

# ข้อควรพิจารณาในการใช้การทดสอบไคสแควร์ กับงานวิจัยทางด้านสังคมศาสตร์

## The Consideration of using Chi - Square Test for Social Science Research

อำเภอ วิจัย

### บทคัดย่อ

การทดสอบไคสแควร์เป็นวิธีการทดสอบสมมติฐานทางสถิติที่มีการใช้กัน อย่างแพร่หลายในงานวิจัยทางด้านสังคมศาสตร์ สามารถใช้ได้ทั้งการทดสอบความแตกต่างระหว่างกลุ่มประชากรและการทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว ในข้อมูลระดับนามบัญญัติหรือสูงกว่านั้นจึงอาจกล่าวได้ว่าการทดสอบไคสแควร์ทดสอบได้กับข้อมูลทุกระดับ เพียงแต่ต้องเข้าใจในการแปลงข้อมูล ข้อควรระวังที่อาจเกิดขึ้นในการทดสอบไคสแควร์ที่ผู้วิจัยควรให้ความสนใจคือ จำนวนความถี่ในเซลล์ ขนาดตัวอย่างขนาดของตารางคูณไขว้ การประมวลผลและการแปรผล ซึ่งจะทำให้การวิจัยหรือวิทยานิพนธ์มีความถูกต้องและมีความน่าเชื่อถือตามหลักและวิธีการทางสถิติ

### Abstract

Chi - Square Test is the Statistics Hypothesis Metrology that used in Social Science Research. It can be used to test about the difference of group and Relationship group test between two variable Data which are Nominal or higher scale ; That is to say Chi - Square Test is suitable for every data level. However, the researcher has to know how to transform data. The Caution is that Chi - Square Test that the researcher should be interested in is the cell frequency, sample size, Crosstabulation, Computing and interpretation, which will be the good and correct research or thesis.

\* ผู้ช่วยศาสตราจารย์, หัวหน้าหมวดวิชาคณิตศาสตร์ประยุกต์ สำนักวิชาการศึกษาทั่วไป มหาวิทยาลัยศรีปทุม

**บทนำ**

การทดสอบความสัมพันธ์ และการทดสอบความแตกต่างของประชากร ในกรณีที่ข้อมูลอยู่ในมาตราวัดระดับ นามบัญญัติ (Nominal Data) หรือข้อมูลอันดับ (Ordinal Data) นั้น มีงานวิจัย วิทยานิพนธ์ สารนิพนธ์ และการค้นคว้าอิสระ ทางด้านสังคมศาสตร์ เป็นจำนวนมากที่ใช้ไคสแควร์ เป็นตัวสถิติในการทดสอบไคสแควร์เป็นตัวสถิติที่เหมาะสมสำหรับข้อมูลเหล่านี้ทั้งการศึกษาในภาพรวมและวิเคราะห์ในส่วนย่อย การใช้สถิติทดสอบไคสแควร์ในการทดสอบความแตกต่าง หรือทดสอบความสัมพันธ์นั้น ถึงแม้จะใช้ได้ดีแต่ก็มีข้อที่ควรระวังหลายประการด้วยกัน เช่น ขนาดของตัวอย่างที่ใช้ความถี่จากการสังเกตของข้อมูล ขนาดของตารางคูณไขว้ การอธิบายค่าในตารางคูณไขว้ การภาพรวมของระดับความคิดเห็น ตลอดจนการนำเสนอผลการวิเคราะห์โดยโปรแกรมสำเร็จรูปมานำเสนอ จากประสบการณ์ที่ผ่านมาผู้เขียนพบว่า มีงานวิจัย วิทยานิพนธ์ สารนิพนธ์ และการค้นคว้าอิสระ บางส่วนที่ขาดความระมัดระวังในเรื่องเหล่านี้ ผู้เขียนจึงขอนำเสนอวิธีการทดสอบและข้อควรระวังในการใช้การทดสอบไคสแควร์

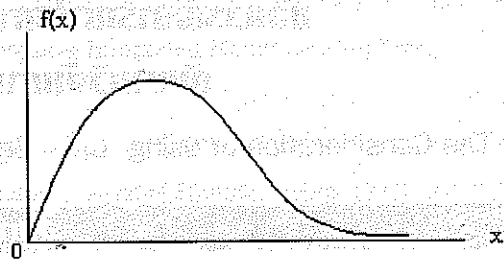
**1. การแจกแจงไคสแควร์ (Chi - Square Distribution)**

การแจกแจงแบบไคสแควร์เป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องรูปแบบหนึ่งที่มีฟังก์ชันดังนี้

ถ้า  $X$  เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง พารามิเตอร์  $V$  แล้ว  $f(x)$  จะมีการแจกแจงแบบไคสแควร์เมื่อ

$$f(x) = \frac{1}{2^{V/2} \Gamma(V/2)} x^{V/2-1} e^{-x/2}, x > 0$$

$$E(X) = V \text{ และ } V(X) = 2V$$



รูปที่ 1 แสดงลักษณะการแจกแจงแบบไคสแควร์

รูปร่างของโค้งจะต่างกันตามค่าพารามิเตอร์  $V$  ซึ่งในทางสถิติเรียกว่า ค่าองศาอิสระ (degree of freedom) และเนื่องจากเป็นฟังก์ชันการแจกแจงความน่าจะเป็นดังนั้นพื้นที่ใต้โค้งจึงมีค่าเท่ากับ 1.0 เสมอ

