วารสารวิทยาศาสตร์บูรพา ปีที่ 26 (ฉบับที่ 2) พฤษภาคม – สิงหาคม พ.ศ. 2564
BURAPHA SCIENCE JOURNAL Volume 26 (No.2) May – August 2021
บทความวิจัย

การเปรียบเทียบประสิทธิภาพของสถิติทดสอบความเป็นอิสระภายใต้การแจกแจงของข้อมูลและระดับความสัมพันธ์ที่ต่างกัน

Efficiency Comparison of Statistical Tests of Independence under the Differences of Data Distributions and Relationships

วิโรจน์ มงคลเทพ*
Wiroj Mongkolthep*
สาขาวิทยาศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยีการเกษตร มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา น่าน
Department of Science, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna Nan

Received : 1 June 2020
Revised : 18 August 2020
Accepted : 2 September 2020

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบความเป็นอิสระภายใต้การแจกแจงของข้อมูลและระดับความสัมพันธ์ที่ต่างกัน โดยพิจารณาจากการควบคุมค่าความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบ 5 วิธี คือ สถิติทดสอบไคกำลังสองของเพียร์สัน สถิติทดสอบของเยตส์ สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง และสถิติทดสอบของเนย์แมน ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล จำนวน 10,000 ครั้ง ภายใต้เงื่อนไขที่ศึกษา คือ ข้อมูลในตารางการจรขนาด 2x2 ที่มีการแจกแจงเอกรูป เบ้ปานกลาง และเบ้มาก ระดับความสัมพันธ์น้อย ปานกลาง และมาก ขนาดตัวอย่างเท่ากับ 20, 30, 40, 50, 100, 150, 200 และ 300 ระดับนัยสำคัญเท่ากับ 0.01 และ 0.05 ผลการศึกษาพบว่า ภายใต้การแจกแจงของข้อมูลและระดับความสัมพันธ์ที่ต่างกันในทุกขนาดตัวอย่างและทุกระดับนัยสำคัญ ส่วนใหญ่สถิติทดสอบของเนย์แมนและสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีประสิทธิภาพดีกว่าสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง สถิติทดสอบไคกำลังสองของเพียร์สัน และสถิติทดสอบของเยตส์ ตามลำดับ และในทุกการแจกแจงของข้อมูล กำลังการทดสอบของสถิติทดสอบทั้ง 5 วิธี จะแปรผันตามระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญ โดยมีกำลังการทดสอบมากขึ้น เมื่อระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญเพิ่มขึ้น

คำสำคัญ : ความผิดพลาดแบบที่ 1 ; กำลังการทดสอบ ; สถิติทดสอบความเป็นอิสระ ; ตารางการจร

Abstract

The purpose of this research is to compare the efficiency of statistical tests of independence under the differences of data distributions and relationships. These tests are considered type I error and power of a test of five statistical tests which are Pearson's chi-squared test, Yates's test, the log likelihood ratio test, the modified log likelihood ratio test and the Neyman's test. The simulation is used Monte Carlo technique 10,000 times. Under the condition of 2x2 contingency tables from uniform, moderately skewed and highly skewed distribution with weak, medium and strong relationships. The sample sizes are 20, 30, 40, 50, 100, 150, 200 and 300. The significance levels are 0.01 and 0.05. The results show that with the differences of data distribution and relationships for all sample sizes and significance levels, the Neyman's test and the log likelihood ratio test are more efficient than the modified log likelihood ratio test, Pearson's chi-squared test and Yates's test, respectively. For all data distributions, power of a test of five statistical tests vary with relationships, sample sizes and significance levels and are more as relationships, sample sizes and significance levels increase.

Keywords : type I error, power of a test, statistical test of independence, contingency table

***Corresponding author. E-mail :** wiroj.mongkolthep@gmail.com

บทนำ

การทดสอบความเป็นอิสระ (test of independence) ของตัวแปร 2 ตัวที่เป็นตัวแปรจำแนกประเภท (categorical variable) ทั้งสองตัวแปรในข้อมูลชนิดตารางสองทางหรือตารางการจรขนาด rxc (rxc contingency table) สถิติทดสอบที่นิยมใช้ในการวิเคราะห์ข้อมูล ได้แก่ สถิติทดสอบไคกำลังสองของเพียร์สัน (Pearson's chi-squared test statistic) ของ Pearson (1900) ซึ่งเป็นผู้ค้นพบคนแรกและเสนอหลักการของทดสอบว่า ผลรวมกำลังสองของค่าความแตกต่างระหว่าง ความถี่ที่สังเกตได้ (observed frequency: O_{ij}) กับความถี่คาดหมาย (expected frequency: E_{ij}) หารด้วยความถี่คาดหมาย โดยมีข้อตกลงเบื้องต้นคือ ข้อมูลที่ใช้ในการทดสอบเป็นข้อมูลที่ได้จากการสุ่ม ข้อมูลจะถูกแจกแจงในกรณีต่างๆ ได้เพียงกรณีเดียว ทั้งแถวและสดมภ์ ต้องมีความถี่คาดหมายในแต่ละเซลล์ของตารางมากกว่าหรือเท่ากับ 5 ในกรณีที่ความถี่คาดหมายน้อยกว่า 5 จะใช้ได้เมื่อความถี่คาดหมายมีค่ามากกว่า 1 และน้อยกว่า 5 แต่ไม่เกินร้อยละ 20 ของจำนวนแถวและสดมภ์ และการแก้ปัญหาในกรณีนี้ให้ทำการรวมกลุ่มย่อยของตัวแปรที่มีลักษณะใกล้เคียงเข้าด้วยกัน (Suwansantisuk, 2009; Chaimay, 2010) และจากงานวิจัยของ Bawornkitiwong & Srisuttiyakorn (2016) พบว่า กรณีตารางการจรขนาด 2x2 ที่มีจำนวน 4 เซลล์ ถ้ามี 2 เซลล์ที่มีความถี่คาดหมายน้อยกว่า 5 และไม่มีเซลล์ใดมีความถี่คาดหมายน้อยกว่า 1 สถิติทดสอบไคกำลังสองของเพียร์สันยังมีความแกร่ง (robustness) แต่เมื่อไรก็ตามที่มีความถี่คาดหมายน้อยกว่า 5 ปรากฏในตารางข้อมูลเพียง 1 ค่าเท่านั้น จะถือว่าผลการทดสอบขาดความน่าเชื่อถือ

