

การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบ
อิงพารามิเตอร์และไม่อิงพารามิเตอร์
ระหว่างค่ากลางของประชากร 2 กลุ่ม สัมพันธ์กัน
ภายใต้การแจกแจงที่มีความโด่งปรกติและโด่งมาก
**Compared the Efficiency of
Parametric and Nonparametric Statistics between
the Central Value of Two Population Groups Related
under Mesokurtic and Leptokurtic Kurtosis Distribution**

สายชล สิ้นสมบุญทอง, ฉัตรพร อันทรบุตร, ณิชกมล วงศ์สุวรรณ,

พิชชา พันิตอุทัย และศุภานัน เพชรโชติ*

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ถนนฉลองกรุง เขตลาดกระบัง กรุงเทพมหานคร 10520

Saichon Sinsomboonthong, Chattaporn Untarabutr, Nitkamon Wongsuwan,

Pitcha Panitutai and Supanun Petchot*

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang,

Chalongkrung Road, Ladkrabang, Bangkok 10520

Received: May 13, 2019; Accepted: July 22, 2019

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวสถิติอิงพารามิเตอร์และไม่อิงพารามิเตอร์ระหว่างค่ากลางของประชากร 2 กลุ่ม สัมพันธ์กัน ภายใต้การแจกแจงที่มีความโด่งปรกติและโด่งมาก โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบในการแจกแจงปรกติหลายตัวแปรและการแจกแจงลาปลาซหลายตัวแปร ค่าสัมประสิทธิ์สหสัมพันธ์ทางบวก ได้แก่ 0.1, 0.5, และ 0.9 ส่วนทางลบ ได้แก่ -0.1, -0.5 และ -0.9 ที่ระดับนัยสำคัญ 0.01 และ 0.05 โดยกำหนดขนาดกลุ่มตัวอย่างเท่ากัน ได้แก่ (20,20), (40,40), (60,60), (80,80) และ (100,100) เมื่อขนาดตัวอย่างเพิ่มมากขึ้นทำให้ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 เพิ่มมากขึ้น และมีกำลังการทดสอบสูงขึ้น พบว่าตัวสถิติทดสอบวิลค็อกซันมีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุด นอกจากนี้ในส่วนของกำลังการทดสอบพบว่าตัวสถิติทดสอบการเรียงสับเปลี่ยน และตัวสถิติทดสอบวิลค็อกซันยังให้กำลังการทดสอบสูงที่สุดเช่นกัน

คำสำคัญ : ความน่าจะเป็นของความผิดพลาดแบบที่ 1; กำลังการทดสอบ; การวัดค่ากลางของข้อมูล

Abstract

The objective of this project was to compare the efficiency of parametric and nonparametric statistics between the central value of two population groups related under mesokurtic and leptokurtic kurtosis distribution. The evaluation was in terms of statistics ability to control the probability of type I error and the power of a test. The runs in the case of multivariate normal distribution and multivariate Laplace distribution, at significant levels of 0.01 and 0.05, correlation coefficients were set to be 0.1, 0.5, 0.9 and -0.1, -0.5, -0.9. The sizes of the samples from the two populations group were the same, as follows: (20,20), (40,40), (60,60), (80,80), (100,100). When the sample size is increased, it has ability to control high probability of type I errors and high power of a test in all cases. It was shown that Wilcoxon matched-pair signed-ranks test had the most control over the probability of type I error. In addition, for power of a test, it was found that randomization test for matched pair or permutation test for paired replicated and Wilcoxon matched-pair signed-ranks test also gave the highest power.

Keywords: probability of type I error; power of a test; measure of central value

1. คำนำ

การวิจัยเป็นกระบวนการแสวงหาศึกษาค้นคว้าหรือพัฒนาองค์ความรู้ใหม่ ๆ ที่มีลักษณะเป็นนัยทั่วไปอย่างมีระบบแบบแผน โดยใช้กระบวนการทางวิทยาศาสตร์ ซึ่งประกอบด้วยขั้นตอนที่สำคัญ ได้แก่ การกำหนดประเด็นปัญหาการตั้งสมมติฐาน การเก็บรวบรวมข้อมูล การวิเคราะห์และการทดสอบสมมติฐาน สรุปผล และยืนยันข้อค้นพบ ดังนั้นสถิติจึงเป็นเทคนิคที่เกี่ยวข้องกับการวิจัยโดยตรง นักวิจัยต้องรู้จักนำสถิติมาประยุกต์ใช้เป็นเครื่องมือของการวิจัย หมายความว่านักวิจัยต้องมีความรู้ความสามารถในการวิเคราะห์ข้อมูล มีการสรุปอ้างอิงที่เชื่อถือได้ จากค่ากลางข้างต้นแสดงให้เห็นว่าสถิติเป็นวิธีการ (method) ซึ่งเป็นส่วนหนึ่งของระเบียบวิธีวิจัย (research methodology) ที่สำคัญในการสรุปอ้างอิงไปถึงค่าพารามิเตอร์ (parameter) ของประชากร โดยอาศัยข้อมูลจากการชักตัวอย่าง ดังนั้นกระบวนการในการเลือกตัวสถิติ

ทดสอบที่เหมาะสมกับแต่ละสถานการณ์ของข้อมูลจึงเป็นสิ่งที่สำคัญอย่างยิ่ง โดยต้องคำนึงถึงข้อกำหนดเบื้องต้น (basic assumption) ของตัวสถิติทดสอบแต่ละชนิดด้วย (มนตรี, 2556)

Blair (1985) ศึกษาเปรียบเทียบกำลังการทดสอบ (power of a test) ของตัวสถิติทดสอบที่ (dependent t test, T) และตัวสถิติทดสอบวิลค็อกซัน (Wilcoxon matched-pair signed-ranks test, W) เมื่อประชากร 2 กลุ่ม มีความสัมพันธ์กัน ภายใต้การแจกแจงของประชากรทั้งหมด 10 การแจกแจงผลการศึกษา คือ T มีประสิทธิภาพมากเมื่อประชากรมีการแจกแจงปกติ ซึ่งเป็นไปตามทฤษฎี อย่างไรก็ตาม เมื่อการแจกแจงของข้อมูลทั้ง 2 กลุ่มไม่ได้มีการแจกแจงปกติ W จะมีประสิทธิภาพมากกว่า T และเมื่อขนาดตัวอย่างใหญ่ขึ้น W จะมีประสิทธิภาพมากขึ้นด้วย

Butar (2008) ทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม ที่ไม่อิสระกัน โดย W

และสถิติทดสอบการเรียงสับเปลี่ยน (permutation test, P) พบว่า P มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 (probability of type I error) และมีกำลังการทดสอบเมื่อข้อมูลอยู่ในมาตรวัดอัตราส่วน และเมื่อข้อมูลมีความเบ้ P ก็ยังคงมีประสิทธิภาพดี

