http://journal.rmutp.ac.th/

การเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน

วรางคณา เรียนสุทธิ์*

คณะวิทยาศาสตร์ มหาวิทยาลัยทักษิณ
222 ตำบลบ้านพร้าว อำเภอป่าพะยอม จังหวัดพัทลุง 93210

รับบทความ 2 กุมภาพันธ์ 2017; ตอบรับบทความ 15 พฤษภาคม 2017

บทคัดย่อ

วัตถุประสงค์ของการศึกษาครั้งนี้ คือ เพื่อคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 กำลังการทดสอบ และเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบทั้งหมด 3 วิธี ได้แก่ ตัวสถิติทดสอบ Z ตัวสถิติทดสอบสติวเดนต์ที และตัวสถิติทดสอบเวลช์ที รวมถึงเพื่อเสนอแนะตัวสถิติทดสอบที่เหมาะสม จำแนกสถานการณ์โดยกำหนดประชากรที่มีการแจกแจง 5 รูปแบบ กำหนดอัตราส่วนค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 ทั้งหมด 3 ระดับ อัตราส่วนความแปรปรวนประชากรกลุ่มที่ 1 และกลุ่มที่ 2 ทั้งหมด 3 ระดับ ขนาดตัวอย่างจากประชากร 2 กลุ่ม 9 ขนาด ระดับนัยสำคัญที่ศึกษา คือ 0.01 และ 0.05 จำลองข้อมูลด้วยการเขียนโปรแกรม SAS กำหนดจำนวนรอบของการทำซ้ำ คือ 1,000 รอบ ผลการศึกษาพบว่า ตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดและมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบอื่นๆ แต่ตัวสถิติทดสอบนี้จำเป็นต้องทราบค่าความแปรปรวนประชากรทั้ง 2 กลุ่ม ดังนั้นจึงควรใช้ตัวสถิติทดสอบที่มีความเหมาะสมรองลงมา คือ ตัวสถิติทดสอบสติวเดนต์ทีหรือเวลช์ที โดยกรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน ตัวสถิติทดสอบสติวเดนต์ทีจะมีความเหมาะสมมากกว่าเวลช์ที แต่กรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน ตัวสถิติทดสอบเวลช์ทีจะมีความเหมาะสมมากกว่า กรณีตัวอย่างทั้ง 2 กลุ่มมีขนาดใหญ่และความแปรปรวนประชากรมีค่าเท่ากัน จะสามารถใช้ตัวสถิติทดสอบใดก็ได้ เนื่องจากมีประสิทธิภาพไม่แตกต่างกัน

คำสำคัญ: การทดสอบความแตกต่างของค่าเฉลี่ย; ความน่าจะเป็นของความผิดพลาดแบบที่ 1; กำลังการทดสอบ

** ผู้นิพนธ์ประสานงาน โทร: +66 88790 8476, ไปรษณีย์อิเล็กทรอนิกส์: warang27@gmail.com*

http://journal.rmutp.ac.th/

An Efficiency Comparison of the Test Statistics for Testing the Mean Difference between Two Independent Populations

Warangkhana Riansut*

Faculty of Science, Thaksin University, Phatthalung
222 Ban Phrao, Papayom, Phatthalung, 93210

Received 2 February 2017; accepted 15 May 2017

Abstract

The objectives of this study were to calculate the probability of type I error, power of a test, and to compare the performance of three test statistics: Z, Student's t, and Welch's t. Including the recommending the appropriate test statistic. Classification of the population according to five distribution. The ratios of mean group 1 per group 2 were three levels. The ratios of variance group 1 per group 2 were three levels. The sample sizes of two populations were nine sizes. 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. 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