ในกรณีที่ความถี่คาดหมายมีขนาดเล็ก Yates (1984) ได้เสนอให้ใช้ค่าปรับแก้ด้วย 0.5 เรียกค่าปรับแก้นี้ว่าค่าปรับให้ต่อเนื่อง (continuity correction) ทั้งนี้ นักสถิติได้มีความคิดเห็นแตกต่างกันหลายกลุ่ม บางกลุ่มเห็นด้วยว่าควรใช้ บางกลุ่มเห็นว่าไม่ควรใช้ เมื่อใช้แล้วยิ่งจะทำให้ผลการทดสอบมีความคลาดเคลื่อนสูง ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ การเสนอค่าปรับแก้ตามสูตรของ Yates ใช้ได้เฉพาะตารางการจรขนาด 2x2 และยังไม่สามารถแก้ปัญหาดังกล่าวได้ (Suwansantisuk, 2009)

จากงานวิจัยของ Wichiencharoen (2000) ได้แนะนำผู้ใช้งานที่ไม่มีความรู้เดิมหรือไม่ทราบการแจกแจงก่อน (prior distribution) เกี่ยวกับความเป็นอิสระของข้อมูล ให้เลือกใช้สถิติทดสอบสำหรับตารางการณ์ขนาด 2x2 ที่มีกำลังการทดสอบสูงสุดในทุกระดับความสัมพันธ์ของข้อมูลและทุกระดับนัยสำคัญ พบว่า เมื่อขนาดตัวอย่างเท่ากับ 50 ควรใช้สถิติทดสอบไคกำลังสองของเพียร์สัน สำหรับขนาดตัวอย่างเท่ากับ 100 และ 150 ควรใช้สถิติทดสอบไคกำลังสองของเพียร์สันหรือสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น (The log likelihood ratio test statistic) และ Sulewski & Motyka (2015) ได้ศึกษากำลังการทดสอบของสถิติทดสอบความเป็นอิสระ กรณีตารางการจรขนาด 2x2 ขนาดตัวอย่างเท่ากับ 100 ในทุกระดับความสัมพันธ์ของข้อมูล พบว่า สถิติทดสอบของเนย์แมน (The Neyman's test statistic) และสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง (The modified log likelihood ratio test statistic) มีกำลังการทดสอบสูงกว่าสถิติทดสอบของฟรีแมนทูคีย์ (Freeman-Tukey's test statistic) สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น (The log likelihood ratio test statistic) สถิติทดสอบของแครซีและรีด (Cressie & Read's test statistic) ตามลำดับ

จากผลสรุปดังกล่าวจะเห็นว่าปัญหาในการทดสอบความเป็นอิสระแต่ละเงื่อนไขต่างๆ นั้น นักสถิติหลายท่านพยายามที่จะนำเสนอและทำการเปรียบเทียบสถิติทดสอบความเป็นอิสระเพื่อให้สามารถควบคุมสถานการณ์ต่างๆ ที่อาจจะเกิดขึ้นในการวิจัยได้ และผลจากการศึกษาเหล่านั้นยังไม่มีข้อสรุปว่าสถิติทดสอบวิธีใดที่มีประสิทธิภาพหรือสามารถแก้ไข

ปัญหาภายใต้เงื่อนไขที่กำหนดได้อย่างชัดเจนสำหรับตารางการจรขนาด 2x2 ดังนั้น การศึกษาครั้งนี้ผู้วิจัยจึงให้ความสนใจในการเปรียบเทียบประสิทธิภาพของสถิติทดสอบความเป็นอิสระ 5 วิธี ภายใต้การแจกแจงของข้อมูลและระดับความสัมพันธ์ที่ต่างกัน คือ สถิติทดสอบไคกำลังสองของเพียร์สัน สถิติทดสอบของเยตส์ สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง และสถิติทดสอบของเนย์แมน ซึ่งส่วนใหญ่สถิติทดสอบเหล่านี้มีกำลังการทดสอบสูงในแต่ละสถานการณ์ โดยพิจารณาจากการควบคุมค่าความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบทั้ง 5 วิธี จำลองข้อมูลด้วยเทคนิคมอนติคาร์โลซึ่งเป็นวิธีที่สามารถควบคุมการทดลองภายใต้เงื่อนไขต่างๆ ตามที่ต้องการได้

วิธีดำเนินการวิจัย

1. สถิติทดสอบความเป็นอิสระที่ศึกษา 5 วิธี มีรายละเอียดดังนี้

ถ้าตัวแปรที่ 1 และตัวแปรที่ 2 เป็นตัวแปรจำแนกประเภท โดยที่ตัวแปรที่ 1 มี r ระดับ(แถว) และตัวแปรที่ 2 มี c ระดับ(สดมภ์) แล้วตารางการจร จะมีขนาด $r \times c$ ถ้าเป็นตารางการจรที่มี 2 แถว และ 2 สดมภ์ จะเรียกว่า ตารางการจรขนาด 2x2 โดยมีสมมุติฐานในการทดสอบความเป็นอิสระของตัวแปร คือ

สมมุติฐานว่าง H_0 : ตัวแปรที่ 1 กับตัวแปรที่ 2 เป็นอิสระต่อกัน ($p_{ij} = p_{i.}p_{.j}$)

สมมุติฐานทางเลือก H_1 : ตัวแปรที่ 1 กับตัวแปรที่ 2 ไม่เป็นอิสระต่อกัน ($p_{ij} \neq p_{i.}p_{.j}$)

เมื่อ p_{ij} เป็นความน่าจะเป็นที่ข้อมูลจะอยู่ในแถวที่ i ของตัวแปรที่ 1 และสดมภ์ที่ j ของตัวแปรที่ 2

$p_{i.}$ เป็นความน่าจะเป็นรวมของข้อมูลจะอยู่ในแถวที่ i ของตัวแปรที่ 1 และทุกสดมภ์ของตัวแปรที่ 2

$p_{.j}$ เป็นความน่าจะเป็นรวมของข้อมูลจะอยู่ในสดมภ์ที่ j ของตัวแปรที่ 2 และทุกแถวของตัวแปรที่ 1

1.1 สถิติทดสอบไคกำลังสองของเพียร์สัน (Pearson's chi-squared test statistic: Chi-squared) เป็นสถิติทดสอบความเป็นอิสระที่เสนอโดย Pearson (1900) คำนวณจาก

$$\text{Chi-squared} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \qquad (1)$$

เมื่อ O_{ij} เป็นความถี่ที่สังเกตได้ของแถวที่ i ของตัวแปรที่ 1 และสดมภ์ที่ j ของตัวแปรที่ 2

E_{ij} เป็นความถี่คาดหมายของแถวที่ i ของตัวแปรที่ 1 และสดมภ์ที่ j ของตัวแปรที่ 2, $E_{ij} = O_{i.}O_{.j} / n$

