การเปรียบเทียบสถิติทดสอบเมื่อประชากร 2 กลุ่มไม่เป็นอิสระกัน
Comparison of Test Statistics Two Dependent Populations

กัลยา บุญหล้า* น้ำเพชร ยอดแสน และ สุรีย์รัตน์ แก้วศรีเมือง
Kanlaya Boonlha* Nampech Yodsaen and Sureerat Keawsrimuang
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร อำเภอเมือง จังหวัดพิษณุโลก 65000
วันที่ส่ง : 30 สิงหาคม 2560 วันที่แก้ไข : 24 ตุลาคม 2560 วันที่ตอบรับ : 22 มิถุนายน 2561

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบสำหรับการทดสอบความแตกต่างระหว่างประชากร 2 กลุ่มที่ไม่อิสระกัน สถิติทดสอบที่ใช้ ได้แก่ การทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) และการทดสอบการเรียงสับเปลี่ยน (Permutation test) ทำการเปรียบเทียบสถิติทดสอบโดยการจำลองข้อมูลโดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำลังการทดสอบของสถิติทดสอบเมื่อประชากรมีการแจกแจงแบบปกติ และแจกแจงแบบเอกรูป กำหนดขนาดตัวอย่าง 5-50 ผลการศึกษาพบว่า การทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) และ การทดสอบการเรียงสับเปลี่ยน (Permutation test) มีกำลังการทดสอบสูง และเมื่อขนาดตัวอย่างใหญ่กำลังการทดสอบจะสูงขึ้นด้วย

คำสำคัญ: การทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน การทดสอบการเรียงสับเปลี่ยน ประชากร 2 กลุ่มที่ไม่อิสระกัน

Abstract

The study aimed to compare the efficient for testing of the two dependent populations. The tests in this study are the Wilcoxon-signed rank test and the Permutation test. The performance of these tests is compare by simulation study according to type I errors and powers of the test under the population assumed to Normal and Uniform distribution. The sample size 5 - 50 was use obtained. Result of the study showed that the power comparisons indicate that the Wilcoxon-signed rank test and the Permutation test have the high-power value of the test. Also, the Wilcoxon-signed rank test and the Permutation test can be preferred for large sample size.

Keywords: Wilcoxon-signed rank test, permutation test, two dependent populations

*ที่อยู่ติดต่อ. โทรศัพท์ : 0892695437 E-mail address: kanlayab@nu.ac.th

1. บทนำ

ในการทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างประชากรสองกลุ่ม กรณีตัวอย่างสุ่ม 2 กลุ่มที่ไม่อิสระกัน จะมีสถิติทดสอบที่นิยมใช้คือ สถิติทดสอบที่อิงพารามิเตอร์ (Parametric Statistic) ได้แก่ สถิติทดสอบที (Dependent t-test) โดยสถิติทดสอบดังกล่าวจะมีข้อตกลงเบื้องต้น คือ ตัวอย่างสุ่มสองกลุ่มไม่อิสระกันต้องมาจากประชากรที่มีการแจกแจงปกติ ข้อมูลอยู่ในมาตราแบบช่วงหรืออัตราส่วน แต่ในทางปฏิบัติบ่อยครั้งที่พบว่าข้อมูลที่นำมาใช้ในการวิเคราะห์นั้นไม่เป็นไปตามข้อตกลงเบื้องต้น คือ ข้อมูลที่ได้มามีลักษณะเบี่ยงเบนจากการแจกแจงปกติ ซึ่งจากผลกระทบดังกล่าวจึงเป็นที่น่าสนใจในการหาวิธีการทดสอบที่เหมาะสมกว่า เมื่อข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น นักสถิติหลายท่านจึงเลือกใช้สถิติทดสอบไม่อิงพารามิเตอร์ (Non-Parametric Statistic) สำหรับใช้ศึกษาเปรียบเทียบความแตกต่างระหว่างประชากร 2 กลุ่มไม่อิสระกัน โดยสถิติทดสอบไม่อิงพารามิเตอร์ ที่นิยมใช้ คือ สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) ซึ่งเป็นสถิติที่ใช้สำหรับเปรียบเทียบความแตกต่างของตัวอย่างสองกลุ่มที่ไม่อิสระกัน โดยจัดลำดับที่ของผลต่างระหว่างคู่สังเกตในแต่ละคู่อันดับที่ไม่คิดเครื่องหมาย จากนั้นคำนวณผลรวมของอันดับที่ ที่มีเครื่องหมายเดียวกันซึ่งเครื่องหมายจะพิจารณาจากผลต่างระหว่างค่าสังเกต นำเสนอโดย Wilcoxon ในปี ค.ศ.1945 และเป็นที่นิยมใช้ในปัจจุบัน สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) มีข้อตกลงเบื้องต้นเกี่ยวกับข้อมูล ได้แก่ กลุ่มตัวอย่างสองกลุ่มได้มาโดยอย่างสุ่ม การแจกแจงของประชากรเป็นแบบสมมาตรและมาตราวัดอันตรภาคเป็นอย่างน้อย

