Journal of Education Studies

Volume 44
Issue 3 *July - September 2016*

Article 14

7-1-2016

ความแกร่งของสถิติทดสอบไคสแควร์ Robustness of the Chi-square Test

สุชาดา บวรกิติวงศ์

สิวะโชติ ศรีสุทธิยากร

วารสารครุศาสตร์
ปีที่ 44 ฉบับที่ 3 (กรกฎาคม-กันยายน) หน้า 212-220

ความแกร่งของสถิติทดสอบไคสแควร์

Robustness of the Chi-square Test

สุชาดา บวรกิติวงศ์ และ สิวะโชติ ศรีสุทธิยากร

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบไคสแควร์แบบเพียร์สัน ผลการวิจัยนำเสนอว่าภายใต้สถานการณ์ใดบ้างที่เพียร์สันไคสแควร์มีความแกร่งสำหรับการทดสอบความเป็นอิสระของตัวแปร 2 ตัว งานวิจัยนี้ใช้การจำลองข้อมูลแบบยูนิฟอร์มที่มีค่าระหว่าง 0 และ 1 จากโปรแกรม SAS ในแต่ละเงื่อนไขใช้นัยสำคัญ 3 ระดับคือ .01, .05, .10 ตัวอย่าง 4 ขนาดคือ 20, 50, 100, 300 สำหรับตารางขนาด 2 × 2, 2 × 4, 2 × 5, 4 × 5 (เฉพาะขนาดตัวอย่าง 100 และ 300) ข้อมูลมีการแจกแจงแบบยูนิฟอร์มจนถึงเบ้มาก แต่ละเงื่อนไขวิเคราะห์ซ้ำ 10,000 ตัวอย่าง

ผลการวิจัยสรุปได้ดังนี้ 1) เพียร์สันไคสแควร์มีความแกร่งเมื่อข้อมูลมีค่า $e_{ij} < 5$ ไม่เกิน 60% และไม่มีค่าใดต่ำกว่า 1 2) เพียร์สันไคสแควร์มีแนวโน้ม ยากต่อการปฏิเสธสมมติฐานหลัก (conservative) เมื่อข้อมูลมีค่า $e_{ij} < 5$ เกินกว่า 60% และไม่มีค่าใดต่ำกว่า 1 3) เพียร์สันไคสแควร์มีแนวโน้มไม่คงเส้นคงวา (inconsistent) เมื่อข้อมูลมีค่า $e_{ij} < 5$ เกินกว่า 60% และมีบางค่าต่ำกว่า 1 ส่วนใหญ่จะมีแนวโน้มง่ายต่อการปฏิเสธสมมติฐานหลัก (liberal) แต่มีบางกรณีที่ยากต่อการปฏิเสธสมมติฐานหลัก (conservative) ในกรณีนี้ไม่แนะนำให้ใช้สถิติทดสอบเพียร์สันไคสแควร์

คำสำคัญ: สถิติทดสอบไคสแควร์ / ความแกร่ง / อัตราความคลาดเคลื่อนประเภทที่ 1

Abstract

The purpose of the current study is to investigate the Type I error rates of the Pearson Chi-square statistic. Results of the study provide recommendations regarding under what conditions the Pearson Chi-square test of independence is robust. For this Monte Carlo study, the data were created using SAS UNIFORM to generate uniform random numbers on the interval from zero to one. For the true null hypotheses, each condition, three significance levels (.01, .05, .10), four sample sizes (20, 50, 100, 300) for 2 × 2, 2 × 4, and 2 × 5 and two sample sizes (100, 300) for 4 × 5 contingency tables were investigated. Three marginal probability distributions ranging from uniform to highly skewed on both rows and columns for each 2 × 2, 2 × 4, and 2 × 5 table and five marginal probability distributions for each 4 × 5 table were examined. For each condition, 10,000 samples were repeated.

Simulation results show that the Pearson Chi-square is robust with respect to Type I error rates, the recommendations on the use of Pearson Chi-square are:(a) the Pearson Chi-square is robust in a sparse table (a table with less than 60% of the eij less than five and none of them less than one). The empirical Type I error rates of the Pearson Chi-square lie within the two standard error range, (b) the Pearson Chi-square tends to be conservative in a moderately sparse table (a table with more than 60% of the e_{ij} less than five and none of them less than one), and (c) the Pearson Chi-square tends to be too liberal and is not recommended in an extremely sparse table (a table with a large variation in the e_{ij} and more than 60% of the e_{ij} less than five and some of them less than one). Therefore, Pearson Chi-square is robust and recommended for a table with less than 60% of $e_{ij} < 5$ and none of them less than one.

