การเปรียบเทียบประสิทธิภาพของสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์ กรณีมีค่านอกกลุ่ม*

The Efficiency Comparisons Among t and Mann-Whitney Test Statistics When the Presence of Outliers

วิภาวรรณ เล้าอรุณ (Wipawan Laoarun)**
ตติยา ลิ้มเทียมรัตน์ (Tatiya Limthiemrat)***
นรมน เรือนเหล็ก (Noramon Ruanlek)***

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์ สำหรับทดสอบความแตกต่างของค่ากลางของประชากร 2 กลุ่ม ในกรณีที่ความแปรปรวนของประชากร 2 กลุ่มไม่แตกต่างกัน ภายใต้การแจกแจงของประชากร 2 กลุ่มเป็นแบบสมมาตร ได้แก่ การแจกแจงแบบปกติ การแจกแจงแบบยูนิฟอร์ม และการแจกแจงแบบโลจิสติก ที่ขนาดตัวอย่าง 2 กลุ่มเท่ากัน คือ 20 40 60 และ 100 โดยทำการศึกษาในกรณีที่ตัวอย่างทั้ง 2 กลุ่มไม่มีค่านอกกลุ่ม และกรณีที่มีค่านอกกลุ่มปะปนอยู่ทั้งในตัวอย่างกลุ่มที่ 1 และปะปนอยู่ในตัวอย่างทั้งสองกลุ่ม ด้วยสัดส่วน 0.05 ในแต่ละกลุ่ม สำหรับข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โล โดยโปรแกรม R และกระทำซ้ำจำนวน 10,000 ครั้งในแต่ละสถานการณ์ เกณฑ์ที่ใช้เปรียบเทียบประสิทธิภาพของสถิติทดสอบพิจารณาจากความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 โดยใช้เกณฑ์ของ Bradley และกำลังของการทดสอบ ที่ระดับนัยสำคัญ 0.05 ถ้าสถิติทดสอบใดมีค่าความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 อยู่ระหว่าง 0.025 – 0.075 ถือว่าสถิติทดสอบนั้นสามารถควบคุมการเกิดความผิดพลาดแบบที่ 1 ได้

ผลการวิจัย สรุปได้ดังนี้

กรณีไม่มีค่านอกกลุ่ม สถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์มีความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ได้ใกล้เคียงกันในทุกขนาดตัวอย่าง และทุกการแจกแจง แต่เมื่อขนาดตัวอย่างทั้งสองกลุ่มเพิ่มขึ้น จะพบว่า กำลังการทดสอบของสถิติทดสอบแมนน์-วิตนีย์จะสูงกว่า

---

* เพื่อเปรียบเทียบประสิทธิภาพในการควบคุมการเกิดความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์ ในการทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม ทั้งในกรณีที่ไม่มีค่านอกกลุ่มและกรณีมีค่านอกกลุ่ม เมื่อข้อมูลมีการแจกแจงแบบต่อเนื่องและสมมาตร และความแปรปรวนของประชากรทั้ง 2 กลุ่มไม่แตกต่างกัน

** ดร. อาจารย์ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร E-mail : lwipawan@gmail.com (Dr., Department of Statistic, Faculty of Science, Silpakorn University)

*** นักศึกษาระดับปริญญาตรี สาขาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร E-mail : tatiya-L@hotmail.com, noramonjoon@gmail.com (Student of Bachelor of Science Program in Statistics, Faculty of Science, Silpakorn University)

Veridian E-Journal, Science and Technology Silpakorn University
Volume 3 Number 5 September– October2016 (ISSN 2408 – 1248)

สำหรับทุกการแจกแจง ในกรณีที่มีค่านอกกลุ่มปะปนอยู่ในตัวอย่างกลุ่มที่ 1 หรืออยู่ในทั้ง 2 กลุ่ม จะพบว่า สถิติทดสอบแมนน์-วิตนีย์จะมีประสิทธิภาพสูงกว่าสถิติทดสอบทีในทุกกรณี และเมื่อขนาดตัวอย่างทั้งสองกลุ่มมากกว่า 60 สถิติทดสอบที จะมีความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ได้ดีขึ้น รวมถึงกำลังการทดสอบที่สูงขึ้น

คำสำคัญ : ค่านอกกลุ่ม ความผิดพลาดแบบที่ 1 กำลังการทดสอบ

Abstract

The purpose of this research is to compare the efficiency among the t and Mann-Whitney test statistics which are used for testing hypothesis about the difference of the two population location parameters. The three symmetric distributions : Normal, Uniform and Logistic distributions with equal variances and sample sizes of 20 40 60 and 100 are considered. In this research, we studied in the case of no outlier and the outliers contaminated in the first sample and all of two samples with proportion 0.05 in each sample. The data in this research are generated through the Monte Carlo simulation technique with 10,000 replications for each situation by the R program. The criterion employed for comparing the efficiency of the two tests are controlling of the probability of type I error according to Bradley criterion and power of the test. The nominal significant level (α) used is 0.05. The test statistic which the estimated probability of type I error value lies between 0.025 and 0.075, it can control the probability of type I error.

The results of this research can be summarized as follows:

For no outlier included in all of two samples, the t test statistic can control the probability of type I error slightly better than the Mann-Whitney test statistic for all sample sizes and distributions. Moreover, when the sample sizes are increased, the power of the test of the Mann-Whitney is higher than of the t test statistic.

For either outliers contaminated in the first sample or in all of two samples, the Mann-Whitney test statistic provides more efficient than the t test statistic. However, the t test statistic has the more performances in controlling the probability of type I error and power of the test when the each sample size of two groups is greater than 60.

Keywords : Outliers, Type I error, Power of the test.

บทนำ

ในการทดสอบสมมติฐานทางสถิติเกี่ยวกับการเปรียบเทียบค่ากลางของประชากร 2 กลุ่ม กรณีตัวอย่างสุ่ม 2 กลุ่มเป็นอิสระกัน จะมีสถิติทดสอบที่นิยมใช้ทั้งสถิติที่อิงพารามิเตอร์ ได้แก่ สถิติทดสอบที โดยสถิติทดสอบดังกล่าวจะมีข้อตกลงเบื้องต้น (assumptions) ที่สำคัญประการหนึ่งคือ ตัวอย่างสุ่ม 2 กลุ่มต้องสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ แต่ในทางปฏิบัติหากข้อมูลที่เก็บรวบรวมมาเพื่อใช้ในการวิเคราะห์มีลักษณะที่เบี่ยงเบนจากการแจกแจงแบบปกติ ผู้วิจัยสามารถใช้สถิติที่ไม่อิงพารามิเตอร์ ได้แก่ สถิติทดสอบแมนน์-วิตนีย์ ในการทดสอบสมมติฐานดังกล่าวได้

