุ่ม ที่เป็นอิสระกัน ซึ่งมีตัวสถิติทดสอบที่ศึกษา 3 การทดสอบ คือ Wilcoxon-Mann-Whitney (WMW), Brunner and Munzel (BM) และ Bootstrap Rank Welch (BRW) โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ ที่ระดับนัยสำคัญเท่ากับ 0.01, 0.05 และ 0.1 เมื่อประชากรทั้งสองกลุ่มมีการแจกแจงเหมือนกันและการแจกแจงต่างกัน ได้แก่ การแจกแจงปรกติ การแจกแจงแกมมา การแจกแจงไคกำลังสอง และการแจกแจงเลขชี้กำลัง โดยกำหนดขนาดตัวอย่างเท่ากัน คือ (10,10), (20,20) และ (50,50) และขนาดตัวอย่างต่างกัน คือ (5,10), (20,30) และ (50,100) และทำซ้ำในแต่ละสถานการณ์จำนวน 1,000 รอบ ผลการวิจัยพบว่าตัวสถิติ BRW มีกำลังการทดสอบสูงสุดในทุกกรณี แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อยที่สุด ในขณะที่ตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีที่การแจงเหมือนกัน ความแปรปรวนเท่ากันและในกรณีการแจกแจงต่างกัน ความแปรปรวนต่างกัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดในทุกกรณี ยกเว้นในกรณีการแจกแจงต่างกัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างต่างกัน และในกรณีการแจกแจงต่างกัน ความแปรปรวนต่างกัน ขนาดตัวอย่างเท่ากัน และตัวสถิติ BM มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีการแจกแจงเหมือนกัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน ในกรณีการแจงแจงเหมือนกัน ความแปรปรวนต่างกัน และในกรณีการแจกแจงต่างกัน ความแปรปรวนเท่ากัน

***ผู้รับผิดชอบบทความ :** gmr.thep@gmail.com

คำสำคัญ : ประชากร 2 กลุ่ม ที่เป็นอิสระกัน; ความน่าจะเป็นของความผิดพลาดแบบที่ 1; กำลังการทดสอบ

Abstract

The objective of this research was to study and to compare the powers of a test for testing means between two independent populations. The test statistics used included Wilcoxon-Mann-Whitney, Brunner and Munzel, and bootstrap rank Welch. These statistical methods were compared by considering their ability to control probability of type I error at 0.01, 0.05, and 0.1 level of significance and their power of test. The two populations were distributed according to normal gamma exponential and Chi-squared distributions. Sample sizes were defined to be equal at (10,10), (20,20), and (50,50) and unequal at (5,10), (20,30), and (50,100), and the type I error and power of test were determined by running the software 1,000 times for each cases. The results revealed that bootstrap rank Welch yielded the highest power of a test for every case but had the lowest ability to control the probability of type I error, whereas Wilcoxon-Mann-Whitney yielded the second highest power of a test in the cases that the distributions were identical and the variances were equal and in the cases that the distributions were not identical and the variances were different. Wilcoxon-Mann-Whitney had the highest ability to control the probability of type I error for every case except for the cases that the distributions were not identical, the variances were equal, and the sample sizes were different as well as the cases that the distributions were not identical, the variances were different, and the sample sizes were the same. Brunner and Munzel yielded the second highest power of a test in the following cases: (1) the distributions were identical, the variances were the same, and the sample sizes were the same, (2) the distributions were identical and the variances were different, (3) the distributions were not identical and the variances were the same.

Keywords: independent population; probability of type I error; power of a test

1. บทนำ

สถิติที่ใช้ในงานวิจัยนั้น ได้รับการพัฒนามาเพื่อใช้ให้เหมาะสมกับสถานการณ์ต่าง ๆ กัน เพื่อให้ผลลัพธ์ที่เชื่อถือได้ในการหาข้อสรุปไปยังประชากร ซึ่งจำเป็นอย่างยิ่งที่ผู้วิจัยต้องรู้จักเลือกใช้ให้เหมาะสมกับคุณลักษณะข้อมูล โดยเฉพาะตัวสถิติทดสอบ (test statistic) ความแตกต่างระหว่างกลุ่มประชากร 2 กลุ่ม ซึ่งเป็นวิธีที่ใช้กันมากในงานวิจัย [1] การทดสอบดังกล่าวมีสถิติทดสอบหลายตัวให้เลือกใช้ตามความเหมาะสมสำหรับวัตถุประสงค์และข้อมูลที่มีอยู่ ถ้าตัวอย่างสองกลุ่มถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปรกติ โดยที่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร การทดสอบสมมติฐานอาจทำโดยการใช้สถิติทดสอบ z (z-test) และเมื่อไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร ควรใช้สถิติทดสอบ t (t-test) โดยที่สถิติทดสอบที่ขึ้นอยู่กับการแจกแจงหรือ

พารามิเตอร์ของการแจกแจงนั้นเรียกว่าสถิติทดสอบอิงพารามิเตอร์ (parametric statistic) แต่ในทางปฏิบัติบ่อยครั้งพบว่าข้อมูลที่นำมาใช้ในการวิเคราะห์นั้นไม่เป็นไปตามข้อสมมติ (assumption) จึงอาจส่งผลให้ข้อสรุปผิดพลาดได้ เพื่อแก้ปัญหานี้สถิติทดสอบไม่อิงพารามิเตอร์ (nonparametric statistic) จึงเป็นทางเลือกหนึ่งที่นักวิจัยเลือกใช้ [2]