KEYWORDS: CHI-SQUARE TEST STATISTIC / ROBUSTNESS / TYPE I ERROR RATE

บทนำ

แม้ว่าการทดสอบไคสแควร์จะถูกใช้มานานนับศตวรรษ แต่ก็ยังมีข้อขัดแย้งในการใช้ กล่าวคือเมื่อความถี่ที่คาดหวังในแต่ละชั้น (กรณีทดสอบภาวะสารูปสนิทดี: goodness of fit test) หรือแต่ละเซล (กรณีทดสอบความเป็นอิสระหรือความเป็นเอกพันธ์: independence or homogeneity test) มีค่าต่ำกว่า 5 การทดสอบไคสแควร์ยังให้ผลที่น่าเชื่อถือหรือไม่

Karl Pearson เป็นคนแรกที่ค้นพบและเสนอการทดสอบไคสแควร์ใน ค.ศ. 1900 เพื่อที่จะเปรียบเทียบความถี่ที่คาดหวังกับความถี่ที่สังเกตได้จากข้อมูล โดยเขาเน้นการประยุกต์ใช้ไคสแควร์ในระยะแรกประกอบด้วย 3 วัตถุประสงค์ คือ (1) การใช้ไคสแควร์เพื่อทดสอบภาวะสารูปสนิทดี (Goodness of Fit Test) เป็นการทดสอบเพื่อดูว่าข้อมูลชุดหนึ่งมีการแจกแจงความถี่เป็นแบบการแจกแจงที่ต้องการทดสอบหรือไม่ (2) การใช้ไคสแควร์เพื่อทดสอบความเป็นเอกพันธ์ของข้อมูล (Homogeneity Test) เป็นการทดสอบเพื่อดูว่าประชากรตั้งแต่ 2 กลุ่มขึ้นไป มีการแจกแจงของสัดส่วนข้อมูลเป็นแบบเดียวกันหรือไม่ (3) การใช้ไคสแควร์เพื่อทดสอบความเป็นอิสระ (Independence Test) เป็นการทดสอบเพื่อดูว่าตัวแปรเชิงคุณภาพ 2 ตัว โดยที่แต่ละตัวถูกจำแนกเป็นหลายระดับ มีความเป็นอิสระกันหรือไม่

ต่อมาในระยะหลังการทดสอบไคสแควร์ได้ถูกประยุกต์ใช้กว้างขวางขึ้น เช่นมีการใช้ไคสแควร์ในการทดสอบตัวแบบโมเดลสมการโครงสร้าง เพื่อทดสอบว่าตัวแบบที่สังเคราะห์ได้จากทฤษฎีมีความสอดคล้องกับข้อมูลเชิงประจักษ์หรือไม่ นอกจากนี้ยังได้มีการประยุกต์ใช้ไคสแควร์ในสถิติที่ไม่ใช้พารามิเตอร์เพิ่มมากขึ้น

งานวิจัยนี้ศึกษาเฉพาะกรณีการทดสอบความเป็นอิสระของตัวแปร 2 ตัว (independence test) หนังสือสถิติส่วนใหญ่จะกล่าวว่าเพียร์สันไคสแควร์ใช้ได้ดีเฉพาะในกรณีที่ค่าคาดหวังในแต่ละเซลของตารางมีค่าเกินกว่า 5 บทความนี้ตั้งใจจะนำเสนอข้อเท็จจริงในกรณีนี้

วัตถุประสงค์

เพื่อตรวจสอบความแกร่งของสถิติทดสอบไคสแควร์แบบเพียร์สันภายใต้สถานการณ์ข้อมูลมีการแจกแจงยูนิฟอร์มจนถึงข้อมูลมีความเบ้มาก

วิธีการวิจัย

1. จำลองข้อมูลให้มีการแจกแจงแบบยูนิฟอร์มที่มีค่าอยู่ระหว่าง 0 และ 1 จากโปรแกรม SAS เพื่อการทดสอบความเป็นอิสระของตัวแปร 2 ตัว ใช้ระดับนัยสำคัญ 3 ระดับคือ .01, .05, และ .10 ขนาดตัวอย่าง 4 ขนาดคือ 20, 50, 100, 300 สำหรับตาราง 4 ขนาดคือ 2×2, 2×4, 2×5 และ 4×5 (เฉพาะขนาดตัวอย่างขนาดใหญ่คือ 100 และ 300) แต่ละกรณีตั้งแต่ไม่มีความเบ้จนถึงมีความเบ้มาก การ ทดสอบแต่ละเงื่อนไขทำซ้ำ 10,000 รอบ