มีผู้วิจัยหลายท่านได้ทำการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบที่ใช้ในการเปรียบเทียบความแตกต่างของค่ากลางของประชากร 2 กลุ่ม เมื่อตัวอย่างสุ่ม 2 กลุ่มอิสระกัน ได้แก่

มนตรี (2556) ศึกษาประสิทธิภาพของสถิติทดสอบอิงพารามิเตอร์และสถิติทดสอบไม่อิงพารามิเตอร์ ได้แก่ t-test, Mann-Whitney U test, Van der Waerden test, Wald Wolfowitz Runs test และ Modified U test จำแนกสถานการณ์โดยกำหนดประชากรที่มีการแจกแจงปกติ การแจกแจงเบ้ซ้ายและความโด่งสูงกว่าปกติ และการแจกแจงเบ้ขวาและโด่งสูงกว่าปกติ ภายใต้ตัวอย่างขนาดเล็ก ((n_1, n_2) = (5,5), (5,10)) ตัวอย่างขนาดกลาง ((n_1, n_2) = (20,20), (25,20)) และตัวอย่างขนาดใหญ่ ((n_1, n_2) = (50,50), (50,100)) ทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ผลการศึกษา พบว่า สถิติทดสอบ t-test และ Van der Waerden test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้และมีกำลังการทดสอบสูงสุด เมื่อตัวอย่างสุ่มทั้ง 2 กลุ่มมีขนาดเล็ก และสุ่มมาจากประชากรที่มีการแจกแจงทุกรูปแบบที่ศึกษา นอกจากนี้สถิติทดสอบ t-test ยังเหมาะกับตัวอย่างสุ่มขนาดใหญ่ที่สุ่มมาจากประชากรที่ รวมทั้งกรณีที่ตัวอย่างสุ่มมีขนาดใหญ่ ที่สุ่มมาจากประชากรที่มีการแจกแจงแบบเบ้ขวาและโด่งสูงกว่าปกติ ในขณะที่สถิติทดสอบ Modified U test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้และมีกำลังการทดสอบสูงสุด เมื่อตัวอย่างสุ่มทั้ง 2 กลุ่มมีขนาดกลางและใหญ่ และสุ่มมาจากประชากรที่มีการแจกแจงปกติ สำหรับสถิติทดสอบ Mann-Whitney U test สามารถควบคุมความผิดพลาดแบบที่ 1 ได้และมีกำลังการทดสอบสูงสุด เมื่อตัวอย่างสุ่มทั้ง 2 กลุ่มมีขนาดกลางและใหญ่ และสุ่มมาจากประชากรที่มีการแจกแจงเบ้ซ้าย(หรือเบ้ขวา) และความโด่งสูงกว่าปกติ

มนตรี (2557) ศึกษาความแกร่งและอำนาจการทดสอบของสถิติทดสอบอิงพารามิเตอร์และสถิติทดสอบไม่อิงพารามิเตอร์ ได้แก่ Z test, t-test, Mann-Whitney U test, Van der Waerden test, Kolmogonov-Smirnov test และ Modified U test ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม สำหรับข้อมูลแบบลิเคิร์ท 5 ระดับ โดยมีเงื่อนไข คือ ประชากรมีการแจกแจงปกติ การแจกแจงเบ้ซ้ายและความโด่งต่ำกว่าปกติ และการแจกแจงเบ้ขวาและความโด่งสูงกว่าปกติ ภายใต้ขนาดตัวอย่างระดับต่างๆ (เล็ก กลาง และใหญ่) ทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 และ 0.01 ผลการศึกษา พบว่า เมื่อประชากรมีการแจกแจงปกติ สถิติทดสอบ t-test, Modified U test และ Mann-Whitney U test มีความเหมาะสม (สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ และมีกำลังการทดสอบสูงสุด) ในกรณีตัวอย่างมีขนาดเล็ก กลาง และใหญ่ ตามลำดับ เมื่อประชากรมีการแจกแจงเบ้ซ้ายและความโด่งต่ำกว่าปกติ และกลุ่มตัวอย่างขนาดเล็ก สถิติทดสอบที่เหมาะสม คือ t-test และ Mann-Whitney U test ตัวอย่างสุ่มทั้ง 2 กลุ่มมีขนาดกลางและ

ขนาดใหญ่ สถิติทดสอบที่เหมาะสม คือ Van der Waerden test และเมื่อประชากรมีการแจกแจงเบ้ขวาและความโด่งสูงกว่าปกติ ตัวอย่างสุ่มทั้ง 2 กลุ่มขนาดเล็ก กลาง และใหญ่ สถิติทดสอบที่เหมาะสม คือ t test, Z test และ Van der Waerden test ตามลำดับ สำหรับกรณีตัวอย่างสุ่มขนาดกลางและขนาดใหญ่ และความแปรปรวนไม่เท่ากัน ตัวสถิติทดสอบดังกล่าวไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้

M. Bakker (2014) ได้ทำการศึกษาผลกระทบของการลบค่านอกกลุ่มและไม่ลบค่านอกกลุ่มออกจากข้อมูลต่อสถิติทดสอบที สถิติทดสอบแมนน์-วิตนีย์ และสถิติทดสอบ Yuen-Welch และกำหนดระดับนัยสำคัญในการทดสอบเท่ากับ 0.05 ขนาดตัวอย่างเท่ากับ 20, 40, 100 และ 500 และทำซ้ำ 100,000 ครั้งในแต่ละสถานการณ์ โดยศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ของ สถิติทดสอบที เมื่อทำการลบค่านอกกลุ่มออกจากข้อมูล และศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบที สถิติทดสอบแมนน์-วิตนีย์ และสถิติทดสอบ Yuen-Welch กรณีไม่มีค่านอกกลุ่ม และกรณีมีค่านอกกลุ่มอยู่ในข้อมูล

โดยกรณีไม่มีค่านอกกลุ่ม M. Bakker (2014) ทำการสุ่มตัวอย่างกลุ่มที่ 1 และ 2 มาจากประชากรที่มีการแจกแจง N(0, 1) สำหรับการศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 และสำหรับการศึกษากำลังการทดสอบ ทำการสุ่มตัวอย่างกลุ่มที่ 1 มาจากประชากรที่มีการแจกแจง N(0, 1) และสุ่มตัวอย่างกลุ่มที่ 2 มาจากประชากรที่มีการแจกแจง N(0.2, 1) , N(0.5, 1) และ N(0.8, 1)