สถิติทดสอบไม่อิงพารามิเตอร์สาหรับใช้ศึกษาเปรียบเทียบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่ม มีหลายวิธี โดยมีผู้ศึกษาไว้หลายท่าน พ.ศ. 2556 ชนาธิป และคณะ [3] ได้ศึกษาความน่าเชื่อถือของตัวสถิติทดสอบที่ไม่อิงพารามิเตอร์สำหรับตัวอย่าง 2 กลุ่มอิสระกัน พบว่าตัวสถิติทดสอบ WMW มีกำลังการทดสอบ (power of a test) สูงสุด และในปีเดียวกัน อุมาพร [4] ได้ศึกษาความน่าเชื่อถือของสถิติทดสอบที่ไม่อิงพารามิเตอร์สำหรับตัวอย่าง 2 กลุ่มอิสระกัน จากโปรแกรมสำเร็จรูปทางสถิติ SPSS พบว่าตัวสถิติ WMW สามารถควบคุมค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 (type I error) ได้เฉพาะกรณีที่ประชากร 2 กลุ่ม มีการแจกแจงเหมือนกัน แต่ในกรณีที่ประชากร 2 กลุ่ม มีการแจกแจงต่างกัน พบว่าสามารถควบคุมค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อย และพบว่ามีกำลังการทดสอบสูง ทั้งในกรณีการแจกแจงเหมือนกันและต่างกัน ในปี ค.ศ. 2005 Reiczigel และคณะ [5] ได้ศึกษาตัวสถิติทดสอบความแตกต่างของตำแหน่งของประชากร 2 กลุ่ม สามารถสรุปได้ว่า BRW สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีและมีกำลังการทดสอบสูงสุดในกรณีที่ 2 ตัวอย่างมีการแจกแจงไม่เป็นปรกติ พ.ศ 2553 นพดล และชินนพงษ์ [6] ได้ศึกษาการเปรียบเทียบตัวสถิติทดสอบไม่อิงพารามิเตอร์ระหว่างประชากร 2 กลุ่ม เมื่อความแปรปรวนไม่เท่ากัน จากผลการวิจัยพบว่า BRW ให้กำลังการทดสอบสูงสุด แต่ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ ในขณะที่ตัวสถิติทดสอบ BM ให้กำลังการทดสอบสูงใกล้เคียงกับตัวสถิติ BRW และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ทุกกรณี จึงควรเลือกใช้ตัวสถิติ BRW ส่วนในปี พ.ศ. 2558 อาภา และคณะ [2] ศึกษาการเปรียบเทียบตัวสถิติทดสอบที่ไม่อิงพารามิเตอร์สำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากรสองกลุ่มที่เป็นอิสระกันเมื่อขนาดตัวอย่างเล็ก จากผลการวิจัยพบว่าตัวสถิติ BM มีกำลังการทดสอบสูงสุด ใน พ.ศ. 2556 วราวัลย์ [7] ได้ศึกษาความแตกต่างของตำแหน่งของประชากร 2 กลุ่ม ของสถิติทดสอบไม่อิงพารามิเตอร์ กรณีความแปรปรวนของประชากรไม่เท่ากัน พบว่าเมื่อตัวอย่างขนาดเล็ก ตัวสถิติ BM สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุด และมีกำลังการทดสอบสูงสุด เมื่อตัวอย่างขนาดปานกลาง ตัวสถิติ WMW, ตัวสถิติ BM สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ใกล้เคียงกัน และมีกำลังการทดสอบสูงสุด

ด้วยเหตุผลดังกล่าว ทำให้ผู้วิจัยสนใจศึกษากำลังการทดสอบของตัวสถิติทดสอบความเท่ากันของค่าเฉลี่ยประชากร 2 กลุ่มที่อิสระกัน 3 ตัว คือ ตัวสถิติ WMW, ตัวสถิติ BM และตัวสถิติ BRW โดยสร้างแบบจำลองจากเทคนิคมอนติคาร์โล (Monte Carlo technique) ในการสร้างข้อมูลตามกรณีต่าง ๆ ที่ต้องการ ซึ่งในแต่ละกรณีจะเลือกตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ มาเปรียบเทียบกำลังการทดสอบเพื่อเป็นแนวทางในการเลือกใช้ตัวสถิติทดสอบได้อย่างเหมาะสมกับลักษณะข้อมูล

2. วิธีการวิจัย

2.1 การวางแผนการวิจัย

การวิจัยครั้งนี้กำหนดสถานการณ์ในการศึกษาเปรียบเทียบ ดังนี้

2.1.1 กำหนดจำนวนประชากร 2 กลุ่ม

2.1.2 กำหนดระดับนัยสำคัญ 0.01, 0.05 และ 0.10

2.1.3 กำหนดขนาดตัวอย่างสุ่มจากแต่ละประชากรเท่ากัน คือ (10,10), (20,20), (50,50) และขนาดตัวอย่างสุ่มจาก แต่ละประชากรต่างกัน คือ (10,15), (20,30), (50,100)

2.1.4 กำหนดข้อมูลที่สุ่มจากประชากรที่มีการแจกแจงปรกติ ด้วยพารามิเตอร์ (μ,σ^2) การแจกแจงแกมมา ด้วยพารามิเตอร์ (α,β) การแจกไคกำลังสอง ด้วยพารามิเตอร์ (ν) และการแจกแจงเลขชี้กำลังด้วยพารามิเตอร์ (β) โดยมีค่าพารามิเตอร์ และกรณีต่าง ๆ ดังตารางที่ 1

(1) ข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน (เช่น ปรกติกับปรกติ แกมมากับแกมมา)

(1.1) ค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน (ตารางที่ 2)

(1.2) ค่าเฉลี่ยเท่ากันและความแปรปรวนต่างกัน (ตารางที่ 3)

(1.3) ค่าเฉลี่ยต่างกันและความแปรปรวนเท่ากัน (ตารางที่ 4)

(1.4) ค่าเฉลี่ยต่างกันและความแปรปรวนต่างกัน (ตารางที่ 5)

