

การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบ ความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม ที่อิสระกัน A Comparison of the Efficiency of the Test Statistics for Testing the Difference between the Two Central Values from Two Independent Populations

วรางคณา เรียนสุทธิ์*

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ
วิทยาเขตพัทลุง ตำบลบ้านพร้าว อำเภอป่าพะยอม จังหวัดพัทลุง 93210

Warangkhan Riansut*

Department of Mathematics and Statistics, Faculty of Science, Thaksin University,
Phatthalung Campus, Ban Phrao, Pa Phayom, Phatthalung 93210

บทคัดย่อ

วัตถุประสงค์ของการศึกษาค้นคว้าครั้งนี้ คือ เพื่อเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบทั้งหมด 6 วิธี ได้แก่ ตัวสถิติทดสอบ Z ตัวสถิติทดสอบสตีเวนสัน ตัวสถิติทดสอบเวลช์ ตัวสถิติทดสอบแมนน์-วิตนีย์ ตัวสถิติทดสอบโดยมัชฌิมา และตัวสถิติทดสอบแวนเดอวอลเด็น รวมถึงเพื่อเสนอแนะตัวสถิติทดสอบที่เหมาะสม จำแนกสถานการณ์โดยกำหนดประชากรที่มีทั้งการแจกแจงปกติ การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ และการแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ กำหนดอัตราส่วนค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 ทั้งหมด 3 ระดับ คือ 1, 0.9 และ 1.2 อัตราส่วนความแปรปรวนประชากรกลุ่มที่ 1 และกลุ่มที่ 2 ทั้งหมด 3 ระดับ คือ 1, 0.25 และ 4 ขนาดตัวอย่างจากประชากร 2 กลุ่ม 9 ขนาด คือ (10, 10), (30, 30), (100, 100), (15, 10), (45, 30), (150, 100), (10, 15), (30, 45) และ (100, 150) ระดับนัยสำคัญที่ศึกษา คือ 0.01 และ 0.05 จำลองข้อมูลด้วยการเขียนโปรแกรม SAS กำหนดจำนวนรอบของการทำซ้ำ คือ 1,000 รอบ ผลการศึกษาพบว่าตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดและมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบอื่นๆ แต่ตัวสถิติทดสอบนี้จำเป็นต้องทราบค่าความแปรปรวนประชากรทั้ง 2 กลุ่ม ดังนั้นจึงควรใช้ตัวสถิติทดสอบที่มีความเหมาะสมรองลงมา คือ ตัวสถิติทดสอบสตีเวนสันหรือเวลช์ โดยกรณีความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าเท่ากัน ตัวสถิติทดสอบสตีเวนสันจะมีความเหมาะสมมากกว่าเวลช์ แต่กรณีความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าไม่เท่ากัน ตัวสถิติทดสอบเวลช์จะมีความเหมาะสมมากกว่า สำหรับตัวสถิติทดสอบไม่อิงพารามิเตอร์ที่ศึกษา มีความเหมาะสมในบางกรณีแต่ยังไม่มีความมั่นใจที่ชัดเจนมากนัก เช่น โดยส่วนใหญ่ตัวสถิติทดสอบแมนน์-วิตนีย์ ตัวสถิติ

ทดสอบโดยมัธยฐาน และตัวสถิติทดสอบแวนเดอวอลเด็นมีความเหมาะสมกับตัวอย่างขนาดเล็กถึงขนาดกลาง และตัวสถิติทดสอบโดยมัธยฐานมีความเหมาะสมกับการทดสอบสมมุติฐานที่กำหนดระดับนัยสำคัญ 0.01 มากกว่าที่ระดับนัยสำคัญ 0.05 กรณีตัวอย่างทั้ง 2 กลุ่ม มีขนาดใหญ่และความแปรปรวนประชากรมีค่าเท่ากัน จะสามารถใช้ตัวสถิติทดสอบใดก็ได้ เนื่องจากมีประสิทธิภาพไม่แตกต่างกัน

คำสำคัญ : การทดสอบความแตกต่างของค่ากลาง; ความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1; กำลังการทดสอบ

Abstract

The objective of this study was to compare the performance of six test statistics: Z, Student's t, Welch's t, Mann-Whitney, Median, and Van der Waerden. Including the recommending the appropriate test statistic. Classification of the population according to normal distribution, left-skewed and leptokurtic kurtosis distribution, left-skewed and platykurtic kurtosis distribution, right-skewed and leptokurtic kurtosis distribution, and right-skewed and platykurtic kurtosis distribution. The ratios of mean group 1 per group 2 were 1, 0.9, and 1.2. The ratios of variance group 1 per group 2 were 1, 0.25, and 4. The sample sizes of two populations were nine sizes: (10, 10), (30, 30), (100, 100), (15, 10), (45, 30), (150, 100), (10, 15), (30, 45), and (100, 150). The significance levels that to study were 0.01 and 0.05. The data were simulated by using SAS programming with 1,000 replicates. Results of the research were Z test statistic could control the probability of type I error and had higher power of a test than other test statistics, but this test is necessary to know the variances of the two groups. Therefore, it was appropriate to apply the other test statistics, i.e., Student's t or Welch's t. If the variances of the two groups were equal, Student's t was a more appropriate test statistic than Welch's t. But, if the variances of the two groups were not equal, Welch's t was a more appropriate. Nonparametric test statistics for this study were appropriate in some cases but no clear, i.e., Mann-Whitney, Median, and Van der Waerden were appropriate for small to medium sample sizes. The median was appropriate to test the hypothesis that given the significance level of 0.01 over the 0.05 level of significance. When both sample sizes were large and the population variances were equal, we could use any tests since there is no difference in the performance.

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

1. บทนำ

การศึกษาวิจัยโดยทั่วไป นักวิจัยอาจไม่สามารถ

ศึกษาประชากรได้ทั้งหมด เพราะมีความยุ่งยากในการ
จัดเก็บข้อมูล สิ้นเปลืองค่าใช้จ่าย และเวลา ฯลฯ ดังนั้น