กรณีมีค่านอกกลุ่มอยู่ในชุดข้อมูล ทำการสุ่มตัวอย่างโดยใช้การแจกแจงแบบปกติปลอมปน สำหรับการศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ทำการสุ่มตัวอย่างกลุ่มที่ 1 และ 2 มาจากประชากรที่มีการแจกแจง N(0, 1) ด้วยความน่าจะเป็นเท่ากับ 0.95 และสุ่มมาจากประชากรที่มีการแจกแจง N(0, 400) ด้วยความน่าจะเป็น 0.05 สำหรับการศึกษากำลังการทดสอบ ทำการสุ่มตัวอย่างกลุ่มที่ 1 มาจากประชากรที่มีการแจกแจง N(0, 1) ด้วยความน่าจะเป็นเท่ากับ 0.95 และสุ่มมาจากประชากรที่มีการแจกแจง N(0, 400) ด้วยความน่าจะเป็น 0.05 ตัวอย่างกลุ่มที่ 2 สุ่มมาจากประชากรที่มีการแจกแจง N(0.2, 1) , N(0.5, 1) , N(0.8, 1) ด้วยความน่าจะเป็นเท่ากับ 0.95 และสุ่มมาจากประชากรที่มีการแจกแจง N(0.2, 400) , N(0.5, 400) , N(0.8, 400) ด้วยความน่าจะเป็น 0.05

เมื่อพิจารณาผลการวิจัยในส่วนที่ 2 พบว่า ในกรณีไม่มีค่านอกกลุ่ม สถิติทดสอบทีมีประสิทธิภาพสูงกว่าการทดสอบแมนน์-วิตนีย์ และสถิติทดสอบแมนน์-วิตนีย์มีประสิทธิภาพสูงกว่าสถิติทดสอบ Yuen-Welch แต่จะมีประสิทธิภาพแตกต่างกันไม่มากนัก และสำหรับกรณีมีค่านอกกลุ่มในข้อมูล สถิติทดสอบทีจะมีประสิทธิภาพต่ำที่สุด และเมื่อขนาดตัวอย่างมากขึ้นจะยิ่งมีประสิทธิภาพต่ำมากเมื่อเทียบกับสถิติทดสอบแมนน์-วิตนีย์ และสถิติทดสอบ Yuen-Welch แต่สถิติทดสอบแมนน์-วิตนีย์และสถิติทดสอบ Yuen-Welch จะมีประสิทธิภาพใกล้เคียงกัน

จากข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงแบบปกติของประชากรทั้ง 2 กลุ่มที่มีผลต่อการเลือกสถิติทดสอบ และสาเหตุที่ข้อมูลไม่ได้มีการแจกแจงแบบปกติอาจเนื่องมาจากการมีค่าผิดปกติหรือค่านอกกลุ่ม (Outlier) ปะปนอยู่ในชุดข้อมูลนั้น Barnett and Lewis (1994) ได้อธิบายว่า ค่านอกกลุ่ม (Outlier) แบ่งออกเป็น 2 ประเภทดังนี้ ประเภทแรก คือ ค่าปลอมปน (Contaminants) เป็นค่าที่มาจากการแจกแจงอื่นที่ไม่ใช่การแจกแจงเดียวกันกับข้อมูลที่เหลือในชุดข้อมูลนั้น ประเภทสอง คือ ค่าสุดขีด (Extreme) เป็นค่าที่สูงสุด

หรือต่ำสุดในชุดข้อมูลแต่ยังคงมาจากการแจกแจงเดียวกันกับข้อมูลที่เหลือในชุดข้อมูลนั้นดังนั้นจึงอาจเกิดปัญหาในการเลือกสถิติทดสอบสำหรับการทดสอบสมมติฐานเมื่อข้อมูลที่จะทำการวิเคราะห์มีค่านอกกลุ่ม (Outlier) ปะปนอยู่ นอกจากนี้ยังพบว่าการแจกแจงแบบปกติเป็นการแจกแจงแบบสมมาตร ถ้าหากพบว่าข้อมูลที่จะทำการวิเคราะห์ไม่ได้มีการแจกแจงแบบปกติแต่ยังคงมีการแจกแจงแบบสมมาตร ผู้วิจัยจึงมีความสนใจที่จะศึกษาว่าการทดสอบอิงพารามิเตอร์จะยังคงมีความเหมาะสมในการวิเคราะห์ข้อมูลชุดดังกล่าวหรือไม่

ดังนั้น ผู้วิจัยจึงมีความสนใจที่จะเปรียบเทียบประสิทธิภาพของสถิติทดสอบทีกับสถิติทดสอบแมนน์-วิตนีย์ที่เป็นสถิติทดสอบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ตามลำดับที่ใช้กันแพร่หลายในโปรแกรมสำเร็จรูปทางสถิติ สำหรับทดสอบความแตกต่างของค่าเฉลี่ย (มัธยฐาน) ของประชากร 2 กลุ่มที่อิสระกันในกรณีที่ข้อมูลถูกสุ่มมาจากประชากรที่มีความแปรปรวนเท่ากัน มีค่านอกกลุ่มปะปนอยู่และมีการแจกแจงแบบสมมาตร โดยพิจารณาจากความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบ

วิธีการศึกษา

1.ขอบเขตในการวิจัย

1.1.ประชากร 2 กลุ่มมีความแปรปรวนเท่ากัน

1.2.การแจกแจงของประชากรทั้ง 2 กลุ่ม เป็นแบบสมมาตรและต่อเนื่อง โดยศึกษาที่การแจกแจงแบบปกติ $N(\mu, \sigma^2)$ การแจกแจงแบบยูนิฟอร์ม U(a,b) และการแจกแจงแบบโลจิสติก Logistic (μ^*, σ^*) ที่มีค่านอกกลุ่มปะปนด้วยสัดส่วน p โดยฟังก์ชันความหนาแน่นน่าจะเป็น (probability density function) ของการแจกแจงแบบปกติ ยูนิฟอร์ม และโลจิสติก กำหนดโดย $f_1(x)$ $f_2(x)$ และ $f_3(x)$ ตามลำดับดังนี้

$$f_1(x) = (1-p)\,N(\mu, \sigma^2) + p\,N(\mu', (\sigma')^2)$$
$$f_2(x) = (1-p)\,U(a, b) + p\,U(a', (b')^2)$$
$$f_3(x) = (1-p)\,\text{Logistic}\,(\mu^*, \sigma^*) + p\,\text{Logistic}\,((\mu^*)', (\sigma^*)')$$

เมื่อ p แทน สัดส่วนของค่านอกกลุ่มที่ปะปนในข้อมูล ในที่นี้กำหนด p = 0 (ไม่มีค่านอกกลุ่ม) และ p = 0.05

โดยแบ่งการศึกษาเป็น 2 ส่วน คือ การศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 และการศึกษากำลังการทดสอบ