2. การจำลองข้อมูลสำหรับตารางการณ์จรขนาด rc สำหรับการทดสอบความเป็นอิสระของตัวแปร 2 ตัว จะมีคุณสมบัติว่า marginal probabilities $p(RC_{ij}) = p(R_i)\ p(C_j)$. ทำให้เลขสุ่มแต่ละตัวจะถูกกำหนดให้อยู่ในแต่ละเซลด้วยการเทียบกับค่าความน่าจะเป็นสะสมดังนี้

$[0, p_1), [p_1, p_1 + p_2), ..., [1-p_{rc}, 1)$ ตามลำดับ

ตัวอย่างการจำลองข้อมูลสำหรับตารางขนาด 2 × 2 ที่มี marginal probability เท่ากัน (กรณีข้อมูลไม่มีความเบ้ ค่าคาดหวังเท่ากันทุกค่า)

.5	.5	
.25	.25	.5
.25	.25	.5

เลขสุ่มที่มีค่าระหว่าง .00 และ .25 จะตกในแถวนอนที่ 1 หลักที่ 1
เลขสุ่มที่มีค่าระหว่าง .26 และ .50 จะตกในแถวนอนที่ 1 หลักที่ 2
เลขสุ่มที่มีค่าระหว่าง .51 และ .75 จะตกในแถวนอนที่ 2 หลักที่ 1
เลขสุ่มที่มีค่าระหว่าง .76 และ 1.00 จะตกในแถวนอนที่ 2 หลักที่ 2

สำหรับการจำลองข้อมูลสำหรับตารางขนาด 2 × 2 ที่มี marginal probability ไม่เท่ากัน (กรณีข้อมูลมีความเบ้มาก)

.9	.1	
.72	.08	.8
.18	.02	.2

เลขสุ่มที่มีค่าระหว่าง .00 และ .72 จะตกในแถวนอนที่ 1 หลักที่ 1
เลขสุ่มที่มีค่าระหว่าง .73 และ .80 จะตกในแถวนอนที่ 1 หลักที่ 2
เลขสุ่มที่มีค่าระหว่าง .81 และ .98 จะตกในแถวนอนที่ 2 หลักที่ 1
เลขสุ่มที่มีค่าระหว่าง .99 และ 1.00 จะตกในแถวนอนที่ 2 หลักที่ 2

จำลองข้อมูลในลักษณะเดียวกันสำหรับข้อมูลในตารางขนาดอื่นๆ

3. ในการประเมินความคลาดเคลื่อนประเภทที่ 1 หากค่าความคลาดเคลื่อนประเภทที่ 1 ที่คำนวณได้มีค่าตกในช่วงความเชื่อมั่นสองเท่าของส่วนเบี่ยงเบนมาตรฐาน (two standard error) กล่าวคือ

$$alpha \pm 2\sqrt{\{alpha(1-alpha)/\,number\ of\ replications\}}$$

จะถือว่าการทดสอบครั้งนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 แสดงว่ามีความแกร่ง

ผลการวิจัย

ผลการจำลองข้อมูลพบว่าสถิติทดสอบเพียร์สันไคสแควร์มีความแกร่ง (robust) สามารถควบคุมค่าความคลาดเคลื่อนประเภทที่ 1 ได้ในเกือบทุกกรณี สรุปได้ดังนี้

1. กรณียูนิฟอร์มคือทุกเซลมีค่าคาดหวังเท่ากัน เพียร์สันไคสแควร์จะมีความแกร่งเมื่อ e_{ij} ทุกตัวมีค่ามากกว่า 5

2. กรณียูนิฟอร์ม ในกรณีที่ e_{ij} ทุกตัวมีค่าต่ำกว่า 5 เช่นมีค่าเป็น 2.5 ทุกตัวในตารางขนาด 2×4 และมีค่าเป็น 2 ทุกตัวในตารางขนาด 2×5 (รายละเอียดแสดงในตารางที่ 1) เพียร์สันไคสแควร์จะมีแนวโน้มเป็น conservative ในทุกกรณี กล่าวคือค่าความคลาดเคลื่อนประเภทที่ 1 ที่คำนวณได้จะต่ำกว่าค่า alpha ที่กำหนด ทำให้สรุปได้ว่าเพียร์สันไคสแควร์ขาดความแกร่งเมื่อค่าคาดหวังทุกค่าต่ำกว่า 5