นักวิจัยจึงควรสุ่มตัวอย่างจากประชากร ตามวิธีการสุ่มตัวอย่างทางสถิติเพื่อให้ได้ตัวอย่างที่เป็นตัวแทนที่ดีของประชากรที่ต้องการศึกษา หลังจากนั้นจึงวิเคราะห์ข้อมูลที่ได้จากตัวอย่าง และใช้หลักการทางสถิติที่เรียกว่าการอนุมาน (inference) เพื่อสรุปคุณลักษณะของประชากร การอนุมานทางสถิติประกอบด้วย การประมาณค่า (estimation) และการทดสอบสมมติฐาน (hypothesis testing) โดยในการทดสอบสมมติฐานแบบอิงพารามิเตอร์ (parametric test) นั้น ตัวสถิติทดสอบที่ใช้จะขึ้นอยู่กับเงื่อนไขต่าง ๆ หรือข้อสมมติ (assumption) เช่น ลักษณะของข้อมูล มาตราวัดของข้อมูล การแจกแจงของข้อมูลจากตัวอย่างที่สุ่มได้ ภาวะความแปรปรวนเท่ากัน (homoscedasticity) เรื่องที่ต้องการศึกษา และจำนวนกลุ่มของข้อมูลหรือจำนวนประชากรที่ต้องการศึกษา เป็นต้น [1] กรณีตัวอย่างที่นำมาศึกษามีขนาดเล็ก ข้อมูลมักไม่เป็นไปตามข้อสมมติ จึงไม่ควรใช้การทดสอบอิงพารามิเตอร์ แต่ควรใช้การทดสอบอีกประเภทหนึ่งที่เรียกว่า การทดสอบไม่อิงพารามิเตอร์ (nonparametric test) ซึ่งบางครั้งอาจเรียกว่า การทดสอบที่ไม่ขึ้นกับการแจกแจง (distribution-free test) ดังนั้นจึงนับเป็นข้อดีของการทดสอบประเภทนี้ คือ ไม่จำเป็นต้องทราบการแจกแจงของประชากร อีกทั้งยังเป็นการทดสอบที่ไม่มีปัญหาในกรณีที่ข้อมูลมีความเบ้ และข้อมูลที่มีค่าผิดปกติ เนื่องจากการทดสอบไม่อิงพารามิเตอร์จะใช้เฉพาะความถี่หรืออันดับที่เท่านั้นในการวิเคราะห์ข้อมูล ไม่ได้ใช้ค่าของข้อมูลจริง [2]

ในงานวิจัยทางการแพทย์ การศึกษา การตลาด หรือทางจิตวิทยา มักมีการทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม ที่อิสระกัน โดยปราศจากการตรวจสอบข้อสมมติ และใช้การทดสอบสตีวเดนส์ (Student's t-test) ทั้ง ๆ ที่มีวิธีการทดสอบที่เหมาะสมมากมายถึง 6 วิธี ได้แก่ การ

ทดสอบ Z (Z-test) การทดสอบสตีวเดนส์ที่ การทดสอบเวลช์ (Welch's t-test) การทดสอบแมนน์-วิตนีย์ (Mann-Whitney test) การทดสอบโดยมัธยฐาน (median test) และการทดสอบแวนเดอวอลเด็น (Van der Waerden test) เป็นต้น ซึ่งผลสรุปที่ได้จากการทดสอบสตีวเดนส์ที่อาจขัดแย้งกับความเป็นจริง และระดับนัยสำคัญ (significance level) หรือความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 (probability of type I error) ไม่เป็นไปตามที่กำหนดไว้ รวมถึงอาจมีกำลังการทดสอบ (power of a test) ต่ำ อย่างไรก็ตาม คำกล่าวเหล่านี้ยังขาดการยืนยันในเชิงทฤษฎี ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีแนวคิดที่จะศึกษาการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม ที่อิสระกันทั้งหมด 6 วิธี จำแนกเป็นการทดสอบอิงพารามิเตอร์ 3 วิธี ได้แก่ การทดสอบ Z การทดสอบสตีวเดนส์ที่ และการทดสอบเวลช์ที่ และการทดสอบไม่อิงพารามิเตอร์อีก 3 วิธี ได้แก่ การทดสอบแมนน์-วิตนีย์ การทดสอบโดยมัธยฐาน และการทดสอบแวนเดอวอลเด็น การเปรียบเทียบประสิทธิภาพจะกระทำภายใต้สถานการณ์ต่าง ๆ ทั้งกรณีการแจกแจง ความแปรปรวน ขนาดตัวอย่าง และระดับนัยสำคัญที่แตกต่างกัน ผู้วิจัยคาดว่าผลการศึกษาที่ได้จะทำให้ทราบความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบทั้งหมด 6 วิธี ได้ทราบว่าตัวสถิติทดสอบวิธีใดมีความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่าง ๆ เพื่อเป็นแนวทางให้กับนักศึกษาคณาจารย์ และบุคคลทั่วไป สำหรับใช้ประโยชน์ในการตัดสินใจเลือกใช้ตัวสถิติทดสอบที่เหมาะสมกับการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2

กลุ่มที่อิสระกัน อีกทั้งผลการวิจัยที่ได้ยังสามารถเผยแพร่ในวารสารวิชาการระดับชาติ/นานาชาติ ซึ่งจะส่งผลต่อการวิจัย การผลิต การพัฒนาครู อาจารย์ และบุคลากรทางการศึกษาให้มีคุณภาพสูง สอดคล้องและรองรับกับแนวทางการพัฒนาประเทศ ต่อไป

2. วิธีการวิจัย

ขั้นตอนการดำเนินการวิจัย แสดงดังนี้

2.1 จำลองข้อมูลของประชากรที่ 1 และ 2 ด้วยโปรแกรม SAS รุ่น 9.00 กำหนดการแจกแจงความน่าจะเป็นของประชากรที่ศึกษา 5 การแจกแจง ได้แก่

- การแจกแจงปกติ โดยใช้คำสั่ง rand ('normal', 0, 1)

- การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ (ความเบ้ = -2, ความโด่ง = 12)

- การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = -0.4, ความโด่ง = 2.2)

- การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ (ความเบ้ = 2, ความโด่ง = 12)

- การแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = 0.4, ความโด่ง = 2.2)

สำหรับการจำลองข้อมูลที่มีลักษณะไม่สมมาตร (เบ้ซ้าย เบ้ขวา ความโด่งสูงกว่าปกติ และความโด่งต่ำกว่าปกติ) จะเริ่มต้นจากการจำลองข้อมูลที่มีการแจกแจงเอกรูปในช่วง (0, 1) โดยใช้คำสั่ง rand('uniform') แล้วแปลงข้อมูลตามวิธีการของ Ramberg และคณะ [3] ซึ่งสร้างตัวแปรสุ่ม X ที่ขึ้นอยู่กับความเบ้(skewness)และความโด่ง(kurtosis) ดังนี้

$$X = \lambda_1 + \frac{U^{\lambda_3} - (1-U)^{\lambda_4}}{\lambda_2}$$

เมื่อ U แทนตัวเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0, 1); λ_1 แทนพารามิเตอร์ตำแหน่ง (location

parameter); λ_2 แทนพารามิเตอร์มาตราส่วน (scale parameter); λ_3, λ_4 แทนพารามิเตอร์รูปร่าง (shape parameter)

ค่าของพารามิเตอร์ $\lambda_1, \lambda_2, \lambda_3$ และ λ_4 ได้จากตารางที่ 1 หน้า 205 หรือตารางที่ 4 หน้า 210-214 ในบทความของ Ramberg และคณะ [3] โดยค่าของพารามิเตอร์ทั้ง 4 ตัวจะให้ข้อมูลที่มีการแจกแจงเบ้ขวา สำหรับข้อมูลที่มีการแจกแจงเบ้ซ้ายให้สลับค่าของ λ_3 และ λ_4 และปรับเครื่องหมายของ λ_1 เป็นตรงข้าม ดังนั้นถ้าต้องการข้อมูลที่มีการแจกแจงตามที่กำหนด ต้องมีค่าของพารามิเตอร์ $\lambda_1, \lambda_2, \lambda_3$ และ λ_4 ดังนี้

- การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปกติ (ความเบ้ = -2, ความโด่ง = 12); $\lambda_1 = 0.579, \lambda_2 = -0.1423, \lambda_3 = -0.0995$ และ $\lambda_4 = -0.0273$

- การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = -0.4, ความโด่ง = 2.2); $\lambda_1 = 1.354, \lambda_2 = 0.2582, \lambda_3 = 0.5683$ และ $\lambda_4 = 0.0129$

- การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ (ความเบ้ = 2, ความโด่ง = 12); $\lambda_1 = -0.579, \lambda_2 = -0.1423, \lambda_3 = -0.0273$ และ $\lambda_4 = -0.0995$

- การแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปกติ (ความเบ้ = 0.4, ความโด่ง = 2.2); $\lambda_1 = -1.354, \lambda_2 = 0.2582, \lambda_3 = 0.0129$ และ $\lambda_4 = 0.5683$

กำหนดให้แต่ละการแจกแจงมีอัตราส่วนค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 (μ_1/μ_2) ทั้งหมด 3 ระดับ คือ 1, 0.9 และ 1.2 ดังนี้ (μ_1, μ_2) = (10, 10), (9, 10) และ (12, 10) กำหนดอัตราส่วนความแปรปรวนประชากรกลุ่มที่ 1 และกลุ่มที่ 2 (σ_1^2/σ_2^2) ทั้งหมด 3 ระดับ คือ 1, 0.25 และ 4 ดังนี้ (σ_1^2, σ_2^2) = (1, 1), (4, 16) และ (16, 4) และกำหนดขนาดตัวอย่าง 9 ขนาด แบ่งเป็นขนาดตัวอย่างทั้ง 2 กลุ่มเท่ากัน 3

ขนาด คือ ขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ได้แก่ (10, 10), (30, 30) และ (100, 100) ตามลำดับ และขนาดตัวอย่างไม่เท่ากันอีก 6 ขนาด แบ่งเป็นขนาดตัวอย่างที่ 1 ใหญ่กว่าตัวอย่างที่ 2 1.5 เท่า 3 ขนาด คือ ขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ได้แก่ (15, 10), (45, 30) และ (150, 100) ตามลำดับ และขนาดตัวอย่างที่ 2 ใหญ่กว่าตัวอย่างที่ 1 1.5 เท่า 3 ขนาด คือ ขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ได้แก่ (10, 15), (30, 45) และ (100, 150) ตามลำดับ จากนั้นแปลงข้อมูลแต่ละประชากรให้มีความเฉลี่ยและความแปรปรวนตามที่กำหนด โดยใช้สูตร

$$Y = \mu + \sigma X$$

เมื่อ X แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1; Y แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2

2.2 คำนวณค่าของตัวสถิติทดสอบและค่าพี (p-value) ของการทดสอบทั้งหมด 6 วิธี คือ การทดสอบ Z [4] การทดสอบสวิตเดนตส์ [4,5] การทดสอบเวลช์ [6] การทดสอบแมนน์-วิตนีย์ [1,7] การทดสอบโดยมัธยฐาน [8] และการทดสอบแวนเดอวอลเด็น [9,10] เพื่อเปรียบเทียบกับระดับนัยสำคัญที่กำหนด คือ 0.01 และ 0.05 โดยถ้าค่าพิน้อยกว่าหรือเท่ากับระดับนัยสำคัญ จะตัดสินใจปฏิเสธสมมุติฐานว่าง (ปฏิเสธ H_0) และถ้าค่าพินากกว่าระดับนัยสำคัญ จะตัดสินใจไม่ปฏิเสธสมมุติฐานว่าง

2.3 คำนวณความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 และกำลังการทดสอบ โดยการนับจำนวนครั้งที่ตัดสินใจปฏิเสธสมมุติฐานว่าง นำมาหารด้วยจำนวนรอบของการทำซ้ำ ซึ่งคือ 1,000 รอบ ดังนี้

- กรณีอัตราส่วนของค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 1 ซึ่งการศึกษาครั้งนี้กำหนดให้เป็น $(\mu_1, \mu_2) = (10, 10)$ จะได้ความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 นั่นคือ $\alpha =$

(จำนวนครั้งที่ตัดสินใจปฏิเสธ H_0 เมื่อ $\mu_1 = \mu_2$) ÷ 1,000 (สมการ 1)

นำค่า α ที่คำนวณได้ไปเปรียบเทียบกับเกณฑ์การทดสอบทวินาม [11] ดังนี้

- กรณีระดับนัยสำคัญเท่ากับ $\alpha_0 = 0.01$ จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ก็ต่อเมื่อค่า α ที่คำนวณได้จากสมการที่ (1) อยู่ในช่วง

$$\alpha_0 - Z_{0.01/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}} < \alpha < \alpha_0 + Z_{0.01/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}}$$

แทนค่า α_0 ที่กำหนด คือ 0.01, $Z_{0.01/2} = Z_{0.005} = 2.576$, n แทนจำนวนรอบของการทำซ้ำ = 1,000 จะได้

$$0.01 - 2.576 \sqrt{\frac{(0.01)(0.99)}{1,000}} < \alpha < 0.01 + 2.576 \sqrt{\frac{(0.01)(0.99)}{1,000}}$$

$$0.0019 < \alpha < 0.0181$$

- กรณีระดับนัยสำคัญเท่ากับ $\alpha_0 = 0.05$ จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ก็ต่อเมื่อค่า α ที่คำนวณได้จากสมการที่ (1) อยู่ในช่วง

$$\alpha_0 - Z_{0.05/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}} < \alpha < \alpha_0 + Z_{0.05/2} \sqrt{\frac{\alpha_0(1-\alpha_0)}{n}}$$

แทนค่า α_0 ที่กำหนด คือ 0.05, $Z_{0.05/2} = Z_{0.025} = 1.96$, n แทนจำนวนรอบของการทำซ้ำ = 1,000 จะได้

$$0.05 - 1.96 \sqrt{\frac{(0.05)(0.95)}{1,000}} < \alpha < 0.05 + 1.96 \sqrt{\frac{(0.05)(0.95)}{1,000}}$$

$$0.0365 < \alpha < 0.0635$$

- กรณีอัตราส่วนของค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 0.9 และ 1.2 ซึ่งการศึกษาครั้งนี้กำหนดให้เป็น $(\mu_1, \mu_2) = (9, 10)$ และ $(12, 10)$ จะได้กำลังการทดสอบ นั่นคือ $1 - \beta =$ (จำนวนครั้งที่ตัดสินใจปฏิเสธ H_0 เมื่อ $\mu_1 \neq \mu_2$) ÷ 1,000 (สมการ 2)

2.4 สรุปผลและเสนอแนะว่าตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่ม ที่อิสระกัน วิธีใดที่สามารถควบคุม

ความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงสุด ภายใต้สถานการณ์ต่าง ๆ ทั้งกรณีการแจกแจง ความแปรปรวนขนาดตัวอย่าง และระดับนัยสำคัญที่แตกต่างกัน