1.3.ขนาดตัวอย่างที่ทำการศึกษาทั้งสองกลุ่มเท่ากัน คือ 20, 40, 60 และ 100

1.4.กำหนดระดับนัยสำคัญของการทดสอบ คือ 0.05

2.วิธีการจำลองข้อมูล ในการดำเนินงานวิจัยครั้งนี้ ผู้วิจัยใช้โปรแกรม R ในการจำลองข้อมูล โดยแบ่งการจำลองข้อมูลเป็น 2 ส่วน คือ

Veridian E-Journal, Science and Technology Silpakorn University
Volume 3 Number 5 September– October2016 (ISSN 2408 – 1248)

2.1 การศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1

กำหนดสถานการณ์ เพื่อทำการศึกษาความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ของสถิติทดสอบแต่ละตัวโดยจะทำการศึกษาภายใต้สมมติฐานว่างเป็นจริง โดยที่สมมติฐานการทดสอบ คือ

สมมติฐานว่าง $H_0 : \mu_1 = \mu_2$ หรือ $H_0 : \tilde{\mu}_1 = \tilde{\mu}_2$

สมมติฐานแย้ง $H_1 : \mu_1 \neq \mu_2$ $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$

โดยที่ μ_1 และ μ_2 แทน ค่าเฉลี่ยของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

$\tilde{\mu}_1$ และ $\tilde{\mu}_2$ แทน ค่ามัธยฐานของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

มีขั้นตอนการจำลองข้อมูลดังนี้

1. ในกรณีไม่มีค่านอกกลุ่ม สุ่มตัวอย่างจากประชากร 2 กลุ่ม ที่มีค่าเฉลี่ยเท่ากัน (= 0) และความแปรปรวนเท่ากัน (= 1) ตามขนาดที่กำหนด จากประชากรที่มีการแจกแจงต่างๆ ตามพารามิเตอร์ที่กำหนด

กรณีที่มีค่านอกกลุ่มเฉพาะในตัวอย่างกลุ่มที่ 1 จะสุ่มตัวอย่างในกลุ่มที่ 1 จำนวนร้อยละ 95 ของขนาดตัวอย่างที่กำหนดในข้อ 1.3 ในแต่ละการแจกแจงตามพารามิเตอร์เช่นเดียวกับกรณีไม่มีค่านอกกลุ่ม และสุ่มค่านอกกลุ่มจากการแจกแจงเดิมที่ความแปรปรวนเท่ากับ 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05 โดยกำหนดพารามิเตอร์ในแต่ละการแจกแจงดังตาราง 1 สำหรับการสุ่มตัวอย่างในกลุ่มที่ 2 จะสุ่มจากประชากรที่มีพารามิเตอร์เช่นเดียวกับกรณีไม่มีค่านอกกลุ่ม

สำหรับกรณีมีค่านอกกลุ่มปะปนในตัวอย่างทั้งสองกลุ่ม จะสุ่มตัวอย่างในแต่ละกลุ่มเช่นเดียวกับกรณีที่มีค่านอกกลุ่มเฉพาะในกลุ่มที่ 1

จากการสุ่มตัวอย่างทั้ง 3 กรณี สรุปค่าพารามิเตอร์ของการแจกแจงในแต่ละกรณี ดังตาราง 1 ดังนี้

ตาราง 1 ค่าพารามิเตอร์ของการแจกแจงของประชากรสำหรับการสุ่มตัวอย่างเพื่อศึกษาความสามารถในการควบคุมความผิดพลาดแบบที่ 1

กรณี	การแจกแจง	ค่าพารามิเตอร์ของการแจกแจง
ไม่มีค่านอกกลุ่ม	N(μ, σ^2)	μ = 0 และ σ^2 = 1
	U(a,b)	a = -1.732050808 และ b = 1.732050808
	Logistic (μ^*, σ^*)	μ^* = 0 และ σ^* = 0.551328895
มีค่านอกกลุ่ม ในตัวอย่าง กลุ่มที่ 1*	N(μ, σ^2)	μ = 0 และ σ^2 = 1 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ μ' = 0 และ $(\sigma')^2$ = 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05
	U(a,b)	a = -1.732050808 และ b = 1.732050808 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ a' = -34.6410162 และ b' = 34.6410162

		ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05
	Logistic (μ^*, σ^*)	μ^* = 0 และ σ^* = 0.551328895 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu^*)'$= 0 และ $(\sigma^*)'$ = 11.02657791 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05
มีค่านอกกลุ่ม ในตัวอย่าง ทั้ง 2 กลุ่ม**	สำหรับทุกการแจกแจงกำหนดเช่นเดียวกับกรณีที่มีค่านอกกลุ่มในตัวอย่างกลุ่มที่ 1	

*กำหนดค่าพารามิเตอร์สำหรับการแจกแจงของประชากรกลุ่มที่ 1

**กำหนดค่าพารามิเตอร์สำหรับการแจกแจงของประชากรทั้งสองกลุ่ม

2.ทำการจำลองข้อมูลตามข้อ 1 ในแต่ละกรณีและขนาดตัวอย่างที่กำหนดจำนวน 10,000 ชุด

3.ในข้อมูลแต่ละชุดทำการคำนวณค่า p-value ของสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์เทียบกับระดับนัยสำคัญ 0.05

4.คำนวณค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ของสถิติทดสอบแต่ละตัว จากการหาสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง กับจำนวนครั้งของการทดสอบทั้งหมด

5.ตัวสถิติทดสอบที่สามารถควบคุมการเกิดความผิดพลาดแบบที่ 1 ที่ระดับนัยสำคัญ 0.05 ภายใต้เกณฑ์ของ Bradley หมายถึง สถิติที่มีค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 อยู่ระหว่าง 0.025 – 0.075 กล่าวคือ สถิติดังกล่าวสามารถควบคุมให้โอกาสการเกิดความผิดพลาดแบบที่ 1 ไม่แตกต่างไปจากที่กำหนด คือ 0.05

2.2 การศึกษากำลังการทดสอบ

กำหนดสถานการณ์เพื่อการศึกษากำลังการทดสอบของสถิติทดสอบแต่ละตัวจะทำการศึกษาภายใต้สมมติฐานแย้งเป็นจริง โดยที่สมมติฐานการทดสอบ คือ

สมมติฐานว่าง $H_0 : \mu_1 = \mu_2$ หรือ $H_0 : \tilde{\mu}_1 = \tilde{\mu}_2$

สมมติฐานแย้ง $H_1 : \mu_1 \neq \mu_2$ $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$

โดยที่ μ_1และ μ_2แทน ค่าเฉลี่ยของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

$\tilde{\mu}_1$และ $\tilde{\mu}_2$ แทน ค่ามัธยฐานของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