**2. การแจกแจงแบบไคสแควร์ที่นำไปประยุกต์ใช้งานวิจัยทางด้านสังคมศาสตร์**

การแจกแจงแบบไคสแควร์ในปัจจุบัน มีการนำไปใช้ประโยชน์ในการทดสอบสมมติฐานที่สามารถแบ่งออกเป็น 2 คือ แบบพาราเมตริก กับแบบนอนพาราเมตริก แบบพาราเมตริก ได้แก่ การทดสอบสมมติฐานเกี่ยวกับความแปรปรวนของประชากร 1 กลุ่ม การประมาณค่าช่วงความเชื่อมั่นของประชากร 1 กลุ่ม ส่วนแบบนอนพาราเมตริก ได้แก่ การทดสอบความเป็นเอกพันธ์ (Homogeneity) การทดสอบความเป็นอิสระ (Independent) การทดสอบภาวะทับกันสนิท (Goodness of fit) และทดสอบความสอดคล้อง (Concordance) ซึ่งการเลือกใช้นั้นจะขึ้นอยู่กับวัตถุประสงค์ คำถามการวิจัย สมมติฐาน และระดับของข้อมูล

ที่มาของไคสแควร์สามารถนำมาประยุกต์ใช้ในการทดสอบสมมติฐานทางสังคมได้เนื่องมาจากนักสถิติท่านหนึ่งคือ Karl Pearson (1857-1936) ได้พบว่าผลรวมของ

(ความ  
มี  
ไคสแควร์จึงมี  
ทางสังคมศ  
ด้านสังคมศ  
ความต่าง  
ความสัมพันธ์  
2  
ระหว่างประ  
ก  
แตกต่างระย  
เชิงปริมาณ  
หรือสัดส่วน  
ตามสัดส่วน  
ชมรายการ  
จำนวนคน  
หรือไม่ เป็น  
นั้น สามารถ  
เป็นตัวสถิติ  
จำนวนไม่ม  
สมมติฐาน  
ในการตีเม  
กรุงเทพฯ  
ดังตาราง

(ความถี่จากการสังเกต-ความถี่คาดหวัง)<sup>2</sup>

ความถี่คาดหวัง

มีการแจกแจงไคสแควร์ ดังนั้นการทดสอบไคสแควร์จึงมีการนำมาใช้เป็นสถิติสำหรับทดสอบสมมติฐานทางสังคมศาสตร์อย่างแพร่หลาย สำหรับการวิจัยทางด้านสังคมศาสตร์การทดสอบไคสแควร์มักใช้ในการทดสอบความแตกต่างระหว่างกลุ่มของประชากรและทดสอบความสัมพันธ์ของตัวแปร 2 ตัว

**2.1 การใช้ไคสแควร์ทดสอบความแตกต่างระหว่างประชากร**

การใช้การทดสอบไคสแควร์ทดสอบความแตกต่างระหว่างกลุ่มประชากรงานวิจัยเชิงสำรวจ หรือเชิงปริมาณส่วนใหญ่ใช้เพื่อเปรียบเทียบ จำนวนร้อยละ หรือสัดส่วนระหว่างกลุ่ม ว่าเท่ากันหรือไม่ หรือเป็นไปตามสัดส่วนที่คาดหวังหรือไม่ เช่น จำนวนคนที่นิยมชมรายการทีวีช่อง 3 และช่อง 7 แตกต่างกันหรือไม่ จำนวนคนที่ดื่มชาเขียวแตกต่างกันระหว่างกลุ่มอายุหรือไม่ เป็นต้น การจะสรุปว่าเท่ากันหรือไม่เท่าให้ชัดเจนนั้น สามารถทำได้โดยใช้การทดสอบไคสแควร์ซึ่งจะช่วยให้เป็นตัวสถิติเชิงอนุมานช่วยในการตัดสินใจ มีงานวิจัยจำนวนไม่น้อยที่มักจะสรุปรายละเอียดโดยไม่ใช้การทดสอบสมมติฐานช่วย เช่น งานวิจัยชิ้นหนึ่งได้สำรวจความนิยมในการดื่มกาแฟของคนกรุงเทพฯ โดยสุ่มตัวอย่างคนกรุงเทพฯ ที่ดื่มกาแฟมาจำนวน 400 คน ได้ข้อมูลดังตาราง

ตารางที่ 1 แสดงจำนวนและร้อยละของปัจจัยที่ผู้บริโภคใช้ในการพิจารณาซื้อกาแฟ

ปัจจัย	จำนวน	ร้อยละ
รสชาติ	83	20.7
กลิ่น	82	20.5
ราคา	80	20.0
บรรจุภัณฑ์	77	19.3
ยี่ห้อ	78	19.5
รวม	400	100.0

จากตารางนี้ ผู้วิจัยได้สรุปว่า "คนกรุงเทพฯ พิจารณาซื้อกาแฟสำหรับบริโภคโดยให้ความสำคัญของรสชาติเป็นปัจจัยหลักรองลงมาเป็น กลิ่นและราคาตามลำดับ..." ซึ่งถ้าพิจารณาโดยผิวเผินก็น่าจะเป็นการสรุปที่ถูกต้อง แต่ถ้าพิจารณาว่าผู้วิจัยใช้สิ่งที่ได้จากตัวอย่างมาอธิบายประชากร ใช้ข้อสรุปที่ได้จากตัวอย่างจำนวน 400 คน อธิบายรายละเอียดของคนจำนวน 6 ล้านคนแค่นั้นแล้วจะพบว่าความน่าเชื่อถือจะลดลง ดังนั้นเพื่อให้ข้อสรุปถูกต้องและน่าเชื่อถือ จึงควรใช้การทดสอบไคสแควร์ทดสอบความแตกต่างระหว่างประชากร แล้วจึงนำข้อสรุปนั้นมาอธิบายผลในตาราง