(2) ข้อมูลทั้งสองกลุ่มที่มีการแจกแจงต่างกัน (เช่น ปรกติกับแกมมา ปรกติกับเลขชี้กำลัง)

(2.1) ค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน (ตารางที่ 6)

ตารางที่ 1 สัญลักษณ์แสดงพารามิเตอร์ของการแจกแจง ค่าเฉลี่ย $(E(X))$ และความแปรปรวน $(Var(X))$ ของประชากร

การแจกแจง	สัญลักษณ์	$E(X)$	$Var(X)$
ปรกติ	$N(\mu,\sigma^2)$	μ	σ^2
แกมมา	$Gamma(\alpha,\beta)$	$\alpha\beta$	$\alpha\beta^2$
เลขชี้กำลัง	$Exponential(\beta)$	β	β^2
ไคกำลังสอง	$Chi\text{-}Squared(\nu)$	ν	2ν

$E(X)$ คือ ค่าเฉลี่ยของประชากร; $Var(X)$ คือ ความแปรปรวนของประชากร

ตารางที่ 2 ค่าพารามิเตอร์ในกรณีการแจกแจงเหมือนกัน ค่าเฉลี่ยเท่ากัน และความแปรปรวนเท่ากัน

การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(100,20)$ และ $X_2 \sim N(100,20)$	100	20	100	20
(2) $X_1 \sim Gamma(10,2)$ และ $X_2 \sim Gamma(10,2)$	20	40	20	40
(3) $X_1 \sim Exponential(5)$ และ $X_2 \sim Exponential(5)$	5	25	5	25
การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(4) $X_1 \sim Chi\text{-}Squared(15)$ และ $X_2 \sim Chi\text{-}Squared(15)$	15	30	15	30

ตารางที่ 3 ค่าพารามิเตอร์ในกรณีการแจกแจงเหมือนกัน ค่าเฉลี่ยเท่ากัน และความแปรปรวนต่างกัน

การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(100,10)$ และ $X_2 \sim N(100,20)$	100	10	100	20
(2) $X_1 \sim Gamma(10,2)$ และ $X_2 \sim Gamma(5,4)$	20	40	20	80

ตารางที่ 4 ค่าพารามิเตอร์ในกรณีการแจกแจงเหมือนกัน ค่าเฉลี่ยต่างกัน และความแปรปรวนเท่ากัน

การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(10,20)$ และ $X_2 \sim N(15,20)$	10	20	15	20
(2) $X_1 \sim Gamma(2,4)$ และ $X_2 \sim Gamma(8,2)$	8	32	16	32

ตารางที่ 5 ค่าพารามิเตอร์ในกรณีการแจกแจงเหมือนกัน ค่าเฉลี่ยต่างกัน และความแปรปรวนต่างกัน

การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(10,10)$ และ $X_2 \sim N(15,20)$	10	10	15	20
(2) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Gamma(10,2)$	10	10	20	40
(3) $X_1 \sim Exponential(5)$ และ $X_2 \sim Exponential(8)$	5	25	8	64
(4) $X_1 \sim Chi\text{-}Squared(10)$ และ $X_2 \sim Chi\text{-}Squared(15)$	10	20	15	30

ตารางที่ 6 ค่าพารามิเตอร์ในกรณีการแจกแจงต่างกัน ค่าเฉลี่ยเท่ากัน และความแปรปรวนเท่ากัน

การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(10,10)$ และ $X_2 \sim Gamma(10,1)$	10	10	10	10
(2) $X_1 \sim N(10,100)$ และ $X_2 \sim Exponential(10)$	10	100	10	100
(3) $X_1 \sim N(10,20)$ และ $X_2 \sim Chi\text{-}Squared(10)$	10	20	10	20

ตารางที่ 7 ค่าพารามิเตอร์ในกรณีการแจกแจงต่างกัน ค่าเฉลี่ยเท่ากัน และความแปรปรวนต่างกัน

การแจกแจงต่างกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(10,2)$ และ $X_2 \sim Gamma(10,1)$	10	2	10	10
การแจกแจงเหมือนกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(2) $X_1 \sim N(10,2)$ และ $X_2 \sim Exponential(10)$	10	2	10	100
(3) $X_1 \sim N(10,2)$ และ $X_2 \sim Chi\text{-}Squared(10)$	10	2	10	20
(4) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Exponential(10)$	10	10	10	100
(5) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Chi\text{-}Squared(10)$	10	10	10	20
(6) $X_1 \sim Exponential(10)$ และ $X_2 \sim Chi\text{-}Squared(10)$	10	100	10	20

ตารางที่ 8 ค่าพารามิเตอร์ในกรณีการแจกแจงต่างกัน ค่าเฉลี่ยต่างกัน และความแปรปรวนเท่ากัน

การแจกแจงต่างกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(15,10)$ และ $X_2 \sim Gamma(10,1)$	15	10	10	10
(2) $X_1 \sim N(10,25)$ และ $X_2 \sim Exponential(5)$	10	25	5	25
(3) $X_1 \sim N(10,40)$ และ $X_2 \sim Chi\text{-}Squared(20)$	10	40	20	40
(4) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Exponential(\sqrt{10})$	10	10	$\sqrt{10}$	10
(5) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Chi\text{-}Squared(5)$	10	10	5	10
(6) $X_1 \sim Exponential(10)$ และ $X_2 \sim Chi\text{-}Squared(50)$	10	100	50	100

ตารางที่ 9 ค่าพารามิเตอร์ในกรณีการแจกแจงต่างกัน ค่าเฉลี่ยต่างกัน และความแปรปรวนต่างกัน