มีขั้นตอนการจำลองข้อมูลดังนี้

1. ในกรณีไม่มีค่านอกกลุ่ม สุ่มตัวอย่างจากประชากรกลุ่มที่ 1 ที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 และจากประชากรกลุ่มที่ 2 ที่มีค่าเฉลี่ยเท่ากับ 1.5 และความแปรปรวนเท่ากับ 1ตามขนาดที่กำหนด จากประชากรที่มีการแจกแจงต่างๆ ตามพารามิเตอร์ที่กำหนด

กรณีที่มีค่านอกกลุ่มเฉพาะในตัวอย่างกลุ่มที่ 1 จะสุ่มตัวอย่างในกลุ่มที่ 1 จำนวนร้อยละ 95 ของขนาดตัวอย่างที่กำหนดในข้อ 1.3 โดยกำหนดการแจกแจงและพารามิเตอร์ที่กำหนดเช่นเดียวกับกรณีไม่มีค่านอกกลุ่ม และสุ่มค่านอกกลุ่มจากการแจกแจงเดิมที่ความแปรปรวนเท่ากับ 400 ด้วยสัดส่วนการปะปนของค่านอก

Veridian E-Journal, Science and Technology Silpakorn University
Volume 3 Number 5 September– October2016 (ISSN 2408 – 1248)

กลุ่ม 0.05 โดยกำหนดพารามิเตอร์ในแต่ละการแจกแจงดังตาราง 2 สำหรับการสุ่มตัวอย่างในกลุ่มที่ 2 จะสุ่มจากประชากรกลุ่มที่ 2 ที่มีพารามิเตอร์เช่นเดียวกับกรณีไม่มีค่านอกกลุ่ม

สำหรับกรณีมีค่านอกกลุ่มปะปนในตัวอย่างทั้งสองกลุ่ม จะสุ่มตัวอย่างจากประชากรกลุ่มที่ 1 และ 2 ในแต่ละการแจกแจงตามที่กำหนดในกรณีไม่มีค่านอกกลุ่มจำนวนร้อยละ 95 ของขนาดตัวอย่างในแต่ละกลุ่ม และสุ่มค่านอกกลุ่มจากการแจกแจงเดิมที่ความแปรปรวนเท่ากับ 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05 ตามที่กำหนดในกรณีมีค่านอกกลุ่มเฉพาะในกลุ่มที่ 1 จากการสุ่มตัวอย่างทั้ง 3 กรณี สรุปค่าพารามิเตอร์ของการแจกแจงในแต่ละกรณี ดังตาราง 2 ดังนี้

ตาราง 2 ค่าพารามิเตอร์ของการแจกแจงของประชากรสำหรับการสุ่มตัวอย่าง เพื่อศึกษากำลังการทดสอบ

กรณี	การแจกแจง	ค่าพารามิเตอร์ของการแจกแจงของประชากรกลุ่มที่ 1	ค่าพารามิเตอร์ของการแจกแจงของประชากรกลุ่มที่ 2
ไม่มีค่านอกกลุ่ม	N(μ, σ^2)	μ_1 = 0 และ σ_1^2 = 1	μ_2 = 1.5 และ σ_2^2 = 1
	U(a,b)	a_1 = -1.732050808 และ b_1 = 1.732050808	a_2= -0.232050808 และ b_2 = 3.232050808
	Logistic (μ^*, σ^*)	μ_1^*= 0 และ σ_1^*= 0.551328895	μ_2^* = 1.5 และ σ_2^*= 0.551328895
มีค่านอกกลุ่มเฉพาะ ในตัวอย่างกลุ่มที่ 1	N(μ, σ^2)	μ_1 = 0 และ σ_1^2 = 1 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu_1)'$ = 0 และ $(\sigma_1')^2$ = 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	μ_2 = 1.5 และ σ_2^2 = 1
	U(a,b)	a_1 = -1.732050808 และ b_1 = 1.732050808 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(a_1)'$ = -34.6410162 และ $(b_1)'$ = 34.6410162 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	a_2= -0.232050808 และ b_2 = 3.232050808
	Logistic (μ^*, σ^*)	μ^* = 0 และ σ^* = 0.551328895	μ_2^* = 1.5 และ σ_2^*= 0.551328895

		ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu^*)' = 0$ และ $(\sigma^*)' =$ 11.02657791 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	
มีค่านอกกลุ่มในตัวอย่างทั้งสองกลุ่ม	N(μ, σ^2)	$\mu_1 = 0$ และ $\sigma_1^2 = 1$ ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu_1)' = 0$ และ $(\sigma_1')^2 =$ 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	$\mu_2 = 1.5$ และ $\sigma_2^2 = 1$ ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu_2)' = 1.5$ และ $(\sigma_2')^2 =$ 400 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05
	U(a,b)	$a_1 =$ -1.732050808 และ $b_1 =$ 1.732050808 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $a_1' =$ -34.6410162 และ $b_1' =$ 34.6410162 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	$a_2 =$ -0.232050808 และ $b_2 =$ 3.232050808 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $a_2' =$ -33.14101615 และ $b_2' =$ 36.14101615 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05

Veridian E-Journal, Science and Technology Silpakorn University
Volume 3 Number 5 September– October2016 (ISSN 2408 – 1248)

ตาราง 2 (ต่อ)

กรณี	การแจกแจง	ค่าพารามิเตอร์ของการแจกแจงของประชากรกลุ่มที่ 1	ค่าพารามิเตอร์ของการแจกแจงของประชากรกลุ่มที่ 2
มีค่านอกกลุ่มในตัวอย่างทั้งสองกลุ่ม	Logistic (μ^*, σ^*)	μ_1^* = 0 และ σ_1^* = 0.551328895 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu_1^*)'$ = 0 และ $(\sigma_1^*)'$ = 11.02657791 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05	μ_2^* = 1.5 และ σ_2^* = 0.551328895 ด้วยสัดส่วน 0.95 ของขนาดตัวอย่าง และ $(\mu_2^*)'$ = 1.5 และ $(\sigma_2^*)'$ = 11.02657791 ด้วยสัดส่วนการปะปนของค่านอกกลุ่ม 0.05

2.ทำการจำลองข้อมูลตามข้อ 1 ในแต่ละกรณีและขนาดตัวอย่างที่กำหนดจำนวน 10,000 ชุด

3.ในข้อมูลแต่ละชุดทำการคำนวณค่า p-value ของสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์เทียบกับระดับนัยสำคัญ 0.05

4.คำนวณค่าประมาณกำลังการทดสอบ ของสถิติทดสอบแต่ละตัว จากการหาสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง กับจำนวนครั้งของการทดสอบทั้งหมด

สถิติทดสอบที่ใช้ในการศึกษา

ในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่ากลางของประชากร 2 กลุ่ม มีสมมติฐานในการทดสอบ คือ