การศึกษาครั้งนี้จะแบ่งผลการศึกษาออกเป็น 2 ส่วน คือ

3.1 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของแต่ละตัวสถิติทดสอบ สรุปได้ดังตารางที่ 1-6

3. ผลการวิจัย

ตารางที่ 1 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของตัวสถิติทดสอบ Z

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
100	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
15	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
45	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
150	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
10	15	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30	45	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

× แทนไม่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 1 พบว่าตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทุกระดับอัตราส่วน

ความแปรปรวนประชากร ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา ยกเว้นเพียงกรณีขนาดตัวอย่างเป็น (10, 15) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงปกติ และระดับนัยสำคัญ 0.05

ตารางที่ 2 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของตัวสถิติทดสอบสตีเวนส์

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	ปกติ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ
10	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	×	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
100	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
15	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	×	×	×	×	×	×	×	×
		16, 4	✓	✓	✓	✓	✓	×	✓	×	×	×
45	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	✓	×	×	×	×	×	×
		16, 4	✓	✓	✓	✓	✓	×	×	×	×	×
150	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	×	✓	×	×	×	×	×	×
		16, 4	×	✓	✓	✓	✓	×	×	×	×	×
10	15	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	✓	×	✓	×
		16, 4	×	×	×	×	×	×	×	×	×	×
30	45	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	×	×	×	×
		16, 4	×	×	×	×	×	×	×	×	×	×
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	×	×	✓	×
		16, 4	×	×	✓	×	✓	×	×	×	×	×

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

×

จากตารางที่ 2 พบว่าตัวสถิติทดสอบ สติวเดนต์ที่สามารถควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบ ทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน ยกเว้นขนาดตัวอย่าง เป็น (10, 10) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ ระดับ นัยสำคัญ 0.01 และขนาดตัวอย่างเป็น (100, 100) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจง ปกติ ระดับนัยสำคัญ 0.05 แต่เมื่อขนาดตัวอย่างไม่

เท่ากัน ตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่ กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่า เท่ากัน และไม่สามารถควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อ ตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากร มากกว่า เช่น ขนาดตัวอย่างเป็น (15, 10) แต่มีความ แปรปรวนประชากรเป็น (4, 16)

ตารางที่ 3 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด ของตัวสถิติทดสอบเวลช์

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
100	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
15	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	×	✓	✓	×	✓	×	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
45	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
150	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
10	15	1, 1	✓	✓	✓	✓	✓	✓	×	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	×	✓	×	×	✓	×	✓	×	×
30	45	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	×	✓	✓	✓	✓	✓	✓
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	×	✓

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

× แทนไม่สามารถควบคุมความน่าจะเป็นที่จะเกิด ความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 3 พบว่าตัวสถิติทดสอบ Welch ที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน ยกเว้นขนาดตัวอย่างเป็น (10, 10) ความแปรปรวนประชากรเป็น (1, 1) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปกติ ระดับนัยสำคัญ 0.05 และขนาดตัวอย่างเป็น (100, 100) ความแปรปรวน

ประชากรเป็น (4, 16) การแจกแจงปกติ ระดับนัยสำคัญ 0.05 โดยตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดดีกว่าตัวสถิติทดสอบสตีเวนส์เมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า

ตารางที่ 4 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของตัวสถิติทดสอบแมนน์-วิตนีย์

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	✓	✓	×	×	×	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
30	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	✓	×	×	×	×
		16, 4	✓	×	×	×	×	×	×	×	×	×
100	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	×	×	×	×	×	×	×	×
		16, 4	✓	×	×	×	×	✓	×	×	×	×
15	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	✓	×	×	×	×	×	×	×
		16, 4	✓	✓	✓	✓	✓	×	×	✓	×	✓
45	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	×	×	×	×	×
		16, 4	✓	×	✓	×	✓	×	×	✓	×	✓
150	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
		4, 16	×	×	×	×	×	×	×	×	×	×
		16, 4	✓	×	×	×	×	×	×	×	×	×
10	15	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		16, 4	✓	×	×	×	×	✓	×	×	×	×
30	45	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	×	✓	✓	×	✓	×	×
		16, 4	×	×	×	×	×	×	×	×	×	×
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	×	×	×	×	×
		16, 4	✓	×	×	×	×	×	×	×	×	×

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

× แทนไม่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 4 พบว่าตัวสถิติทดสอบแมนน์-วิตนีย์สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากันเกือบทุกขนาดตัวอย่าง ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา ยกเว้นเพียงกรณีขนาดตัวอย่างเป็น (150, 100) การแจกแจงเบ้ขวาที่มีความโด่งสูง

กว่าปกติ และระดับนัยสำคัญ 0.05 สำหรับกรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน ตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเพียงบางกรณี เช่น กรณีตัวอย่างขนาดเล็ก โดยตัวอย่างที่มีขนาดเล็กกว่าต้องมีความแปรปรวนประชากรน้อยกว่า

ตารางที่ 5 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของตัวสถิติทดสอบโดยมีฐาน

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	1, 1	×	×	×	×	×	×	×	×	×	×
		4, 16	×	×	×	×	×	✓	✓	✓	×	×
		16, 4	×	×	×	×	×	×	✓	✓	×	×
30	30	1, 1	✓	✓	✓	✓	✓	×	×	×	✓	×
		4, 16	✓	✓	✓	✓	✓	×	×	×	×	×
		16, 4	✓	×	✓	✓	✓	×	×	×	×	×
100	100	1, 1	✓	✓	✓	✓	✓	×	✓	×	✓	×
		4, 16	✓	×	✓	×	✓	×	×	×	×	×
		16, 4	✓	×	×	×	✓	×	×	×	×	×
15	10	1, 1	✓	✓	✓	×	✓	×	✓	×	×	✓
		4, 16	✓	✓	✓	✓	✓	✓	×	✓	×	✓
		16, 4	✓	✓	✓	✓	✓	✓	×	×	✓	✓
45	30	1, 1	✓	✓	✓	✓	✓	✓	×	×	✓	×
		4, 16	✓	×	✓	×	✓	×	×	×	×	×
		16, 4	✓	×	✓	×	✓	×	×	×	×	×
150	100	1, 1	×	✓	✓	✓	✓	✓	✓	✓	✓	×
		4, 16	×	×	×	×	×	×	×	×	×	×
		16, 4	✓	×	×	×	×	✓	×	×	×	×
10	15	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	✓	✓	✓	×	×	✓	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
30	45	1, 1	✓	✓	✓	✓	✓	✓	×	✓	✓	✓
		4, 16	✓	×	×	✓	✓	×	×	×	×	×
		16, 4	✓	×	✓	×	×	×	×	×	×	×
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	✓	×	×	×	×
		16, 4	×	×	×	×	×	✓	×	×	×	×

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

× แทนไม่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 5 พบว่าตัวสถิติทดสอบโดยมีฐานสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดน้อยกว่าตัวสถิติทดสอบอื่น ๆ โดยตัวสถิติทดสอบนี้สามารถ

ควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดที่ระดับนัยสำคัญ 0.01 มากกว่าที่ระดับนัยสำคัญ 0.05