สำหรับการศึกษาเปรียบเทียบความแตกต่างระหว่างประชากร 2 กลุ่ม กรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นมีผู้ศึกษาไว้หลายท่าน ในปี ค.ศ. 2008 Butar [1] ได้ทำการทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่มที่ไม่อิสระกัน โดยสถิติทดสอบผลบวกลำดับที่ของวิลค็อกซัน (Wilcoxon rank sum test) และสถิติทดสอบการเรียงสับเปลี่ยน (Permutation test) พบว่า สถิติทดสอบการเรียงสับเปลี่ยนมีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และมีกำลังการทดสอบเมื่อข้อมูลอยู่ในมาตรอัตราส่วนและเมื่อข้อมูลมีความเบ้ สถิติทดสอบการเรียงสับเปลี่ยนก็ยังคงมีประสิทธิภาพดี ต่อมาในปี พ.ศ.2554 ลลิตา [2] ได้ศึกษาเปรียบเทียบสถิติทดสอบที (t-test) และสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) ในการทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่มไม่อิสระกัน พบว่า สถิติทั้ง 2 มีความสามารถในการควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 และมีกำลังการทดสอบสูง เมื่อค่าผลต่างมีการแจกแจงปกติ แต่เมื่อผลต่างมีการแจกแจงแบบเอกรูป สถิติทั้งสองมีกำลังการทดสอบค่อนข้างต่ำ เมื่อตัวอย่างมีขนาดเล็ก

ดังนั้นผู้วิจัยจึงได้สนใจศึกษาเปรียบเทียบสถิติทดสอบไม่อิงพารามิเตอร์สำหรับทดสอบความแตกต่างระหว่างประชากร 2 กลุ่มที่ไม่อิสระกันด้วยสถิติทดสอบ 2 วิธี คือ สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) และ สถิติทดสอบการเรียงสับเปลี่ยน (Permutation test) โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำลังการทดสอบ สำหรับการแจกแจงที่เลือกใช้ในการศึกษา คือ การแจกแจงปกติและการแจกแจงเอกรูป