ลลิตา (2554) ศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบ T และ W ในการทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระกัน พบว่า T และ W มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อผลต่างของข้อมูล 2 กลุ่ม มีการแจกแจงปกติและเอกรูปในทุกขนาดตัวอย่าง แต่ W ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในกรณีผลต่างของข้อมูล 2 กลุ่ม มีการแจกแจงกัมเบล ส่วนกำลังการทดสอบ T มีกำลังการทดสอบสูงกว่า W ในทุกกรณี แต่เมื่อผลต่างของข้อมูล 2 กลุ่ม มีการแจกแจงเอกรูป ตัวสถิติทดสอบทั้งสองมีกำลังการทดสอบค่อนข้างต่ำ โดยเฉพาะอย่างยิ่งเมื่อตัวอย่างมีขนาดเล็ก

อานนท์ (2555) ศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบค่ากลางของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระ พบว่า T มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ดีที่สุดในทุกกรณี รองลงมา คือ W ส่วนตัวสถิติทดสอบด้วยเครื่องหมาย (sign test for two related sample, S) มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 น้อยที่สุด ส่วนในกรณีของกำลังการทดสอบ T มีกำลังการทดสอบสูงที่สุดในทุกกรณี รองลงมา คือ W ส่วน S มีกำลังการทดสอบต่ำที่สุด และเมื่อขนาดของความสัมพันธ์มีค่ามากขึ้น และตัวอย่างมีขนาดใหญ่ขึ้น กำลังการทดสอบของตัวสถิติทดสอบทั้ง 3 วิธี จะมีค่าเพิ่มขึ้นและมีค่าใกล้เคียงกัน

Porschanc และคณะ (2014) ศึกษาความสัมพันธ์ระหว่าง P และ T พบว่า P จะให้ผลการทดสอบใกล้เคียงกับ T เมื่อการแจกแจงแบบสุมเข้าใกล้การแจกแจงปกติ

กมลวรรณ และคณะ (2560) ศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบตัวสถิติทดสอบไม่อิงพารามิเตอร์ระหว่างค่ามัธยฐานของประชากร 2 กลุ่ม สัมพันธ์กัน พบว่า W มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุดในทุกกรณี และกำลังการทดสอบสูงที่สุดในทุกกรณี ในขณะที่ S มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากกว่า P ในทุกกรณี แต่ทว่า S และ P มีกำลังการทดสอบรองลงมาจาก W ในกรณีที่แตกต่างกัน อีกทั้งกำลังการทดสอบของตัวสถิติทดสอบทั้ง 3 การทดสอบที่ได้มีค่าน้อยมาก ในกรณีที่ข้อมูลทั้ง 2 กลุ่ม มีการแจกแจงแกมมาหลายตัวแปร ซึ่งอาจทำให้ตัวสถิติทดสอบทั้ง 3 การทดสอบ นี้เหมาะสมกับการใช้ ในกรณีที่ข้อมูลที่มีการแจกแจงปกติหลายตัวแปร มากกว่าการแจกแจงแกมมาหลายตัวแปร

กัลยา และคณะ (2561) ศึกษางานวิจัยเกี่ยวกับการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างระหว่างประชากร 2 กลุ่ม ที่ไม่เป็นอิสระกัน พบว่า W และ P มีกำลังการทดสอบสูง และเมื่อตัวอย่างขนาดใหญ่ กำลังการทดสอบจะสูงขึ้นด้วย

ดังนั้นงานวิจัยนี้จึงศึกษาการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ ได้แก่ T ไม่เป็นอิสระกัน S สำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน W และ P กำหนดระดับนัยสำคัญ 0.01 และ 0.05 และขนาดของความสัมพันธ์ 6 ระดับ ทั้งทางบวกและทางลบ ได้แก่ 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9 ระหว่างประชากร 2 กลุ่ม ในการวิเคราะห์ความแตกต่างของค่าเฉลี่ยหรือค่ามัธยฐานของประชากร 2 กลุ่มที่

สัมพันธ์กัน โดยพิจารณาจากความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ โดยสุ่มข้อมูลจากการแจกแจงที่มีค่าสัมประสิทธิ์ความโค้งที่มีค่าปกติและโค้งมาก ได้แก่ การแจกแจงปกติหลายตัวแปร (multivariate normal distribution) และการแจกแจงลาปลาซหลายตัวแปร (multivariate Laplace distribution) ตามลำดับ เพื่อให้ทราบถึงตัวสถิติทดสอบที่เหมาะสมสำหรับข้อมูลในแต่ละสถานการณ์ ซึ่งการจำลองและวิเคราะห์ข้อมูลทำได้โดยใช้โปรแกรมอาร์ (R)

2. ขอบเขตของการวิจัย

งานวิจัยนี้ศึกษาการเปรียบเทียบตัวสถิติทดสอบทั้ง 4 วิธี ของตัวสถิติทดสอบที่สัมพันธ์กัน ได้แก่ ตัวสถิติทดสอบที่ไม่เป็นอิสระกัน ตัวสถิติทดสอบด้วยเครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน ตัวสถิติทดสอบวิลค็อกซัน และตัวสถิติทดสอบการเรียงสับเปลี่ยน ซึ่งแบ่งขนาดของตัวอย่าง (n_1, n_2) ออกเป็น $(20, 20)$, $(40, 40)$, $(60, 60)$, $(80, 80)$ และ $(100, 100)$ เพื่อต้องการให้ได้ผลที่ครอบคลุมและครบถ้วน การแจกแจงที่นำมาศึกษา คือ การแจกแจงปกติหลายตัวแปร และการแจกแจงลาปลาซหลายตัวแปร มีการทำซ้ำ 1,000 ครั้ง ในแต่ละกรณี และใช้เกณฑ์การทดสอบของแบรดลีย์ (Bradley, 1978) ซึ่งเป็นเกณฑ์ที่ใช้ควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 กำหนดระดับนัยสำคัญที่ใช้ในการทดสอบ 2 ระดับ คือ 0.01 และ 0.05 กำหนดค่าของสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9 โดยได้นำข้อเสนอแนะของงานวิจัยที่ศึกษามาพัฒนาเนื่องจากงานวิจัยส่วนใหญ่ยังไม่มีมีการกำหนดค่าสัมประสิทธิ์สหสัมพันธ์เป็นทางลบ และใช้โปรแกรมอาร์ในการจำลองและวิเคราะห์ข้อมูล

3. สถิติและการแจกแจงที่ใช้ในงานวิจัย

3.1 ตัวสถิติที่ใช้ในงานวิจัย

3.1.1 ตัวสถิติทดสอบที่ไม่เป็นอิสระกัน

การทดสอบสมมติฐานแบบจับคู่ในกรณีที่หน่วยตัวอย่างคนละหน่วยกัน ให้แบ่งหน่วยตัวอย่างเป็น 2 กลุ่ม เท่า ๆ กัน โดยที่หน่วยตัวอย่างของแต่ละกลุ่มที่มาจับคู่กันจะต้องมีลักษณะต่าง ๆ คล้ายคลึงกันหรือเหมือนกันให้มากที่สุด ส่วนหน่วยตัวอย่างต่างคู่กันจะมีลักษณะคล้ายคลึงกันหรือแตกต่างกันก็ได้ แล้วสุ่มสิ่งที่ต้องการทดสอบให้กับแต่ละกลุ่ม ส่วนการทดสอบสมมติฐานแบบจับคู่ในกรณีที่หน่วยตัวอย่างเดียวกันแต่วัดผล 2 ครั้ง มักใช้กับการทดสอบผลต่างของผลของการไม่ใช้สิ่งที่ต้องการทดสอบกับผลของใช้สิ่งที่ต้องการทดสอบ (สายชล, 2551)