ตารางที่ 6 ความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของตัวสถิติทดสอบแวนเดอวอลเด็น

n_1	n_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
			ปกติ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	ปกติ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ	เบี่ยงโด่งสูง	เบี่ยงโด่งต่ำ
10	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	×	✓	✓	×	×	×	✓
		16, 4	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
30	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	×	✓	✓	×	×	×	✓
		16, 4	✓	×	✓	×	✓	✓	×	✓	×	✓
100	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	✓	×	×	×	×
		16, 4	✓	×	✓	×	×	✓	×	×	×	×
15	10	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	✓	×	✓	×	×	×	×	×
		16, 4	✓	✓	✓	✓	×	×	✓	✓	✓	×
45	30	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	×	✓	×	×	×	×
		16, 4	✓	×	✓	×	✓	×	×	✓	×	✓
150	100	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	×	×	×	×	×	✓	×	×	×	×
		16, 4	✓	×	✓	×	×	×	×	×	×	×
10	15	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	✓	✓	✓	✓	×	✓	×	✓	×
		16, 4	✓	×	✓	×	×	✓	×	×	×	×
30	45	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	✓	×	✓	×	×	×	×	✓
		16, 4	×	×	×	×	×	✓	×	×	×	×
100	150	1, 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		4, 16	✓	×	×	×	✓	×	×	✓	×	×
		16, 4	✓	×	×	×	×	✓	×	×	×	×

✓ แทนสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

× แทนไม่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

จากตารางที่ 6 พบว่าตัวสถิติทดสอบ แวนเดอวอลเด็นสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน ทุกขนาดตัวอย่าง ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา สำหรับกรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน ตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์

ที่กำหนดเพียงบางกรณี โดยส่วนใหญ่จะควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อข้อมูลมีการแจกแจงปกติ

3.2 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด สรุปได้ดังตารางที่ 7

ตารางที่ 7 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	9, 10	1, 1	Z (0.392)	MW (0.383)	Z (0.382)	MW (0.414)	Z (0.357)	Z (0.640)	VW (0.685)	Z (0.604)	MW (0.703)	Z (0.591)
			4, 16	St (0.038)	St (0.078)	St (0.043)	Z (0.032)	Z (0.032)	St (0.116)	St (0.190)	St (0.125)	Z (0.094)	Z (0.122)
			16, 4	St (0.030)	Z (0.036)	Z (0.030)	MW (0.081)	St (0.042)	St (0.106)	Z (0.103)	Z (0.102)	St (0.157)	St (0.128)
		12, 10	1, 1	Z (0.974)	Z (0.964)	Z (0.967)	Z (0.976)	Z (0.971)	Z (0.994)	Z (0.990)	Z (0.990)	Z (0.995)	Z (0.993)
			4, 16	Z (0.128)	Z (0.131)	Z (0.132)	Wt (0.201)	St (0.118)	Z (0.285)	Z (0.285)	Z (0.292)	St (0.383)	Z (0.293)
			16, 4	Z (0.137)	MW (0.212)	Z (0.118)	Z (0.126)	Z (0.118)	Z (0.293)	MW (0.470)	Z (0.302)	Z (0.289)	Z (0.294)
		9, 10	1, 1	Z (0.920)	MW (0.957)	Z (0.924)	MW (0.970)	Z (0.900)	Z (0.978)	VW (0.991)	Z (0.972)	MW (0.995)	VW (0.967)
			4, 16	St (0.100)	Med (0.217)	St (0.107)	Z (0.091)	Z (0.086)	MW (0.250)	St (0.331)	St (0.234)	Z (0.220)	Z (0.227)
			16, 4	MW (0.080)	Z (0.103)	Z (0.094)	Med (0.167)	St (0.087)	St (0.237)	Z (0.224)	Z (0.238)	St (0.288)	VW (0.242)
30	30	12, 10	1, 1	All (1.000)	All (1.000)	*	All (1.000)	*	*	*	*	All (1.000)	*
			4, 16	Z (0.459)	Z (0.428)	Z (0.457)	Med (0.551)	Z (0.470)	Z (0.685)	St (0.711)	Z (0.695)	Z (0.688)	Z (0.687)
			16, 4	Z (0.420)	St (0.483)	Z (0.470)	St (0.420)	Z (0.430)	Z (0.674)	Z (0.699)	Z (0.692)	St (0.728)	Z (0.685)
		9, 10	1, 1	Z (0.920)	MW (0.957)	Z (0.924)	MW (0.970)	Z (0.900)	Z (0.978)	VW (0.991)	Z (0.972)	MW (0.995)	VW (0.967)
			4, 16	St (0.100)	Med (0.217)	St (0.107)	Z (0.091)	Z (0.086)	MW (0.250)	St (0.331)	St (0.234)	Z (0.220)	Z (0.227)
			16, 4	MW (0.080)	Z (0.103)	Z (0.094)	Med (0.167)	St (0.087)	St (0.237)	Z (0.224)	Z (0.238)	St (0.288)	VW (0.242)

ตารางที่ 7 (ต่อ)