สูตรไคสแควร์สำหรับการทดสอบความแตกต่างระหว่างกลุ่มคือ

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

- k คือ จำนวนกลุ่มของข้อมูลที่ได้จัดแบ่งทั้งหมด
- $O_i$  คือ ความถี่ของการทั้งหมดสังเกตในกลุ่มที่ i
- $E_i$  คือ ความถี่จากการคาดหวังของกลุ่มที่ i
- $E_i = NP_i$
- N คือ จำนวนความถี่ทั้งหมดของข้อมูล
- $P_i$  คือ ความน่าจะเป็นของกลุ่มที่ i

ตารางที่ 2 แสดงจำนวนคนที่ดื่มชาเขียวจากการสังเกต และจากค่าคาดหวังจำแนกตามระดับการศึกษา

ระดับการศึกษา	จำนวน (O <sub>ij</sub> )	ร้อยละ	Pi	Ei	O <sub>i</sub> - E <sub>i</sub>	$\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$
ประถมศึกษา	47	18.8	1/5	50	-3	0.18
มัธยมต้น	52	20.8	1/5	50	2	0.08
มัธยมปลาย/ปวช.	48	19.2	1/5	50	-2	0.08
อนุปริญญา	53	21.2	1/5	50	3	0.18
ปริญญาตรีหรือสูงกว่า	50	20.0	1/5	50	0	0.00
<b>รวม</b>	<b>250</b>	<b>100.0</b>	<b>1.00</b>	<b>250</b>	<b>0</b>	<b>0.52</b>

ตัวอย่าง การทดสอบไคสแควร์ เช่น งานวิจัยชิ้นหนึ่ง มีวัตถุประสงค์ เพื่อเปรียบเทียบ จำนวนคนที่ดื่มชาเขียวจำแนกตามระดับการศึกษา จึงสุ่มตัวอย่างคนที่ดื่มชาเขียวมา 250 คน ได้ผลดังตาราง

จากตารางพบว่า จำนวนคนที่ดื่มชาเขียวเมื่อจำแนกตามระดับการศึกษาแล้วมีจำนวนใกล้เคียงกัน แต่ยังไม่สามารถหาข้อสรุปได้ในระดับประชากร จึงต้องทดสอบโดยใช้สถิติทดสอบไคสแควร์ โดยสมมติฐานหลัก (Null Hypothesis : H<sub>0</sub>) คือ จำนวนคนที่ดื่มชาเขียวไม่แตกต่างกันระหว่างระดับการศึกษา สมมติฐานทางเลือก (Alternative Hypothesis : H<sub>1</sub>) คือ จำนวนคนที่ดื่มชาเขียวแตกต่างกันระหว่างระดับการศึกษา จากตารางที่ 2 ระดับการศึกษามี 5 กลุ่ม (k = 5) สมมติฐานตั้งว่าไม่แตกต่างกัน สัดส่วนหรือความน่าจะเป็น (P<sub>i</sub>) ของทุกกลุ่มจึงมีค่าเท่ากับ 1/5 จำนวนคนดื่มชาเขียวคาดหวังจึงมีค่าเท่ากับ 50 ทุกกลุ่ม (E<sub>i</sub> = NP<sub>i</sub>) วิธีคำนวณค่าไคสแควร์พิจารณาได้จากตารางที่ 2 จากตารางได้ค่าไคสแควร์เท่ากับ 0.52 ซึ่งจะหาข้อสรุปโดยเปรียบเทียบกับค่าไคสแควร์ที่เปิดจากตาราง ค่าไคสแควร์จากตารางจะเปิดได้ เมื่อกำหนดระดับนัยสำคัญ หรือ a (significant level) และองศาอิสระหรือ n (degree of freedom,

n = k - 1) เช่นเมื่อกำหนด a = .05 เปิดตารางไคสแควร์ที่ n = 4 ได้ 9.488 ข้อสรุปก็คือ ค่าไคสแควร์ที่ได้จากการคำนวณน้อยกว่าค่าไคสแควร์ที่ได้จากการเปิดตารางอธิบายผลได้ว่า จำนวนคนที่ดื่มชาเขียวไม่แตกต่างกันตามระดับการศึกษา

### 2.2 การใช้ไคสแควร์ทดสอบความสัมพันธ์

การใช้การทดสอบไคสแควร์ทดสอบความสัมพันธ์หรือทดสอบความเป็นอิสระ (Independent test) ของตัวแปร 2 ตัว ซึ่งมีการใช้กันอย่างแพร่หลายในงานวิจัยทางด้านสังคม เช่น ศึกษาความสัมพันธ์ระหว่างเพศและกลิ่นของน้ำหอมที่ใช้ เพศ และความนิยมในรสชาติกาแฟ ระดับการศึกษากับความพึงพอใจในการบริหารงานของบริษัท ระดับอายุกับความพึงพอใจในการให้บริการของร้านอาหารฟาสฟู๊ด เป็นต้น

สูตรทางคณิตศาสตร์สำหรับการทดสอบความสัมพันธ์ คือ

$$\chi^2 = \sum_{j=1}^k \sum_{i=1}^h \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

k, h คือ จำนวน  
O<sub>ij</sub> คือ ความ  
E<sub>ij</sub> คือ ความ

จากสูตร  
กำลังสองระหว่าง  
การคำนวณโดย  
ตัวแปรทั้งสองตัว  
จะต้องมีค่าใกล้เคียง  
ซึ่งใกล้เท่าใดจะยอม  
ได้จากค่าไคสแคว  
(กรณีใช้โปรแกรม

เพื่อให้  
ตัวอย่างประกอบ  
ความสัมพันธ์ระห  
จึงสุ่มตัวอย่างมา

ตารางที่ 3 แสดง  
การใ

กลิ่นน้ำหอม	
A	110
B	55
C	45
รวม	210

การศึกษา

$$\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

0.18

0.08

0.08

0.18

0.00

0.52

k, h คือ จำนวนกลุ่มของตัวแปรที่ 1 และตัวแปรที่ 2  
 $O_{ij}$  คือ ความถี่ของค่าสังเกตกลุ่มที่ ij  
 $E_{ij}$  คือ ความถี่จากการคาดหวังของกลุ่มที่ ij