$O_{i.}$ เป็นผลรวมของความถี่ที่สังเกตได้ของแถวที่ i ของตัวแปรที่ 1 และทุกสดมภ์ของตัวแปรที่ 2

$O_{.j}$ เป็นผลรวมของความถี่ที่สังเกตได้ของสดมภ์ที่ j ของตัวแปรที่ 2 และทุกแถวของตัวแปรที่ 1

n เป็นผลรวมของความถี่ที่สังเกตได้ทั้งหมดหรือจำนวนตัวอย่างทั้งหมด

r เป็นจำนวนแถวของตัวแปรที่ 1

c เป็นจำนวนสดมภ์ของตัวแปรที่ 2

1.2 สถิติทดสอบไคกำลังสองที่ใช้ค่าปรับแก้ของเยตส์หรือสถิติทดสอบของเยตส์ (Yates's chi-squared test statistic หรือ Yates's test statistic: Yates) เป็นสถิติทดสอบความเป็นอิสระที่นำเสนอโดย Yates (1984) ซึ่งได้ปรับปรุงสถิติทดสอบไคกำลังสองของเพียร์สันโดยใช้ค่าปรับแก้ด้วยค่าคงที่ 0.5 เรียกค่านี้ว่า ค่าปรับให้ต่อเนื่อง (continuity correction) คำนวณจาก

$$\text{Yates} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(\left|O_{ij}-E_{ij}\right|-0.5\right)^2}{E_{ij}} \qquad (2)$$

1.3 สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น (The log likelihood ratio test statistic: G^2) เป็นสถิติทดสอบความเป็นอิสระที่ถูกพัฒนามาจากสถิติทดสอบของ Cressie & Read (1984) จากสูตร

$$D^{\lambda} = \frac{2}{\lambda(\lambda+1)}\sum_{i=1}^{r}\sum_{j=1}^{c}O_{ij}\left[\left(\frac{O_{ij}}{E_{ij}}\right)^{\lambda}-1\right] \qquad (3)$$

ถ้า λ เป็นจำนวนจริงใดๆ ที่มีค่าเข้าใกล้ 0 แล้วสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น คำนวณจาก (Sulewski & Motyka, 2015)

$$G^2 = 2\sum_{i=1}^{r}\sum_{j=1}^{c}O_{ij}\ln\left(\frac{O_{ij}}{E_{ij}}\right) \qquad (4)$$

1.4 สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง (The modified log likelihood ratio test statistic: Modified G^2) เป็นสถิติทดสอบความเป็นอิสระที่ปรับปรุงจากสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น โดยที่ λ มีค่าเข้าใกล้ –1 ในสมการ (3) คำนวณจาก (Sulewski & Motyka, 2015)

$$\text{Modified } G^2 = 2\sum_{i=1}^{r}\sum_{j=1}^{c}E_{ij}\ln\left(\frac{E_{ij}}{O_{ij}}\right) \qquad (5)$$

1.5 สถิติทดสอบไคกำลังสองที่ปรับปรุงของเนย์แมนหรือสถิติทดสอบของเนย์แมน (The Neyman modified chi-squared test statistic หรือ The Neyman's test statistic: Neyman) เป็นสถิติทดสอบความเป็นอิสระที่นำเสนอโดย Neyman (1949) ซึ่งได้ปรับปรุงสถิติทดสอบไคกำลังสองของเพียร์สัน คำนวณจาก

$$\text{Neyman} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij} - E_{ij}\right)^2}{O_{ij}} \qquad (6)$$

การสรุปผลการทดสอบสมมุติฐานของสถิติทดสอบความเป็นอิสระทั้ง 5 วิธี พิจารณาจากค่าสถิติทดสอบที่คำนวณได้นั้นมีค่ามากกว่าค่าวิกฤติ ณ ระดับนัยสำคัญที่กำหนด ($\chi^2_{0.01,1}$ = 6.635, $\chi^2_{0.05,1}$ = 3.841) จะทำการปฏิเสธสมมุติฐานว่างจึงสรุปได้ว่า ตัวแปรทั้งสองไม่เป็นอิสระต่อกันหรือมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ณ ระดับนัยสำคัญที่กำหนด

2. ข้อมูลที่ใช้ในการศึกษาได้จากการจำลองข้อมูล (simulation) ด้วยเทคนิคมอนติคาร์โล (monte carlo technique) บนโปรแกรม R เวอร์ชัน 3.6.2 ภายใต้เงื่อนไขที่ศึกษา ดังนี้

2.1 ข้อมูลในตารางการจรที่ศึกษา ขนาด 2x2 ที่มีการแจกแจงเอกรูป เบ้ปานกลาง และเบ้มาก

2.2 ความสัมพันธ์ของข้อมูล 3 ระดับ คือ น้อย ปานกลาง และมาก

2.3 ขนาดตัวอย่างที่ศึกษาเท่ากับ 20, 30, 40, 50, 100, 150, 200 และ 300

2.4 ระดับนัยสำคัญของการทดสอบ (α) 2 ระดับ คือ 0.01 และ 0.05

2.5 การทดลองซ้ำจำนวน 10,000 ครั้งในแต่ละสถานการณ์ที่กำหนด

3. การจำลองข้อมูลตารางการจรขนาด 2x2 สำหรับการทดสอบความเป็นอิสระของตัวแปร 2 ตัว จะมีคุณสมบัติว่าความน่าจะเป็นตามขอบ (marginal probability) $p_{ij} = p_{i.}p_{.j}$ ทำให้เลขสุ่มแต่ละตัวถูกกำหนดให้อยู่ในแต่ละเซลล์ด้วยการเทียบกับค่าความน่าจะเป็นสะสม คือ $[0, p_{11}), [p_{11}, p_{11}+p_{12}), [p_{11}+p_{12}, p_{11}+p_{12}+p_{21}), [p_{11}+p_{12}+p_{21}, 1)$ ตามลำดับ ถ้าความน่าจะเป็นตามขอบเท่ากัน จะทำให้ข้อมูลมีการแจกแจกเอกรูป นั่นคือ ข้อมูลไม่มีความเบ้และมีค่าคาดหมายเท่ากันทุกค่า หากความน่าจะเป็นตามขอบไม่เท่ากัน ($p_{ij} \neq p_{i.}p_{.j}$) จะทำให้ข้อมูลมีการแจกแจกแบบเบ้และมีค่าคาดหมายไม่เท่ากัน (Bawornkitiwong & Srisuttiyakorn, 2016) และภายใต้ระดับความสัมพันธ์ของข้อมูล 3 ระดับ (τ) คือ 0.25 (น้อย), 0.50 (ปานกลาง) และ 0.75 (มาก) พิจารณาจากค่า Goodman and Kruskal's tau จากสูตร (7) (Wichiencharoen, 2000)

$$\tau = \frac{\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij}^2 \Big/ p_{i.} - \sum_{j=1}^{c} p_{.j}^2}{1 - \sum_{j=1}^{c} p_{.j}^2} \qquad (7)$$