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
100	100	9, 10	1, 1	*	*	*	All	*	*	All	*	All	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
			4, 16	Z	St	Z	Z	Z	Z	Z	Z	St, Wt	Z
				(0.403)	(0.396)	(0.381)	(0.378)	(0.359)	(0.614)	(0.621)	(0.599)	(0.623)	(0.598)
		12, 10	16, 4	Z	Z	Z	Wt	Z	St, Wt	St	Z	Z	Z
				(0.378)	(0.347)	(0.331)	(0.399)	(0.383)	(0.628)	(0.620)	(0.562)	(0.621)	(0.611)
			1, 1	All	All	All	All	All	*	All	*	All	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
			4, 16	Z	St, Wt	St, Wt	Z	Z	Z	St, Wt	St, Wt	Z	Z
				(0.967)	(0.984)	(0.979)	(0.962)	(0.974)	(0.991)	(0.998)	(1.000)	(0.988)	(0.994)
15	10	9, 10	16, 4	Z	Z	Z	St	Z	Z, St, Wt	Z	Z	St, Wt	St, Wt
				(0.979)	(0.970)	(0.964)	(0.979)	(0.980)	(0.994)	(0.989)	(0.994)	(0.998)	(0.994)
			1, 1	All	All	All	All	All	*	All	*	All	*
		12, 10	4, 16	Z	St, Wt	St, Wt	Z	Z	Z	St, Wt	St, Wt	Z	Z
				(0.967)	(0.984)	(0.979)	(0.962)	(0.974)	(0.991)	(0.998)	(1.000)	(0.988)	(0.994)
			16, 4	Z	Z	Z	St	Z	Z, St, Wt	Z	Z	St, Wt	St, Wt
				(0.979)	(0.970)	(0.964)	(0.979)	(0.980)	(0.994)	(0.989)	(0.994)	(0.998)	(0.994)
		9, 10	1, 1	Z	MW	Z	MW	Z	Z	VW	Z	MW	Z
				(0.455)	(0.514)	(0.436)	(0.509)	(0.434)	(0.706)	(0.809)	(0.681)	(0.781)	(0.655)
			4, 16	Z	Med	MW	Z	Z	Z	Z	Wt	Z	Z
45	30	9, 10		(0.026)	(0.032)	(0.046)	(0.052)	(0.041)	(0.104)	(0.101)	(0.126)	(0.107)	(0.124)
			16, 4	Wt	Z	Z	MW	Wt	Z	Z	Z	Wt	Wt
				(0.034)	(0.040)	(0.047)	(0.100)	(0.047)	(0.114)	(0.109)	(0.138)	(0.185)	(0.135)
		12, 10	1, 1	Z	Z	Z	Z	Z	Z	Med	Z	Z	Z, St
				(0.989)	(0.990)	(0.992)	(0.988)	(0.990)	(0.997)	(0.998)	(0.999)	(0.998)	(1.000)
			4, 16	Z	Z	Z	Z	Z	Z	Z	Z	Z	Z
				(0.134)	(0.134)	(0.121)	(0.127)	(0.108)	(0.307)	(0.285)	(0.295)	(0.340)	(0.292)
		9, 10	16, 4	Z	Med	Z	Z	Z	Z	VW	Z	Z	Z
				(0.160)	(0.341)	(0.189)	(0.176)	(0.162)	(0.367)	(0.493)	(0.396)	(0.352)	(0.357)
			1, 1	Z	VW	Z	MW	Z	Z	MW, VW	Z, VW	MW, VW	VW
45	30	9, 10		(0.957)	(0.988)	(0.967)	(0.991)	(0.958)	(0.992)	(0.997)	(0.992)	(1.000)	(0.996)
			4, 16	St	Wt	Z	St	Z	Z	Wt	Z	Z	Z
				(0.153)	(0.167)	(0.097)	(0.112)	(0.091)	(0.238)	(0.299)	(0.257)	(0.231)	(0.238)
		12, 10	16, 4	Z	Z	Z	Wt	MW	Z	Z	Z	Wt	MW
				(0.106)	(0.119)	(0.135)	(0.168)	(0.129)	(0.285)	(0.287)	(0.296)	(0.337)	(0.330)
			1, 1	All	All	All	*	*	All	*	*	*	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
		9, 10	4, 16	St	Z	Z	Wt	Z	Z	Wt	Z	Z	Z
				(0.594)	(0.474)	(0.472)	(0.515)	(0.494)	(0.710)	(0.753)	(0.721)	(0.738)	(0.716)
			16, 4	Z	Z	Z	Wt	Z	Z	Z	Z	Wt	Z
45	30	9, 10		(0.627)	(0.638)	(0.622)	(0.618)	(0.618)	(0.827)	(0.803)	(0.830)	(0.827)	(0.809)
			1, 1	All	All	All	*	*	All	*	*	*	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
		12, 10	4, 16	St	Z	Z	Wt	Z	Z	Wt	Z	Z	Z
				(0.594)	(0.474)	(0.472)	(0.515)	(0.494)	(0.710)	(0.753)	(0.721)	(0.738)	(0.716)
			16, 4	Z	Z	Z	Wt	Z	Z	Z	Z	Wt	Z
				(0.627)	(0.638)	(0.622)	(0.618)	(0.618)	(0.827)	(0.803)	(0.830)	(0.827)	(0.809)
		9, 10	1, 1	All	All	All	*	*	All	*	*	*	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
			4, 16	St	Z	Z	Wt	Z	Z	Wt	Z	Z	Z
45	30	9, 10		(0.594)	(0.474)	(0.472)	(0.515)	(0.494)	(0.710)	(0.753)	(0.721)	(0.738)	(0.716)
			16, 4	Z	Z	Z	Wt	Z	Z	Z	Z	Wt	Z
				(0.627)	(0.638)	(0.622)	(0.618)	(0.618)	(0.827)	(0.803)	(0.830)	(0.827)	(0.809)
		12, 10	1, 1	All	All	All	*	*	All	*	*	*	*
				(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
			4, 16	St	Z	Z	Wt	Z	Z	Wt	Z	Z	Z
				(0.594)	(0.474)	(0.472)	(0.515)	(0.494)	(0.710)	(0.753)	(0.721)	(0.738)	(0.716)
		9, 10	16, 4	Z	Z	Z	Wt	Z	Z	Z	Z	Wt	Z
				(0.627)	(0.638)	(0.622)	(0.618)	(0.618)	(0.827)	(0.803)	(0.830)	(0.827)	(0.809)
			1, 1	All	All	All	*	*	All	*	*	*	*

ตารางที่ 7 (ต่อ)

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปฏิบัติ	เข้าข้างสูง	เข้าข้างต่ำ	เข้าข้างสูง	เข้าข้างต่ำ	ปฏิบัติ	เข้าข้างสูง	เข้าข้างต่ำ	เข้าข้างสูง	เข้าข้างต่ำ
150	100	9, 10	1, 1	*	All	*	All	*	All	All	*	*	*
			4, 16	Z	Wt	Z	St	Z	Z	Z	Z	Wt	Z
			16, 4	Z	Wt	Z	Z	Z	Z	Wt	Z	Z	Z
		12, 10	1, 1	*	All	All	All	All	All	All	All	*	*
			4, 16	Z	Wt	Wt	Z	Z	Z, Wt	Wt	Wt	Z	Z
			16, 4	Z	Z	Z, Wt	Z, St, Wt	Z	Z, Wt	Z	Z, Wt	Z	Z, Wt
		9, 10	1, 1	Z	MW	Z	MW	Z	Z	MW	Z	VW	Z
			4, 16	Z	MW	Z	Z	Z	Med	MW	Z	Z	Z
			16, 4	MW, VW	Z	Z	Med	Z	MW	Z	Z	Z	Med
10	15	9, 10	1, 1	Z	Z	Z	Z	Z	Z, St, Wt	Z	Z, St	Z, MW	Z
			4, 16	Z	Z	Z	MW	Z	Wt	Wt	Z	MW	Z
			16, 4	Z	Med	Z	Z	Z	Med	Z	Z	Z	Z
		12, 10	1, 1	Z	Z	Z	Z	Z	Z	Z	Z	Z	Z
			4, 16	Z	Z	Z	MW	Z	Wt	Wt	Z	MW	Z
			16, 4	Z	Med	Z	Z	Z	Med	Z	Z	Z	Z
		9, 10	1, 1	Z	MW	Z	VW	Z	Z	MW, VW	Wt	VW	Z
			4, 16	Z	Wt	MW	Z	Z	Z	Wt	MW	Z	Z
			16, 4	Z	Z	Z	Z	Wt	VW	Z	Z	Wt	Wt
30	45	9, 10	1, 1	All	*	*	All	All	All	*	All	All	All
			4, 16	Z	Wt	Z	Med	Z	Z	Wt	Z, Wt	Z	Z
			16, 4	Z	Wt	Z	Z	Z	Z	Z	Z	Wt	Z
		12, 10	1, 1	All	*	*	All	All	All	*	All	All	All
			4, 16	Z	Wt	Z	Med	Z	Z	Wt	Z, Wt	Z	Z
			16, 4	Z	Wt	Z	Z	Z	Z	Z	Z	Wt	Z
		9, 10	1, 1	Z	MW	Z	VW	Z	Z	MW, VW	Wt	VW	Z
			4, 16	Z	Wt	MW	Z	Z	Z	Wt	MW	Z	Z
			16, 4	Z	Z	Z	Z	Wt	VW	Z	Z	Wt	Wt