จากสูตรค่าไคสแควร์ คือค่าที่เกิดจากผลต่างกำลังสองระหว่างความถี่จากการสังเกตกับค่าความถี่จากการคำนวณโดยสูตร (หรือตามทฤษฎี) ที่กำหนดว่าตัวแปรทั้งสองตัวเป็นอิสระกัน ถ้า  $O_{ij}$  และ ค่า  $E_{ij}$  จะต้องมีค่าใกล้เคียงกัน ผลต่างกำลังสองจึงมีค่าเข้าสู่ศูนย์ซึ่งใกล้เคียงที่จะยอมรับและปฏิเสธสมมติฐานนั้นพิจารณาได้จากค่าไคสแควร์ที่เปิดจากตาราง หรือ ค่า P-value (กรณีใช้โปรแกรมสำเร็จรูปประมวลผล)

เพื่อให้เกิดภาพที่ชัดเจนยิ่งขึ้นลองพิจารณาตัวอย่างประกอบ เช่น นักวิจัยท่านหนึ่งต้องการทดสอบความสัมพันธ์ระหว่างเพศและรสนิยมในกลิ่นของน้ำหอม จึงสุ่มตัวอย่างมา 300 คน ดังตาราง

ตารางที่ 3 แสดงจำนวนและร้อยละของรสนิยมในการใช้น้ำหอมจำแนกตามเพศและกลิ่น

กลิ่นน้ำหอม	เพศ		รวม
	ชาย	หญิง	
A	110 (52.38)	40 (44.44)	150
	105	45	
B	55 (26.19)	15 (16.67)	70
	49	21	
C	45 (21.43)	35 (38.89)	80
	56	24	
รวม	210 (100.0)	90 (100.0)	300

จากข้อมูลในตารางนี้ ผู้วิจัยส่วนใหญ่จะนำเสนอค่าสังเกตที่ได้ในรูปจำนวนและร้อยละ ไปพร้อมๆ กับการทดสอบความสัมพันธ์โดยใช้ สถิติทดสอบไคสแควร์ การเปรียบเทียบร้อยละนั้นมีข้อสงสัยหรือมีคำถามอยู่บ่อยๆ เช่นกันว่าจะเปรียบเทียบตามแนวนอนหรือแนวตั้ง หลักการง่ายๆ คือเปรียบเทียบตามแนวนอนจะรวมร้อยละตามแนวตั้ง เปรียบเทียบตามแนวตั้งรวมร้อยละตามแนวนอน ส่วนการคำนวณค่าไคสแควร์นั้นเริ่มด้วยการคำนวณ ค่าความถี่จากการคาดหวังหรือจากทฤษฎีความน่าจะเป็นของเหตุการณ์ที่เป็นอิสระต่อกันนั้นคือ

$$E_{ij} = \frac{r_i \cdot c_j}{n}$$

$r_i$  คือ ผลรวมค่าความถี่จากการสังเกตของแถวที่ i  
 $c_j$  คือ ผลรวมของค่าความถี่จากการสังเกตของแถวที่ j  
 $n$  คือ ความถี่ทั้งหมดจากการสังเกต เช่น

$$E_{ij} = \frac{r_i \cdot c_j}{n} = \frac{150 \times 210}{300} = 105$$

เมื่อกำหนดได้ครบทุกค่าดังตารางแล้ว เราจึงสามารถคำนวณค่า ไคสแควร์ได้ เช่น

$$\chi^2 = \frac{(110 - 105)^2}{105} + \frac{(40 - 45)^2}{45} + \frac{(55 - 49)^2}{49} + \frac{(15 - 21)^2}{21} + \frac{(45 - 56)^2}{56} + \frac{(35 - 24)^2}{24} = 5.99$$

เปรียบเทียบกับค่าที่เปิดจากตารางที่

$$d.f. = (2 - 1)(3 - 1) = 2 \quad \chi^2_{0.05, 2} = 5.99$$

จะเห็นว่าค่าที่คำนวณได้มากกว่าที่เปิดจากตารางสรุปได้ว่าระดับการศึกษาที่มีความสัมพันธ์กับความนิยมในการบริโภคชาเขียว

### 3. การใช้ไคสแควร์ทดสอบความสัมพันธ์ของข้อมูลที่เป็นผลรวมเฉลี่ย

ในการศึกษาความสัมพันธ์บางครั้งผู้วิจัยอาจต้องการศึกษาความสัมพันธ์ของตัวแปรที่ 1 ตัวหรือ 2 ตัวที่เป็นผลรวมเฉลี่ยของระดับความคิดเห็นของผู้ตอบแบบ

5 ระดับ (อาจมีมากกว่าหรือน้อยกว่า 5 ระดับก็ได้) จำนวนหลายๆ ข้อ เช่น การศึกษาความสัมพันธ์ระหว่างระดับการศึกษาของลูกคำกับความพึงพอใจต่อระดับบริการที่จัดให้ของร้านอาหารแห่งหนึ่ง การวัดความพึงพอใจต่อบริการที่จัดให้ อาจต้องวัดในหลายๆ ด้านแล้วนำผลของแต่ละด้านมารวมกัน เช่น ด้านความมโนชนสัมพันธ์ของพนักงานที่ให้บริการอาจวัดออกมาเป็น 4 ข้อดังตาราง

3) นำค่าที่ได้แปลงเป็น 3 ระดับ ดังตาราง เมื่อแปลงแล้วก็สามารถนำไปใช้ศึกษาความสัมพันธ์ได้ตามวัตถุประสงค์ของผู้วิจัย