** Corresponding Author. Tel.: +66 88790 8476, E-mail Address: warang27@gmail.com*

1. บทนำ

ในการศึกษาวิจัยโดยทั่วไปนั้น นักวิจัยอาจไม่สามารถทำการศึกษาประชากรได้ทั้งหมด เพราะมีความยุ่งยากในการจัดเก็บข้อมูล สิ้นเปลืองค่าใช้จ่ายและเวลา ฯลฯ ดังนั้นนักวิจัยจึงควรทำการสุ่มตัวอย่างจากประชากรตามวิธีการสุ่มตัวอย่างทางสถิติเพื่อให้ได้ตัวอย่างที่เป็นตัวแทนที่ดีของประชากรที่ต้องการศึกษา หลังจากนั้นจึงวิเคราะห์ข้อมูลที่ได้จากตัวอย่าง และใช้หลักการทางสถิติที่เรียกว่า การอนุมาน (Inference) เพื่อสรุปถึงคุณลักษณะของประชากร การอนุมานทางสถิติประกอบด้วยการประมาณค่า (Estimation) และการทดสอบสมมุติฐาน (Hypothesis Testing) สำหรับการทดสอบสมมุติฐานจะแบ่งออกเป็นการทดสอบอิงพารามิเตอร์ (Parametric Test) และการทดสอบไม่อิงพารามิเตอร์ (Nonparametric Test) โดยการทดสอบสมมุติฐานแบบอิงพารามิเตอร์นั้น ตัวสถิติทดสอบที่ใช้จะขึ้นอยู่กับเงื่อนไขต่าง ๆ หรือข้อสมมุติ (Assumption) เช่น ลักษณะของข้อมูล มาตรวัดของข้อมูล การแจกแจงของข้อมูลจากตัวอย่างที่สุ่มได้ ภาวะความแปรปรวนเท่ากัน (Homoscedasticity) เรื่องที่ต้องการศึกษา และจำนวนกลุ่มของข้อมูลหรือจำนวนประชากรที่ต้องการศึกษา เป็นต้น [1] ในงานวิจัยทางการแพทย์ การศึกษา การตลาด หรือทางด้านจิตวิทยา มักมีการทดสอบสมมุติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน โดยปราศจากการตรวจสอบข้อสมมุติ และใช้การทดสอบสติวเดนต์ที (Student's t-Test) ทั้งๆ ที่มีวิธีการทดสอบที่เหมาะสมอื่น ๆ เช่น การทดสอบ Z (Z-Test) และการทดสอบเวลช์ที (Welch's t-Test) เป็นต้น ซึ่งผลสรุปที่ได้จากการทดสอบสติวเดนต์ทีอาจขัดแย้งกับความเป็นจริง และระดับนัยสำคัญ (Significance Level) หรือ ความน่าจะเป็นของความผิดพลาดแบบที่ 1 (Probability of Type I Error) ไม่เป็นไปตามที่กำหนดไว้ รวมถึงอาจมีกำลังการทดสอบ (Power of a Test) ต่ำ อย่างไรก็ตาม คำกล่าวเหล่านี้ยังขาดการยืนยันในเชิงทฤษฎี ด้วยเหตุผลดังกล่าว ผู้วิจัยจึงมีแนวคิดที่จะศึกษาการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกันทั้งหมด 3 วิธี ได้แก่ การทดสอบ Z การทดสอบสติวเดนต์ที และการทดสอบเวลช์ที การเปรียบเทียบประสิทธิภาพจะกระทำภายใต้สถานการณ์ต่างๆ ทั้งกรณีการแจกแจง ความแปรปรวน ขนาดตัวอย่าง และระดับนัยสำคัญที่แตกต่างกัน ผู้วิจัยคาดว่าผลการศึกษาที่ได้จะทำให้ทราบความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบทั้งหมด 3 วิธี จึงทำให้ทราบว่าตัวสถิติทดสอบวิธีใดมีความสามารถในการควบคุม ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่าง ๆ เพื่อเป็นแนวทางให้กับนิสิต คณาจารย์ และบุคคลทั่วไป สำหรับใช้ประโยชน์ในการตัดสินใจเลือกใช้ตัวสถิติทดสอบที่เหมาะสมกับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน อีกทั้งผลการวิจัยที่ได้ยังสามารถเผยแพร่ในวารสารวิชาการระดับชาติ/นานาชาติ ซึ่งจะส่งผลดีต่อการวิจัย การผลิตการพัฒนาครู อาจารย์ และบุคลากรทางการศึกษาให้มีคุณภาพสูง สอดคล้อง และรองรับกับแนวทางการพัฒนาประเทศ ต่อไป

2. วิธีการศึกษา

ขั้นตอนการดำเนินการวิจัย แสดงดังนี้

1. จำลองข้อมูลของประชากรที่ 1 และ 2 ด้วยโปรแกรม SAS รุ่น 9.00 กำหนดการแจกแจงความน่าจะเป็นของประชากรที่ศึกษา 5 รูปแบบ ได้แก่ การแจกแจงปรกติ โดยใช้คำสั่ง rand ('normal', 0, 1) การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปรกติ (ความเบ้ = -2, ความโด่ง = 12) การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปรกติ (ความเบ้ = -0.4, ความโด่ง = 2.2) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปรกติ (ความเบ้

= 2, ความโด่ง = 12) และการแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปรกติ (ความเบ้ = 0.4, ความโด่ง = 2.2) สำหรับการจำลองข้อมูลที่มีลักษณะไม่สมมาตร (เบ้ซ้าย เบ้ขวา ความโด่งสูงกว่าปรกติ และความโด่งต่ำกว่าปรกติ) จะเริ่มต้นจากการจำลองข้อมูลที่มีการแจกแจงเอกรูปในช่วง (0, 1) โดยใช้คำสั่ง rand('uniform') แล้วแปลงข้อมูลตามวิธีการของ Ramberg และคณะ [2] 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 λ_1, λ_2, λ_3 และ λ_4 ได้จากตารางที่ 1 หน้า 205 หรือตารางที่ 4 หน้า 210 – 214 ในบทความของ Ramberg และคณะ [2] โดยค่าของพารามิเตอร์ทั้ง 4 ตัวจะให้ข้อมูลที่มีการแจกแจงเบ้ขวา สำหรับข้อมูลที่มีการแจกแจงเบ้ซ้ายให้สลับค่าของ λ_3 และ λ_4 และปรับเครื่องหมายของ λ_1 เป็นตรงข้าม ดังนั้นถ้าต้องการข้อมูลที่มีการแจกแจงตามที่กำหนด ต้องมีค่าของพารามิเตอร์ λ_1, λ_2, λ_3 และ λ_4 ดังนี้

• การแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปรกติ (ความเบ้ = -2, ความโด่ง = 12)

λ_1 = 0.579, λ_2 = -0.1423, λ_3 = -0.0995 และ λ_4 = -0.0273

• การแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปรกติ (ความเบ้ = -0.4, ความโด่ง = 2.2)

λ_1 = 1.354, λ_2 = 0.2582, λ_3 = 0.5683 และ λ_4 = 0.0129

• การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปรกติ (ความเบ้ = 2, ความโด่ง = 12)

λ_1 = -0.579, λ_2 = -0.1423, λ_3 = -0.0273 และ λ_4 = -0.0995

• การแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปรกติ (ความเบ้ = 0.4, ความโด่ง = 2.2)

λ_1 = -1.354, λ_2 = 0.2582, λ_3 = 0.0129 และ λ_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 กลุ่ม เท่ากัน 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 จากนั้นแปลงข้อมูลแต่ละประชากรให้มีค่าเฉลี่ยและความแปรปรวนตามที่กำหนด โดยใช้สูตร

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

เมื่อ X แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

Y แทนแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2

2. คำนวณค่าของตัวสถิติทดสอบและค่าพี (p-value) ของการทดสอบทั้งหมด 3 วิธี คือ การทดสอบ Z การทดสอบสติวเดนต์ที และการทดสอบเวลช์ที [3] เพื่อเปรียบเทียบกับระดับนัยสำคัญที่กำหนด คือ 0.01 และ 0.05 โดยถ้าค่าพีน้อยกว่าหรือเท่ากับระดับนัยสำคัญ จะตัดสินใจปฏิเสธสมมุติฐานว่าง (ปฏิเสธ H_0) และถ้าค่าพีมากกว่าระดับนัยสำคัญ จะตัดสินใจไม่ปฏิเสธสมมุติฐานว่าง

3. คำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบ โดยการนับจำนวนครั้งที่ตัดสินใจปฏิเสธสมมุติฐานว่าง นำมาหารด้วยจำนวนรอบของการทำซ้ำ ซึ่งคือ 1,000 รอบ ดังนี้

• กรณีอัตราส่วนของค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 1 ซึ่งการศึกษาครั้งนี้กำหนดให้เป็น (μ_1, μ_2) = (10, 10) จะได้ความน่าจะเป็นของความผิดพลาดแบบที่ 1 นั่นคือ

$$\alpha = \frac{\text{จำนวนครั้งที่ตัดสินใจปฏิเสธ } H_O \text{ เมื่อ } \mu_1, \mu_2}{1{,}000} \qquad (1)$$

นำค่า α ที่คำนวณได้ไปเปรียบเทียบกับเกณฑ์การทดสอบทวินาม [4] ดังนี้

- กรณีระดับนัยสำคัญเท่ากับ α_0 = 0.01 จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 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, $Z_{0.01/2} = Z_{0.0005}$ = 2.576, n แทนจำนวนรอบของการทำซ้ำ = 1,000 จะได้

$$0.0019 < \alpha\ 0.0181$$

- กรณีระดับนัยสำคัญเท่ากับ α_0 = 0.05 จะกล่าวว่าตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 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

$$0.0365 < \alpha\ 0.0635$$

• กรณีอัตราส่วนของค่าเฉลี่ยประชากรกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 0.9 และ 1.2 ซึ่งการศึกษาครั้งนี้กำหนดให้เป็น (μ_1, μ_2) = (9, 10) และ (12, 10) จะได้กำลังการทดสอบ นั่นคือ

$$1-\beta = \frac{\text{จำนวนครั้งที่ตัดสินใจปฏิเสธ } H_O \text{ เมื่อ } \mu_1, \mu_2}{1{,}000} \qquad (2)$$

4. สรุปผลและเสนอแนะว่าตัวสถิติทดสอบสำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน วิธีใดที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และมีกำลังการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่าง ๆ ทั้งกรณีการแจกแจง ความแปรปรวน ขนาดตัวอย่าง และระดับนัยสำคัญที่แตกต่างกัน

3. ผลการศึกษาและอภิปรายผล

การศึกษาครั้งนี้จะแบ่งผลการศึกษาออกเป็น 2 ส่วน คือ

1. ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด สรุปได้ดังตารางที่ 1 – 5

จากตารางที่ 1 – 5 พบว่า ตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทุกระดับอัตราส่วนความ

แปรปรวนประชากร ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา ยกเว้นเพียงกรณีขนาดตัวอย่างเป็น (10, 15) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงปรกติ และระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบสติวเดนต์ทีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน ยกเว้นขนาดตัวอย่างเป็น (10, 10) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปรกติ ระดับนัยสำคัญ 0.01 และขนาดตัวอย่างเป็น (100, 100) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงปรกติ ระดับนัยสำคัญ 0.05 แต่เมื่อขนาดตัวอย่างไม่เท่ากัน ตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน และไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า เช่น ขนาดตัวอย่างเป็น (15, 10) แต่มีความแปรปรวนประชากรเป็น (4, 16) และตัวสถิติทดสอบเวลช์ทีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน ยกเว้นขนาดตัวอย่างเป็น (10, 10) ความแปรปรวนประชากรเป็น (1, 1) การแจกแจงเบ้ขวาที่มีความโด่งสูงกว่า

ตารางที่ 1 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงปรกติ

n_1	n_2	σ_1^2/σ_2^2	ระดับนัยสำคัญ 0.01			ระดับนัยสำคัญ 0.05		
			Z	Student's t	Welch's t	Z	Student's t	Welch's t
10	10	1, 1	0.015*	0.017*	0.015*	0.059*	0.054*	0.061*
		4, 16	0.010*	0.010*	0.006*	0.042*	0.060*	0.043*
		16, 4	0.009*	0.010*	0.009*	0.047*	0.043*	0.040*
30	30	1, 1	0.004*	0.006*	0.006*	0.055*	0.045*	0.043*
		4, 16	0.006*	0.006*	0.005*	0.053*	0.053*	0.048*
		16, 4	0.010*	0.010*	0.010*	0.061*	0.058*	0.057*
100	100	1, 1	0.015*	0.013*	0.013*	0.045*	0.044*	0.044*
		4, 16	0.013*	0.014*	0.014*	0.058*	0.069	0.069
		16, 4	0.008*	0.007*	0.006*	0.053*	0.053*	0.053*
15	10	1, 1	0.008*	0.007*	0.004*	0.039*	0.040*	0.041*
		4, 16	0.014*	0.033	0.012*	0.051*	0.097	0.062*
		16, 4	0.006*	0.003*	0.007*	0.043*	0.023	0.040*
45	30	1, 1	0.011*	0.011*	0.013*	0.056*	0.062*	0.063*
		4, 16	0.006*	0.015*	0.007*	0.041*	0.076	0.044*
		16, 4	0.008*	0.002*	0.006*	0.051*	0.026	0.044*
150	100	1, 1	0.011*	0.010*	0.010*	0.051*	0.050*	0.050*
		4, 16	0.012*	0.028	0.010*	0.055*	0.094	0.054*
		16, 4	0.005*	0.001	0.004*	0.041*	0.025	0.044*
10	15	1, 1	0.011*	0.013*	0.012*	0.041*	0.043*	0.044*
		4, 16	0.011*	0.008*	0.009*	0.036	0.025	0.038*
		16, 4	0.011*	0.027	0.013*	0.044*	0.074	0.044*
30	45	1, 1	0.014*	0.011*	0.012*	0.056*	0.056*	0.057*
		4, 16	0.014*	0.007*	0.014*	0.059*	0.031	0.056*
		16, 4	0.011*	0.028	0.014*	0.045*	0.075	0.047*
100	150	1, 1	0.007*	0.010*	0.009*	0.040*	0.041*	0.044*
		4, 16	0.007*	0.002*	0.008*	0.041*	0.025	0.043*
		16, 4	0.006*	0.026	0.005*	0.053*	0.082	0.054*

* แทนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 2 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ซ้ายที่มีความโด่งสูงกว่าปรกติ

n_1	n_2	σ_1^2/σ_2^2	ระดับนัยสำคัญ 0.01			ระดับนัยสำคัญ 0.05		
			Z	Student's t	Welch's t	Z	Student's t	Welch's t
10	10	1, 1	0.015*	0.004*	0.004*	0.057*	0.045*	0.048*
		4, 16	0.016*	0.014*	0.016*	0.011*	0.063*	0.062*
		16, 4	0.013*	0.016*	0.011*	0.057*	0.057*	0.050*
30	30	1, 1	0.009*	0.006*	0.006*	0.054*	0.055*	0.052*
		4, 16	0.018*	0.015*	0.014*	0.061*	0.060*	0.059*
		16, 4	0.010*	0.010*	0.010*	0.047*	0.048*	0.048*
100	100	1, 1	0.015*	0.011*	0.011*	0.053*	0.050*	0.050*
		4, 16	0.008*	0.009*	0.008*	0.040*	0.054*	0.054*
		16, 4	0.013*	0.009*	0.009*	0.058*	0.051*	0.051*
15	10	1, 1	0.014*	0.011*	0.009*	0.054*	0.055*	0.048*
		4, 16	0.018*	0.033	0.029	0.045*	0.110	0.077
		16, 4	0.018*	0.009*	0.010*	0.046*	0.042*	0.060*
45	30	1, 1	0.010*	0.006*	0.006*	0.046*	0.039*	0.039*
		4, 16	0.013*	0.027	0.015*	0.051*	0.086	0.053*
		16, 4	0.013*	0.006*	0.010*	0.060*	0.027	0.049*
150	100	1, 1	0.013*	0.011*	0.013*	0.056*	0.048*	0.045*
		4, 16	0.013*	0.021	0.010*	0.046*	0.086	0.053*
		16, 4	0.009*	0.002*	0.008*	0.047*	0.022	0.044*
10	15	1, 1	0.013*	0.004*	0.004*	0.043*	0.041*	0.036
		4, 16	0.012*	0.007*	0.006*	0.056*	0.037*	0.050*
		16, 4	0.017*	0.031	0.020	0.054*	0.095	0.079
30	45	1, 1	0.011*	0.010*	0.010*	0.047*	0.043*	0.045*
		4, 16	0.011*	0.004*	0.008*	0.063*	0.028	0.050*
		16, 4	0.009*	0.021	0.013*	0.055*	0.078	0.048*
100	150	1, 1	0.013*	0.007*	0.007*	0.061*	0.063*	0.060*
		4, 16	0.014*	0.004*	0.010*	0.052*	0.030	0.057*
		16, 4	0.011*	0.024	0.007*	0.043*	0.087	0.058*

* แทนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 3 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ซ้ายที่มีความโด่งต่ำกว่าปรกติ

n_1	n_2	σ_1^2/σ_2^2	ระดับนัยสำคัญ 0.01			ระดับนัยสำคัญ 0.05		
			Z	Student's t	Welch's t	Z	Student's t	Welch's t
10	10	1, 1	0.011*	0.010*	0.010*	0.053*	0.049*	0.047*
		4, 16	0.013*	0.009*	0.006*	0.052*	0.057*	0.053*
		16, 4	0.009*	0.015*	0.012*	0.057*	0.058*	0.054*
30	30	1, 1	0.014*	0.013*	0.013*	0.052*	0.054*	0.054*
		4, 16	0.007*	0.011*	0.009*	0.054*	0.057*	0.057*
		16, 4	0.010*	0.014*	0.014*	0.051*	0.053*	0.050*
100	100	1, 1	0.006*	0.008*	0.008*	0.051*	0.044*	0.044*
		4, 16	0.013*	0.011*	0.010*	0.060*	0.062*	0.061*
		16, 4	0.012*	0.012*	0.011*	0.049*	0.056*	0.056*
15	10	1, 1	0.012*	0.011*	0.016*	0.048*	0.047*	0.045*
		4, 16	0.011*	0.019	0.012*	0.045*	0.086	0.047*
		16, 4	0.007*	0.007*	0.009*	0.050*	0.029	0.052*
45	30	1, 1	0.012*	0.011*	0.011*	0.059*	0.058*	0.058*
		4, 16	0.015*	0.027	0.016*	0.056*	0.099	0.062*
		16, 4	0.012*	0.006*	0.013*	0.050*	0.027	0.052*
150	100	1, 1	0.012*	0.014*	0.011*	0.050*	0.053*	0.057*
		4, 16	0.007*	0.029	0.010*	0.063*	0.099	0.059*
		16, 4	0.013*	0.006*	0.012*	0.056*	0.031	0.054*
10	15	1, 1	0.009*	0.013*	0.015*	0.053*	0.058*	0.057*
		4, 16	0.011*	0.008*	0.011*	0.047*	0.036	0.052*
		16, 4	0.008*	0.027	0.014*	0.042*	0.089	0.055*
30	45	1, 1	0.012*	0.011*	0.012*	0.058*	0.059*	0.059*
		4, 16	0.009*	0.004*	0.009*	0.046*	0.032	0.051*
		16, 4	0.010*	0.025	0.012*	0.046*	0.078	0.053*
100	150	1, 1	0.008*	0.010*	0.009*	0.048*	0.050*	0.050*
		4, 16	0.012*	0.003*	0.014*	0.058*	0.036	0.060*
		16, 4	0.012*	0.017*	0.010*	0.041*	0.090	0.048*

* แทนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 4 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ขวาที่มีความโด่งสูงกว่าปรกติ

n_1	n_2	σ_1^2/σ_2^2	ระดับนัยสำคัญ 0.01			ระดับนัยสำคัญ 0.05		
			Z	Student's t	Welch's t	Z	Student's t	Welch's t
10	10	1, 1	0.018*	0.006*	0.007*	0.054*	0.047*	0.034
		4, 16	0.014*	0.020	0.018*	0.048*	0.062*	0.055*
		16, 4	0.012*	0.018*	0.016*	0.043*	0.059*	0.056*
30	30	1, 1	0.010*	0.007*	0.007*	0.051*	0.054*	0.053*
		4, 16	0.010*	0.018*	0.016*	0.046*	0.051*	0.051*
		16, 4	0.010*	0.010*	0.010*	0.049*	0.060*	0.059*
100	100	1, 1	0.013*	0.014*	0.014*	0.055*	0.055*	0.055*
		4, 16	0.015*	0.012*	0.012*	0.054*	0.055*	0.054*
		16, 4	0.006*	0.011*	0.009*	0.048*	0.046*	0.046*
15	10	1, 1	0.015*	0.011*	0.011*	0.054*	0.041*	0.042*
		4, 16	0.014*	0.033	0.022	0.060*	0.097	0.065
		16, 4	0.008*	0.009*	0.011*	0.038*	0.035	0.046*
45	30	1, 1	0.010*	0.004*	0.010*	0.045*	0.039*	0.054*
		4, 16	0.010*	0.017*	0.013*	0.050*	0.074	0.060*
		16, 4	0.013*	0.002*	0.011*	0.061*	0.033	0.051*
150	100	1, 1	0.007*	0.010*	0.010*	0.045*	0.050*	0.053*
		4, 16	0.010*	0.016*	0.008*	0.043*	0.086	0.050*
		16, 4	0.014*	0.004*	0.009*	0.050*	0.017	0.037*
10	15	1, 1	0.015*	0.005*	0.006*	0.049*	0.041*	0.038*
		4, 16	0.011*	0.009*	0.008*	0.063*	0.039*	0.059*
		16, 4	0.017*	0.029	0.023	0.051*	0.090	0.065
30	45	1, 1	0.011*	0.006*	0.008*	0.045*	0.045*	0.054*
		4, 16	0.012*	0.006*	0.010*	0.048*	0.029	0.046*
		16, 4	0.010*	0.027	0.020	0.046*	0.085	0.061*
100	150	1, 1	0.009*	0.014*	0.009*	0.061*	0.049*	0.054*
		4, 16	0.013*	0.003*	0.010*	0.043*	0.039*	0.051*
		16, 4	0.007*	0.025	0.003*	0.042*	0.081	0.028

* แทนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ตารางที่ 5 ความน่าจะเป็นของความผิดพลาดแบบที่ 1 เมื่อประชากรมีการแจกแจงเบ้ขวาที่มีความโด่งต่ำกว่าปรกติ

n_1	n_2		ระดับนัยสำคัญ 0.01			ระดับนัยสำคัญ 0.05		
			Z	Student's t	Welch's t	Z	Student's t	Welch's t
10	10	1, 1	0.009*	0.012*	0.011*	0.043*	0.051*	0.049*
		4, 16	0.008*	0.016*	0.013*	0.045*	0.056*	0.055*
		16, 4	0.005*	0.017*	0.013*	0.040*	0.054*	0.051*
30	30	1, 1	0.004*	0.004*	0.004*	0.044*	0.040*	0.040*
		4, 16	0.015*	0.016*	0.015*	0.045*	0.047*	0.046*
		16, 4	0.007*	0.009*	0.009*	0.041*	0.042*	0.042*
100	100	1, 1	0.015*	0.015*	0.015*	0.047*	0.051*	0.051*
		4, 16	0.011*	0.011*	0.010*	0.055*	0.060*	0.060*
		16, 4	0.012*	0.012*	0.012*	0.062*	0.059*	0.059*
15	10	1, 1	0.005*	0.007*	0.007*	0.044*	0.042*	0.049*
		4, 16	0.011*	0.027	0.018*	0.046*	0.084	0.054*
		16, 4	0.008*	0.003*	0.005*	0.051*	0.022	0.050*
45	30	1, 1	0.011*	0.013*	0.012*	0.055*	0.051*	0.053*
		4, 16	0.015*	0.030	0.012*	0.048*	0.083	0.056*
		16, 4	0.012*	0.004*	0.010*	0.049*	0.035	0.047*
150	100	1, 1	0.005*	0.004*	0.005*	0.040*	0.039*	0.040*
		4, 16	0.010*	0.022	0.008*	0.044*	0.082	0.044*
		16, 4	0.014*	0.007*	0.016*	0.059*	0.032	0.059*
10	15	1, 1	0.004*	0.005*	0.009*	0.045*	0.046*	0.050*
		4, 16	0.009*	0.004*	0.012*	0.055*	0.035	0.060*
		16, 4	0.014*	0.039	0.020	0.056*	0.099	0.066
30	45	1, 1	0.014*	0.015*	0.011*	0.061*	0.062*	0.061*
		4, 16	0.009*	0.006*	0.010*	0.040*	0.027	0.046*
		16, 4	0.013*	0.027	0.016*	0.062*	0.098	0.055*
100	150	1, 1	0.013*	0.011*	0.011*	0.053*	0.053*	0.053*
		4, 16	0.014*	0.004*	0.015*	0.055*	0.034	0.054*
		16, 4	0.005*	0.016*	0.002*	0.047*	0.077	0.046*

* แทนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

ปรกติ ระดับนัยสำคัญ 0.05 และขนาดตัวอย่างเป็น (100, 100) ความแปรปรวนประชากรเป็น (4, 16) การแจกแจงปรกติ ระดับนัยสำคัญ 0.05 โดยตัวสถิติทดสอบนี้สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดดีกว่าตัวสถิติทดสอบสติวเดนต์ทีเมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า

2. ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด สรุปได้ดังตารางที่ 6

ตารางที่ 6 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปรกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปรกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
10	10	9, 10	1, 1	Z (0.392)	St (0.362)	Z (0.382)	St (0.386)	Z (0.357)	Z (0.640)	St (0.621)	Z (0.604)	St (0.645)	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)	St (0.071)	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)	St (0.207)	Z (0.118)	Z (0.126)	Z (0.118)	Z (0.293)	St (0.392)	Z (0.302)	Z (0.289)	Z (0.294)
30	30	9, 10	1, 1	Z (0.920)	Z (0.879)	Z (0.924)	Z (0.909)	Z (0.900)	Z (0.978)	Z (0.959)	Z (0.972)	Z (0.977)	All (0.963)
			4, 16	St (0.100)	St (0.172)	St (0.107)	Z (0.091)	Z (0.086)	St (0.241)	St (0.331)	St (0.234)	Z (0.220)	Z (0.227)
			16, 4	Z (0.079)	Z (0.103)	Z (0.094)	St (0.137)	St (0.087)	St (0.237)	Z (0.224)	Z (0.238)	St (0.288)	St (0.229)
		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 (0.459)	Z (0.428)	Z (0.457)	St (0.493)	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)

ตารางที่ 6 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด (ต่อ)

n_1	n_2	μ_1, μ_2	σ_1^2, σ_2^2	ระดับนัยสำคัญ 0.01					ระดับนัยสำคัญ 0.05				
				ปรกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ	ปรกติ	เบ้ซ้ายโด่งสูง	เบ้ซ้ายโด่งต่ำ	เบ้ขวาโด่งสูง	เบ้ขวาโด่งต่ำ
100	100	9, 10	1, 1	All (1.000)	Z (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 (0.403)	St (0.396)	Z (0.381)	Z (0.378)	Z (0.359)	Z (0.614)	Z (0.621)	Z (0.599)	St, Wt (0.623)	Z (0.598)
			16, 4	Z (0.378)	Z (0.347)	Z (0.331)	Wt (0.399)	Z (0.383)	St, Wt (0.628)	St (0.620)	Z (0.562)	Z (0.621)	Z (0.611)
		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 (0.967)	St, Wt (0.984)	St, Wt (0.979)	Z (0.962)	Z (0.974)	Z (0.991)	St, Wt (0.998)	St, Wt (1.000)	Z (0.988)	Z (0.994)
			16, 4	Z (0.979)	Z (0.970)	Z (0.964)	St (0.979)	Z (0.980)	All (0.994)	Z (0.989)	Z (0.994)	St, Wt (0.998)	St, Wt (0.994)
15	10	9, 10	1, 1	Z (0.455)	Z (0.453)	Z (0.436)	Z (0.431)	Z (0.434)	Z (0.706)	Z (0.713)	Z (0.681)	St (0.667)	Z (0.655)
			4, 16	Z (0.026)	Z (0.027)	Wt (0.038)	Z (0.052)	Z (0.041)	Z (0.104)	Z (0.101)	Wt (0.126)	Z (0.107)	Z (0.124)
			16, 4	Wt (0.034)	Z (0.040)	Z (0.047)	Wt (0.073)	Wt (0.047)	Z (0.114)	Z (0.109)	Z (0.138)	Wt (0.185)	Wt (0.135)
		12, 10	1, 1	Z (0.989)	Z (0.990)	Z (0.992)	Z (0.988)	Z (0.990)	Z (0.997)	Z (0.996)	Z (0.999)	Z (0.998)	Z, St (1.000)
			4, 16	Z (0.134)	Z (0.134)	Z (0.121)	Z (0.127)	Z (0.108)	Z (0.307)	Z (0.285)	Z (0.295)	Z (0.340)	Z (0.292)
			16, 4	Z (0.160)	Wt (0.220)	Z (0.189)	Z (0.176)	Z (0.162)	Z (0.367)	Wt (0.414)	Z (0.396)	Z (0.352)	Z (0.357)

ตารางที่ 6 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด (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
45	30	9, 10	1, 1	Z (0.957)	Z (0.951)	Z (0.967)	Z (0.946)	Z (0.958)	Z (0.992)	Z (0.987)	Z (0.992)	Z (0.984)	Z (0.994)
			4, 16	St (0.153)	Wt (0.167)	Z (0.097)	St (0.112)	Z (0.091)	Z (0.238)	Wt (0.299)	Z (0.257)	Z (0.231)	Z (0.238)
			16, 4	Z (0.106)	Z (0.119)	Z (0.135)	Wt (0.168)	Wt (0.120)	Z (0.285)	Z (0.287)	Z (0.296)	Wt (0.337)	Z (0.294)
		12, 10	1, 1	All (1.000)	All (1.000)	All (1.000)	Z (1.000)	All (1.000)	All (1.000)	All (1.000)	All (1.000)	Z, St (1.000)	All (1.000)
			4, 16	St (0.594)	Z (0.474)	Z (0.472)	Wt (0.515)	Z (0.494)	Z (0.710)	Wt (0.753)	Z (0.721)	Z (0.738)	Z (0.716)
			16, 4	Z (0.627)	Z (0.638)	Z (0.622)	Wt (0.618)	Z (0.618)	Z (0.827)	Z (0.803)	Z (0.830)	Wt (0.827)	Z (0.809)
150	100	9, 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 (0.399)	Wt (0.440)	Z (0.415)	St (0.525)	Z (0.352)	Z (0.639)	Z (0.649)	Z (0.647)	Wt (0.676)	Z (0.605)
			16, 4	Z (0.519)	Wt (0.509)	Z (0.543)	Z (0.515)	Z (0.527)	Z (0.734)	Wt (0.773)	Z (0.763)	Z (0.732)	Z (0.754)
		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 (0.982)	Wt (0.991)	Wt (0.982)	Z (0.978)	Z (0.983)	Z, Wt (0.997)	Wt (0.999)	Wt (0.997)	Z (0.990)	Z (0.996)
			16, 4	Z (0.997)	Z (0.992)	Z, Wt (0.996)	All (0.993)	Z (1.000)	Z, Wt (1.000)	Z (1.000)	Z, Wt (1.000)	Z (1.000)	Z, Wt (1.000)