ตารางที่ 7 (ต่อ)

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
100	150	9, 10	1, 1	All (1.000)	All (1.000)	*	*	*	All (1.000)	All (1.000)	*	All (1.000)	*
			4, 16	Z (0.531)	Wt (0.536)	Z (0.498)	Wt (0.514)	Z (0.509)	Z (0.751)	Z (0.754)	Z (0.760)	Wt (0.741)	Wt (0.765)
			16, 4	MW (0.448)	Z (0.391)	St (0.488)	Wt (0.426)	St (0.501)	Z (0.661)	Wt (0.643)	Z (0.638)	Z (0.620)	Z (0.658)
		12, 10	1, 1	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)
			4, 16	Z, Wt (0.996)	Z (0.996)	Z, Wt (0.997)	Z (0.997)	Z, Wt (0.995)	Z, Wt (1.000)	Z, Wt (0.999)	Wt (1.000)	Z (1.000)	Z, Wt (1.000)
			16, 4	Z (0.984)	Z (0.975)	St (0.989)	Wt (0.992)	St (0.991)	Z, Wt (0.996)	Z (0.990)	Z, Wt (0.996)	Z (0.998)	Wt (0.995)

Z แทนตัวสถิติทดสอบ Z; St แทนตัวสถิติทดสอบสตีเวนส์; Wt แทนตัวสถิติทดสอบเวลช์; MW แทนตัวสถิติทดสอบแมนน์-วิตนีย์; Med แทนตัวสถิติทดสอบโดยมัธยฐาน; VW แทนตัวสถิติทดสอบแวนเดอวอลเด็น; All แทนทุกตัวสถิติทดสอบ; * แทนเกือบทุกตัวสถิติทดสอบ; ค่าในวงเล็บแทนกำลังการทดสอบ

จากตารางที่ 7 พบว่าผลการศึกษเกี่ยวกับกำลังการทดสอบสามารถสรุปได้เป็นข้อ ๆ ดังนี้

3.2.1 กรณีขนาดตัวอย่างเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย

3.2.2 กรณีค่าเฉลี่ยประชากรมีค่าแตกต่างกันมาก $(\mu_1, \mu_2) = (12, 10)$ กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีค่าเฉลี่ยประชากรที่มีค่าแตกต่างกันน้อย $(\mu_1, \mu_2) = (9, 10)$

3.2.3 กรณีความแปรปรวนประชากรมีค่าเท่ากัน $(\sigma_1^2, \sigma_2^2) = (1, 1)$ กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีความแปรปรวนประชากรที่มีค่าไม่เท่ากัน $(\sigma_1^2, \sigma_2^2) = (4, 16)$ และ $(16, 4)$

3.2.4 กรณีระดับนัยสำคัญเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย

3.2.5 โดยส่วนใหญ่ตัวสถิติทดสอบ Z มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์

ที่กำหนด

3.2.6 กรณีตัวอย่างขนาดกลาง $(n_1, n_2) = (30, 30), (30, 45)$ และ $(45, 30)$ ที่มีค่าเฉลี่ยประชากรแตกต่างกันมาก $(\mu_1, \mu_2) = (12, 10)$ แต่มีความแปรปรวนประชากรเท่ากัน $(\sigma_1^2, \sigma_2^2) = (1, 1)$ และกรณีตัวอย่างขนาดใหญ่ $(n_1, n_2) = (100, 100), (150, 100)$ และ $(100, 150)$ ที่มีความแปรปรวนประชากรเท่ากัน $(\sigma_1^2, \sigma_2^2) = (1, 1)$ พบว่า เกือบทุกตัวสถิติทดสอบมีกำลังการทดสอบสูงพอๆ กัน หรือมีประสิทธิภาพไม่แตกต่างกัน

4. สรุปและวิจารณ์ผล

จากผลการศึกษเกี่ยวกับความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของแต่ละตัวสถิติทดสอบ พบว่าตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

เกือบทุกกรณี ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทุกระดับอัตราส่วนความแปรปรวนประชากร ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา ตัวสถิติทดสอบสตีวเดนต์ที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน แต่เมื่อขนาดตัวอย่างไม่เท่ากัน จะควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้เมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน และไม่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้เมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า ตัวสถิติทดสอบเวลช์ที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน และควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ดีกว่าตัวสถิติทดสอบสตีวเดนต์ที่เมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า ตัวสถิติทดสอบแมนน์-วิตนีย์สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน เกือบทุกขนาดตัวอย่าง ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา แต่กรณีความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าไม่เท่ากัน จะควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้เพียงบางกรณี เช่น กรณีตัวอย่างขนาดเล็ก โดยตัวอย่างที่มีขนาดเล็กกว่าต้องมีความแปรปรวนประชากรน้อยกว่า ตัวสถิติทดสอบโดยมีฐานสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดน้อยกว่าตัวสถิติทดสอบอื่นๆ และควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ที่ระดับนัยสำคัญ 0.01 มากกว่าที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบแวนเดอวอลเด็นสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาด

แบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าเท่ากัน ทุกขนาดตัวอย่าง ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา แต่กรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน จะควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้เพียงบางกรณี โดยส่วนใหญ่จะควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้เมื่อข้อมูลมีการแจกแจงปกติ และจากผลการศึกษาเกี่ยวกับกำลังการทดสอบของแต่ละตัวสถิติทดสอบ พบว่า กรณีขนาดตัวอย่างเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย กรณีค่าเฉลี่ยประชากรมีค่าแตกต่างกันมาก กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีค่าเฉลี่ยประชากรที่มีค่าแตกต่างกันน้อย กรณีความแปรปรวนประชากรมีค่าเท่ากัน กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีความแปรปรวนประชากรที่มีค่าไม่เท่ากัน กรณีระดับนัยสำคัญเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย โดยส่วนใหญ่ตัวสถิติทดสอบ Z มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และกรณีตัวอย่างขนาดกลางไม่ว่าขนาดตัวอย่างจะเท่ากันหรือไม่ก็ตาม ที่มีค่าเฉลี่ยประชากรแตกต่างกันมาก แต่มีความแปรปรวนประชากรเท่ากัน และกรณีตัวอย่างขนาดใหญ่ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ที่มีความแปรปรวนประชากรเท่ากัน พบว่า เกือบทุกตัวสถิติทดสอบมีกำลังการทดสอบสูงพอ ๆ กัน หรือมีประสิทธิภาพไม่แตกต่างกัน ผู้วิจัยสามารถอภิปรายผลได้ดังนี้ ตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดและมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบอื่นๆ อย่างไรก็ตาม ตัวสถิติทดสอบ Z อาจไม่สามารถดำเนินการได้จริงในการทดสอบความแตกต่าง