ด้านมนุษยสัมพันธ์	ระดับความคิดเห็น					x̄	S.D.	ความหมาย
	5	4	3	2	1			
1. พนักงานยิ้มแย้มแจ่มใส	11	3	3	2	1	4.05	1.28	มาก
2. พนักงานให้บริการด้วยความสุภาพ	7	5	3	3	2	3.60	1.39	ปานกลาง
3. พนักงานมีเต็มใจในการให้บริการ	6	7	4	2	1	3.75	1.16	มาก
4. พนักงานมีความชำนาญในงาน	4	10	3	1	2	3.65	1.18	ปานกลาง
รวม	28	25	13	8	6	3.76	1.22	มาก

จากตารางถ้าผู้วิจัยต้องการศึกษาโดยแบ่งระดับของการให้บริการออกเป็น 3 ระดับก็สามารถทำได้โดยการแปลงผลรวมระดับความคิดเห็นเฉลี่ยออกเป็น 3 ระดับหรือ 3 ชั้น โดยทำดังนี้คือ

- 1) คำนวณหาความกว้างของชั้น จากสูตร ความกว้าง

$$= \frac{\text{ค่าสูงสุด}-\text{ค่าต่ำสุด}}{\text{จำนวนชั้น}} = \frac{5-1}{3} = 1.33$$

- 2) หาขีดจำกัดล่างและขีดจำกัดบนของแต่ละชั้น

ชั้นที่	ช่วงคะแนน	ความหมาย
1	1.00 - 2.33	น้อย
2	2.34 - 3.67	ปานกลาง
3	3.68 - 5.00	มาก

#### 4. ข้อควรระวังในการทดสอบไคสแควร์

ข้อควรระวังในการใช้การทดสอบไคสแควร์ที่ไม่ควรมองข้ามมีอยู่ 6 ประการ ดังนี้ คือ

4.1 ปัญหาจำนวนความถี่ในบางเซลล์ น้อยกว่า 5 มากกว่า 1 ใน 5 เซลล์ หรือ 2 ใน 10 เซลล์ ซึ่งจะทำให้ผลที่ได้คลาดเคลื่อน ปัญหานี้เป็นปัญหาที่พบเห็นได้บ่อยๆ ในงานวิจัยหรือวิทยานิพนธ์ของนักศึกษา ทางด้านสังคมศาสตร์ บริหารธุรกิจ นิเทศศาสตร์ ฯลฯ ซึ่งอาจจะเป็นเพราะการขาดความระมัดระวังของผู้ใช้ ในกรณีนี้ถ้าหากเกิดขึ้นกับงานวิจัยของเรา เราสามารถที่จะแก้ไขได้ดังนี้ 1) เก็บข้อมูลเพิ่มเติมเพื่อให้จำนวนความถี่ในแต่ละเซลล์สูงขึ้น 2) ไม่ทดสอบความสัมพันธ์ประเด็นนั้น อธิบายเฉพาะค่าร้อยละเท่านั้น 3) ยุบเซลล์รวมกัน ที่เมื่อยุบแล้วความหมายไม่เปลี่ยนแปลง เช่น ระดับการศึกษาในการเก็บข้อมูลผู้วิจัยอาจจำแนกกลุ่ม เป็นระดับประถม

ระดับ
ระดับ
มัธยม
ปวช.
มัธยม
ปวส.
ปริญญา
ปริญญา
ปริญญา

ศึกษา มัธยม  
ปวส. ปริญญา  
อาจจะยุบเป็น  
และระดับอุดม

หรือ  
ตัวแปรสองตัว  
2 ตัว เช่น ค  
ได้รับเดือนกั  
จาก  
น้อยกว่า 5 จั  
ไม่ควรทำเพ  
ปกติอันเป็นผ

ความพึงพ ในเงินเดี
มากที่สุด
มาก
ปานกลาง
น้อย
น้อยที่สุด

ระดับ ดังตาราง  
นำไปใช้ศึกษา  
ทุกประสงค์ของ

ระดับการศึกษา	จำนวน
ระดับประถมศึกษา	20
มัธยมศึกษาตอนต้น	10
ปวช.	3
มัธยมศึกษาตอนปลาย	20
ปวส.	12
ปริญญาตรี	7
ปริญญาโท	2
ปริญญาเอก	1

⇒ ยุบรวมกลุ่มเป็น

ระดับการศึกษา	จำนวน
ภาคบังคับ	30
เตรียมอุดม	23
อุดมศึกษา	22

**ความหมาย**

- มาก
- ปานกลาง
- มาก
- ปานกลาง
- มาก

ศึกษา มัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย ปวช. ปวส. ปริญญาตรี ปริญญาโท และ ปริญญาเอก อาจจะถูกเป็นระดับการศึกษาภาคบังคับ เตรียมอุดม และระดับอุดมศึกษา

ผลให้เกิดการปฏิเสธสมมติฐานหลักและยอมรับสมมติฐานทางเลือก ยอมรับว่ามีความสัมพันธ์กันเรา จึงควรมีการยุบเซลล์ ที่ยุบแล้วสามารถอธิบายความหมายได้ เช่น เปลี่ยนจาก 5 ระดับ เป็น 3 ระดับคือ น้อย ปานกลาง มาก ดังตาราง

หรือกรณีที่เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรสองตัวที่ข้อมูลเป็นระดับความคิดเห็น 5 ระดับทั้ง 2 ตัว เช่น ความสัมพันธ์ระหว่างความพึงพอใจในเงินที่ได้รับเดือนกับความภักดีต่อองค์กร

ความพึงพอใจ ในเงินเดือน	ความภักดีต่อองค์กร		
	มาก	ปานกลาง	น้อย
มาก	125	65	10
ปานกลาง	50	43	12
น้อย	48	20	20