2. วิธีการดำเนินการ

ผู้วิจัยทำการศึกษาเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อน ประเภทที่ 1 และกำลังการทดสอบของการทดสอบความแตกต่างกันระหว่างประชากร 2 กลุ่ม ที่ไม่อิสระกันด้วยสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) และสถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) โดยการสร้างข้อมูลประชากรให้มี การแจกแจงปกติ (Normal Distribution) กำหนดประชากรที่มีการแจกแจง $N(0,1)$ สำหรับควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำหนดประชากรที่มีการแจกแจง $N(1,1), N(2,1)$ และ $N(3,1)$ สำหรับการทดสอบกำลังการทดสอบ นอกจากนี้ผู้วิจัยกำหนดข้อมูลให้มีการแจกเอกรูป (Uniform Distribution) กำหนดประชากรให้มีการแจกแจง $U(-4,4)$ สำหรับควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และกำหนดประชากรที่มีการแจกแจง $U(-4,6), U(-4,8)$ และ $U(-4,10)$ สำหรับการทดสอบกำลังการทดสอบ สุ่มตัวอย่างข้อมูล 2 กลุ่ม ขนาด(n_1,n_2) เท่ากับ (5,5), (10,10), (20,20), (30,30), (40,40) และ (50,50) จากประชากรที่มีการแจกแจงปกติ (Normal Distribution) และการแจกแจงเอกรูป (Uniform Distribution) เพื่อนำมาคำนวณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่ากำลังการทดสอบ กำหนดระดับนัยสำคัญ (α) เท่ากับ 0.01, 0.05 และ 0.10 กำหนดจำนวนรอบของการสุ่มตัวอย่างและคำนวณค่าสถิติทดสอบ 10,000 ครั้ง ทำการจำลองข้อมูลตามสถานการณ์ต่างๆ ด้วยวิธีมอนติคาร์โล โดยใช้โปรแกรม R เวอร์ชั่น 3.3.2 ทำการเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 โดยใช้การทดสอบทวินาม (Binomial Test) [3] ซึ่งสถิติทดสอบ คือ

$$Z = \frac{\alpha_* - \alpha_0}{\sqrt{\frac{\alpha_0(1-\alpha_0)}{M}}}$$

โดยที่ α_* แทน ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง

α_0 แทน ระดับนัยสำคัญที่กำหนด

M แทน จำนวนครั้งที่ทดลอง

กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.01 สถิติทดสอบจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อ ค่า α_* อยู่ในช่วง [0.0074, 0.0126] และในทำนองเดียวกัน เมื่อกำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.05 และ 0.10 สถิติทดสอบจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อ สถิติทดสอบใดที่คำนวณค่า α_* อยู่ในช่วง [0.00457, 0.0543] และ [0.0951, 0.1049] ตามลำดับ และคำนวณค่ากำลังการทดสอบในแต่ละของสถิติทดสอบลำดับที่โดยเครื่องหมายของ

วิลค็อกซัน (Wilcoxon-signed rank test) และ สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) พิจารณากำลังการทดสอบสูงสุดจะถือว่าเป็นสถิติทดสอบที่มีประสิทธิภาพ

3. สถิติทดสอบ

3.1 สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test)

สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon-signed rank test) [4] มีข้อกำหนดดังนี้ กลุ่มตัวอย่างทั้งสองกลุ่มที่ได้มาโดยสุ่ม การแจกแจงของประชากรสมมาตร มาตราวัดข้อมูลเป็นอันตราภาคขึ้นไป การทดสอบมีวิธีการ คือ จัดอันดับความแตกต่าง (D_i) โดยไม่คำนึงถึงเครื่องหมายจากค่าน้อยที่สุดไปหาค่ามากที่สุด ในกรณีที่ผลต่างกรณีที่ผลต่างมีค่าเท่ากันหลายคู่ให้ใช้อันดับเฉลี่ย สำหรับข้อมูลที่มีค่าเท่ากันซึ่งผลต่างจะมีค่าเท่ากับ 0 ให้ตัดทิ้งไปไม่นำมาวิเคราะห์ มีผลทำให้ขนาดของกลุ่มตัวอย่างลดลง หลังการจัดอันดับเรียบร้อยแล้ว ให้นำเครื่องหมาย + และ – ที่อยู่หน้าผลต่างระหว่างข้อมูล (D_i) มากำกับไว้หน้าอันดับนั้น ๆ คำนวณหาค่า T^+ ซึ่งเท่ากับผลรวมของอันดับที่มีเครื่องหมายเป็นบวก ทดสอบความมีนัยสำคัญของค่า T^+ ตามขนาดของกลุ่มตัวอย่างโดยจะจำแนกตามขนาดตัวอย่าง ดังนี้