ตารางที่ 6 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด (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
10	15	9, 10	1, 1	Z (0.454)	St (0.437)	Z (0.440)	Wt (0.452)	Z (0.440)	Z (0.695)	St (0.686)	Z (0.674)	Z (0.687)	Z (0.687)
			4, 16	Z (0.035)	Wt (0.066)	Z (0.038)	Z (0.041)	Z (0.036)	Wt (0.111)	Wt (0.172)	Z (0.122)	Z (0.117)	Z (0.117)
			16, 4	Z (0.021)	Z (0.044)	Z (0.032)	Z (0.023)	Z (0.028)	Wt (0.094)	Z (0.123)	Z (0.096)	Z (0.090)	Z (0.112)
		12, 10	1, 1	Z (0.993)	Z (0.990)	Z (0.989)	Z (0.992)	Z (0.994)	All (0.999)	Z (0.999)	Z, St (0.997)	Z (0.998)	Z (1.000)
			4, 16	Z (0.168)	Z (0.168)	Z (0.194)	Wt (0.209)	Z (0.190)	Wt (0.324)	Wt (0.349)	Z (0.381)	Wt (0.400)	Z (0.395)
			16, 4	Z (0.128)	Z (0.107)	Z (0.145)	Z (0.145)	Z (0.130)	Z (0.327)	Z (0.326)	Z (0.316)	Z (0.278)	Z (0.308)
30	45	9, 10	1, 1	Z (0.948)	Z (0.957)	Z (0.953)	Z (0.950)	Z (0.951)	Z (0.989)	Z (0.993)	Wt (0.989)	Z (0.987)	Z (0.994)
			4, 16	Z (0.126)	Wt (0.182)	Z (0.129)	Z (0.116)	Z (0.129)	Z (0.313)	Wt (0.342)	Z (0.296)	Z (0.268)	Z (0.289)
			16, 4	Z (0.094)	Z (0.095)	Z (0.075)	Z (0.096)	Wt (0.112)	Z (0.266)	Z (0.245)	Z (0.250)	Wt (0.323)	Wt (0.276)
		12, 10	1, 1	All (1.000)	Z, St (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 (0.623)	Wt (0.609)	Z (0.622)	Z (0.639)	Z (0.634)	Z (0.812)	Wt (0.843)	Z, Wt (0.828)	Z (0.831)	Z (0.825)
			16, 4	Z (0.509)	Wt (0.504)	Z (0.488)	Z (0.446)	Z (0.465)	Z (0.708)	Z (0.728)	Z (0.711)	Wt (0.748)	Z (0.708)

ตารางที่ 6 ตัวสถิติทดสอบที่มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด (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)	Z, St (1.000)	All (1.000)	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	Z (0.413)	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 แทนตัวสถิติทดสอบสติวเดนต์ที
Wt แทนตัวสถิติทดสอบเวลช์ที
All แทนทุกตัวสถิติทดสอบ
ค่าในวงเล็บแทนกำลังการทดสอบ

จากตารางที่ 6 พบว่า ผลการศึกษาเกี่ยวกับกำลังการทดสอบสามารถสรุปได้เป็นข้อ ๆ ดังนี้

1. กรณีขนาดตัวอย่างเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย

2. กรณีค่าเฉลี่ยประชากรมีค่าแตกต่างกันมาก (μ_1, μ_2) = (12, 10) กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีค่าเฉลี่ยประชากรที่มีค่าแตกต่างกันน้อย (μ_1, μ_2) = (9, 10)

3. กรณีความแปรปรวนประชากรมีค่าเท่ากัน (σ_1^2, σ_2^2) = (1, 1) กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีความแปรปรวนประชากรที่มีค่าไม่เท่ากัน (σ_1^2, σ_2^2) = (4, 16) และ (16, 4)

4. กรณีระดับนัยสำคัญเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย

5. โดยส่วนใหญ่ตัวสถิติทดสอบ Z มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด

6. กรณีตัวอย่างขนาดกลาง = (30, 30), (30, 45) และ (45, 30) ที่มีค่าเฉลี่ยประชากรแตกต่างกันมาก (μ_1, μ_2) = (12, 10) แต่มีความแปรปรวนประชากรเท่ากัน (σ_1^2, σ_2^2) = (1, 1) และกรณีตัวอย่างขนาดใหญ่ (n_1, n_2) = (100, 100), (150, 100) และ (100, 150) ที่มีความแปรปรวนประชากรเท่ากัน (σ_1^2, σ_2^2)= (1, 1) พบว่า เกือบทุกตัวสถิติทดสอบมีกำลังการทดสอบสูงพอ ๆ กัน หรือมีประสิทธิภาพไม่แตกต่างกัน