จากตารางจะพบว่า มีจำนวนเซลล์ ที่มีค่าความถี่น้อยกว่า 5 จำนวน 10 เซลล์ การทดสอบความสัมพันธ์จึงไม่ควรทำเพราะว่า ค่าไคสแควร์ที่คำนวณได้จะสูงกว่าปกติอันเป็นผลต่อเนื่องมาจากเซลล์ที่มีค่าต่ำกว่า 5 จะส่ง

ความพึงพอใจ ในเงินเดือน	ความภักดีต่อองค์กร				
	มากที่สุด	มาก	ปานกลาง	น้อย	น้อยที่สุด
มากที่สุด	39	43	50	1	4
มาก	4	39	15	2	3
ปานกลาง	16	34	43	1	11
น้อย	9	35	6	5	3
น้อยที่สุด	0	4	14	12	0

ตารางที่ 4 แสดงผลการทดสอบไคสแควร์เมื่อขนาดตัวอย่างมีขนาดใหญ่

กลุ่ม	1	2	3	4	รวม
1	100	200	250	300	850
2	450	650	800	630	2530
3	1200	1000	770	365	3335
รวม	1750	1850	1820	1295	6715
ค่าไคสแควร์เท่ากับ 589.61					

กลุ่ม	1	2	3	4	รวม
1	150	30	110	210	500
2	50	270	180	200	700
3	100	50	300	450	900
รวม	300	350	590	860	2100
ค่าไคสแควร์เท่ากับ 481.24					

4.2 ปัญหาจำนวนตัวอย่างมีขนาดใหญ่ ค่าไคสแควร์จะเป็นสัดส่วนกับขนาดตัวอย่าง จึงไม่ควรใช้การทดสอบไคสแควร์ทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว เมื่อตัวอย่างมีขนาดใหญ่มากๆ เพราะส่วนใหญ่ผลการทดสอบจะปฏิเสธสมมติฐานหลักเสมอ จึงไม่สามารถบอกความสัมพันธ์ได้ ดังนั้น ถ้าจำเป็นต้องทดสอบความสัมพันธ์ของตัวอย่างที่มีขนาดใหญ่อาจใช้วิธีการสุ่มตัวอย่างศึกษาเพียงบางส่วนเท่านั้น

จากตารางทั้ง 4 จะพบว่าค่าไคสแควร์ทั้ง 2 ตัว มีค่ามากกว่า 12.59 (ค่าไคสแควร์ที่เปิดจากตาราง) ทุกตัว แนวทางในการแก้ปัญหาอีกวิธีหนึ่งก็คือให้ใช้ค่าที่มีการปรับขนาดตัวอย่างแล้วได้แก่ Contingency Coefficient Phi และ Cramer's V คำนวณได้จากสูตรดังนี้

Contingency Coefficient (C)

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} ; 0 < C < 1$$

ถ้าค่า C = 0 หมายถึงตัวแปรทั้ง 2 ตัว ไม่มีความสัมพันธ์กัน

C เข้าสู่อันดับ 1 หมายถึงตัวแปรทั้ง 2 ตัว มีความสัมพันธ์กัน

Phi and Cramer's V

$$\phi = \sqrt{\frac{\chi^2}{n}} ; \phi > 1$$

Cramer's V

$$V = \sqrt{\frac{\Phi^2}{\min\{(r-1), (c-1)\}}} ; 0 < V < 1$$

ถ้าค่า  $\phi$  และ V มีค่ามาก หมายถึงตัวแปรทั้ง 2 ตัว มีความสัมพันธ์กัน

4.3 การทดสอบความสัมพันธ์กรณีตารางคูณได้เป็นแบบ 2x2 หรือ d.f. มีค่าเท่ากับ 1

ก) การคำนวณค่าไคสแควร์ที่มีค่าความถี่ไม่เกิน 5 จะทำให้ได้ค่าไคสแควร์สูงกว่าค่าที่เป็นจริง การศึกษาความสัมพันธ์สามารถได้โดยวิธีการของ yate (Yate collection for Continuity) โดยใช้ 0.5 เป็นตัวปรับค่าสูตรคือ

$$\chi^2 = \sum_{j=1}^2 \sum_{i=1}^2 \frac{(O_{ij} - E_{ij} - 0.5)^2}{E_{ij}}$$

ซึ่งการคำนวณนั้นปัจจุบันทำได้ไม่ยาก ถ้าใช้โปรแกรมสำเร็จรูปตัวโปรแกรมจะปรับสูตรให้โดยอัตโนมัติ

เพศ
ชาย
หญิง
รวม
ค่าไคส

คุณไขว้เป็นแบบ การทดสอบของ Fisher แสดงให้

Row 1	
Row 2	
รวม	

ให้ T 1 และหลักที่ 1 T

$$P(T = x) =$$

และถ้า  $p_1$  คือ แถวที่ 1 และหลัก  $p_2$  คือ แถวที่ 2 และหลัก



4	รวม
210	500
270	700
450	900
660	2100

เพศ	ความนิยมในการดื่มชาเขียว		
	ชอบ	ไม่ชอบ	รวม
ชาย	4	10	14
หญิง	7	8	15
รวม	11	18	29

ค่าไคสแควร์จากการคำนวณ = 0.39

ข) การทดสอบความสัมพันธ์ที่ตารางคูณไขว้เป็นแบบ 2x2 มีความถี่ทั้งหมดน้อยๆ ควรใช้การทดสอบของ Fisher's exact แทนการทดสอบของ Fisher แสดงได้ดังนี้

	Column 1	Column 2	รวม
Row 1	x	r - x	R
Row 2	c - x	N - r - c + x	N - r
รวม	c	N - c	N