สมมติฐานว่าง $H_0 : \mu_1 = \mu_2$ หรือ $H_0 : \tilde{\mu}_1 = \tilde{\mu}_2$

สมมติฐานแย้ง $H_1 : \mu_1 \neq \mu_2$ $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$

โดยที่ μ_1 และ μ_2 แทน ค่าเฉลี่ยของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

$\tilde{\mu}_1$ และ $\tilde{\mu}_2$ แทน ค่ามัธยฐานของประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

โดยในที่นี้ทำการศึกษาสถิติทดสอบสำหรับทดสอบสมมติฐานดังกล่าว 2 ตัว คือ

1. สถิติทดสอบที (t-test statistic)

เป็นสถิติทดสอบอิงพารามิเตอร์ใช้ทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่ม การใช้สถิติทดสอบที ข้อมูลต้องเป็นไปตามข้อสมมติ คือ ตัวอย่าง 2 กลุ่มได้จากการสุ่มอย่างเป็นอิสระกันจากประชากรที่มีการแจกแจงปกติ ตัวอย่างทั้ง 2 กลุ่มมีขนาดเล็ก และไม่ทราบความแปรปรวนของประชากรทั้ง 2 กลุ่ม (ล้วน สายยศ และอังคณา สายยศ, 2540, น. 242) โดยที่ความแปรปรวนของประชากรทั้งสองกลุ่มอาจจะเท่ากันหรือไม่เท่ากันก็ได้ ในงานวิจัยนี้ผู้วิจัยศึกษาเฉพาะกรณีที่ ไม่ทราบค่า σ_1^2และ σ_2^2 แต่ทราบว่า $\sigma_1^2 = \sigma_2^2 = \sigma^2$

สถิติทดสอบ คือ

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim t_{n_1 + n_2 - 2}$$

โดยที่ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_1 - 2}$

s_p^2 คือ ค่าความแปรปรวนร่วมที่ได้จากตัวอย่างทั้งสองชุด

$\bar{x}_1,\ \bar{x}_2$ คือ ค่าเฉลี่ยของตัวอย่างของกลุ่มที่ 1 และกลุ่มที่ 2

$\mu_1,\ \mu_2$ คือ ค่าเฉลี่ยของประชากรของกลุ่มที่ 1 และกลุ่มที่ 2

s_1^2, s_2^2 คือ ค่าความแปรปรวนของตัวอย่างของกลุ่มที่ 1 และกลุ่มที่ 2

จากการคำนวณค่าสถิติทดสอบจะปฏิเสธสมมติฐานว่าง เมื่อ $|t| > t_{1-\alpha/2,\, n_1 + n_1 - 2}$ หรือ p-value มีค่าต่ำกว่าระดับนัยสำคัญในการทดสอบ โดยที่ $t_{\alpha/2}$ คือ ค่าเปอร์เซ็นต์ไทล์ส่วนบนที่ $\alpha/2$ ของการแจกแจงแบบ *t*

2. สถิติทดสอบแมนน์-วิตนีย์ (Mann-Whitney test statistic)

สถิติทดสอบแมนน์-วิตนีย์จะใช้การให้อันดับของค่าสังเกตกับตัวอย่างทั้งสองกลุ่ม โดยมีข้อสมมติต่างๆ ดังนี้

1.ตัวอย่าง 2 กลุ่มต้องทำการสุ่มอย่างเป็นอิสระต่อกัน โดยทำการสุ่มตัวอย่างขนาด *n* จากประชากรกลุ่มที่ 1 และสุ่มตัวอย่างขนาด *m* จากประชากรกลุ่มที่ 2

2. มาตราวัดของตัวแปรที่ศึกษาต้องเป็นมาตราอันดับหรือสูงกว่า

3. ถ้าประชากรมีความแตกต่างกัน หมายถึง มีความแตกต่างกันเฉพาะมัธยฐานที่แตกต่างกันเท่านั้น (สถิติแมนน์-วิตนีย์ พิจารณาแต่มัธยฐานเท่านั้น ไม่ได้พิจารณาการกระจายของประชากรทั้ง 2 กลุ่ม)

สถิติทดสอบแมนน์-วิตนีย์ สามารถแยกออกเป็นกรณีย่อยได้ดังต่อไปนี้

1) กรณีขนาดตัวอย่าง $n \leq 20$ และ $m \leq 20$

$$\mathrm{T} = \mathrm{S} - \frac{n(n+1)}{2}$$

เมื่อ *n* คือ ขนาดตัวอย่างที่สุ่มจากประชากรกลุ่มที่ 1 (X)

S คือ ผลรวมอันดับค่าสังเกตของตัวอย่างจากประชากรกลุ่มที่ 1

Veridian E-Journal, Science and Technology Silpakorn University
Volume 3 Number 5 September– October2016 (ISSN 2408 – 1248)

2)กรณีขนาดตัวอย่าง n >20 และ m >20

$$Z = \frac{T - \frac{nm}{2}}{\sqrt{\frac{(nm)(n+m+1)}{12}}}$$

เมื่อ T, S , *n* และ *m* มีความหมายเช่นเดียวกับที่กำหนดในกรณีที่ 1)

การปฏิเสธสมมติฐานว่าง

1) กรณีขนาดตัวอย่าง $n \leq 20$ และ $m \leq 20$

จะใช้ค่าสถิติทดสอบ T เปรียบเทียบกับค่าวิกฤติจากการเปิดตารางค่าวิกฤตของการทดสอบแมนน์-วิตนีย์ (ตารางค่าควอนไทล์ของการทดสอบแมนน์-วิตนีย์) ซึ่งให้ค่า w_{α}ที่ระดับนัยสำคัญ α ถ้าค่าสถิติทดสอบ T ที่คำนวณได้น้อยกว่า$w_{\alpha/2}$หรือมากกว่า $w_{1-\alpha/2}$ เมื่อ $w_{1-\alpha/2} = (n \times m) - w_{\alpha/2}$ จะทำการปฏิเสธสมมติฐานว่าง

2) กรณีขนาดตัวอย่าง n >20 และ m >20

จะใช้ค่าสถิติทดสอบ Z เปรียบเทียบกับค่าวิกฤติจากการเปิดตารางค่าวิกฤตของตัวแปรสุ่มปกติมาตรฐาน (ตาราง Z) ถ้าค่า $|z|$> $Z_{\frac{\alpha}{2}}$ จะปฏิเสธสมมติฐานว่าง (คณาจารย์ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2555, น.351 - 354)