การแจกแจงต่างกัน	$E(X_1)$	$Var(X_1)$	$E(X_2)$	$Var(X_2)$
(1) $X_1 \sim N(15,2)$ และ $X_2 \sim Gamma(10,1)$	15	2	10	10
(2) $X_1 \sim N(15,2)$ และ $X_2 \sim Exponential(10)$	15	2	10	100
(3) $X_1 \sim N(15,2)$ และ $X_2 \sim Chi\text{-}Squared(20)$	15	2	20	40
(4) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Exponential(20)$	10	10	20	400
(5) $X_1 \sim Gamma(10,1)$ และ $X_2 \sim Chi\text{-}Squared(20)$	10	10	20	40
(6) $X_1 \sim Exponential(10)$ และ $X_2 \sim Chi\text{-}Squared(20)$	10	100	20	40

(2.2) ค่าเฉลี่ยเท่ากันและความแปรปรวนต่างกัน (ตารางที่ 7)

(2.3) ค่าเฉลี่ยต่างกันและความแปรปรวนเท่ากัน (ตารางที่ 8)

(2.4) ค่าเฉลี่ยต่างกันและความแปรปรวนต่างกัน (ตารางที่ 9)

2.1.5 คำนวณความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของตัวสถิติทดสอบตามเกณฑ์ของ Bradley ในแต่ละสถานการณ์

2.1.6 คำนวณกำลังการทดสอบของตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley [8] ใน แต่ละสถานการณ์

2.1.7 เปรียบเทียบกำลังการทดสอบของตัวสถิติทดสอบและความน่าจะเป็นของความผิดพลาดแบบที่ 1 แล้วเลือกตัวสถิติทดสอบ โดยมีเกณฑ์ดังนี้

(1) ถ้ากำลังการทดสอบมีค่าสูงที่สุดและสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley ได้มากที่สุดหรือปานกลาง จะเลือกตัวสถิติทดสอบนั้น

(2) ถ้ากำลังการทดสอบมีค่าสูงที่สุด แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley ได้น้อยที่สุด จะพิจารณาตัวสถิติทดสอบที่มีกำลังการทดสอบรองลงมาแทน (กำลังการทดสอบปานกลาง)

(3) ถ้ากำลังการทดสอบมีค่าต่ำที่สุดจะไม่นำตัวสถิติทดสอบนั้นมาพิจารณา

2.2 วิธีการดำเนินการวิจัย

งานวิจัยครั้งนี้จะดำเนินงานตามขั้นตอน โดยแบ่งออกเป็น 2 ขั้นตอน ดังนี้

2.2.1 การคำนวณความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 (α) ของตัวสถิติทดสอบ

(1) จำลองข้อมูลในแต่ละขนาดจากประชากรที่มีการแจกแจงปรกติ การแจกแจงแกมมา การแจกแจงไคกำลังสอง และการแจกแจงเลขชี้กำลัง ให้มีพารามิเตอร์ตามที่ต้องการ โดยกำหนดค่า seeding number = 10 เพื่อให้การสุ่มข้อมูลในครั้งต่อไปได้ค่าเท่าเดิม โดยโปรแกรมอาร์ (R) เวอร์ชั่น 3.3.2 [9]

(2) คำนวณตัวสถิติทดสอบทั้ง 3 การทดสอบ ได้แก่ ตัวสถิติ WMW, ตัวสถิติ BM และตัวสถิติ BRW โดยใช้คำสั่งจากโปรแกรมอาร์ (R) เวอร์ชั่น 3.3.2 ซึ่งตัวสถิติทดสอบทั้ง 3 มีรายละเอียดดังนี้

(2.1) Wilcox [10] ได้เสนอตัวสถิติ WMW เป็นการทดสอบที่ใช้ผลรวมของลำดับ โดย ตัวสถิติ WMW มีประสิทธิภาพในการทดสอบสูง โดยจะมีความไวต่อการปฏิเสธสมมติฐานหลัก (H_0)

สมมติฐานการทดสอบ

H_0: $M_X = M_Y$

H_1: $M_X \neq M_Y$

ตัวสถิติทดสอบ คือ $T = S - \frac{m(n+1)}{2}$

เมื่อ S คือ ผลรวมลำดับที่ของตัวอย่างขนาด m ในข้อมูลที่เรียงลำดับแล้ว, m คือ ประชากรของกลุ่มที่ 1, n คือ ประชากรของกลุ่มที่ 1

ใช้ตารางค่าแสดงเขตวิกฤต คือ $T < W_{\alpha/2}$ หรือ $T > W_{1-\alpha/2}$ เมื่อ $W_{1-\alpha/2} = n_1 n_2 - W_{\alpha/2}$

(2.2) Brunner และ Munzel [11] ได้เสนอตัวสถิติ BM เป็นตัวสถิติที่ใช้ลำดับที่ในการคำนวณ ซึ่งได้ปรับปรุงมาจากตัวสถิติ WMW ตัวสถิติทดสอบนี้จึงเหมาะสำหรับตัวอย่างที่มีขนาดเล็ก $(m, n \le 10)$

สมมติฐานการทดสอบ

$H_0 : P(X < Y) = P(X > Y)$ หรือ

H_0: ข้อมูลทั้ง 2 มีลักษณะเหมือนกัน

$H_1 : P(X < Y) \neq P(X > Y)$

H_1: ข้อมูลทั้ง 2 มีลักษณะต่างกัน

ตัวสถิติทดสอบ คือ

$$BM = \frac{mn(\bar{R}_Y - \bar{R}_X)}{(m+n)\sqrt{ms_X^2 + ns_Y^2}}$$

โดยที่ $s_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(R_{X_i} - W_{X_i} - \bar{R}_X + \frac{m+1}{2}\right)^2$,

$$s_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(R_{Y_i} - W_{Y_i} - \bar{R}_Y + \frac{n+1}{2}\right)^2,$$