ตัวอย่างการจำลองข้อมูลเมื่อตัวแปรทั้งสองเป็นอิสระต่อกัน และมีการแจกแจงเอกรูป เมื่อเลขสุ่ม (x) มีการแจกแจงเอกรูปที่มีค่าอยู่ระหว่าง 0 และ 1 จากโปรแกรม R จะได้ค่า p_{ij} เป็น

0.5	0.5		
0.25	0.25	0.5	$0 \le x < 0.25$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 1
0.25	0.25	0.5	$0.25 \le x < 0.5$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 2
			$0.5 \le x < 0.75$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 1
			$0.75 \le x < 1$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 2

สำหรับการจำลองข้อมูลเมื่อตัวแปรทั้งสองไม่เป็นอิสระต่อกัน มีการแจกแจกแบบเบ้ปานกลาง และมีความสัมพันธ์ระดับปานกลาง (τ = 0.50) จะได้ค่า p_{ij} เป็น

0.4	0.6		
0.2787	0.0213	0.3	$0 \le x < 0.2787$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 1
0.1213	0.5787	0.7	$0.2787 \le x < 0.3$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 2
			$0.3 \le x < 0.4213$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 1
			$0.4213 \le x < 1$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 2

กรณีการจำลองข้อมูลเมื่อตัวแปรทั้งสองไม่เป็นอิสระต่อกัน มีการแจกแจกแบบเบ้มาก และมีความสัมพันธ์ระดับน้อย (τ = 0.25) จะได้ค่า p_{ij} เป็น

0.1	0.9		
0.055	0.045	0.1	$0 \le x < 0.055$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 1
0.045	0.855	0.9	$0.055 \le x < 0.1$ ค่าของ x จะตกอยู่ในแถวที่ 1 สดมภ์ที่ 2
			$0.1 \le x < 0.145$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 1
			$0.145 \le x < 1$ ค่าของ x จะตกอยู่ในแถวที่ 2 สดมภ์ที่ 2

4. การเปรียบเทียบประสิทธิภาพของสถิติทดสอบ พิจารณาจากค่าความผิดพลาดแบบที่ 1 และกำลังการทดสอบของสถิติทดสอบ มีวิธีการดังนี้

4.1 ความสามารถในการควบคุมค่าความผิดพลาดแบบที่ 1 จากข้อมูลที่จำลองขึ้นภายใต้สมมุติฐานว่างที่ว่าตัวแปรทั้งสองเป็นอิสระต่อกัน แล้วทดสอบสมมุติฐานเพื่อหาค่าความผิดพลาดแบบที่ 1 จากอัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมุติฐานว่างในการทดลองซ้ำจำนวน 10,000 ครั้ง เทียบกับเกณฑ์ของ Bradley (1978) อยู่ในช่วง [0.5α, 1.5α] สำหรับการทดสอบที่ระดับนัยสำคัญ (α) ดังนั้น สถิติทดสอบที่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ที่ระดับนัยสำคัญ 0.01 และ 0.05 จะต้องมีค่าอยู่ในช่วง [0.005, 0.015] และ [0.025, 0.075] ตามลำดับ

4.2 หากำลังการทดสอบจากข้อมูลที่จำลองขึ้นภายใต้สมมุติฐานทางเลือกที่ว่า ตัวแปรทั้งสองไม่เป็นอิสระต่อกัน แล้วทดสอบสมมุติฐานเพื่อหากำลังการทดสอบ จากอัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมุติฐานว่างในการทดลองซ้ำจำนวน 10,000 ครั้ง แล้วทำการเปรียบเทียบกำลังการทดสอบของสถิติทดสอบที่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้เท่านั้น ดังนั้น สถิติทดสอบที่มีกำลังการทดสอบสูงสุดจะเป็นสถิติทดสอบที่ดีที่สุด

ผลการวิจัย

1. การควบคุมค่าความผิดพลาดแบบที่ 1 ของสถิติทดสอบความเป็นอิสระ

พิจารณาตารางที่ 1 ตามเกณฑ์การควบคุมค่าความผิดพลาดแบบที่ 1 ของ Bradley พบว่า เมื่อข้อมูลมีการแจกแจงเอกรูป สถิติทดสอบไคกำลังสองของเพียร์สันและสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในทุกกรณี ส่วนสถิติทดสอบของเยตส์ที่ไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในกรณีขนาดตัวอย่างไม่เกิน 50 ที่ระดับนัยสำคัญ 0.01 กับขนาดตัวอย่างไม่เกิน 40 ที่ระดับนัยสำคัญ 0.05 สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงไม่สามารถควบคุมความผิดพลาดแบบที่ 1 ได้ในกรณีขนาดตัวอย่างไม่เกิน 30 ที่ระดับนัยสำคัญ 0.01 และสถิติทดสอบของเนย์แมนไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ในกรณีขนาดตัวอย่างไม่เกิน 50 ที่ระดับนัยสำคัญ 0.01 กับขนาดตัวอย่างเท่ากับ 20 ที่ระดับนัยสำคัญ 0.05

เมื่อข้อมูลมีการแจกแจงแบบเบ้ปานกลาง สถิติทดสอบไคกำลังสองของเพียร์สัน สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น และสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ในทุกกรณี ส่วนใหญ่สถิติทดสอบของเยตส์ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ในกรณีขนาดตัวอย่างไม่เกิน 100 ที่ระดับนัยสำคัญ 0.01 กับขนาดตัวอย่างไม่เกิน 50 ที่ระดับนัยสำคัญ 0.05 และสถิติทดสอบของเนย์แมนไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ในกรณีขนาดตัวอย่างไม่เกิน 50 ที่ระดับนัยสำคัญ 0.01

สำหรับข้อมูลที่มีการแจกแจงแบบเบ้มาก สถิติทดสอบทั้ง 5 วิธี สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้แตกต่างกันออกไป โดยส่วนใหญ่สถิติทดสอบไคกำลังสองของเพียร์สัน สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้เมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไปที่ระดับนัยสำคัญ 0.05 ส่วนสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นสามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ในทุกกรณี ยกเว้นกรณีขนาดตัวอย่างเท่ากับ 20 ที่ระดับนัยสำคัญ 0.01 และสถิติทดสอบของเนย์แมนสามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้กรณีขนาดตัวอย่างไม่เกิน 50 ในทุกระดับนัยสำคัญ