อย่างไรก็ตาม หาก p-value มีค่าต่ำกว่าระดับนัยสำคัญในการทดสอบจะปฏิเสธสมมติฐานว่าง

ผลการวิจัย

ผู้วิจัยได้นำเสนอผลการวิจัยเป็น 2 ส่วนดังนี้

1.ค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 แสดงดังตาราง 3

ตาราง 3 ค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ของสถิติทดสอบที และสถิติทดสอบแมนน์-วิตนีย์ ในกรณีไม่มีค่านอกกลุ่ม มีค่านอกกลุ่มเฉพาะในกลุ่มที่ 1 และ มีค่านอกกลุ่มทั้ง 2 กลุ่ม ที่ระดับนัยสำคัญ α = 0.05

การแจกแจง	ขนาดตัวอย่าง	ไม่มีค่านอกกลุ่ม		มีค่านอกกลุ่มเฉพาะในกลุ่มที่ 1		มีค่านอกกลุ่มทั้ง 2 กลุ่ม	
		สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์	สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์	สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์
ปกติ	20	0.0519	0.0499	0.0184*	0.0509	0.0105*	0.0505
	40	0.0514	0.0509	0.0191*	0.0523	0.0232*	0.0510
	60	0.0494	0.0506	0.0224*	0.0519	0.0347	0.0514
	100	0.0523	0.0524	0.0340	0.0526	0.0429	0.0525
ยูนิฟอร์ม	20	0.0500	0.0474	0.0128*	0.0479	0.0050*	0.0478
	40	0.0505	0.0490	0.0115*	0.0483	0.0263	0.0483
	60	0.0484	0.0478	0.0221*	0.0486	0.0431	0.0486
	100	0.0485	0.0478	0.0452	0.0477	0.0481	0.0481
โลจิสติก	20	0.0479	0.0474	0.0187*	0.0480	0.0096*	0.0480
	40	0.0493	0.0490	0.0168*	0.0483	0.0197*	0.0484
	60	0.0474	0.0478	0.0221*	0.0482	0.0321	0.0487
	100	0.0451	0.0478	0.0341	0.0475	0.0445	0.0483

*หมายถึง ค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ที่อยู่นอกเกณฑ์ในการพิจารณาของ Bradley (กล่าวคือ ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 แตกต่างจากที่กำหนด (0.05))

จากตาราง 3 พบว่า ในกรณีไม่มีค่านอกกลุ่ม ที่ระดับนัยสำคัญเท่ากับ 0.05 สถิติทดสอบที และสถิติทดสอบแมนน์-วิตนีย์ สามารถควบคุมค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ให้อยู่ในเกณฑ์ของ Bradley ได้ สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง แต่โดยภาพรวมสถิติทดสอบทีจะมีค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ใกล้เคียงกับค่าระดับนัยสำคัญ 0.05 ที่กำหนดมากกว่าสถิติทดสอบแมนน์-วิตนีย์

ในกรณีมีค่านอกกลุ่มเฉพาะในตัวอย่างกลุ่มที่ 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 พบว่า ในทุกการแจกแจง ความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ของสถิติทดสอบทีจะลดลงอย่างเห็นได้ชัด โดยจะสามารถควบคุมการเกิดความผิดพลาดแบบที่ 1 เมื่อขนาดตัวอย่างเท่ากับ 100 กล่าวคือ เมื่อขนาดตัวอย่างต่ำกว่า 100 ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 จะแตกต่างไปจากค่าที่กำหนด (0.05) ในขณะที่สถิติ

ทดสอบแมนน์-วิตนีย์สามารถควบคุมค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ให้อยู่ในเกณฑ์ของ Bradley ได้ สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง

ในกรณีมีค่านอกกลุ่มในตัวอย่างทั้งสองกลุ่ม ที่ระดับนัยสำคัญเท่ากับ 0.05 สถิติทดสอบแมนน์-วิตนีย์สามารถควบคุมค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ให้อยู่ในเกณฑ์ของ Bradley ได้ สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง แต่สถิติทดสอบทีจะสามารถควบคุมได้กรณีขนาดตัวอย่างเท่ากับ 60 และ 100 สำหรับการแจกแจงแบบปกติ และแบบโลจิสติก และกรณีขนาดตัวอย่างตั้งแต่ 40 ขึ้นไป สำหรับการแจกแจงแบบยูนิฟอร์ม

2. ค่าประมาณกำลังการทดสอบของการทดสอบที และการทดสอบแมนน์-วิตนีย์ แสดงดังตาราง 4

ตาราง 4 ค่าประมาณกำลังการทดสอบของสถิติทดสอบที และสถิติทดสอบแมนน์-วิตนีย์ ในกรณีไม่มีค่านอกกลุ่มมีค่านอกกลุ่มเฉพาะในกลุ่มที่ 1 และ มีค่านอกกลุ่มทั้ง 2 กลุ่ม ที่ระดับนัยสำคัญ α = 0.05

การแจกแจง	ขนาดตัวอย่าง	ไม่มีค่านอกกลุ่ม		มีค่านอกกลุ่มในกลุ่มที่ 1		มีค่านอกกลุ่มทั้ง 2 กลุ่ม	
		สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์	สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์	สถิติทดสอบที	สถิติทดสอบแมนน์-วิตต์นีย์
ปกติ	20	0.9963	0.9945	0.5587*	0.9887	0.3379*	0.9796
	40	1.0000	1.0000	0.6494*	1.0000	0.4173*	1.0000
	60	1.0000	1.0000	0.7404*	1.0000	0.5131	1.0000
	100	1.0000	1.0000	0.8666	1.0000	0.6602	1.0000
ยูนิฟอร์ม	20	0.9969	0.9891	0.5195*	0.9799	0.3121*	0.9665
	40	1.0000	1.0000	0.5791*	1.0000	0.3640	0.9996
	60	1.0000	1.0000	0.7075*	1.0000	0.4714	1.0000
	100	1.0000	1.0000	0.8730	1.0000	0.6481	1.0000
โลจิสติก	20	0.9942	0.9964	0.6004*	0.9915	0.3790*	0.9862
	40	1.0000	1.0000	0.6858*	1.0000	0.4588*	1.0000
	60	1.0000	1.0000	0.7588*	1.0000	0.5407	1.0000
	100	1.0000	1.0000	0.8707	1.0000	0.6772	1.0000

*หมายถึง ค่าประมาณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 ที่อยู่นอกเกณฑ์ในการพิจารณาของ Bradley (กล่าวคือ ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 แตกต่างจากที่กำหนด (0.05))