$\bar{R}_X = \frac{1}{m}\sum_{i=1}^{m} R_{X_i}$ และ $\bar{R}_Y = \frac{1}{n}\sum_{i=1}^{n} R_{Y_i}$

เมื่อ $\bar{R}_X$ คือ ค่าเฉลี่ยของลำดับที่ในตัวอย่างที่ 1, $\bar{R}_Y$ คือ ค่าเฉลี่ยของลำดับที่ในตัวอย่างที่ 2, R_{X_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 1 เมื่อนำข้อมูลทั้ง 2 กลุ่มมารวมกัน, R_{Y_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 2 เมื่อนำข้อมูลทั้ง 2 กลุ่มมารวมกัน, W_{X_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 1 เมื่อนำข้อมูลกลุ่มที่ 1 มาเรียงลำดับ, W_{Y_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 2 เมื่อนำข้อมูลกลุ่มที่ 2 มาเรียงลำดับ

ซึ่งจะปฏิเสธสมมติฐานว่าง (H_0) ที่ระดับนัยสำคัญ α เมื่อตัวสถิติ BM มีค่าน้อยกว่า $-t_{\left(\frac{\alpha}{2}, df_{BW}\right)}$ หรือมีค่ามากกว่า $t_{\left(\frac{\alpha}{2}, df_{BW}\right)}$

(2.3) Reiczigel และคณะ [5] ได้เสนอตัวสถิติ BRW เพื่อใช้ทดสอบความแตกต่างของตำแหน่ง ของประชากร 2 กลุ่ม จะให้ค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในระดับที่ดี และมี

กำลังการทดสอบสูง เมื่อตัวอย่างมีขนาดกลางและขนาดใหญ่ $(m, n \geq 30)$

สมมติฐานการทดสอบ

$H_0 : P(X < Y) = P(X > Y)$ หรือ

H_0: ข้อมูลทั้ง 2 มีลักษณะเหมือนกัน

$H_1 : P(X < Y) \neq P(X > Y)$

H_1: ข้อมูลทั้ง 2 มีลักษณะต่างกัน

ตัวสถิติทดสอบ $RW = \dfrac{\bar{R}_Y - \bar{R}_X}{\sqrt{\dfrac{S_X^2}{m} + \dfrac{S_Y^2}{n}}}$

โดยที่ $\bar{R}_X = \frac{1}{m}\sum_{i=1}^{m} R_{X_i}$, $\bar{R}_Y = \frac{1}{n}\sum_{i=1}^{n} R_{Y_i}$,

$S_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}(R_{X_i} - \bar{R}_X)^2$ และ $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(R_{Y_i} - \bar{R}_Y)^2$,

$$df_{RW} = \frac{(ns_X^2 + ms_Y^2)^2}{\dfrac{(ns_X^2)^2}{m-1} + \dfrac{(ms_Y^2)^2}{n-1}}$$

เมื่อ $\bar{R}_X$ คือ ค่าเฉลี่ยของลำดับที่ในตัวอย่างที่ 1, $\bar{R}_Y$ คือ ค่าเฉลี่ยของลำดับที่ในตัวอย่างที่ 2, R_{X_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 1 เมื่อนำข้อมูลทั้ง 2 กลุ่มมารวมกัน, R_{Y_i} คือ ลำดับที่ของหน่วยตัวอย่างที่ i ของกลุ่มที่ 2 เมื่อนำข้อมูลทั้ง 2 กลุ่มมารวมกัน

ซึ่งจะปฏิเสธสมมติฐานว่าง (H_0) ที่ระดับนัยสำคัญ α เมื่อตัวสถิติทดสอบ RW มีค่าน้อยกว่า $-t_{\left(\frac{\alpha}{2}, df_{RW}\right)}$ หรือมีค่ามากกว่า $t_{\left(\frac{\alpha}{2}, df_{RW}\right)}$

ขั้นตอนการดำเนินงาน

- สุ่มตัวอย่างทีละ 1 ค่า โดยใช้การสุ่มแบบใส่คืนจำนวน m ครั้ง จากชุดของตัวอย่าง $X_1, X_2, ..., X_m$ จะได้ $X_1^*, X_2^*, ..., X_m^*$

- สุ่มตัวอย่างทีละ 1 ค่า โดยใช้การสุ่มแบบใส่คืนจำนวน n ครั้ง จากชุดของตัวอย่าง $Y_1, Y_2, ..., Y_n$ จะได้ $Y_1^*, Y_2^*, ..., Y_n^*$

- นำ $X_1, X_2, ..., X_m$ และ $X_1^*, X_2^*, ..., X_m^*$ มาคำนวณค่าตัวสถิติทดสอบแรงค์เวลช์ (rank Welch test)

- เมื่อได้ค่าตัวสถิติทดสอบแรงค์เวลช์เรียบร้อยแล้ว ให้กลับไปทำซ้ำข้อ 1-3 ทำเช่นนี้จนครบจำนวนครั้งของการบูทสแตรปทั้งหมด 1,000 ครั้ง

- หาค่าเฉลี่ยของค่าตัวสถิติ RW ทั้ง 1,000 ครั้ง และนำมาเปรียบเทียบกับอาณาเขตวิกฤต

(3) นำค่าตัวสถิติทดสอบที่คำนวณได้เทียบกับค่าวิกฤตเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานว่าง ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10 และบันทึกจำนวนครั้งของการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นจริง ทำซ้ำจนครบ 1,000 ครั้ง

(4) หาความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง (H_0) คือ ความน่าจะเป็นของความผิดพลาดแบบที่ 1 = จำนวนครั้งของการปฏิเสธ H_0 เมื่อ H_0 เป็นจริง ÷ 1,000