ให้ T คือความถี่จากการสังเกตในเซลล์ แถวที่ 1 และหลักที่ 1 T จะมีการแจกแจงแบบ ไฮเปอร์จีโอเมตริก

$$P(T = x) = \frac{\binom{r}{x} \binom{N-r}{c-x}}{\binom{N}{c}} \quad x = 0, 1, \dots, \min(r, c)$$

และถ้าให้

$p_1$  คือความน่าจะเป็นของค่าสังเกต ในเซลล์ของแถวที่ 1 และหลักที่ 1

$p_2$  คือความน่าจะเป็นของค่าสังเกตในเซลล์ของแถวที่ 2 และหลักที่ 1 สมมติฐานที่ได้คือ

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2$$

ยกตัวอย่าง เช่น ต้องการศึกษาว่าเพศชายและเพศหญิงมีลักษณะการถอนเงินแตกต่างกันหรือไม่จึงสุ่มตัวอย่างคนมากลุ่มหนึ่งจำนวน 14 คนได้ผลดังตาราง

เพศ	ลักษณะการเบิกเงิน		
	บัญชี	เครื่อง ATM	รวม
ชาย	1	9	10
หญิง	3	1	4
รวม	4	10	14

$P(T = 1) = 0.041$

ค่าสถิติทดสอบคำนวณได้จาก

$$P(T = 1) = \frac{\binom{10}{1} \binom{4}{3}}{\binom{14}{4}} = \frac{10 \cdot 4}{14 \cdot 4} = \frac{10}{14} = \frac{5}{7} \approx 0.714$$

$$p(T = 1) = \frac{1}{1001} + \frac{40}{1001} = 0.041$$

ปฏิเสธ  $H_0$  ยอมรับ  $H_1$

### 5. การทดสอบไคสแควร์โดยใช้โปรแกรมสำเร็จรูป SPSS

การวิเคราะห์ไคสแควร์ในปัจจุบันนี้สามารถทำได้ง่าย ๆ โดยใช้โปรแกรมสำเร็จรูปที่มีอยู่ทั่วไป

#### 5.1 การใช้โปรแกรมสำเร็จรูป SPSS วิเคราะห์

วิเคราะห์ความแตกต่างระหว่างสัดส่วนของประชากรเมื่อเราสร้างแฟ้มข้อมูลเสร็จเรียบร้อยแล้ว ใช้คำสั่ง Analyze / Nonparametric Test / Chi - Square เลือกตัวแปรที่ต้องการทดสอบ เลือก Expected Value ซึ่งจะมีให้เลือก 2 แบบ คือทดสอบว่าเท่ากัน กับทดสอบตามสัดส่วนที่ต้องการ สมมติว่าเราทดสอบว่าเท่ากัน จากนั้นโปรแกรมจะประมวลผลแล้วให้ผลดังตาราง

EDUC			Test Statistics		
	Observed N	Expected N	Residual		EDUC
ประถมศึกษา	47	50.0	-3.0	Chi-Square <sup>a</sup>	.520
มัธยมศึกษาตอนต้น	52	50.0	2.0	df	4
มัธยมศึกษาตอนปลาย	48	50.0	-2.0	Asymp. Sig.	.972
อนุสัญญา	53	50.0	3.0	a. 0 cells (.0%) have expected frequencies less than .5. The minimum expected cell frequency is 50.0.	
ปริญญาตรี/สูงกว่า	50	50.0	.0		
Total	250				

จาก Output ผลการทดสอบได้ค่า ไคสแควร์เท่ากับ 0.52 มีค่า P-value หรือ Asymp. Sig เท่ากับ .972 ซึ่งการสรุปผลของการทดสอบจะพิจารณาจากค่า P-value ถ้าน้อยกว่า .05 จะปฏิเสธสมมติฐานหลัก นั่นคือสรุปว่าสัดส่วนไม่แตกต่างกัน หลักในการพิจารณาการยอมรับและปฏิเสธสมมติฐานคือ ถ้า ค่า P-value มากกว่า ระดับนัยสำคัญที่กำหนด จะยอมรับสมมติฐานหลัก ถ้าน้อยกว่าระดับนัยสำคัญที่กำหนดปฏิเสธสมมติฐานหลัก

### 5.2 การใช้โปรแกรม SPSS ทดสอบความสัมพันธ์

การทดสอบความสัมพันธ์ระหว่างเพศและความนิยมในกลิ่นน้ำหอม เมื่อเราสร้างแฟ้มข้อมูลเสร็จเรียบร้อยแล้ว วิเคราะห์โดยใช้คำสั่งจากเมนูคำสั่ง Analyze / Descriptive Statistics / Crosstabs เลือกตัวแปรที่ต้องการศึกษาความสัมพันธ์ด้าน Row(s) และ Column(s) และ Statistics chi-square เลือกให้ตารางแสดงร้อยละตามที่ต้องการใน Percentages ซึ่งมีให้เลือก 3 แบบคือ Row Column และ Total อาจเลือกทั้ง 3 แบบก็ได้ เมื่อเลือกเสร็จแล้วจะให้ผลดังตาราง

กลิ่น \* ชาติ Crosstabulation

			เพศ		Total
			ชาย	หญิง	
กลิ่น	A	Count	110	45	155
		% within ชาติ	52.4%	50.0%	51.7%
B	Count	55	21	76	
	% within ชาติ	26.2%	23.3%	25.3%	
C	Count	45	24	69	
	% within ชาติ	21.4%	26.7%	23.0%	
Total	Count	210	90	300	
	% within ชาติ	100.0%	100.0%	100.0%	

<sup>1</sup> ถ้าขนาดตัวอย่างขนาดเล็กให้ใช้ Exact ... Test ซึ่งรายละเอียดอ่านได้ในหนังสือการใช้โปรแกรมสำเร็จรูป SPSS ทั่วๆ ไป