จากตาราง 4 พบว่า ในกรณีไม่มีค่านอกกลุ่ม ที่ระดับนัยสำคัญเท่ากับ 0.05 เมื่อขนาดตัวอย่างเท่ากับ 20 สำหรับการแจกแจงแบบปกติและแบบยูนิฟอร์ม สถิติทดสอบทีจะมีค่าประมาณกำลังการทดสอบสูงกว่าสถิติทดสอบแมนน์-วิตนีย์ แต่ที่ขนาดตัวอย่างตั้งแต่ 40 ขึ้นไปสถิติทดสอบทีและสถิติทดสอบแมนน์-วิตนีย์จะมีค่าประมาณกำลังการทดสอบไม่แตกต่างกัน สำหรับทุกการแจกแจง

ในกรณีมีค่านอกกลุ่มเฉพาะในตัวอย่างกลุ่มที่ 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 โดยผลจากตารางที่ 3 พบว่า สถิติแมนวิตนีย์สามารถควบคุมความน่าจะเป็นการเกิดความผิดพลาดแบบที่ 1 ได้ดีกว่าสถิติทดสอบที และเพื่อพิจารณาที่กำลังการทดสอบ จะพบว่าสถิติทดสอบแมนน์-วิตนีย์ จะมีค่าประมาณกำลังการทดสอบสูงกว่าสถิติทดสอบที โดยมีค่าเข้าใกล้ 1 สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง

ในกรณีมีค่านอกกลุ่มในตัวอย่างทั้งสองกลุ่ม ที่ระดับนัยสำคัญเท่ากับ 0.05 โดยผลจากตารางที่ 3 พบว่า สถิติทดสอบแมนน์-วิตนีย์ สามารถควบคุมความน่าจะเป็นการเกิดความผิดพลาดแบบที่ 1 ได้ดีกว่าสถิติทดสอบที และเพื่อพิจารณาที่กำลังการทดสอบ จะพบว่าสถิติทดสอบแมนน์-วิตนีย์จะมีค่าประมาณกำลังการทดสอบสูงกว่าสถิติทดสอบที โดยมีค่าเข้าใกล้ 1 สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง ส่วนสถิติทดสอบทีโดยภาพรวมจะมีค่าประมาณกำลังการทดสอบค่อนข้างต่ำเมื่อเทียบกับสถิติทดสอบแมนน์-วิตนีย์ พบว่ามีค่าอยู่ในช่วง 0.37 – 0.67 สำหรับทุกการแจกแจงและทุกขนาดตัวอย่าง

สรุปผลการวิจัย

ในการเปรียบเทียบความแตกต่างของค่ากลางของประชากร 2 กลุ่ม กรณีที่ความแปรปรวนเท่ากัน พบว่า จากการศึกษาในกรณีที่การแจกแจงของประชากรทั้ง 2 กลุ่มเป็นแบบสมมาตร ในกรณีไม่มีค่านอกกลุ่ม สถิติทดสอบทีและแมนน์-วิตนีย์ จะมีความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ได้ใกล้เคียงกันในทุกขนาดตัวอย่างและทุกการแจกแจง แต่เมื่อขนาดตัวอย่างทั้งสองกลุ่มมีค่ามากขึ้น กำลังการทดสอบของสถิติทดสอบแมนน์-วิตนีย์จะสูงกว่าสถิติทดสอบที ซึ่งผลการวิจัยสอดคล้องกับผลการวิจัยของ มนตรี (2556, 2557) ในกรณีประชากรมีการแจกแจงแบบปกติ และ ผลการวิจัยของ M. Bakker (2014)

กรณีที่ตัวอย่างสุ่มกลุ่มที่ 1 หรือทั้งสองกลุ่มมีค่าผิดปกติปลอมปนอยู่ จะพบว่า สถิติทดสอบแมนน์-วิตนีย์จะมีประสิทธิภาพสูงกว่าสถิติทดสอบทีในทุกสถานการณ์ และเมื่อขนาดตัวอย่างเพิ่มมากขึ้น สถิติทดสอบทีจะมีความสามารถในการควบคุมการเกิดความผิดพลาดแบบที่ 1 ได้ดีขึ้น รวมถึงกำลังการทดสอบมีแนวโน้มที่สูงขึ้น ซึ่งผลการวิจัยสอดคล้องกับผลการวิจัยของ M. Bakker (2014)

จากผลการวิจัยจะเห็นว่าค่านอกกลุ่มในชุดข้อมูลส่งผลประสิทธิภาพของสถิติทดสอบทีมากกว่าสถิติทดสอบแมนน์-วิตนีย์ ดังนั้นหากในการทำงานในสถานการณ์จริงที่เกี่ยวข้องกับการทดสอบเพื่อเปรียบเทียบความแตกต่างของค่ากลางของประชากร 2 กลุ่ม เมื่อตัวอย่างสุ่มทั้ง 2 กลุ่มอิสระกัน ผู้วิจัยควรตระหนักถึงการตรวจสอบค่านอกกลุ่มในชุดข้อมูลเพิ่มเติมจากการตรวจสอบข้อสมมติเบื้องต้นของการใช้สถิติทดสอบทีหรือแมนน์-วิตนีย์ เสียก่อน เพื่อที่จะได้เลือกใช้สถิติทดสอบที่เหมาะสม ซึ่งจะนำไปสู่การสรุปผลการวิเคราะห์ที่มีความน่าเชื่อถือต่อไป

เอกสารอ้างอิง

ภาษาไทย

คณาจารย์ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร. (2555). **สถิติพื้นฐาน 1**. นครปฐม : โรงพิมพ์มหาวิทยาลัยศิลปากร.

มนตรี สังข์ทอง. (2556). “การศึกษาประสิทธิภาพของสถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม”**วารสารวิทยาศาสตร์ มข.** 41,1 (2556): 226-238, เข้าถึงเมื่อ 8 ธันวาคม 2558, เข้าถึงได้จาก http://scijournal.kku.ac.th/files/Vol_41_No_1_P_226-238.pdf

มนตรี สังข์ทอง. (2557). “ความแกร่งและอำนาจการทดสอบของสถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากรสองกลุ่ม สำหรับข้อมูลแบบลิเคิร์ท 5 ระดับ” **วารสารวิทยาศาสตร์และเทคโนโลยี** 22, 5 (2557): 605-619, เข้าถึงเมื่อ 8 ธันวาคม 2558, เข้าถึงได้จาก http://www.tcithaijo.org/index.php/tstj/article/ view/20618/17912.

ล้วน สายยศ และอังคณา สายยศ. (2540). **สถิติวิทยาทางการวิจัย**. กรุงเทพฯ : สุวีริยาสาส์น. น.242.

ภาษาต่างประเทศ

Marjan Bakker. (2014) “Outlier removal, sum scores, and the inflation of the Type I error rate in independent samples t test :The power of alternatives and recommendations” in **Good Science, Bad Science Questioning Research Practices in Psychological Research** (2014) : 63-92, accessed February 20, 2016, available from http://hdl.handle.net/11245/2.138854