สมมติฐาน (hypothesis)

การทดสอบแบบสองหาง

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

การทดสอบแบบหางเดียวด้านขวา

$$H_0: \mu_d = 0$$

$$H_1: \mu_d > 0$$

การทดสอบแบบหางเดียวด้านซ้าย

$$H_0: \mu_d = 0$$

$$H_1: \mu_d < 0$$

ตัวสถิติทดสอบ (test statistics)

$$t = \frac{\bar{d}}{S_d / \sqrt{n}}$$

$$\text{โดยที่ } \bar{d} = \frac{\sum_{i=1}^n d_i}{n}$$

$$S_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n - 1}$$

เมื่อ $\bar{d}$ คือ ค่าเฉลี่ยของ d_i โดยที่ $i = 1, 2, \dots, n$; d_i คือ ผลต่างระหว่างข้อมูลคู่ที่ i ;

S_d คือ ค่าส่วนเบี่ยงเบนมาตรฐานของ d_i ; n คือ ขนาดตัวอย่าง

โดยตัวสถิติทดสอบที่นี้จะมีการแจกแจงที่ด้วยองศาเสรีเท่ากับ $n-1$

อาณาเขตวิกฤต (critical region)

การทดสอบแบบสองหาง จะปฏิเสธสมมติฐาน H_0 ถ้า $t > t_{\frac{\alpha}{2}; n-1}$ หรือ $t < -t_{\frac{\alpha}{2}; n-1}$

การทดสอบแบบหางเดียวด้านขวา

จะปฏิเสธสมมติฐาน H_0 ถ้า $t > t_{\alpha; n-1}$

การทดสอบแบบหางเดียวด้านซ้าย

จะปฏิเสธสมมติฐาน H_0 ถ้า $t < -t_{\alpha; n-1}$

3.1.2 ตัวสถิติทดสอบด้วยเครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน

การทดสอบเครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน จะมีประโยชน์มากถ้ามาตรวัดสำหรับค่าสังเกตจะเป็นเพียงมาตรวัดอันดับ ทั้งนี้เพราะว่าในแต่ละคู่ของการวัดจะพิจารณาแต่เพียงว่าค่าหนึ่งใหญ่กว่าอีกค่าหนึ่งหรือไม่เท่านั้น และถ้ามีค่าใดใหญ่กว่าจริงจะหาค่าใหญ่ ๆ นั้น ถ้าข้อมูลที่ได้จากมาตรวัดที่สูงกว่า เช่น มาจากมาตรวัดแบบอันดับภาคหรืออัตราส่วน ควรใช้แบบทดสอบที่มีกำลังการทดสอบมากกว่า ข้อมูลประกอบด้วยค่าสังเกตของตัวอย่างคู่ที่มีสองตัวแปร นั่นคือ $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ เมื่อ n แทนจำนวนคู่ของค่าสังเกต ซึ่งคู่ของค่าสังเกตควรมีความสัมพันธ์กันทางธรรมชาติบางอย่าง ภายในแต่ละคู่ (X_i, Y_i) จะมีการเปรียบเทียบระหว่างค่าสังเกตทั้งสอง โดยกำหนดว่าจะจัดคู่อยู่ในกลุ่มเครื่องหมายบวก (+) ถ้า $X_i < Y_i$ จะจัดอยู่ในกลุ่มเครื่องหมายลบ (-) ถ้า $X_i > Y_i$ และจะจัดอยู่ในกลุ่มศูนย์ (0) ถ้า $X_i = Y_i$ ดังนั้นมาตรวัดต้องการแต่เพียงแบบอันดับเท่านั้น (สยชล, 2552)

สมมติฐาน

การทดสอบแบบสองหาง

$$H_0 : M_1 = M_2$$

$$H_1 : M_1 \neq M_2$$

การทดสอบแบบหางเดียวด้านขวา

$$H_0 : M_1 = M_2$$

$$H_1 : M_1 > M_2$$

การทดสอบแบบหางเดียวด้านซ้าย

$$H_0 : M_1 = M_2$$

$$H_1 : M_1 < M_2$$

ตัวสถิติทดสอบ

ตัวอย่างมีขนาดเล็ก ($n \leq 20$)

$$T = \text{ผลรวมของเครื่องหมายบวก}$$

บวก

การแจกแจงของตัวสถิติทดสอบ T

ภายใต้ H_0 มีการแจกแจงทวินามด้วย $p = \frac{1}{2}$ และ n เท่ากับจำนวนคู่ที่ตัวแปรทั้งสองค่าไม่เท่ากัน นั่นคือ ไม่นับจำนวนคู่ที่มีค่าของตัวแปรทั้งสองเท่ากัน

ตัวอย่างมีขนาดใหญ่ ($n > 20$)

เมื่อสุ่มตัวอย่างขนาดใหญ่ จะประมาณการแจกแจงทวินามด้วยการแจกแจงปกติซึ่งทำให้

$$t = \frac{1}{2}(n - Z_{\frac{\alpha}{2}}\sqrt{n})$$

เมื่อ $Z_{\frac{\alpha}{2}}$ หาได้จากตารางแจกแจง

ปรกติมาตรฐาน

อาณาเขตวิกฤต

ตัวอย่างมีขนาดเล็ก ($n \leq 20$)

การทดสอบแบบสองหาง

อาณาเขตวิกฤต คือ $P(X \leq t) =$

$\frac{\alpha}{2}$ จะปฏิเสธสมมติฐาน H_0 ถ้า $T \leq t$ หรือ $T > n - t$

การทดสอบแบบหางเดียวด้านขวา

อาณาเขตวิกฤต คือ $P(X \leq t)$

$= \alpha$ จะปฏิเสธสมมติฐาน H_0 ถ้า $T \geq n - t$

การทดสอบแบบหางเดียวด้านซ้าย

อาณาเขตวิกฤต คือ $P(X \leq t)$

$= \alpha$ จะปฏิเสธสมมติฐาน H_0 ถ้า $T \leq t$

ตัวอย่างมีขนาดใหญ่ ($n > 20$)