ตารางที่ 1 ค่าความผิดพลาดแบบที่ 1 ของสถิติทดสอบ จำแนกตามการแจกแจงของข้อมูล ขนาดตัวอย่าง และระดับนัยสำคัญ

การแจกแจงของข้อมูล	n	α = 0.01					α = 0.05				
		Chi-squared	Yates	G^2	Modified G^2	Neyman	Chi-squared	Yates	G^2	Modified G^2	Neyman
เอกรูป	20	0.0098	**0.0020***	0.0099	**0.0166***	**0.0336***	0.0470	**0.0103***	0.0511	0.0627	**0.0873***
	30	0.0107	**0.0020***	0.0138	**0.0185***	**0.0227***	0.0476	**0.0202***	0.0539	0.0599	0.0738
	40	0.0080	**0.0036***	0.0103	0.0148	**0.0187***	0.0481	**0.0197***	0.0483	0.0532	0.0700
	50	0.0084	**0.0031***	0.0104	0.0139	**0.0185***	0.0568	0.0280	0.0573	0.0586	0.0633
	100	0.0108	0.0051	0.0117	0.0120	0.0129	0.0543	0.0322	0.0544	0.0560	0.0573
	150	0.0122	0.0081	0.0122	0.0127	0.0131	0.0491	0.0329	0.0505	0.0525	0.0543
	200	0.0099	0.0072	0.0105	0.0107	0.0116	0.0562	0.0406	0.0563	0.0567	0.0573
	300	0.0105	0.0074	0.0106	0.0107	0.0110	0.0528	0.0409	0.0530	0.0539	0.0547
เบ้ปานกลาง	20	0.0078	**0.0007***	0.0071	0.0090	**0.0211***	0.0270	**0.0048***	0.0313	0.0427	0.0616
	30	0.0063	**0.0010***	0.0071	0.0117	**0.0264***	0.0392	**0.0116***	0.0431	0.0592	0.0693
	40	0.0072	**0.0017***	0.0089	0.0148	**0.0255***	0.0476	**0.0160***	0.0520	0.0599	0.0733
	50	0.0088	**0.0027***	0.0116	0.0150	**0.0246***	0.0504	**0.0223***	0.0551	0.0627	0.0740
	100	0.0087	**0.0047***	0.0092	0.0108	0.0139	0.0517	0.0291	0.0531	0.0564	0.0612
	150	0.0099	0.0050	0.0104	0.0112	0.0128	0.0527	0.0341	0.0536	0.0557	0.0585
	200	0.0093	0.0057	0.0093	0.0102	0.0114	0.0509	0.0369	0.0518	0.0535	0.0549
	300	0.0100	0.0065	0.0101	0.0107	0.0113	0.0536	0.0376	0.0538	0.0549	0.0559
เบ้มาก	20	**0.0175***	**0.0002***	**0.0016***	**0.0016***	0.0050	**0.1604***	**0.0010***	0.0373	**0.0168***	0.0376
	30	**0.0272***	**0.0020***	0.0093	**0.0035***	0.0107	**0.0879***	0.0260	0.0517	0.0421	0.0421
	40	**0.0468***	**0.0047***	0.0063	0.0063	0.0084	**0.1196***	0.0261	0.0685	0.0411	0.0411
	50	**0.0437***	0.0096	0.0124	0.0081	0.0082	**0.1110***	0.0270	0.0545	0.0332	0.0347
	100	**0.0262***	0.0091	0.0095	**0.0045***	**0.0032***	0.0703	0.0254	0.0375	**0.0197***	**0.0159***
	150	**0.0156***	0.0064	0.0051	**0.0026***	**0.0021***	0.0557	0.0254	0.0339	**0.0205***	**0.0159***
	200	**0.0155***	0.0063	0.0059	**0.0023***	**0.0024***	0.0435	**0.0208***	0.0269	**0.0182***	**0.0216***
	300	0.0136	0.0067	0.0060	**0.0044***	**0.0169***	0.0368	**0.0194***	0.0278	0.0339	0.0639

* *หมายถึง* สถิติทดสอบที่ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley

2. กำลังการทดสอบของสถิติทดสอบความเป็นอิสระ

พิจารณาตารางที่ 2 พบว่า กรณีข้อมูลมีการแจกแจงเอกรูป และข้อมูลมีความสัมพันธ์กันน้อย ส่วนใหญ่สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุด เมื่อขนาดตัวอย่างตั้งแต่ 30 ขึ้นไป ที่ระดับนัยสำคัญ 0.05 ในขณะที่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นและสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุด ที่ระดับนัยสำคัญ 0.01 เมื่อขนาดตัวอย่าง 20–30 และ 40–50 ตามลำดับ และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากัน เมื่อขนาด

ตัวอย่างตั้งแต่ 150 ขึ้นไปในทุกระดับนัยสำคัญ ส่วนกรณีข้อมูลมีความสัมพันธ์กันปานกลาง ส่วนใหญ่สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุด เมื่อขนาดตัวอย่างตั้งแต่ 30 ขึ้นไป ที่ระดับนัยสำคัญ 0.05 โดยที่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นและสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุด ที่ระดับนัยสำคัญ 0.01 เมื่อขนาดตัวอย่าง 20–30 และ 40–50 ตามลำดับ และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากันและสูงที่สุดเท่ากับ 1 เมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไปในทุกระดับนัยสำคัญ และกรณีข้อมูลมีความสัมพันธ์กันมาก ส่วนใหญ่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีกำลังการทดสอบสูงสุดในทุกขนาดตัวอย่างที่ระดับนัยสำคัญ 0.01 ส่วนสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุดในทุกขนาดตัวอย่างที่ระดับนัยสำคัญ 0.05 และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากันและสูงที่สุดเท่ากับ 1 เมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไปในทุกระดับนัยสำคัญ

พิจารณาตารางที่ 3 พบว่า กรณีข้อมูลมีการแจกแจงแบบเบ้ปานกลาง และข้อมูลมีความสัมพันธ์กันน้อยในทุกขนาดตัวอย่าง ส่วนใหญ่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.05 และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากันและสูงที่สุดเท่ากับ 1 เมื่อขนาดตัวอย่างตั้งแต่ 200 ขึ้นไปในทุกระดับนัยสำคัญ ส่วนกรณีข้อมูลมีความสัมพันธ์กันปานกลางในทุกขนาดตัวอย่าง สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.05 และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากันและสูงที่สุดเท่ากับ 1 เมื่อขนาดตัวอย่างตั้งแต่ 150 ขึ้นไปในทุกระดับนัยสำคัญ และกรณีข้อมูลมีความสัมพันธ์กันมากในทุกขนาดตัวอย่าง สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดที่ระดับนัยสำคัญ 0.05 และสถิติทดสอบทั้ง 5 วิธีมีกำลังการทดสอบเท่ากันและสูงที่สุดเท่ากับ 1 เมื่อขนาดตัวอย่างตั้งแต่ 150 ขึ้นไปในทุกระดับนัยสำคัญ

พิจารณาตารางที่ 4 พบว่า กรณีข้อมูลมีการแจกแจงแบบเบ้มาก และข้อมูลมีความสัมพันธ์กันน้อย ที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 50 และส่วนใหญ่สถิติทดสอบของเยตส์มีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 150 ขึ้นไป ส่วนที่ระดับนัยสำคัญ 0.05 ส่วนใหญ่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 50 และสถิติทดสอบไคกำลังสองของเพียร์สันมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป ส่วนกรณีข้อมูลมีความสัมพันธ์กันปานกลาง ที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 40 และสถิติทดสอบของเยตส์มีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป ส่วนที่ระดับนัยสำคัญ 0.05 ส่วนใหญ่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 50 และสถิติทดสอบเพียร์สันไคกำลังสองมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป และกรณีข้อมูลมีความสัมพันธ์กันมาก ที่ระดับนัยสำคัญ 0.01 สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 40 และสถิติทดสอบของเยตส์มีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป ส่วนที่ระดับนัยสำคัญ 0.05 ส่วนใหญ่สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างไม่เกิน 50 และสถิติทดสอบไคกำลังสองของเพียร์สันมีกำลังการทดสอบสูงสุดเมื่อขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป

ตารางที่ 2 กำลังการทดสอบของสถิติทดสอบ เมื่อข้อมูลมีการแจกแจงเอกรูป จำแนกตามระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญ

ระดับความสัมพันธ์	n	α = 0.01					α = 0.05				
		Chi-squared	Yates	G^2	Modified G^2	Neyman	Chi-squared	Yates	G^2	Modified G^2	Neyman
น้อย	20	0.3148	-	**0.3158***	-	-	0.5695	-	0.5782	**0.6138***	-
(τ=0.25)	30	0.5829	-	**0.6225***	-	-	0.8017	-	0.8128	0.8228	**0.8422***
	40	0.7431	-	0.7611	**0.7973***	-	0.9146	-	0.9148	0.9209	**0.9366***
	50	0.8564	-	0.7729	**0.8687***	-	0.9609	0.9284	0.9611	0.9620	**0.9660***
	100	0.9970	0.9942	0.9972	0.9975	**0.9977***	0.9992	0.9988	0.9992	0.9992	**0.9993***
	150	**0.9999***	**0.9999***	**0.9999***	**0.9999***	**0.9999***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
ปานกลาง	20	0.7798	-	**0.7851***	-	-	0.9360	-	0.9384	**0.9489***	-
(τ=0.50)	30	0.9700	-	**0.9760***	-	-	0.9952	-	0.9958	0.9961	**0.9969***
	40	0.9967	-	0.9974	**0.9984***	-	0.9998	-	0.9998	0.9998	**0.9999***
	50	0.9997	-	**0.9998***	**0.9998***	-	**0.9999***	**0.9999***	**0.9999***	**0.9999***	**0.9999***
	100	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	150	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
มาก	20	0.9460	-	**0.9460***	-	-	0.9910	-	0.9911	**0.9931***	-
(τ=0.75)	30	0.9992	-	**0.9995***	-	-	**0.9999***	-	**0.9999***	**0.9999***	**0.9999***
	40	**1.0000***	-	**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***	**1.0000***	**1.0000***
	100	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	150	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***

- *หมายถึง* ไม่พิจารณากำลังการทดสอบของสถิติทดสอบที่ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้

***** *หมายถึง* สถิติทดสอบที่มีกำลังการทดสอบสูงสุดในแต่ละสถานการณ์

ตารางที่ 3 กำลังการทดสอบของสถิติทดสอบ เมื่อข้อมูลมีการแจกแจงแบบเบ้ปานกลาง จำแนกตามระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญ

ระดับความสัมพันธ์	n	α = 0.01					α = 0.05				
		Chi-squared	Yates	G^2	Modified G^2	Neyman	Chi-squared	Yates	G^2	Modified G^2	Neyman
น้อย	20	0.2690	-	0.2516	**0.2948***	-	0.5063	-	0.5186	0.5591	**0.6083***
(τ=0.25)	30	0.5259	-	0.5356	**0.5812***	-	0.7791	-	0.7778	0.8001	**0.8133***
	40	0.7319	-	0.7289	**0.7491***	-	0.8969	-	0.8967	0.9004	**0.9058***
	50	0.8574	-	0.8505	**0.8604***	-	0.9487	-	0.9465	0.9475	**0.9504***
	100	0.9961	-	0.9958	0.9961	**0.9963***	**0.9996***	0.9985	**0.9996***	**0.9996***	**0.9996***
	150	**1.0000***	0.9997	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
ปานกลาง	20	0.6114	-	0.5788	**0.6304***	-	0.8406	-	0.8477	0.8568	**0.8947***
(τ=0.50)	30	0.9200	-	0.9179	**0.9359***	-	0.9801	-	0.9794	0.9811	**0.9839***
	40	0.9985	-	0.9984	**0.9987***	-	0.9981	-	0.9982	0.9982	**0.9983***
	50	**1.0000***	-	**1.0000***	**1.0000***	-	**1.0000***	-	**1.0000***	**1.0000***	**1.0000***
	100	**1.0000***	-	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	150	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
มาก	20	0.9333	-	0.9211	**0.9394***	-	0.9893	-	0.9896	0.9891	**0.9939***
(τ=0.75)	30	0.9989	-	0.9990	**0.9992***	-	0.9998	-	0.9998	0.9998	**0.9999***
	40	**1.0000***	-	**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***	**1.0000***
	100	**1.0000***	-	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	150	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	200	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
	300	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***

- *หมายถึง* ไม่พิจารณากำลังการทดสอบของสถิติทดสอบที่ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้

* *หมายถึง* สถิติทดสอบที่มีกำลังการทดสอบสูงสุดในแต่ละสถานการณ์

ตารางที่ 4 กำลังการทดสอบของสถิติทดสอบ เมื่อข้อมูลมีการแจกแจงแบบเบ้มาก จำแนกตามระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญ

ระดับความสัมพันธ์	n	α = 0.01					α = 0.05				
		Chi-squared	Yates	G^2	Modified G^2	Neyman	Chi-squared	Yates	G^2	Modified G^2	Neyman
น้อย	20	-	-	-	-	**0.1015***	-	-	0.3463	-	**0.3488***
(τ=0.25)	30	-	-	0.2849	-	**0.2929***	-	0.4190	**0.5215***	0.4826	0.4828
	40	-	-	0.3698	0.3696	**0.3763***	-	0.5443	**0.6910***	0.6055	0.6055
	50	-	0.5219	0.5496	0.4885	**0.4899***	-	0.6631	**0.7580***	0.6791	0.6749
	100	-	0.8770	**0.8775***	-	-	**0.9704***	0.9323	0.9468	-	-
	150	-	**0.9713***	0.9674	-	-	**0.9952***	0.9877	0.9900	-	-
	200	-	**0.9931***	0.9922	-	-	**0.9986***	-	0.9977	-	-
	300	**0.9999***	**0.9999***	**0.9999***	-	-	**0.9999***	-	0.9998	0.9997	0.9994
ปานกลาง	20	-	-	-	-	**0.1916***	-	-	0.4968	-	**0.4981***
(τ=0.50)	30	-	-	0.5115	-	**0.5148***	-	0.6580	**0.7105***	0.6900	0.6900
	40	-	-	0.6350	0.6350	**0.6374***	-	0.7763	**0.8758***	0.8102	0.8102
	50	-	0.8059	**0.8138***	0.7553	0.7553	-	0.8751	**0.9185***	0.8777	0.8777
	100	-	**0.9898***	0.9875	-	-	**0.9989***	0.9944	0.9964	-	-
	150	-	**0.9993***	0.9990	-	-	**0.9998***	0.9994	0.9997	-	-
	200	-	**0.9998***	0.9997	-	-	**1.0000***	-	**1.0000***	-	-
	300	**1.0000***	**1.0000***	**1.0000***	-	-	**1.0000***	-	**1.0000***	**1.0000***	**1.0000***
มาก	20	-	-	-	-	**0.2448***	-	-	0.5736	-	**0.5739***
(τ=0.75)	30	-	-	0.2160	-	**0.6729***	-	0.7486	**0.7637***	0.7568	0.7568
	40	-	-	0.8000	0.8000	**0.8003***	-	0.8741	**0.9502***	0.8823	0.8823
	50	-	0.9168	**0.9174***	0.8765	0.8765	-	0.9321	**0.9691***	0.9324	0.9324
	100	-	**0.9989***	0.9986	-	-	**0.9999***	0.9991	0.9993	-	-
	150	-	**0.9999***	**0.9999***	-	-	**1.0000***	**1.0000***	**1.0000***	-	-
	200	-	**1.0000***	**1.0000***	-	-	**1.0000***	-	**1.0000***	-	-
	300	**1.0000***	**1.0000***	**1.0000***	-	-	**1.0000***	-	**1.0000***	**1.0000***	**1.0000***

- หมายถึง ไม่พิจารณากำลังการทดสอบของสถิติทดสอบที่ไม่สามารถควบคุมค่าความผิดพลาดแบบที่ 1 ได้

* หมายถึง สถิติทดสอบที่มีกำลังการทดสอบสูงสุดในแต่ละสถานการณ์

วิจารณ์ผลการวิจัย

ประสิทธิภาพของสถิติทดสอบความเป็นอิสระ 5 วิธี สำหรับตารางการจรขนาด 2x2 ได้แก่ สถิติทดสอบไคกำลังสองของเพียร์สัน สถิติทดสอบของเยตส์ สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็น สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง และสถิติทดสอบของเนย์แมน เมื่อพิจารณาเงื่อนไขภายใต้การแจกแจงของข้อมูลและระดับความสัมพันธ์ที่ต่างกันในทุกขนาดตัวอย่างและทุกระดับนัยสำคัญ พบว่า ไม่มีสถิติทดสอบวิธีใดได้ผลดีที่สุดในทุกๆ กรณีที่ศึกษา โดยส่วนใหญ่สถิติทดสอบของเนย์แมนและสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีประสิทธิภาพดีกว่าสถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง สถิติทดสอบไคกำลังสองของเพียร์สัน และสถิติทดสอบของเยตส์ ตามลำดับ ซึ่งผลการวิจัยนี้สอดคล้องกับการศึกษาของ Sulewski & Motyka (2015) ที่พบว่าในทุกระดับความสัมพันธ์ของข้อมูล สถิติทดสอบของเนย์แมนมีกำลังการทดสอบสูงกว่าสถิติทดสอบวิธีอื่นๆ และสอดคล้องกับการศึกษาของ Wichiencharoen (2000) พบว่าในทุกระดับความสัมพันธ์ของข้อมูลและทุกระดับนัยสำคัญ สถิติทดสอบอัตราส่วนล็อกภาวะน่าจะเป็นมีกำลังการทดสอบสูงสุด สำหรับขนาดตัวอย่างเท่ากับ 100 และ 150

สรุปผลการวิจัย

จากผลการศึกษาครั้งนี้พบว่า ในทุกการแจกแจงของข้อมูล กำลังการทดสอบของสถิติทดสอบทั้ง 5 วิธี จะแปรผันตามระดับความสัมพันธ์ ขนาดตัวอย่าง และระดับนัยสำคัญ โดยมีกำลังการทดสอบมากขึ้น เมื่อระดับความสัมพันธ์ ขนาดตัวอย่างและระดับนัยสำคัญเพิ่มขึ้น ดังนั้น การนำผลการวิจัยครั้งนี้ไปประยุกต์ใช้กับงานวิจัยเพื่อเลือกสถิติทดสอบความเป็นอิสระที่เหมาะสมและมีประสิทธิภาพสำหรับตารางการจรขนาด 2x2 ผู้วิจัยได้สรุปสถิติทดสอบความเป็นอิสระที่มีประสิทธิภาพสูงสุดโดยให้พิจารณาจากการแจกแจงของข้อมูลหรือค่าความน่าจะเป็นตามขอบ (marginal probability) $p_{ij} = p_{i.}p_{.j}$ ว่าข้อมูลมีการแจกแจงแบบใด (ข้อมูลมีการแจกแจงเอกรูป เบ้ปานกลาง หรือเบ้มาก) และพิจารณาระดับความพันธ์ของข้อมูลว่าอยู่ในระดับใด (ข้อมูลมีความสัมพันธ์น้อย ปานกลาง หรือมาก) ซึ่งอาจวัดค่าความสัมพันธ์ได้จากสัมประสิทธิ์ฟี (phi coefficient) หรือค่าแลมบ์ดาของกูดแมนและครัสคัล (Goodman and Kruskal's lambda) เมื่อตัวแปรทั้งสองเป็นตัวแปรที่อยู่ในมาตรานามบัญญัติ (nominal scale) (Khamis, 2008) พร้อมกับระบุขนาดตัวอย่างและระดับนัยสำคัญที่ต้องการทดสอบ จึงจะได้สถิติทดสอบความเป็นอิสระที่มีประสิทธิภาพในการทดสอบสมมุติฐาน ดังแสดงในตารางที่ 5 สำหรับการศึกษาครั้งต่อไปเพื่อเป็นการขยายขอบเขตของการศึกษาสถิติทดสอบความเป็นอิสระให้กว้างขึ้น ควรมีการเปรียบเทียบประสิทธิภาพของสถิติทดสอบความเป็นอิสระภายใต้การแจกแจงของข้อมูลที่มีความเบ้ที่แตกต่างกันมากขึ้น และมีระดับความสัมพันธ์ของข้อมูลละเอียดมากขึ้น และขยายตารางการจรที่ทำการศึกษาให้มีขนาดมากกว่า 2x2 เป็นต้น