กรณีขนาดตัวอย่างมีขนาดเล็ก (มีขนาดน้อยกว่าเท่ากับ 20)

$$T = \min\left|\sum R_i^+, \sum R_i^-\right|$$

เมื่อ $\sum R_i^+$ และ $\sum R_i^-$ คือ ผลรวมของอันดับของ D_i ที่มีเครื่องหมายบวกและลบตามลำดับ

T คือ ค่าของผลรวมของอันดับที่มีค่าน้อยกว่า (ไม่คิดเครื่องหมาย) ระหว่างอันดับที่มีเครื่องหมายบวกและอันดับที่มีเครื่องหมายลบ

เกณฑ์ในการตัดสินค่าที่ได้จากการคำนวณจะปฏิเสธสมมติฐาน H_0 เมื่อค่า T ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าวิกฤต T ที่ได้จากการเปิดตาราง Wilcoxon Matched Pairs Sign-Rank Test

กรณีขนาดตัวอย่างมีขนาดใหญ่

การแจกแจงของ T^+ จะเข้าใกล้การแจกแจงปกติซึ่งมีค่าเฉลี่ย และความแปรปรวน ดังนี้

$$\mu_{T^+} = \frac{N(N+1)}{4} \quad \text{และ} \quad \sigma^2_{T^+} = \sqrt{\frac{N(N+1)(2N+1)}{24}}$$

สถิติทดสอบคือ

$$Z = \frac{T^+ - \mu_{T^+}}{\sqrt{\sigma^2_{T^+}}}$$

3.2 สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test)

สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) [5] ใช้ในการทดสอบความแตกต่างของประชากร 2 กลุ่มที่ไม่อิสระกัน ความแตกต่างของข้อมูลแต่ละคู่ (D_i) ทุกค่าเป็นอิสระต่อกัน การแจกแจงของ D_i มีลักษณะสมมาตร โดยไม่จำเป็นต้องมีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของประชากร ซึ่งในบางครั้งเรียกว่า การทดสอบแบบสุ่ม (Randomization test) ข้อมูลที่ทำการศึกษาต้องอยู่ในมาตร

อันตราภาคเป็นอย่างต่ำ การทดสอบทำได้โดยคำนวณค่าผลต่างของข้อมูลแต่ละคู่ พิจารณาเครื่องหมายของความแตกต่าง หาจำนวนรูปแบบทั้งหมดที่เป็นไปได้ของการเรียงสับเปลี่ยนรูปแบบเครื่องหมายดังกล่าว เกณฑ์ในการตัดสินใจกรณีการทดสอบสมมติฐานทางเดียวให้คำนวณจำนวนรูปแบบของการสลับเครื่องหมายของ D_i ถ้าน้อยกว่า $n=\alpha(2^n)$ จะทำการปฏิเสธสมมติฐาน สำหรับการทดสอบสมมติฐานสองทางพิจารณาจำนวนรูปแบบของการสลับเครื่องหมายของ D_i ถ้าน้อยกว่า $n=0.5\alpha(2^n)$ จะทำการปฏิเสธสมมติฐาน