ถ้าความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของการทดสอบสำหรับแต่ละสถานการณ์มีค่าอยู่ในช่วงที่ได้กำหนดไว้ในเกณฑ์ของการเปรียบเทียบวิธีการทดสอบ ได้แก่ เกณฑ์ของ Bradley จะถือว่าตัวสถิติทดสอบนั้นมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ โดยมีรายละเอียดของเกณฑ์ของ Bradley ดังนี้ คือ ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองอยู่ในช่วง (0.005,0.015) สำหรับการทดสอบที่ระดับนัยสำคัญ 0.01 ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองอยู่ในช่วง (0.025,0.075) สำหรับการทดสอบที่ระดับนัยสำคัญ 0.05 และถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองอยู่ในช่วง

(0.05,0.15) สำหรับการทดสอบที่ระดับนัยสำคัญ 0.1 จะสรุปได้ว่าตัวสถิติทดสอบนั้นสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

2.2.2 การคำนวณกำลังการทดสอบ $(1-\beta)$ ของตัวสถิติทดสอบ

สุ่มตัวอย่างตามขอบเขตของงานวิจัย คำนวณค่าตัวสถิติทดสอบเทียบกับค่าวิกฤตและสรุปผล ทำซ้ำจนครบ 1,000 ครั้ง แล้วนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง (H_0) ดังนี้ กำลังการทดสอบ = จำนวนครั้งของการปฏิเสธ H_0 เมื่อ H_0 ไม่เป็นจริง ÷ 1,000

3. ผลการวิจัย

เมื่อพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของตัวสถิติทดสอบที่ไม่อิงพารามิเตอร์ระหว่างเฉลี่ยของประชากร 2 กลุ่มที่อิสระกัน สำหรับข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน และข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน สามารถสรุปผลดังตารางที่ 10-11

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน พบว่าตัวสถิติ WMW มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุดในทุกกรณี

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน พบว่า ตัวสถิติ WMW มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุดในกรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน และกรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนต่างกัน ขนาดตัวอย่างต่างกัน สำหรับ ตัวสถิติ BM มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุดในกรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างต่างกัน และกรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนต่างกัน ขนาดตัวอย่างเท่ากัน

ตารางที่ 10 ตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุดตามลำดับจากมากไปน้อย เมื่อจำลองข้อมูลมาจากประชากรทั้งสองกลุ่มที่มีการแจกแจงเหมือนกันโดยมีค่าพารามิเตอร์ต่าง ๆ ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10

กรณี	ข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน					
	ขนาดตัวอย่างเท่ากัน			ขนาดตัวอย่างต่างกัน		
	ระดับนัยสำคัญ			ระดับนัยสำคัญ		
	0.01	0.05	0.10	0.01	0.05	0.10
ค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน	1. WMW 2. BM 3. BRW	1. WMW, BM 2. BRW	1. WMW, BM และ BRW	1. WMW 2. BM 3. BRW	1. WMW, BM และ BRW	1. WMW, BM และ BRW
ค่าเฉลี่ยเท่ากันและความแปรปรวนต่างกัน	1. WMW 2. BM 3. BRW	1. WMW, BM 2. BRW	1. WMW, BM 2. BRW	1. WMW 2. BM 3. BRW	1. WMW, BM 2. BRW	1. WMW, BM 2. BRW

WMW หมายถึง ตัวสถิติทดสอบวิลคอกสัน-แมน-วิทนีย์; BM หมายถึง ตัวสถิติทดสอบบรุนเนอร์-มุนเซล; BRW หมายถึง ตัวสถิติทดสอบบูทสแตรปแรงค์เวลซ์

ตารางที่ 11 ตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุดตามลำดับจากมากไปน้อย เมื่อจำลองข้อมูลมาจากประชากรทั้งสองกลุ่มที่มีการแจกแจงต่างกันโดยมีค่าพารามิเตอร์ต่าง ๆ ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10

กรณี	ข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน					
	ขนาดตัวอย่างเท่ากัน			ขนาดตัวอย่างเท่ากัน		
	ระดับนัยสำคัญ			ระดับนัยสำคัญ		
	0.01	0.01	0.01	0.01	0.01	0.01
ค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน	1. WMW 2. BM 3. BRW	1. WMW 2. BM 3. BRW	1. WMW 2. BM 3. BRW	1. WMW 2. BM 3. BRW	1. WMW 2. BM 3. BRW	1. WMW 2. BM 3. BRW
ค่าเฉลี่ยเท่ากันและความแปรปรวนต่างกัน	1. WMW 2. BM	1. WMW 2. BM	1. WMW 2. BM	1. WMW 2. BM	1. WMW 2. BM	1. WMW 2. BM

เมื่อพิจารณาเปรียบเทียบกำลังการทดสอบของตัวสถิติทดสอบที่ไม่อิงพารามิเตอร์ระหว่างค่าเฉลี่ยของประชากร 2 กลุ่มที่อิสระกัน สำหรับข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกันและข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน สามารถสรุปผลดังตารางที่ 12-13

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน พบว่าตัวสถิติ BRW ให้กำลังการทดสอบมากที่สุดในทุกกรณี โดยที่ตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีที่ความแปรปรวนเท่ากัน แต่ตัวสถิติ BM ให้กำลังการทดสอบรองลงมาจาก ตัวสถิติ BRW ในกรณีความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน และในกรณีความแปรปรวนต่างกัน

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน พบว่าตัวสถิติ BRW ให้กำลังการทดสอบมากที่สุดในทุกกรณี โดยที่ตัวสถิติ WMW ให้ค่ากำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีความแปรปรวนต่างกัน แต่ตัวสถิติ BM ให้ค่ากำลังการทดสอบรองลงมาจาก ตัวสถิติ BRW ในกรณีความแปรปรวนเท่ากัน