การทดสอบแบบสองหาง จะปฏิเสธ

สมมติฐาน H_0 ถ้า $T \leq t$ หรือ $T \geq n - t$

การทดสอบแบบหางเดียวด้านขวา

จะปฏิเสธสมมติฐาน H_0 ถ้า $T \geq n - t$

การทดสอบแบบหางเดียวด้านซ้าย

จะปฏิเสธสมมติฐาน H_0 ถ้า $T \leq t$

3.1.3 ตัวสถิติทดสอบวิลค็อกซัน

เป็นวิธีการทดสอบที่พัฒนามาจากตัวสถิติทดสอบด้วยเครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน แต่นำเอาขนาดความแตกต่างของข้อมูลแต่ละคู่มาพิจารณาด้วย นอกจากจะพิจารณาเครื่องหมายในแต่ละคู่แล้ว โดยยึดหลักที่ว่าถ้าค่ามัธยฐานของประชากรทั้ง 2 กลุ่ม ไม่แตกต่างกันแล้ว ผลรวมของลำดับที่ของความแตกต่างที่มีเครื่องหมายบวก (+) และเครื่องหมายลบ (-) น่าจะใกล้เคียงกัน ส่วนการทดสอบโดยใช้เครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กันไม่ได้คำนึงถึงขนาดของความแตกต่างของข้อมูลแต่ละคู่ โดยเฉพาะเครื่องหมายของความแตกต่างภายในคู่เท่านั้น คือ ดูว่า X_i มีค่ามากกว่า Y_i จะให้เครื่องหมายบวก (+) และ X_i มีค่าน้อยกว่า Y_i จะให้เครื่องหมายลบ (-) ดังนั้นการทดสอบแบบนี้จะมีกำลังการทดสอบมากกว่าและมีความไว (sensitive) ในการทดสอบมากกว่าการทดสอบเครื่องหมายสำหรับตัวอย่าง 2 กลุ่ม ที่มีความสัมพันธ์กัน เนื่องจากใช้ข้อมูลได้สมบูรณ์กว่า คือใช้ทั้งขนาดและเครื่องหมายของความแตกต่างของข้อมูล ข้อมูลอยู่ในมาตรวัดอย่างน้อยเป็นแบบอันดับภาค (สายซล, 2552)

สมมติฐาน

การทดสอบแบบสองหาง

$$H_0: M_1 = M_2$$

$$H_1: M_1 \neq M_2$$

การทดสอบแบบหางเดียวด้านขวา

$$H_0: M_1 = M_2$$

$$H_1: M_1 > M_2$$

การทดสอบแบบหางเดียวด้านซ้าย

$$H_0: M_1 = M_2$$

$$H_1: M_1 < M_2$$

เมื่อ M_1 คือ มัธยฐานของประชากรที่ตัวอย่างขนาด n_1 ; M_2 คือ มัธยฐานของประชากรที่ตัวอย่างขนาด n_2

ตัวสถิติทดสอบ

ตัวอย่างมีขนาดเล็ก ($n \leq 30$)

การทดสอบแบบสองหาง

$$T = \min(T_{(+)}, T_{(-)})$$

การทดสอบแบบหางเดียวด้านขวา

$$T = T_{(-)}$$

การทดสอบแบบหางเดียวด้านซ้าย

$$T = T_{(+)}$$

ตัวอย่างมีขนาดใหญ่ ($n > 30$)

เมื่อตัวอย่างมีขนาดใหญ่ จากทฤษฎีบทขีดจำกัดส่วนกลาง (central limit theorem) จะได้ว่าตัวสถิติทดสอบ T ภายใต้ H_0 จะมีการแจกแจงโดยประมาณใกล้เคียงกับการแจกแจงปกติ โดยมีค่าเฉลี่ยและความแปรปรวนดังนี้

$$\mu_T = E(T) = \frac{n(n+1)}{4}$$

$$\sigma_T^2 = Var(T) = \frac{n(n+1)(2n+1)}{24}$$

แปลงให้เป็นค่า Z จะได้ $Z = \frac{T - \mu_T}{\sigma_T} \sim N(0,1)$

การทดสอบแบบสองหาง

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

การทดสอบแบบหางเดียวด้านขวา

$$Z = \frac{T_{(-)} - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

การทดสอบแบบหางเดียวด้านซ้าย

$$Z = \frac{T_{(+)} - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

อาณาเขตวิกฤต

ตัวอย่างมีขนาดเล็ก ($n \leq 30$)

การทดสอบแบบสองหาง จะปฏิเสธ

สมมติฐาน H_0 ถ้า $T \leq T_{\frac{\alpha}{2}, n}$

การทดสอบแบบหางเดียวด้านขวา

จะปฏิเสธสมมติฐาน H_0 ถ้า $T_{(-)} \leq T_{\alpha, n}$

การทดสอบแบบหางเดียวด้านซ้าย

จะปฏิเสธสมมติฐาน H_0 ถ้า $T_{(+)} \leq T_{\alpha, n}$

ตัวอย่างมีขนาดใหญ่ ($n > 30$)

การทดสอบแบบสองหาง จะปฏิเสธ

สมมติฐาน H_0 ถ้า $Z > Z_{\frac{\alpha}{2}}$ หรือ $Z < -Z_{\frac{\alpha}{2}}$

การทดสอบแบบหางเดียวด้านขวา

จะปฏิเสธสมมติฐาน H_0 ถ้า $Z > Z_{\alpha}$

การทดสอบแบบหางเดียวด้านซ้าย

จะปฏิเสธสมมติฐาน H_0 ถ้า $Z < -Z_{\alpha}$

3.1.4 ตัวสถิติทดสอบการเรียงสับเปลี่ยน

การทดสอบนี้เป็นทดสอบประชากรแบบ

จับคู่ (matched paired) โดยมีสมมติฐานว่าตัวอย่างทั้ง 2 กลุ่ม มีความแตกต่างกันหรือไม่ การทดสอบนี้เป็น การทดสอบที่แสดงให้เห็นถึงแนวคิดและวิธีการของสถิติไม่อิงพารามิเตอร์ เนื่องจากการหาอาณาเขตวิกฤตและระดับนัยสำคัญที่แท้จริงสามารถหาได้จากข้อมูลตัวอย่างโดยตรงโดยไม่จำเป็นต้องมีข้อกำหนดเบื้องต้นเกี่ยวกับการแจกแจงปกติ หรือความแปรปรวนของประชากร 2 กลุ่ม ต้องเท่ากัน

การทดสอบนี้มักถูกนำเสนอในช่วงต้น ๆ ของหนังสือสถิติไม่อิงพารามิเตอร์ เพื่อแสดงถึงแนวคิดและวิธีการที่ไม่จำเป็นต้องอาศัยรูปร่างของโค้งการแจกแจง การทดสอบนี้ภายใต้เงื่อนไขหนึ่งจะมีกำลังการทดสอบสูงสุดของสถิติไม่อิงพารามิเตอร์ (อุมพร, 2542)