4. ผลการศึกษา

การศึกษาเปรียบเทียบการทดสอบการเรียงสับเปลี่ยน (Permutation Test) และสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test) สำหรับประชากร 2 กลุ่มที่ไม่อิสระกัน ผลการศึกษาเป็นดังนี้ เมื่อผลต่างของประชากร 2 กลุ่มมีการแจกแจงแบบปกติ ความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 เมื่อกำหนดให้ค่าเฉลี่ยของผลต่างของประชากร 2 กลุ่มเท่ากับ 0 และค่าความแปรปรวนของผลต่างเท่ากับ 1 ที่ระดับนัยสำคัญ 0.01 0.05 และ 0.10 พบว่า สถิติทดสอบทั้ง 2 วิธีสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ยกเว้นกรณีขนาดตัวอย่างเท่ากับ 5 ดังแสดงในตารางที่ 1 ค่ากำลังการทดสอบ เมื่อกำหนดให้ค่าเฉลี่ยของผลต่างของประชากร 2 กลุ่มเท่ากับ 1, 2 และ 3 ค่าความแปรปรวนเท่ากับ 1 ที่ระดับนัยสำคัญ 0.01 0.05 และ 0.10 พบว่าสถิติทดสอบทั้ง 2 วิธี มีค่ากำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และที่ขนาดตัวอย่าง 5 - 30 สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) จะมีค่ากำลังการทดสอบสูงกว่าสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test) ดังแสดงในตารางที่ 2

เมื่อผลต่างของประชากร 2 กลุ่มมีการแจกแจงแบบเอกรูปความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 เมื่อกำหนดให้ค่าเฉลี่ยของผลต่างประชากร 2 กลุ่มเท่ากับ 0 ที่ระดับนัยสำคัญ 0.01 0.05 และ 0.10 พบว่า สถิติทดสอบทั้งสองวิธีสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ยกเว้นกรณีขนาดตัวอย่างเท่ากับ 5 ดังแสดงในตารางที่ 3 ค่ากำลังการทดสอบ เมื่อกำหนดให้ค่าเฉลี่ยของผลต่างประชากร 2 กลุ่มเท่ากับ 1, 2 และ 3 ที่ระดับนัยสำคัญ 0.01 0.05 และ 0.10 พบว่า สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) มีกำลังการทดสอบดีกว่าสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test) โดยมีค่ากำลังการทดสอบเพิ่มขึ้นเมื่อขนาดทดสอบเพิ่มขึ้น ดังแสดงในตารางที่ 4

ตารางที่ 1. ค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงปกติ N(0,1)

ระดับนัยสำคัญ	ขนาดตัวอย่าง	สถิติทดสอบ	
		W	P
0.01	5	0.0000	0.0618
	10	0.0094*	0.0107*
	20	0.0080*	0.0082*
	30	0.0095*	0.0103*
	40	0.0073*	0.0075*
	50	0.0077*	0.0094*
0.05	5	0.0000	0.0607
	10	0.0459*	0.0498*
	20	0.0438*	0.0462*
	30	0.0512*	0.0512*
	40	0.0475*	0.0466*
	50	0.0474*	0.0477*
0.1	5	0.0607	0.1253
	10	0.0831*	0.1016*
	20	0.0911*	0.0960*
	30	0.0949*	0.0982*
	40	0.0946*	0.0958*
	50	0.0969*	0.0980*

หมายเหตุ *สถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 ได้
W แทน สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test)
P แทน สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test)

ตารางที่ 2. ค่ากำลังการทดสอบเมื่อประชากรมีการแจกแจงแบบปกติ

ระดับนัยสำคัญ	ขนาดตัวอย่าง	Normal(1,1)		Normal(2,1)		Normal(3,1)	
		W	P	W	P	W	P
0.01	5	-	-	-	-	-	-
	10	0.4684	0.4958	0.9863	0.9894	1.0000	1.0000
	20	0.4457	0.5027	1.0000	1.0000	1.0000	1.0000
	30	0.9946	0.9957	1.0000	1.0000	1.0000	1.0000
	40	0.9997	0.9998	1.0000	1.0000	1.0000	1.0000
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.05	5	-	-	-	-	-	-
	10	0.7889	0.8072	0.9998	1.0000	1.0000	1.0000
	20	0.9863	0.9894	1.0000	1.0000	1.0000	1.0000
	30	0.9994	0.9996	1.0000	1.0000	1.0000	1.0000
	40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.1	5	-	-	-	-	-	-
	10	0.8639	0.8991	1.0000	1.0000	1.0000	1.0000
	20	0.9936	0.9947	1.0000	1.0000	1.0000	1.0000
	30	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