ตารางที่ 5 สถิติทดสอบที่สามารถควบคุมความผิดพลาดแบบที่ 1 และให้กำลังการทดสอบสูงสุดสำหรับตารางการจร ขนาด 2x2 ในแต่ละสถานการณ์ที่ศึกษา

การแจกแจงของข้อมูล	n	α = 0.01			α = 0.05		
		τ=0.25	τ=0.50	τ=0.75	τ=0.25	τ=0.50	τ=0.75
เอกรูป	20	G^2	G^2	G^2	Modified G^2	Modified G^2	Modified G^2
	30	G^2	G^2	G^2	Neyman	Neyman	Chi-squared, G^2, Modified G^2, Neyman
	40	Modified G^2	Modified G^2	Chi-squared, G^2, Modified G^2	Neyman	Neyman	Chi-squared, G^2, Modified G^2, Neyman
	50	Modified G^2	G^2, Modified G^2	Chi-squared, G^2, Modified G^2	Neyman	All	All
	100	Neyman	All	All	Neyman	All	All
	150	All	All	All	All	All	All
	200	All	All	All	All	All	All
	300	All	All	All	All	All	All
เบ้ปานกลาง	20	Modified G^2	Modified G^2	Modified G^2	Neyman	Neyman	Neyman
	30	Modified G^2	Modified G^2	Modified G^2	Neyman	Neyman	Neyman
	40	Modified G^2	Modified G^2	Chi-squared, G^2, Modified G^2	Neyman	Neyman	Chi-squared, G^2, Modified G^2, Neyman
	50	Modified G^2	Chi-squared, G^2, Modified G^2	Chi-squared, G^2, Modified G^2	Neyman	Chi-squared, G^2, Modified G^2, Neyman	Chi-squared, G^2, Modified G^2, Neyman
	100	Neyman	Chi-squared, G^2, Modified G^2, Neyman	Chi-squared, G^2, Modified G^2, Neyman	G^2, Modified G^2, Neyman	All	All
	150	Chi-squared, G^2, Modified G^2, Neyman	All	All	All	All	All
	200	All	All	All	All	All	All
	300	All	All	All	All	All	All
เบ้มาก	20	Neyman	Neyman	Neyman	Neyman	Neyman	Neyman
	30	Neyman	Neyman	Neyman	G^2	G^2	G^2
	40	Neyman	Neyman	Neyman	G^2	G^2	G^2
	50	Neyman	G^2	G^2	G^2	G^2	G^2
	100	G^2	Yates	Yates	Chi-squared	Chi-squared	Chi-squared
	150	Yates	Yates	Yates, G^2	Chi-squared	Chi-squared	Chi-squared, G^2, Modified G^2
	200	Yates	Yates	Yates, G^2	Chi-squared	Chi-squared, G^2	Chi-squared, G^2
	300	Chi-squared, Yates, G^2	Chi-squared, Yates, G^2	Chi-squared, Yates, G^2	Chi-squared	Chi-squared, G^2, Modified G^2, Neyman	Chi-squared, G^2, Modified G^2, Neyman

All หมายถึง สถิติทดสอบทั้ง 5 วิธี (Chi-squared, Yates, G^2, Modified G^2, Neyman)

กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จลงได้ด้วยการสนับสนุนจากกลุ่มวิจัยทางสถิติ วิจัย และประเมินผลการศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา น่าน และขอขอบคุณสาขาวิทยาศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยีการเกษตร มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา น่าน สำหรับการเอื้อเฟื้อทรัพยากรต่างๆ และสถานที่ในการดำเนินการวิจัยจนสำเร็จ

เอกสารอ้างอิง

Bradley, J. V. (1978). Robustness. *The British Journal of Mathematical and Statistical Psychology*, *31*(2), 144-152.

Bawornkitiwong, S. & Srisuttiyakorn, S. (2016). Robust of chi-square test. *Journal of Education Studies*, *44*(3), 212-220. (in Thai)

Chaimay, P. (2010). Notice of using chi-square test in health science researches. *Thaksin University Journal*, *13*(1), 55-58. (in Thai)

Cressie, N. & Read, T. (1984). Multinomial goodness of fit tests. *Journal of the Royal Statistical Society*, *Series B (Methodological)*, *46*(3). 440–464.

Khamis, H. (2008). Measures of association: How to choose. *Journal of Diagnostic Medical Sonography*, *24*(3). 155–162.

Neyman, J. (1949). Contribution to the theory of the chi-square test. In *Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability*. (pp. 239-273). Berkeley, CA: University of California Press.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophy Magazine Series*, *1900*(50), 157–175.

Sulewski, P. & Motyka, R. (2015). Power analysis of independence testing for contingency tables. *Scientific Journal of Polish Naval Academy*, *56*(1), 37-46.

Suwansantisuk, A. (2009). Development of the chi-square correction for contingency tables with more than 20% of small expected frequencies. *Academic Journal Uttaradit Rajabhat University*, *4*(10), 206-223. (in Thai)

Wichiencharoen, N. (2000). *Bayesian test of independence for multinomial distribution using independence prior*. Master's Thesis, Faculty of Commerce and Accountancy, Chulalongkorn University. (in Thai)

Yates, F. (1984). Tests of significance for 2x2 contingency tables. *Journal of the Royal Statistical Society*, *147*(3), 426–463.