สมมติฐาน

การทดสอบแบบสองหาง

$$H_0 : M_1 = M_2$$

$$H_1 : M_1 \neq M_2$$

การทดสอบแบบด้านเดียวข้างขวา

$$H_0 : M_1 \leq M_2$$

$$H_1 : M_1 > M_2$$

การทดสอบแบบด้านเดียวข้างซ้าย

$$H_0 : M_1 \geq M_2$$

$$H_1 : M_1 < M_2$$

โดยที่ M_1 และ M_2 แทนค่ามัธยฐานของประชากรของกลุ่มที่ 1 และ 2 ตามลำดับ

ตัวสถิติทดสอบ

$$P = \sum_{i=1}^n d_i$$

อาณาเขตวิกฤต

การหาอาณาเขตวิกฤตสำหรับการทดสอบสมมติฐานสองหางเมื่อกำหนดระดับนัยสำคัญเท่ากับ α จะประกอบด้วย $\alpha 2^n$ วิธี ซึ่งมีค่า $\sum_{i=1}^n d_i$ ที่มาก โดยแบ่งออกเป็นทางบวกและทางลบด้วยจำนวนวิธีที่เท่ากัน ถ้าตัวอย่างที่สุ่มมามีค่า d_i และ $\sum_{i=1}^n d_i$ ที่เป็นส่วนหนึ่งของอาณาเขตวิกฤต แล้วจะปฏิเสธ H_0 และยอมรับ H_1 ส่วนการทดสอบแบบหางเดียว อาณาเขตวิกฤตจะประกอบด้วยค่า $\sum_{i=1}^n d_i$ ที่มีค่ามากหรือน้อยเพียงทางเดียวตามใน H_1

3.2 การแจกแจงที่ใช้ในงานวิจัย

3.2.1 การแจกแจงปกติหลายตัวแปร (น้อมจิต, 2558)

นิยามที่ 1 ให้ $X_1, \dots, X_p$ เป็นตัวแปรสุ่มที่มีการแจกแจงปกติหลายตัวแปร ซึ่งใช้สัญลักษณ์ว่า $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ ถ้า $f(\underline{x}; \underline{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})' \Sigma^{-1}(\underline{x}-\underline{\mu})}$; $-\infty < X_i < \infty, -\infty < \mu_i < \infty$

3.2.2 การแจกแจงลาปลาซหลายตัวแปร (Singmann, 2018)

นิยามที่ 2 ให้ $X_1, \dots, X_p$ เป็นตัวแปรสุ่มที่มีการแจกแจงลาปลาซหลายตัวแปร ซึ่งใช้สัญลักษณ์ว่า $\underline{X} \sim L_p(\underline{\mu}, \Sigma)$ ถ้า $f(\underline{x}; \underline{\mu}, \Sigma) = \frac{2}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \left(\frac{\pi}{2\sqrt{2(\underline{x}-\underline{\mu})' \Sigma^{-1}(\underline{x}-\underline{\mu})}} \right)^{\frac{1}{2}} e^{-\sqrt{2(\underline{x}-\underline{\mu})' \Sigma^{-1}(\underline{x}-\underline{\mu})}}$; $-\infty < X_i < \infty$
 $-\infty < \mu_i < \infty$

3.3 การวิเคราะห์สหสัมพันธ์เชิงเส้นอย่างง่าย

ตัวสถิติที่ใช้วัดความสัมพันธ์ระหว่าง X และ Y ว่ามากหรือน้อยนั้นเรียกว่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ซึ่งในกรณีที่ค่าของ Y ขึ้นอยู่กับค่าของ X เพียงตัวเดียวเรียกว่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (simple correlation coefficient) โดยที่ ρ จะไม่มีหน่วย จึงสามารถใช้วัดความสัมพันธ์ระหว่าง Y และ X ได้ว่ามีความสัมพันธ์มากหรือน้อยเพียงใด เนื่องจากค่า ρ จะมีค่าสูงสุดเป็น 1 และต่ำสุดเป็น -1 (กัลยา, 2554)

การใช้ข้อมูลตัวอย่างจึงประมาณค่า ρ ด้วยค่า r โดยที่ r คือ สัมประสิทธิ์สหสัมพันธ์ตัวอย่าง โดยที่ $r = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{\sqrt{(\sum_{i=1}^n X_i^2 - n \bar{X}^2)(\sum_{i=1}^n Y_i^2 - n \bar{Y}^2)}}$

$$\text{เมื่อ } s_{xx} = \sum_{i=1}^n X_i^2 - n \bar{X}^2; s_{yy} = \sum_{i=1}^n Y_i^2 - n \bar{Y}^2;$$

$$s_{xy} = \sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}$$

3.4 การวัดความโด่ง (peakness) ของข้อมูลตัวสถิติ

ความโด่งเป็นค่าลักษณะเฉพาะมิติที่ 4 ของการพรรณนาเชิงสถิติ ใช้บอกลักษณะยืดขยายของข้อมูลสถิติทางแนวตั้ง โดยเทียบความโด่งกับเส้นโค้งสมมาตรหรือโค้งปกติ (ณททัย, 2554)

การวัดความโด่งคำนวณค่า Kurtosis ของข้อมูล สามารถคำนวณดังนี้

$$\text{ข้อมูลไม่จัดหมวดหมู่ } Ku = \frac{\sum_{i=1}^n (X_i - \mu)^4}{(n-1)\sigma^4}$$

$$\text{ข้อมูลจัดหมวดหมู่ } Ku = \frac{\sum_{i=1}^k (X_i - \mu)^4 f_i}{(n-1)\sigma^4}$$

ถ้าค่าความโด่งเท่ากับ 3 แสดงว่าการแจกแจงของข้อมูลมีลักษณะการแจกแจงสมมาตร ถ้าค่าความโด่งมีค่ามากกว่า 3 แสดงว่าข้อมูลมีลักษณะเป็นเส้นโค้งสูงแหลมหรือโด่งมาก ถ้ามีค่าน้อยกว่า 3 แสดงว่าข้อมูลมีลักษณะเป็นเส้นโค้งแบนหรือโด่งน้อย

งานวิจัยพิจารณาใช้การแจกแจงปกติและการแจกแจงลาปลาซ โดยที่การแจกแจงปกติมีค่าความโด่งเท่ากับ 3 ซึ่งมีความโด่งปกติ (mesokurtic) และการแจกแจงลาปลาซมีค่าความโด่งเท่ากับ 6 ซึ่งมีความโด่งมาก (leptokurtic) (สายชล, 2558)

4. ขั้นตอนของการวิจัย

เปรียบเทียบตัวสถิติทดสอบทั้ง 4 วิธี โดยพิจารณาจากความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ มีขั้นตอนของการวิจัยดังนี้

4.1 กำหนดขนาดกลุ่มตัวอย่าง

กำหนดขนาดกลุ่มตัวอย่างเท่ากัน ได้แก่ (20,20), (40,40), (60,60), (80,80) และ (100,100) ในกรณีที่ข้อมูลมีการแจกแจงปกติหลายตัวแปร และการแจกแจงลาปลาซหลายตัวแปร กำหนดค่าเฉลี่ย ค่าความแปรปรวน และค่าสัมประสิทธิ์สหสัมพันธ์ทางบวก ได้แก่ 0.1, 0.5, 0.9 และทางลบ ได้แก่ -0.1, -0.5 และ -0.9 ที่ระดับนัยสำคัญ 0.01 และ 0.05

4.2 กำหนดความแตกต่างของความแปรปรวน

ความแตกต่างของความแปรปรวนสำหรับการเปรียบเทียบกำลังการทดสอบโดยใช้ค่าพารามิเตอร์ไม่ศูนย์กลาง θ (วราฤทธิ์, 2549)

$$\text{คำนวณจาก } \theta = \sqrt{\frac{\sum_{i=1}^n (\sigma_i^2 - \bar{\sigma}^2)^2}{2\sigma_I^2}}$$