หมายเหตุ W แทน สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test)
P แทน สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test)

ตารางที่ 3. ค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงเอกรูป U(0,1)

ระดับนัยสำคัญ	ขนาดตัวอย่าง	สถิติทดสอบ	
		W	P
0.01	5	0.0000	0.0655
	10	0.0096*	0.0106*
	20	0.0093*	0.0089*
	30	0.0109*	0.0111*
	40	0.0105*	0.0097*
	50	0.0074*	0.0080*
0.05	5	0.0000	0.0655
	10	0.0539*	0.0548*
	20	0.0475*	0.0494*
	30	0.0480*	0.0490*
	40	0.0527*	0.0533*
	50	0.0466*	0.0476*
0.1	5	0.0655	0.1297
	10	0.0893*	0.1071*
	20	0.0950*	0.0987*
	30	0.0925*	0.0951*
	40	0.1007*	0.1047*
	50	0.0951*	0.0957*

หมายเหตุ W แทน สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test)
P แทน สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test)

ตารางที่ 4. ค่ากำลังการทดสอบเมื่อประชากรมีการแจกแจงแบบเอกรูป

ระดับนัยสำคัญ	ขนาดตัวอย่าง	Uniform(-4,6)		Uniform(-4,8)		Uniform(-4,10)	
		W	P	W	P	W	P
0.01	5	-	-	-	-	-	-
	10	0.0327	0.0363	0.0841	0.0931	0.1628	0.1767
	20	0.0971	0.1058	0.3341	0.3834	0.5636	0.6453
	30	0.1710	0.1890	0.5617	0.6424	0.8254	0.9036
	40	0.2681	0.2996	0.7350	0.8208	0.9377	0.9786
	50	0.3502	0.4079	0.8407	0.9181	0.9778	0.9954
0.05	5	-	-	-	-	-	-
	10	0.1422	0.1483	0.3090	0.3288	0.4660	0.5058
	20	0.2605	0.2885	0.6097	0.6760	0.8209	0.8867
	30	0.3975	0.4359	0.8045	0.8685	0.9492	0.9816
	40	0.5079	0.5588	0.9008	0.9493	0.9875	0.9970
	50	0.6111	0.6680	0.9592	0.9837	0.9975	0.9999
0.1	5	-	-	-	-	-	-
	10	0.2083	0.2474	0.4120	0.4854	0.5935	0.6831
	20	0.4006	0.4348	0.7317	0.7994	0.8965	0.9479
	30	0.5308	0.5796	0.8785	0.9320	0.9753	0.9928
	40	0.6336	0.6861	0.9457	0.9783	0.9944	0.9989
	50	0.7331	0.7789	0.9801	0.9929	0.9985	0.9999

หมายเหตุ W แทน สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test)
P แทน สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test)

5. สรุป

การศึกษาเปรียบเทียบการทดสอบการเรียงสับเปลี่ยน (Permutation Test) และสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test) สำหรับประชากร 2 กลุ่มที่ไม่อิสระกัน สรุปได้ว่า เมื่อผลต่างของประชากรทั้ง 2 กลุ่มมีการแจกแจงแบบปกติและการแจกแจงแบบเอกรูป ที่ระดับนัยสำคัญ 0.01 0.05 และ 0.10 ขนาดตัวอย่างเท่ากับ (10,10) , (20,20) , (30,30) , (40,40) , (50,50) สถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) และ สถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน(Wilcoxon Signed-Rank Test) มีความสามารถในการควบคุมความน่าจะเป็นของความ