ตารางที่ 12 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 2 ลำดับแรก เมื่อจำลองข้อมูลมาจากประชากรทั้งสองกลุ่มที่มีการแจกแจงเหมือนกันโดยมีค่าพารามิเตอร์ต่าง ๆ ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10

กรณี	ข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน					
	ขนาดตัวอย่างเท่ากัน			ขนาดตัวอย่างต่างกัน		
	ระดับนัยสำคัญ			ระดับนัยสำคัญ		
	0.01	0.05	0.10	0.01	0.05	0.10
ค่าเฉลี่ยต่างกันและความแปรปรวนเท่ากัน	1. BRW 2. BM	1. BRW 2. WMW	1. BRW 2. BM, WMW	1. BRW 2. WMW	1. BRW 2. WMW	1. BRW 2. WMW
ค่าเฉลี่ยต่างกันและความแปรปรวนต่างกัน	1. BRW 2. BM	1. BRW 2. BM	1. BRW 2. BM, WMW	1. BRW 2. BM	1. BRW 2. BM	1. BRW 2. BM

ตารางที่ 13 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 2 ลำดับแรก เมื่อจำลองข้อมูลมาจากประชากรที่มีข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกันที่มีค่าพารามิเตอร์ต่าง ๆ ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10

กรณี	ข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน					
	ขนาดตัวอย่างเท่ากัน			ขนาดตัวอย่างต่างกัน		
	ระดับนัยสำคัญ			ระดับนัยสำคัญ		
	0.01	0.05	0.10	0.01	0.05	0.10
ค่าเฉลี่ยต่างกันและความแปรปรวนเท่ากัน	1. BRW 2. BM	1. BRW 2. WMW	1. BRW 2. BM	1. BRW 2. BM	1. BRW 2. BM, WMW	1. BRW 2. BM
ค่าเฉลี่ยต่างกันและความแปรปรวนต่างกัน	1. BRW 2. WMW	1. BRW 2. WMW	1. BRW 2. WMW	1. BRW 2. BM	1. BRW 2. WMW	1. BRW 2. WMW

ตารางที่ 14 ตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุด 2 ลำดับแรกและตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 2 ลำดับแรก

กรณี	ความผิดพลาดแบบที่ 1		กำลังการทดสอบ	
	ขนาดตัวอย่างเท่ากัน	ขนาดตัวอย่างต่างกัน	ขนาดตัวอย่างเท่ากัน	ขนาดตัวอย่างต่างกัน
การแจกแจงเหมือนกันและความแปรปรวนเท่ากัน	1. WMW 2. BM	1. WMW 2. BM	1. BRW 2. WMW และ BM	1. BRW 2. WMW
การแจกแจงเหมือนกันและความแปรปรวนต่างกัน	1. WMW 2. BM	1. WMW 2. BM	1. BRW 2. BM	1. BRW 2. BM
การแจกแจงต่างกันและความแปรปรวนเท่ากัน	1. WMW 2. BM	1. BM 2. WMW	1. BRW 2. BM	1. BRW 2. BM
การแจกแจงต่างกันและความแปรปรวนต่างกัน	1. BM 2. WMW	1. WMW 2. BM	1. BRW 2. WMW	1. BRW 2. WMW

เมื่อพิจารณาตัวสถิติทดสอบที่สามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีที่สุด 2 ลำดับแรกและตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 2 ลำดับแรกสำหรับข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกันและข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน สามารถสรุปผลดังตารางที่ 14

ตัวสถิติ BRW มีกำลังการทดสอบสูงสุดในทุกกรณี แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อยที่สุด ในขณะที่ตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีที่การแจงเหมือนกัน ความแปรปรวนเท่ากัน และในกรณีการแจกแจงต่างกัน ความแปรปรวนต่างกัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดในทุกกรณี ยกเว้นในกรณีการแจกแจงต่างกัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างต่างกัน และในกรณีการแจกแจงต่างกัน ความ

แปรปรวนต่างกัน ขนาดตัวอย่างเท่ากัน สำหรับตัวสถิติ BM มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีการแจกแจงเหมือนกัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน ในกรณีการแจงแจงเหมือนกัน ความแปรปรวนต่างกัน และในกรณีการแจกแจงต่างกัน ความแปรปรวนเท่ากัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดในกรณีการแจกแจงต่างกัน ความแปรปรวนเท่ากัน ขนาดตัวอย่างต่างกัน และในกรณีการแจกแจงต่างกัน ความแปรปรวนต่างกัน ขนาดตัวอย่างเท่ากัน

4. สรุป วิจารณ์ และข้อเสนอแนะ

4.1 สรุปผลการวิจัย

ตารางที่ 5 พบว่า BRW มีกำลังการทดสอบมากที่สุดในทุกกรณี แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อยที่สุดในทุกกรณี ในขณะที่ตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุด ดังนั้นจึงควรเลือกใช้ตัวสถิติ WMW ในการเปรียบเทียบสถิติทดสอบที่ไม่อิงพารามิเตอร์ระหว่างค่าเฉลี่ยของประชากร 2 กลุ่มที่เป็นอิสระกัน

4.2 วิจารณ์ผลการวิจัย

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงเหมือนกัน พบว่า BRW มีกำลังการทดสอบมากที่สุดในทุกกรณี แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อยที่สุดในทุกกรณี ซึ่งตรงกับงานวิจัยของ นพดล และชินนพงษ์ [6] ที่กล่าวว่าตัวสถิติ BRW มีกำลังการทดสอบมากที่สุด แต่ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ ส่วนในงานวิจัยนี้ยังพบอีกว่าตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีที่ความแปรปรวนเท่ากัน แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุด ในขณะที่ตัวสถิติ BM มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน และในกรณีความแปรปรวนต่างกัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 รองลงมาจากตัวสถิติ WMW ทุกกรณี

เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงต่างกัน พบว่าตัวสถิติ BRW มีกำลังการทดสอบมากที่สุดในทุกกรณี แต่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้น้อยที่สุดในทุกกรณี ซึ่งขัดแย้งกับงานวิจัยของ ชนาธิป และคณะ [3] ที่กล่าวว่าตัวสถิติ WMW มีกำลังการทดสอบสูงสุดในทุกกรณี เนื่องจากงานวิจัยไม่ได้ทดสอบตัวสถิติ BM และตัวสถิติ BRW ส่วนในงานวิจัยนี้ยังพบอีกว่าตัวสถิติ WMW มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีความแปรปรวนต่างกัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดในกรณี ความแปรปรวนเท่ากัน ขนาดตัวอย่างเท่ากัน และความแปรปรวนต่างกัน ขนาดตัวอย่างต่างกัน ในขณะที่ตัวสถิติ BM มีกำลังการทดสอบรองลงมาจากตัวสถิติ BRW ในกรณีความแปรปรวนเท่ากัน และสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดในกรณีความแปรปรวนเท่ากัน ขนาดตัวอย่างต่างกัน และความแปรปรวนต่างกัน ขนาดตัวอย่างเท่ากัน

4.3 ข้อเสนอแนะ

4.3.1 ควรศึกษาเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบไม่อิงพารามิเตอร์ระหว่างค่าเฉลี่ยของประชากร 2 กลุ่ม ด้วยตัวสถิติทดสอบอื่น ๆ ได้แก่ O’Gorman adaptive test, adjusted oooled t-test และ modified intrinsically ties adjusted Mann-Whitney U test เป็นต้น

4.3.2 ควรเพิ่มขนาดตัวอย่างให้ครอบคลุม

มากยิ่งขึ้น เพื่อที่จะสามารถสรุปผลในกรณีตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ได้

4.3.3 ควรศึกษาในกรณีที่ข้อมูลสุ่มมาจากประชากรที่มีการแจกแจงอื่น ๆ เช่น การแจกแจงเบ้ซ้าย การแจกแจงที่มีความโด่งน้อยและโด่งมาก หรือการแจกแจงที่แตกต่างกันในแต่ละประชากร

5. รายการอ้างอิง

[1] กุสุมา อุมารี, ชลิตา รัตนวรสุทธิ์ และรมิตา ศรีภากร, 2558, การเปรียบเทียบประสิทธิภาพของสถิติทดสอบบาร์ตเลต สถิติทดสอบเลวีน สถิติทดสอบบราวน์-ฟอร์สิตี และสถิติทดสอบเลห์เมนสำหรับการทดสอบความเท่ากันของความแปรปรวนในกรณี 3 ประชากร โดยใช้โปรแกรมอาร์, โครงงานพิเศษ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ.

[2] อาภา วงศ์จินดา, อภิญญา หิรัญวงษ์ และบุญอ้อม โฉมที, 2558, การเปรียบเทียบสถิติทดสอบไม่อิงพารามิเตอร์สำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากรสองกลุ่มที่เป็นอิสระกันเมื่อขนาดตัวอย่างเล็ก, ว.วิทยาศาสตร์และเทคโนโลยี 23(4): 558-567.

[3] ชนาธิป ภาสุรกุล, ณัฐิมา ศาลาวงศ์, ปทิตตา ศิลาหม้อม และพัสตราภรณ์ ไชยโคตร, 2556, ความน่าเชื่อถือของสถิติทดสอบไม่อิงพารามิเตอร์สำหรับตัวอย่าง 2 กลุ่มอิสระจากโปรแกรมสำเร็จรูปทางสถิติ SPSS ในกรณีการวัดค่ากลาง, รายงานวิจัย คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ.

[4] อุมาพร จันทศร, 2556, ความน่าเชื่อถือของผลการวิเคราะห์ด้วยสถิติทดสอบวิลคอกซัน-แมนวิทนีย์เมื่อคำนึงถึงข้อกำหนดเบื้องต้นจากโปรแกรมสำเร็จรูป SPSS และ MINITAB, ว.วิทยาศาสตร์ลาดกระบัง 22(2): 1-16.

[5] Reiczigel, J., Zak0arias, I. and Rozsa, L., 2005, A bootstrap test of stochastic equality of two populations, Amer. Stat. 59: 156-161.

[6] นพดล วันชนะชัย และชินนพงษ์ บำรุงทรัพย์, 2553, การเปรียบเทียบตัวสถิติทดสอบไม่อิงพารามิเตอร์ระหว่างประชากร 2 กลุ่ม เมื่อความแปรปรวนไม่เท่ากัน, ว.วิทยาศาสตร์และเทคโนโลยี 18(3): 60-66.

[7] วราวัลย์ นิลพัทธ์, 2556, สถิติทดสอบไม่อิงพารามิเตอร์สำหรับทดสอบความแตกต่างของตำแหน่งของประชากร 2 กลุ่ม กรณีความแปรปรวนของประชากรไม่เท่ากัน, วิทยานิพนธ์ปริญญาโท, สถาบันบัณฑิตพัฒนาบริหารศาสตร์, กรุงเทพฯ.

[8] Bradley, J.V., 1978, Robustness, J. Math. Stat. Psychol. 31: 144-152.

[9] วราฤทธิ์ พาณิชกิจโกศสกุล, 2557, การใช้โปรแกรม R ในงานวิจัยด้านสถิติและสถิติประยุกต์, ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต, ปทุมธานี.

[10] Wilcox, R.R., 1990, Comparing the means of two independent groups, Biomet. J. 62: 771-780.

[11] Brunner, E. and Munzel, U. 2000. The nonparametric Behrens-Fisher problem: Asymptotic theory and a small sample approximation, Biomet. J. 42: 17-25.