เมื่อ σ_i^2 เป็นค่าความแปรปรวนของประชากรที่มีค่าต่ำสุด; σ_i^2 เป็นค่าความแปรปรวนของประชากรที่ i โดย $i=1,2,...,n$; $\bar{\sigma}^2$ เป็นค่าความแปรปรวนเฉลี่ยของประชากรทั้ง 2 กลุ่ม

ตารางที่ 1 เกณฑ์ของค่าความแปรปรวนที่ใช้ในการศึกษา

ระดับความแตกต่างของความแปรปรวน	ความแปรปรวนของแต่ละประชากร	θ
มีความแตกต่างกันน้อย ($0 < \theta \leq 1.5$)	(4,5.6)	0.4
	(4,8.6)	1.15
มีความแตกต่างกันปานกลาง ($1.5 < \theta \leq 3.0$)	(4,11)	1.75
	(4,14)	2.5
มีความแตกต่างกันมาก ($\theta > 3.0$)	(4,18)	3.5
	(4,32)	7

4.3 รูปแบบที่ใช้การศึกษา

การแจกแจงและความแตกต่างของความแปรปรวนดังกล่าวสามารถกำหนดรูปแบบการศึกษาดังนี้

4.3.1 ศึกษาข้อมูลที่สุ่มมาจากการแจกแจงปกติหลายตัวแปรด้วยพารามิเตอร์ μ และ Σ

4.3.2 ศึกษาข้อมูลที่สุ่มการแจกแจงลาปลาซหลายตัวแปรด้วยพารามิเตอร์ μ และ Σ

โดยแบ่งเป็น 4 กรณีเท่ากันทั้งสองการแจกแจง คือ กรณีที่ 1 เมื่อค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน กรณีที่ 2 เมื่อค่าเฉลี่ยเท่ากัน แต่ความแปรปรวนไม่เท่ากัน กรณีที่ 3 เมื่อค่าเฉลี่ยไม่เท่ากัน แต่ความแปรปรวนเท่ากัน และกรณีที่ 4 เมื่อค่าเฉลี่ยไม่เท่ากันและความแปรปรวนไม่เท่ากัน

4.4 คำนวณความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 (α)

4.4.1 จำลองข้อมูลในแต่ละขนาดตัวอย่างสุ่มจากแต่ละประชากรที่มีการแจกแจงปกติหลายตัวแปรและการแจกแจงลาปลาซหลายตัวแปร ให้ข้อมูลมีค่าเฉลี่ยหรือค่ามัธยฐานเท่ากันทั้ง 2 กลุ่ม มีค่าพารามิเตอร์และค่าของสัมประสิทธิ์สหสัมพันธ์เป็นไปตามที่กำหนดไว้ โดยใช้โปรแกรมอาร์

4.4.2 คำนวณตัวสถิติทดสอบทั้ง 4 การทดสอบ ได้แก่ T, S, W และ P

4.4.3 นำค่าตัวสถิติทดสอบที่คำนวณได้เทียบกับค่าวิกฤตเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานว่าง ที่ระดับนัยสำคัญ 0.01 และ 0.05 และบันทึกจำนวนครั้งของการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นจริง ทำซ้ำจนครบ 1,000 ครั้ง

4.4.4 หาความน่าจะเป็นของ ความผิดพลาดแบบที่ 1 โดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่างดังนี้ ความน่าจะเป็นของความผิดพลาดแบบที่ 1 = จำนวนครั้งของการปฏิเสธ H_1 เมื่อ H_0 เป็นจริง ÷ 1,000

ถ้าความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของการทดสอบสำหรับแต่ละสถานการณ์ มีค่าอยู่ในช่วงที่ได้กำหนดไว้ในเกณฑ์ของการเปรียบเทียบวิธีการทดสอบ ได้แก่ เกณฑ์ของแบรดลีย์ จะถือว่าตัวสถิติทดสอบนั้น มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยนับตัวสถิติทดสอบทุกตัวว่าผ่านเกณฑ์ของแบรดลีย์ก็ครั้งตัวสถิติทดสอบที่ผ่านเกณฑ์ของแบรดลีย์มากที่สุดจะเป็นตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด มีรายละเอียดของเกณฑ์ของแบรดลีย์ดังนี้ เกณฑ์ของแบรดลีย์ ที่ระดับนัยสำคัญ 0.05 ถ้าค่าที่ได้มีค่าในช่วง (0.025-0.075) ที่ระดับนัยสำคัญ 0.01 ถ้าค่าที่ได้มีค่าในช่วง (0.005-0.015) จะสรุปได้ว่าตัวสถิติทดสอบนั้นควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

4.5 จำนวนกำลังการทดสอบ ($1 - \beta$)

4.5.1 จำลองข้อมูลในแต่ละขนาดตัวอย่างสุ่มจากแต่ละประชากรที่มีการแจกแจงปกติหลายตัวแปร และการแจกแจงลาปลาซหลายตัวแปร ให้ข้อมูลมีค่าเฉลี่ยหรือค่ามัธยฐานไม่เท่ากันทั้ง 2 กลุ่ม มีค่าพารามิเตอร์และค่าของสัมประสิทธิ์

สหสัมพันธ์เป็นไปตามที่กำหนดไว้โดยใช้โปรแกรมอาร์

4.5.2 จำนวนตัวสถิติทดสอบทั้ง 4 การทดสอบ ได้แก่ T, S, W และ P

4.5.3 นำค่าตัวสถิติทดสอบที่คำนวณได้เปรียบเทียบกับค่าวิกฤตเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานว่าง ที่ระดับนัยสำคัญ 0.01 และ 0.05 และบันทึกจำนวนครั้งของการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นเท็จ ทำซ้ำจนครบ 1,000 ครั้ง

4.5.4 หากำลังการทดสอบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 โดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง ดังนั้น กำลังการทดสอบ = จำนวนครั้งของการปฏิเสธ H_0 เมื่อ H_1 เป็นจริง $\div 1,000$ จะสรุปได้ว่าตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของแบรดลีย์ และมีกำลังการทดสอบสูงที่สุดจะเป็นตัวสถิติทดสอบที่ดีที่สุด

5. ผลการวิจัย

5.1 ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

ตารางที่ 2 จำนวนของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ผ่านเกณฑ์ของแบรดลีย์ เมื่อข้อมูลทั้ง 2 กลุ่ม มีการแจกแจงปกติหลายตัวแปรและมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9

ระดับนัยสำคัญ ตัวสถิติทดสอบ	0.01						0.05					
	ค่าสัมประสิทธิ์สหสัมพันธ์						ค่าสัมประสิทธิ์สหสัมพันธ์					
	0.1	0.5	0.9	-0.1	-0.5	-0.9	0.1	0.5	0.9	-0.1	-0.5	-0.9
T	47	46	50	46	46	47	50	50	50	50	50	50
S	31	35	36	36	35	34	44	45	46	49	45	48
W	46	49	48	48	47	45	55	50	50	50	50	50
P	47	47	50	46	46	47	50	50	50	50	50	50