3. กรณีที่ข้อมูลมีความเบ้ระดับปานกลาง (moderately skewed) เพียร์สันไคสแควร์มีแนวโน้ม conservative เมื่อข้อมูลมีค่า $e_{ij} < 5$ เกินกว่า 60% และไม่มีข้อมูลในเซลใดมีค่าต่ำกว่า 1 (รายละเอียดดูจากตารางที่ 2) กล่าวคือค่าความ

คลาดเคลื่อนประเภทที่ 1 ที่คำนวณได้มีแนวโน้มต่ำกว่าค่า alpha ที่กำหนด

4. กรณีที่ข้อมูลมีความเบ้ระดับมาก (extremely skewed) คือข้อมูลมีค่า $e_{ij} < 5$ เกินกว่า 60% และมีบางตัวต่ำกว่า 1 เพียร์สันไคสแควร์จะ inconsistent กล่าวคือบางครั้ง conservative บางครั้ง liberal ในกรณีนี้เพียร์สันไคสแควร์จะขาดความแกร่งและไม่แนะนำให้ใช้ (รายละเอียดดูจากตารางที่ 3 และ 4)

จากข้อค้นพบทำให้สามารถสรุปได้ว่าเพียร์สันไคสแควร์มีความแกร่งสามารถใช้ได้ดีในทุกกรณีที่มีค่าคาดหวัง (e_{ij}) ต่ำกว่า 5 น้อยกว่าร้อยละ 60 และไม่มีเซลใดมีค่าต่ำกว่า 1 ในกรณีที่ค่าคาดหวังต่ำกว่า 5 ทุกค่า (ร้อยละ 100) เพียร์สันไคสแควร์จะมีแนวโน้มยากต่อการปฏิเสธสมมติฐานหลัก (conservative)

ในกรณีที่ข้อมูลเบ้มากคือมีค่าคาดหวังต่ำกว่า 5 เกินกว่าร้อยละ 60 และมีบางค่าต่ำกว่า 1 เพียร์สัน ไคสแควร์จะขาดความแกร่ง กล่าวคือบางครั้งอาจจะ conservative บางครั้งอาจจะ liberal ข้อมูลในลักษณะเช่นนี้จึงไม่แนะนำให้ใช้เพียร์สันไคสแควร์ในการทดสอบสมมติฐาน ตัวอย่างเช่น ข้อมูลในตารางที่ 4 สำหรับตารางการณ์จรขนาด 4 × 5 กรณีที่ 5 เมื่อขนาดตัวอย่างเท่ากับ 100 ข้อมูลมีค่าคาดหวังต่ำกว่า 5 ร้อยละ 80 ความคลาดเคลื่อนประเภทที่ 1 คำนวณได้ .032 (ในขณะที่ alpha = .01) .084 (ในขณะที่ alpha = .05) .133 (ในขณะที่ alpha = .10) เพียร์สันไคสแควร์ขาดความแกร่งคือไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ในช่วงความเชื่อมั่นสองเท่าความคลาดเคลื่อนมาตรฐาน (two standard error confident interval) ได้ทั้ง 3 ระดับนัยสำคัญ นั่นคือค่าความคลาดเคลื่อนไม่ตกในช่วงความเชื่อมั่นดังกล่าว ในกรณีนี้ทำให้เพียร์สันไคสแควร์ขาดความแกร่ง เรียกว่า too liberal คือมีแนวโน้มง่ายต่อการปฏิเสธคือการทดสอบสมมติฐานมีโอกาสปฏิเสธสมมติฐานหลักสูงกว่าค่า alpha ที่ระบุไว้

ในขณะเดียวกัน ตัวอย่างข้อมูลในตารางที่ 4 สำหรับตารางการณ์จรขนาด 4 × 5 กรณีที่ 3 เมื่อขนาดตัวอย่างเท่ากับ 100 ข้อมูลมีค่าคาดหวังต่ำกว่า 5 ร้อยละ 65 ความคลาดเคลื่อนประเภทที่ 1 คำนวณได้ .093 (ในขณะที่ alpha = .10) เพียร์สันไคสแควร์ขาดความแกร่งคือไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ให้มีค่าอยู่ในช่วงความเชื่อมั่นสองเท่าความคลาดเคลื่อนมาตรฐาน (two standard error confident interval) ในกรณีนี้ทำให้เพียร์สันไคสแควร์ขาดความแกร่ง เรียกว่า conservative คือมีแนวโน้มยากต่อการปฏิเสธคือการทดสอบสมมติฐานมีโอกาสปฏิเสธสมมติฐานหลักน้อยกว่าค่า alpha ที่ระบุไว้