คลาดเคลื่อนประเภทที่ 1 ได้ เมื่อพิจารณาค่ากำลังการทดสอบพบว่าเมื่อผลต่างของประชากรทั้ง 2 มีการแจกแจงแบบปกติ เมื่อขนาดตัวอย่างเพิ่มขึ้นค่ากำลังการทดสอบจะมีค่าเพิ่มขึ้น และในทำนองเดียวกันเมื่อมีระดับนัยสำคัญของการทดสอบเพิ่มขึ้นค่ากำลังการทดสอบก็จะมากขึ้นเช่นกัน เมื่อผลต่างของประชากรทั้ง 2 มีการแจกแจงแบบเอกรูป ค่าเฉลี่ยของผลต่างเท่ากับ 1, 2 และ 3 กำลังการทดสอบของสถิติทดสอบการเรียงสับเปลี่ยน (Permutation Test) มีค่ากำลังการทดสอบดีกว่าสถิติทดสอบลำดับที่โดยเครื่องหมายของวิลค็อกซัน (Wilcoxon Signed-Rank Test) และเมื่อมีขนาดตัวอย่างมากขึ้นค่ากำลังการทดสอบก็จะสูงขึ้นตามขนาดตัวอย่าง ในการทำการวิจัยนี้ผู้วิจัยได้ศึกษาเฉพาะกรณีที่ความแปรปรวนของประชากรที่ 2 กลุ่มเท่ากัน ผู้ที่สนใจศึกษาอาจศึกษาในกรณีที่ความแปรปรวนไม่เท่ากัน นอกจากนี้ผู้สนใจอาจเลือกศึกษาเปรียบเทียบสถิติทดสอบอื่นๆ เช่น สถิติ Baumgartner-Wei β-Schindler สถิติ Brunner and Munzel และ สถิติ Fisher-Pitman permutation รวมถึงศึกษากรณีข้อมูลมีการแจกแจงอื่น ๆ เช่น การแจกแจงแกมมา การแจกแจงไวบูลล์ เป็นต้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนจากทุนเพื่อการศึกษาค้นคว้าด้วยตนเอง สำหรับนิสิตระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร ประจำปีงบประมาณ 2560

เอกสารอ้างอิง (References)

[1] Butar, B.F. 2008. Permutation Tests for Comparing Two Populations. *Journal of Mathematical Sciences & Mathematics Education*, 3(2), 19-30.

[2] ลลิตา โชติธนากิจ. 2554. การเปรียบเทียบประสิทธิภาพของสถิติทดสอบทีและสถิติทดสอบวิลค็อกซันสำหรับประชากรสองไม่อิสระกัน. วิทยานิพนธ์. มหาลัยบูรพา. [Lalita Chotthanakit. 2011. Efficiency Comparison of t-Test and Wilcoxon Signed-Rank Test for Two Dependent Populations. Thesis. Burapha University. (in Thai)]

[3] อาภา วงศ์จินดา, อภิญญา หิรัญวงษ์ และบุญอ้อม โฉมที. 2558. การเปรียบเทียบสถิติทดสอบไม่อิงพารามิเตอร์สำหรับกรทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากรสองกลุ่มที่เป็นอิสระกันเมื่อขนาดตัวอย่างเล็ก. *วารสารวิทยาศาสตร์และเทคโนโลยี*, 4, 558-568. [Arpha Vongjinda, Apinya Hirunwong and Boonorm Chomtee. 2015. A Comparison of Nonparametric Statistic for Testing the Mean Difference between Two Independent Populations with Small Sample Sizes. *Journal of Science and Technology*, 4, 558-568. (in Thai)]

[4] สุวิมล ติรกานันท์. 2553. สถิตินันพาราเมตริก. พิมพ์ครั้งที่ 1, กรุงเทพมหานคร จุฬาลงกรณ์มหาวิทยาลัย. [Suwimol Tirakanant. 2010. Nonparametric Statistic. 1st, Bangkok Chulalongkorn University. (in Thai)]

[5] Good, P. 1994**.** Permutation Tests: A practical guide to resampling methods for testing hypothesis. Springer. New York.