4. สรุป

จากผลการศึกษาเกี่ยวกับความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดของแต่ละตัวสถิติทดสอบ พบว่า ตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณี ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากัน ทุกระดับอัตราส่วนความแปรปรวนประชากร ทุกการแจกแจง และทุกระดับนัยสำคัญที่ศึกษา ตัวสถิติทดสอบสติวเดนต์ทีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน แต่เมื่อขนาดตัวอย่างไม่เท่ากัน จะควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้เมื่อความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน และไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้เมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า ตัวสถิติทดสอบเวลช์ทีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดเกือบทุกกรณีเมื่อขนาดตัวอย่างเท่ากัน และควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ดีกว่าตัวสถิติทดสอบสติวเดนต์ทีเมื่อตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า และจากผลการศึกษาเกี่ยวกับกำลังการทดสอบของแต่ละตัวสถิติทดสอบ พบว่า กรณีขนาดตัวอย่างเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย กรณีค่าเฉลี่ยประชากรมีค่าแตกต่างกันมาก กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีค่าเฉลี่ยประชากรที่มีค่าแตกต่างกันน้อย กรณีความแปรปรวนประชากรมีค่าเท่ากัน กำลังการทดสอบของทุกตัวสถิติทดสอบจะสูงกว่ากรณีความแปรปรวนประชากรที่มีค่าไม่เท่ากัน กรณีระดับนัยสำคัญเพิ่มขึ้น กำลังการทดสอบของทุกตัวสถิติทดสอบจะเพิ่มขึ้นด้วย โดยส่วนใหญ่ตัวสถิติทดสอบ Z มีกำลังการทดสอบสูงที่สุดในกลุ่มที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนด และกรณีตัวอย่างขนาดกลางไม่ว่าขนาดตัวอย่างจะเท่ากันหรือไม่ก็ตาม ที่มีค่าเฉลี่ยประชากรแตกต่างกันมาก แต่มีความแปรปรวนประชากรเท่ากัน และกรณีตัวอย่างขนาดใหญ่ทั้งกรณีขนาดตัวอย่างเท่ากันและไม่เท่ากันที่มีความแปรปรวนประชากรเท่ากัน พบว่า เกือบทุกตัวสถิติทดสอบมีกำลังการทดสอบสูงพอ ๆ กัน หรือมีประสิทธิภาพไม่แตกต่างกัน ดังนั้นจากผลการศึกษาครั้งนี้สรุปได้ว่า ตัวสถิติทดสอบ Z สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ที่กำหนดและมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบอื่นๆ อย่างไรก็ตาม ตัวสถิติทดสอบ Z อาจไม่สามารถดำเนินการได้จริงในการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน เนื่องจากตัวสถิติทดสอบ Z จะอยู่ภายใต้ข้อสมมุติว่า ต้องทราบค่าความแปรปรวนประชากรทั้ง 2 กลุ่ม ดังนั้นตัวสถิติทดสอบที่มีความเหมาะสมรองลงมา ควรจะเป็นตัวสถิติทดสอบสติวเดนต์ทีหรือตัวสถิติทดสอบเวลช์ที ที่ไม่มีข้อสมมุติเกี่ยวกับความแปรปรวนประชากร โดยกรณีความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าเท่ากัน ตัวสถิติทดสอบสติวเดนต์ทีจะมีความเหมาะสมมากกว่าตัวสถิติทดสอบเวลช์ที แต่กรณีตัวอย่างที่มีขนาดเล็กกว่ามีความแปรปรวนประชากรมากกว่า หรือความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน ตัวสถิติทดสอบเวลช์ทีจะมีความเหมาะสมมากกว่า ทั้งนี้เพราะตัว

สถิติทดสอบเวลช์ทีเป็นตัวสถิติทดสอบที่พัฒนาขึ้นมาเพื่อใช้ในกรณีที่ความแปรปรวนประชากรทั้ง 2 กลุ่มมีค่าไม่เท่ากัน กรณีการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่อิสระกัน โดยที่ตัวอย่างทั้ง 2 กลุ่มมีขนาดใหญ่และความแปรปรวนประชากรมีค่าเท่ากัน จะสามารถใช้ตัวสถิติทดสอบใดก็ได้ระหว่างตัวสถิติทดสอบ Z ตัวสถิติทดสอบสติวเดนต์ที หรือตัวสถิติทดสอบเวลช์ที เนื่องจากผลการศึกษาครั้งนี้พบว่า ตัวสถิติทดสอบเหล่านี้มีกำลังการทดสอบสูงหรือมีประสิทธิภาพไม่แตกต่างกัน

การศึกษาครั้งนี้ได้พิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยใช้เกณฑ์การทดสอบทวินามเพียงเกณฑ์เดียว ดังนั้นในการวิจัยครั้งต่อไป จึงควรศึกษาเกณฑ์อื่น ๆ เพิ่มเติม เช่น เกณฑ์ของแบรดลีย์ (Bradley) ซึ่งผลการศึกษาที่ได้อาจมีความแตกต่างกัน การวิจัยครั้งต่อไป ควรศึกษาประสิทธิภาพของสถิติทดสอบไม่อิงพารามิเตอร์ เช่น การทดสอบแมนน์-วิตนีย์ (Mann-Whitney Test) การทดสอบโดยมัธยฐาน (Median Test) การทดสอบแวนเดอวอลเด็น (Van der Waerden Test) การทดสอบคอลโมโกรอฟ-สมีร์นอฟ (Kolmogorov-Smirnov Test) การทดสอบรันส์ของวัลด์-วอลโฟวิตซ์ (Wald-Wolfowitz Runs Test) และการทดสอบบรุนเนอร์-มุนเซล (Brunner-Munzel Test) เป็นต้น เพื่อเปรียบเทียบกับผลของสถิติทดสอบอิงพารามิเตอร์ที่ได้นำเสนอในการวิจัยครั้งนี้ รวมถึงควรศึกษาประสิทธิภาพของการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่ไม่อิสระกัน และการทดสอบความแตกต่างของค่าเฉลี่ยประชากรตั้งแต่ 2 กลุ่มที่อิสระกันและไม่อิสระกัน

5. กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ กองทุนวิจัยมหาวิทยาลัยทักษิณ ประจำปีงบประมาณ พ.ศ. 2559

6. เอกสารอ้างอิง

[1] W. Riansut, Nonparametric Statistics. Phatthalung: Thaksin University, 2015.

[2] J.S. Ramberg, P.R. Tadikamalla, E.J. Dudewicz and E.F. Mykytka, “A probability distribution and its uses in fitting data,” Technometrics, Vol. 21, no. 2, pp. 201-214, 1979.

[3] K. Vanichbuncha, Principal Statistics, 7th ed. Bangkok: Chulalongkorn University Printing Houses, 2002.

[4] J. Seksan, “Effect of failure to meet the assumptions of homogeneity of variances on type I error in two population tests of difference,” M.S. thesis, Dept. Stat., Kasetsart University, Thailand, 2008.