สรุปว่าในกรณีข้อมูลเบ้มากคือมีค่าคาดหวังต่ำกว่า 5 เกินกว่าร้อยละ 60 และมีค่าคาดหวังบางตัวต่ำกว่า 1 สถิติทดสอบเพียร์สันไคสแควร์จะขาดคุณสมบัติความคงเส้นคงวา (inconsistent) กล่าวคือผลการทดสอบบางกรณีอาจจะ liberal บางกรณีอาจจะ conservative และในบางกรณีอาจจะ robust ก็ได้ ในกรณีนี้จึงไม่แนะนำให้ใช้

อภิปรายผล

ผลการวิจัยมีความขัดแย้งกับงานวิจัยของ Slakter (1966) ที่กล่าวว่าเพียร์สันไคสแควร์ยังคงสามารถใช้ได้ดีเมื่อค่าคาดหวังทุกตัวมีค่าเท่ากันและต่ำกว่า 5 ผลการศึกษาครั้งนี้ระบุชัดเจนว่าเพียร์สันไคสแควร์จะขาดความแกร่งและมีแนว

ตาราง 1 ค่าความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงยูนิฟอร์ม

ตาราง	N	e_{ij}	Alpha		
			.01	.05	.10
2 × 2	20	5	.010	.050	.121
	50	12.5	.009	.057	.108
	100	25	.012	.056	.101
	300	75	.010	.051	.097
2 × 4	20	2.5	.004	.042	.096
	50	6.25	.009	.048	.104
	100	12.5	.008	.051	.100
	300	37.5	.011	.051	.105
2 × 5	20	2	.004	.034	.093
	50	5	.008	.050	.104
	100	10	.009	.048	.099
	300	30	.010	.051	.100
4 × 5	100	5	.008	.045	.094
	300	15	.009	.050	.096

โน้ม conservative เมื่อค่าคาดหวังทุกตัวมีค่าเท่ากันและต่ำกว่า 5 ผลการวิเคราะห์ข้อมูลแสดงรายละเอียดในตารางที่ 1ในกรณีที่ e_{ij} ทุกตัวมีค่าเป็น 2.5 สำหรับตารางขนาด 2 × 4 และมีค่าเป็น 2 ทุกตัวในตารางขนาด 2 × 5 เพียร์สันไคสแควร์จะมีแนวโน้มเป็น conservative นอกจากนี้ยังพบอีกว่าเพียร์สันไคสแควร์จะมีความแกร่งเมื่อข้อมูลมีค่าคาดหวังต่ำกว่า 5 น้อยกว่าร้อยละ 60 ไม่ว่าค่าคาดหวังจะเท่าหรือต่างกัน

จะอย่างไรก็ตามผลการวิจัยนี้สนับสนุนข้อค้นพบของ Lewontin and Felsentein (1965); Camilli and Hopkins (1978) และ Bradley, Bradley, McGrath, and Cutcomb (1979) ที่กล่าวว่าเมื่อข้อมูลมีค่าคาดหวังต่ำกว่า 5 เกินว่าร้อยละ 60 เพียร์สันไคสแควร์มีแนวโน้มยากต่อการปฏิเสธสมมติฐานหลัก (conservative) แต่ผลการวิจัยครั้งนี้ไม่ได้สนับสนุนความคิดดังกล่าวทั้งหมดกล่าวคือสนับสนุนเฉพาะในกรณีที่ข้อมูลมีค่าคาดหวังต่ำกว่า 5 เกินกว่าร้อยละ 60 แต่ต้องเพิ่มเงื่อนไขว่าไม่มีค่าคาดหวังในเซลใดมีค่าต่ำกว่า 1 เนื่องจากหากมีบางเซลมีค่าคาดหวังต่ำกว่า 1 เพียร์สันไคสแควร์จะ inconsistent และมีแนวโน้ม too liberal (ง่ายต่อการปฏิเสธสมมติฐานหลัก) ไม่ใช่มีแนวโน้ม conservative อย่างที่เขาทั้งสามกล่าวไว้

สรุปว่าเพียร์สันไคสแควร์สามารถใช้ได้ดีมีความแกร่งในหลายกรณี ใช้ได้กว้างกว่าที่คนทั่วไปเข้าใจกันกล่าวคือคนทั่วไปเข้าใจกันว่า