ตัวเลขในตารางที่เป็นตัวหนา คือ จำนวนครั้งของตัวสถิติทดสอบนั้น ๆ ที่ผ่านเกณฑ์ของแบรดลีย์และมีจำนวนครั้งที่ผ่านเกณฑ์มากที่สุดที่ระดับนัยสำคัญและค่าสัมประสิทธิ์สหสัมพันธ์ที่กำหนด

ตารางที่ 3 จำนวนของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ผ่านเกณฑ์ของเบรตลีย์ เมื่อข้อมูลทั้ง 2 กลุ่ม มีการแจกแจงลาปลาซหลายตัวแปรและมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9

ระดับนัยสำคัญ ตัวสถิติทดสอบ	0.01						0.05					
	ค่าสัมประสิทธิ์สหสัมพันธ์						ค่าสัมประสิทธิ์สหสัมพันธ์					
	0.1	0.5	0.9	-0.1	-0.5	-0.9	0.1	0.5	0.9	-0.1	-0.5	-0.9
T	45	44	47	47	45	49	50	50	50	50	50	50
S	35	32	37	39	29	35	48	48	47	47	45	45
W	46	45	48	46	50	46	50	50	5	50	50	50
P	48	42	44	46	50	48	50	50	50	50	50	50

ตัวเลขในตารางที่เป็นตัวหนา คือ จำนวนครั้งของตัวสถิติทดสอบนั้น ๆ ที่ผ่านเกณฑ์ของเบรตลีย์และมีจำนวนครั้งที่ผ่านเกณฑ์มากที่สุดที่ระดับนัยสำคัญและค่าสัมประสิทธิ์สหสัมพันธ์ที่กำหนด

เมื่อพิจารณาตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุด 4 ลำดับ สำหรับ

ข้อมูลทั้ง 2 กลุ่ม มีการแจกแจงเหมือนกัน สามารถสรุปผลได้ดังตารางที่ 4

ตารางที่ 4 ตัวสถิติทดสอบที่มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด 4 ลำดับ

ระดับ นัยสำคัญ	การแจกแจง	ค่าสัมประสิทธิ์สหสัมพันธ์					
		0.1	0.5	0.9	-0.1	-0.5	-0.9
0.01	การแจกแจงปรกติ หลายตัวแปร	1. T, P 2. W 3. S	1. W 2. P 3. T 4. S	1. T, P 2. W 3. S	1. W 2. T, P 3. S	1. W 2. T, P 3. S	1. T, P 2. W 3. S
	การแจกแจง ลาปลาซ หลายตัวแปร	1. P 2. W 3. T 4. S	1. W 2. T 3. P 4. S	1. W 2. T 3. P 4. S	1. T 2. W, P 3. S	1. W, P 2. T 3. S	1. T 2. P 3. W 4. S
0.05	การแจกแจงปรกติ หลายตัวแปร	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S
	การแจกแจง ลาปลาซ หลายตัวแปร	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S	1. T, W, P 2. S

ตารางที่ 4 การแจกแจงปรกติหลายตัวแปร ที่ระดับนัยสำคัญ 0.01 และ 0.05 ค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกและลบ พบว่า T, W, และ P มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุดใกล้เคียงกัน ส่วนกรณีการแจกแจงลาปลาซหลายตัวแปร ที่ระดับนัยสำคัญ 0.01 และ 0.05 ค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกและลบ พบว่า W มีความ

สามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด

5.2 กำลังการทดสอบ

เมื่อพิจารณาตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 4 ลำดับ สำหรับข้อมูลทั้ง 2 กลุ่ม มีการแจกแจงเหมือนกันสามารถสรุปผลดังตารางที่ 7

ตารางที่ 5 จำนวนของกำลังการทดสอบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของแบรดลีย์ เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงปรกติหลายตัวแปรและมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9

ระดับนัยสำคัญ ตัวสถิติทดสอบ	0.01						0.05					
	ค่าสัมประสิทธิ์สหสัมพันธ์						ค่าสัมประสิทธิ์สหสัมพันธ์					
	0.1	0.5	0.9	-0.1	-0.5	-0.9	0.1	0.5	0.9	-0.1	-0.5	-0.9
T	39	36	46	30	30	29	34	37	47	33	34	30
S	0	7	25	0	0	0	1	10	34	1	0	0
W	9	19	37	5	1	4	11	28	43	11	3	1
P	25	37	47	32	31	25	43	43	48	35	35	39

ตัวเลขในตารางที่เป็นตัวหนา คือ ตัวสถิติทดสอบที่ให้กำลังการทดสอบสูงที่สุดและมีจำนวนครั้งมากที่สุดที่ระดับนัยสำคัญและค่าสัมประสิทธิ์สหสัมพันธ์ที่กำหนด

ตารางที่ 6 จำนวนของกำลังการทดสอบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของแบรดลีย์ เมื่อข้อมูลทั้งสองกลุ่มมีการแจกแจงลาปลาซหลายตัวแปรและมีค่าสัมประสิทธิ์สหสัมพันธ์เป็น 0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9

ระดับนัยสำคัญ ตัวสถิติทดสอบ	0.01						0.05					
	ค่าสัมประสิทธิ์สหสัมพันธ์						ค่าสัมประสิทธิ์สหสัมพันธ์					
	0.1	0.5	0.9	-0.1	-0.5	-0.9	0.1	0.5	0.9	-0.1	-0.5	-0.9
T	3	13	34	4	0	0	10	22	40	5	1	0
S	11	18	32	8	4	7	14	21	37	11	5	4
W	44	43	46	43	47	36	50	49	49	48	48	49
P	5	12	32	6	1	6	10	22	40	6	1	0

ตัวเลขในตารางที่เป็นตัวหนา คือ ตัวสถิติทดสอบที่ให้กำลังการทดสอบสูงที่สุดและมีจำนวนครั้งมากที่สุดที่ระดับนัยสำคัญและค่าสัมประสิทธิ์สหสัมพันธ์ที่กำหนด

ตารางที่ 7 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุด 4 ลำดับ

ระดับ นัยสำคัญ	การแจกแจง	ค่าสัมประสิทธิ์สหสัมพันธ์					
		0.1	0.5	0.9	-0.1	-0.5	-0.9
0.01	การแจกแจงปกติ หลายตัวแปร	1. T	1. P	1. P	1. P	1. P	1. T
		2. P	2. T	2. T	2. T	2. T	2. P
		3. W	3. W	3. W	3. W	3. W	3. W
		4. S	4. S	4. S	4. S	4. S	4. S
	การแจกแจงลา ปลาหลายตัวแปร	1. W	1. W	1. W	1. W	1. W	1. W
		2. S	2. S	2. T	2. S	2. S	2. S
		3. P	3. T	3. S, P	3. P	3. P	3. P
		4. T	4. P		4. T	4. T	4. T
0.05	การแจกแจงปกติ หลายตัวแปร	1. P	1. P	1. P	1. P	1. P	1. P
		2. T	2. T	2. T	2. T	2. T	2. T
		3. W	3. W	3. W	3. W	3. W	3. W
		4. S	4. S	4. S	4. S	4. S	4. S
	การแจกแจงลา ปลาหลายตัวแปร	1. W	1. W	1. W	1. W	1. W	1. W
		2. S	2. T, P	2. T, P	2. S	2. S	2. S
		3. T, P	3. S	3. S	3. P	3. T, P	3. T, P
					4. T		