ของค่ากลางระหว่างประชากร 2 กลุ่มที่อิสระกัน เนื่องจากตัวสถิติทดสอบ Z จะอยู่ภายใต้ข้อสมมุติว่า ต้องทราบค่าความแปรปรวนประชากรทั้ง 2 กลุ่ม ดังนั้นตัวสถิติทดสอบที่มีความเหมาะสมรองลงมาควรจะเป็นตัวสถิติทดสอบสตีวเดนตส์หรือตัวสถิติทดสอบเวลช์ที่ ที่ไม่มีข้อสมมุติเกี่ยวกับความแปรปรวนประชากร โดยกรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน ตัวสถิติทดสอบสตีวเดนตส์จะมีความเหมาะสมมากกว่าตัวสถิติทดสอบเวลช์ที่ แต่กรณีตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า หรือความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าไม่เท่ากัน ตัวสถิติทดสอบเวลช์ที่ จะมีความเหมาะสมมากกว่า ทั้งนี้เพราะตัวสถิติทดสอบเวลช์ที่เป็นตัวสถิติทดสอบที่พัฒนาขึ้นมาเพื่อใช้ในกรณีที่ความแปรปรวนประชากรทั้ง 2 กลุ่ม มีค่าไม่เท่ากัน สำหรับตัวสถิติทดสอบไม่อิงพารามิเตอร์ที่ศึกษา คือ ตัวสถิติทดสอบแมนน์-วิตนีย์ ตัวสถิติทดสอบโดยมัชฐาน และตัวสถิติทดสอบแวนเดอวอลเด็น มีความเหมาะสมในบางกรณี แต่ยังไม่มีแนวโน้มที่ชัดเจนมากนัก เช่น โดยส่วนใหญ่ ตัวสถิติทดสอบแมนน์-วิตนีย์ ตัวสถิติทดสอบโดยมัชฐาน และตัวสถิติทดสอบแวนเดอวอลเด็นมีความเหมาะสมกับตัวอย่างขนาดเล็กถึงขนาดกลาง และตัวสถิติทดสอบโดยมัชฐานมีความเหมาะสมกับการทดสอบสมมุติฐานที่กำหนดระดับนัยสำคัญ 0.01 มากกว่าที่ระดับนัยสำคัญ 0.05 กรณีการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่มที่อิสระกัน โดยที่ตัวอย่างทั้ง 2 กลุ่มมีขนาดใหญ่และความแปรปรวนประชากรมีค่าเท่ากัน จะสามารถใช้ตัวสถิติทดสอบใดก็ได้ระหว่างตัวสถิติทดสอบ Z ตัวสถิติทดสอบสตีวเดนตส์ ตัวสถิติทดสอบเวลช์ที่ ตัวสถิติทดสอบแมนน์-วิตนีย์ ตัวสถิติทดสอบโดยมัชฐาน หรือตัวสถิติทดสอบแวนเดอวอลเด็น เนื่องจากผลการศึกษาครั้งนี้ พบว่าตัวสถิติทดสอบเหล่านี้มีกำลังการ

ทดสอบสูงหรือมีประสิทธิภาพไม่แตกต่างกัน

การศึกษาค้นคว้าได้พิจารณาความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดแบบที่ 1 โดยใช้เกณฑ์การทดสอบทวินามเพียงเกณฑ์เดียว ดังนั้นในการวิจัยครั้งต่อไป จึงควรศึกษาเกณฑ์อื่นๆ เพิ่มเติม เช่น เกณฑ์ของแบรดลีย์ (Bradley) ซึ่งผลการศึกษาที่ได้อาจมีความแตกต่างกัน จากการศึกษาที่พบว่า ตัวสถิติทดสอบไม่อิงพารามิเตอร์ที่ศึกษา มีความเหมาะสมในบางกรณีแต่ยังไม่มีแนวโน้มที่ชัดเจนมากนัก ดังนั้นในการวิจัยครั้งต่อไป จึงควรศึกษาสถานการณ์ที่มีความละเอียดมากกว่านี้ เพื่อให้ทราบผลการศึกษาที่ชัดเจน อีกทั้งควรศึกษาประสิทธิภาพของตัวสถิติทดสอบไม่อิงพารามิเตอร์อื่น ๆ เพิ่มเติมในการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่มที่อิสระกัน เช่น ตัวสถิติทดสอบคอลโมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov test) ตัวสถิติทดสอบรันส์ของวัลด์-วอลโฟวิทซ์ (Wald-Wolfowitz runs test) และตัวสถิติทดสอบบรุนเนอร์-มุนเซล (Brunner-Munzel test) เป็นต้น รวมถึงควรศึกษาประสิทธิภาพของการทดสอบความแตกต่างของค่ากลางระหว่างประชากร 2 กลุ่มที่ไม่อิสระกัน และการทดสอบความแตกต่างของค่ากลางประชากรตั้งแต่ 2 กลุ่มที่อิสระกันและไม่อิสระกัน

5. กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ กองทุนวิจัยมหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2559

6. รายการอ้างอิง

- [1] วรางคณา เรียนสุทธิ, 2558, สถิติไม่อิงพารามิเตอร์, เอกสารคำสอน, มหาวิทยาลัยทักษิณ, พัทลุง.

- [2] สำนักงานสถิติแห่งชาติ, การเลือกใช้ตัวทดสอบสถิติ, แหล่งที่มา : <http://service.nso.go.th/nso/nsopublish/known/estat04.pdf>, 21 ธันวาคม 2558.
- [3] Ramberg, J.S., Tadikamalla, P.R., Dudewicz, E.J. and Mykytka, E.F., 1979, A Probability Distribution and Its Uses in Fitting Data, *Technometrics* 21(2): 201-214.
- [4] กัลยา วาณิชย์บัญชา, 2545, หลักสถิติ, พิมพ์ครั้งที่ 7, โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.
- [5] Wikipedia, Student's t-Test, Available Source: https://en.wikipedia.org/wiki/Student%27s_t-test, January 2, 2016.
- [6] Wikipedia, Welch's t-Test, Available Source: https://en.wikipedia.org/wiki/Welch%27s_t_test, January 2, 2016.
- [7] Wikipedia, Mann-Whitney U Test, Available Source: https://en.wikipedia.org/wiki/Mann%2E%80%93Whitney_U_test, January 2, 2016.
- [8] SAS Help, Chapter 47: The NPAR1WAY Procedure, Available Source: <http://www.okstate.edu/sas/v8/saspdf/stat/cha p47.pdf>, August 31, 2016.
- [9] Van der Waerden, B.L., 1952, Order Tests for the two-sample problem and their power, *Indagat. Math. Proc.* 55: 453-458.
- [10] Wikipedia, Van der Waerden Test, Available Source: https://en.wikipedia.org/wiki/Van_der_Waerden_test, January 2, 2016.
- [11] จริญญา เสกสรรค์, 2551, ผลกระทบของการฝ่าฝืนข้อกำหนดเบื้องต้นเกี่ยวกับความเป็นเอกภาพของความแปรปรวนต่อความคลาดเคลื่อนประเภทที่ 1 ในการทดสอบความแตกต่างระหว่างประชากรสองกลุ่ม, *วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ*, 109 น.