ตาราง 2 ค่าความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงแบบเบ้ปานกลาง

ตาราง	N	$\%e_{ij}<5$	Alpha		
			.01	.05	.10
2 × 2	20	75	.011	.044	.097
	50	25	.009	.046	.100
	100	0	.009	.049	.100
	300	0	.010	.051	.106
2 × 4	20	87.5	.006	.044	.095
	50	37.5	.007	.046	.099
	100	0	.008	.052	.100
	300	0	.009	.048	.098
2 × 5	20	100	.004	.037	.095
	50	60	.007	.044	.093
	100	10	.007	.047	.096
	300	0	.010	.049	.098
4 × 5	100	65	.008	.047	.097
	300	0	.009	.046	.096

ตาราง 3 ค่าความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงแบบเบ้มาก

ตาราง	N	$\%e_{ij}<5$	Alpha		
			.01	.05	.10
2 × 2	20	75	.020	.057	.076
	50	50	.012	.041	.075
	100	25	.010	.040	.090
	300	0	.009	.048	.101
2 × 4	20	87.5	.012	.055	.102
	50	50	.010	.044	.087
	100	37.5	.010	.047	.098
	300	0	.009	.052	.103
2 × 5	20	90	.012	.051	.105
	50	70	.011	.048	.094
	100	40	.010	.049	.100
	300	0	.010	.050	.097

ตาราง 4 ค่าความคลาดเคลื่อนประเภทที่ 1 เมื่อข้อมูลมีการแจกแจงแบบเบ้มาก (ตาราง 4 × 5)

กรณ	N	$\%e_{ij} < 5$	Alpha		
			.01	.05	.10
3	100	65	.010	.045	.093
	300	10	.009	.044	.095
4	100	60	.012	.051	0.97
	300	40	.011	.046	.092
5	100	80	.032	.084	.133
	300	55	.019	.059	.103

เมื่อไรก็ตามที่มีค่าคาดหวังต่ำกว่า 5 ปรากฏในตารางข้อมูลแม้เพียง 1 ค่าเท่านั้นจะถือว่าผลการทดสอบอาจขาดความน่าเชื่อถือ ผลการวิจัยนี้สรุปว่าสำหรับตารางการณ์จรทุกขนาดที่มีค่าคาดหวังต่ำกว่า 5 รวมอยู่ในปริมาณที่น้อยกว่าร้อยละ 60 และไม่มีค่าคาดหวังใดมีค่าน้อยกว่า 1 เพียร์สันไคสแควร์ยังมีความแกร่งสามารถใช้ได้ดี ดังนั้นในกรณีเป็นตารางการณ์จรขนาด 2 × 2 จะมีความถี่จำนวน 4 เซลถ้ามี 2 เซลที่มีค่าคาดหวังต่ำกว่า 5 และไม่มีเซลใดมีค่าคาดหวังต่ำกว่า 1 ผลการวิจัยนี้ระบุว่าเพียร์สันไคสแควร์ยังคงมีความแกร่ง สามารถใช้ได้ในการทดสอบสมมติฐาน เนื่องจากยังมีค่าความคลาดเคลื่อนประเภทที่ 1 ตกในช่วงความเชื่อมั่นสองเท่าของความคลาดเคลื่อนมาตรฐาน (two standard error confident interval) เป็นต้น

รายการอ้างอิง

Bradley, D. R., Bradley, T. D., McGrath, S. G., & Cutcomb, S. D. (1979). Type I error rate of the chi-square test of independence in r × c tables that have small expected frequencies. *Psychological Bulletin, 86,* 1290-1297.

Camilli, G., & Hopkins, K. D. (1978). Aplicability of chi-square to 2 × 2 contingency tables with small expected cell frequencies. *Psychological Bulletin, 85,* 163-167.

Lewontin, R. C., & Felsenstein, J. (1965). The robustness of homogeneity tests in 2 × N tables. *Biometrics, 21,* 19-33.

Slakter, M. J. (1966). Comparative validity of the chi-square and two modified chi-square goodness-of-fit tests for small but equal expected frequencies. *Biometrika, 53,* 619-622.

ผู้เขียน

รองศาสตราจารย์ ดร.สุชาดา บวรกิติวงศ์ ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร อีเมล: Suchada.B@chula.ac.th

อาจารย์ ดร.สิวะโชติ ศรีสุทธิยากร ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร อีเมล: choat.cu@gmail.com