ตารางที่ 7 การแจกแจงปกติหลายตัวแปร ที่ระดับนัยสำคัญ 0.01 และ 0.05 ค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกและลบ พบว่า P ให้กำลังการทดสอบสูงที่สุด ส่วนกรณีการแจกแจงลาปลาหลายตัวแปร ที่ระดับนัยสำคัญ 0.01 และ 0.05 ค่าสัมประสิทธิ์สหสัมพันธ์ทางบวกและลบ พบว่า W ให้กำลังการทดสอบสูงที่สุด

6. สรุปผลการวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

เมื่อข้อมูลทั้ง 2 กลุ่ม มีขนาดตัวอย่างเท่ากับ (20,20), (40,40), (60,60), (80,80) และ (100,100) จากการแจกแจงปกติหลายตัวแปรและการแจกแจงลาปลาหลายตัวแปรที่ระดับนัยสำคัญ 0.01 และ 0.05 ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ

0.1, 0.5, 0.9, -0.1, -0.5 และ -0.9 เมื่อขนาดตัวอย่างเพิ่มมากขึ้น ทำให้ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 เพิ่มมากขึ้น และมีกำลังการทดสอบสูงขึ้น พบว่า W มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด รองลงมา คือ T และ P นอกจากนี้ W มีกำลังการทดสอบสูงที่สุดเช่นกัน รองลงมา คือ P ซึ่งสอดคล้องกับงานวิจัยของกมลวรรณ และคณะ (2560) ที่กล่าวว่า W มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด และมีกำลังการทดสอบสูงที่สุดในทุกกรณี และยังให้ผลใกล้เคียงกับงานวิจัยของอานนท์ (2555) ที่กล่าวว่า T มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากที่สุด และมีกำลังการทดสอบสูงที่สุดในทุก

กรณี แต่งานวิจัยนี้พบว่า T มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และมีกำลังการทดสอบรองลงมาจาก W

6.2 ข้อเสนอแนะ

6.2.1 ควรศึกษาเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบอื่น ๆ ที่มีข้อกำหนดที่สามารถใช้เงื่อนไข 2 กลุ่ม สัมพันธ์กัน เช่น McNemar test

6.2.2 ควรใช้ขนาดตัวอย่างอื่น ๆ มาเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ ด้วยขนาดตัวอย่างอื่น ๆ ได้แก่ 35, 45 และ 55 เป็นต้น

6.2.3 ควรศึกษาในกรณีที่มีลักษณะของความสัมพันธ์ในทางบวกและทางลบที่มากขึ้น ได้แก่ 0.2, 0.4, 0.6, 0.8, -0.2, -0.4, -0.6, -0.8 เป็นต้น

7. รายการอ้างอิง

- กมลวรรณ ศรีเพชร, จิรายุ ไทยพิทักษ์, เจษฎา จงธนโชติ และสุภาพร พวงชะอุ่ม, 2560, การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบไม่อิงพารามิเตอร์ระหว่างค่ามัธยฐานของประชากร 2 กลุ่ม สัมพันธ์กัน, ปัญหาพิเศษระดับปริญญาตรี, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ.
- กัลยา บุญหล้า, น้ำเพชร ยอดแสน และสุริย์รัตน์ แก้วศรีเมือง, 2561, การเปรียบเทียบสถิติทดสอบเมื่อประชากร 2 กลุ่ม ไม่เป็นอิสระกัน, ว.วิทยาศาสตร์ลาดกระบัง 27(1): 78-87.
- กัลยา วานิชย์บัญชา, 2554, หลักสถิติ, บริษัท ธรรมสาร, กรุงเทพฯ.
- น้อมจิต กิตติโชติพาณิชย์, 2558, สถิติคณิตศาสตร์, เอกสารประกอบการสอน, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ.
- มนตรี สังข์ทอง, 2556, การศึกษาประสิทธิภาพของสถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ในการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม, KKU Sci. J. 41(1): 226-238.
- ณทัย ราตรี, 2554, สถิติเบื้องต้น : แนวคิดและทฤษฎี, บริษัท วี.พรินท์, กรุงเทพฯ.
- ลลิตา โชติชนากิจ, 2554, การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบที่และตัวสถิติทดสอบวิลคอกซันสำหรับประชากรสองกลุ่มไม่เป็นอิสระกัน, ปัญหาพิเศษระดับปริญญาตรี, มหาวิทยาลัยบูรพา, ชลบุรี.
- วราฤทธิ์ พานิชกิจโกศลกุล, 2549, ขนาดตัวอย่างที่น้อยที่สุดสำหรับการประมาณค่าแบบช่วงของผลต่างระหว่างค่าเฉลี่ยของสองประชากร, ว.วิจัยและพัฒนา มจร. 29(2): 173-184.
- สายชล สันสมบุรณ์ทอง, 2551, สถิติเบื้องต้น, บริษัท จามจุรีโปรดักส์, กรุงเทพฯ.
- สายชล สันสมบุรณ์ทอง, 2552, สถิติที่ไม่ใช้พารามิเตอร์, บริษัท จามจุรีโปรดักส์, กรุงเทพฯ.
- สายชล สันสมบุรณ์ทอง, 2558, การแจกแจงเชิงสถิติ, บริษัท จามจุรีโปรดักส์, กรุงเทพฯ.
- อานนท์ แก้วปลั่ง, 2555, การเปรียบเทียบประสิทธิภาพของสถิติทดสอบสำหรับการทดสอบค่ากลางของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระกันเมื่อข้อมูลอยู่ในมาตราอันดับ, ปัญหาพิเศษระดับปริญญาตรี, มหาวิทยาลัยบูรพา, ชลบุรี.
- อุมาพร จันทกร, 2542, สถิติที่ไม่ใช้พารามิเตอร์, บริษัท ฟิสิกส์เซ็นเตอร์, กรุงเทพฯ.
- Blair, R.C. and Higgins, J.J., 1985, Comparison of the power of the paired samples t test to that of Wilcoxon's signed-ranks test under various population shapes, Psychol. Bull. 97(1): 119-128.
- Butar, B.F., 2008, Permutation tests for comparing two populations, J. Math. Sci.

- Math. Edu. 3(2): 19-30.
- Bradley, J. V. , 1978, Robustness, Br. J. Math. Stat. Psychol. 31: 144-151.
- Proschan, M., Glimm, E. and Posch, M., 2014, Connection between permutation and t tests: Relevance to adaptive methods, Stat. Med. 33: 4734-4742.
- Singmann, H. , 2018, Multivariate Laplace Distribution, Available Source: <https://www.rdocumentation.org/packages/LaplacesDemon/versions/16.1.1/topics/dist.Multivariate.Laplace>, April 30, 2019.