จาก Output ให้เลือกค่า chi-square ที่พิจารณาว่าจะเปรียบเทียบจากค่าที่กำหนดไว้ เช่น กำหนดไว้ .05 ให้ยอมรับสมมติฐานหลัก .05 ให้ยอมรับสมมติฐานหลักมากกว่า .05 ความงามนิยมใน

6. บทสรุป

สถิติที่... การนำไปใช้ประโยชน์แพร่หลายทั้งทางทางด้านสังคมศาสตร์นำไปใช้ในการทดสอบความสัมพันธ์ที่มีมาตรวัดตั้งระดับ อัตราส่วน

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.024 <sup>a</sup>	2	.599
Likelihood Ratio	1.007	2	.604
Linear-by-Linear Association	.549	1	.459
N of Valid Cases	300		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 20.70.

ทดสอบความ

ระหว่างเพศและ  
 ว่างเพิ่มข้อมูล  
 ึ่งจากเมนูคำสั่ง  
 os / Crosstabs  
 ้นด้าน Row(s)  
 5 chi-square<sup>1</sup>  
 ักใน Percent-  
 umn และ Total  
 ึ่งแล้วจะให้ผล

จาก Output ให้เลือกใช้ค่า Pearson Chi - Square ซึ่งมีค่า chi-square เท่ากับ 1.024 มีค่า P-value เท่ากับ .599 การพิจารณาว่าจะยอมรับหรือปฏิเสธสมมติฐานหลักให้เปรียบเทียบจากค่า P-value กับค่า ระดับนัยสำคัญที่กำหนดไว้ เช่น กำหนดไว้ที่ .05 ถ้าค่า P-value มากกว่า .05 ให้ยอมรับสมมติฐานหลัก ถ้าค่า P-value น้อยกว่า .05 ให้ยอมรับสมมติฐานหลัก สำหรับตัวอย่างนี้ค่า P-value มากกว่า .05 จึงยอมรับ สมมติฐานหลักสรุปได้ว่า ความงามนิยมในกลุ่มน้ำหอมไม่ขึ้นอยู่กับเพศ

6. บทสรุป

สถิติทดสอบไคสแควร์เป็นสถิติตัวหนึ่งที่มีการนำไปใช้ประโยชน์ในการทดสอบสมมติฐานกันอย่างแพร่หลายทั้งทางด้านวิทยาศาสตร์และด้านสังคมศาสตร์ ทางด้านสังคมศาสตร์ สถิติทดสอบไคสแควร์สามารถนำไปใช้ในการทดสอบความแตกต่างระหว่างกลุ่ม และทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว ในข้อมูลที่มีมาตรวัดตั้งแต่ นามบัญญัติ (Nominal Scale) จนถึงระดับ อัตราส่วน เพราะข้อมูลระดับอัตราส่วน ข้อมูล

ระดับอันดับภาค และข้อมูลระดับอันดับสามารถแปลงให้เป็นข้อมูลระดับความถี่หรือข้อมูลนามบัญญัติได้ ความสามารถมักจะคู่กับข้อจำกัดเสมอ ดังนั้นการที่จะนำสถิติทดสอบไคสแควร์ไปใช้ผู้วิจัยควรทราบข้อกำหนดเบื้องต้น (Basic Assumption) ของการทดสอบไคสแควร์ ต้องทราบข้อควรระวังต่างๆ ได้แก่ จำนวนเซลล์ที่ว่างหรือมีค่าความถี่ต่ำกว่า 5 ไม่ควรจะมีมากกว่า 1 ใน 5 เซลล์ ถ้ามากกว่าต้องมีการปรับยุบเซลล์รวมกันโดยยังคงมีความหมายในการวิจัย หรืออาจต้องสุ่มตัวอย่างเพิ่มกรณีเป็นตารางคูณไขว้แบบ 2 x 2 ความถี่น้อยๆ ต้องใช้วิธีการของ Yate เพื่อลดค่าไคสแควร์ที่สูงเกินจริง หรือใช้วิธีการของ Fisher exact การศึกษาความสัมพันธ์กรณีที่มีขนาดตัวอย่าง มีขนาดใหญ่เกินไปไม่สามารถทำได้ ต้องใช้วิธีการสุ่มตัวอย่างมาบางส่วน เป็นต้น ส่วนเรื่อง การคำนวณค่าสถิติทดสอบไคสแควร์ในปัจจุบันนี้ทำได้ไม่ยาก โดยใช้โปรแกรมสำเร็จรูปต่างๆ เพียงแต่เราทราบวิธีการป้อนข้อมูล วิธีการใช้คำสั่ง วิธีการแปลผล และนำเสนอผลให้ถูกต้องเท่านั้น

155  
 174  
 76  
 134  
 69  
 104  
 100  
 174

ss ทัว ไป

## บรรณานุกรม

- กัลยา วาณิชย์บัญชา, การใช้ SPSS for Windows ในการวิเคราะห์ข้อมูล. พิมพ์ครั้งที่ 6. กรุงเทพฯ : ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2546.
- วรชัย เขียวป่าณี, โปรแกรมสำเร็จรูปทางสถิติ SPSS/PC + 4.0. กรุงเทพฯ : โอเดียนสโตร์, 2542.
- สุชาติ ประสิทธิ์รัฐสินธุ์, เทคนิคการวิเคราะห์หลายตัวแปรสำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์. พิมพ์ครั้งที่ 4, กรุงเทพฯ : เลียงเชียง, 2540.
- อุทุมพร จามรมาน, ไคสแควร์ การทดสอบทางสถิติ. กรุงเทพมหานคร : ฟีนีเฟ้นบลิสซิ่ง, 2535.
- W.J. Conover. *Practical Nonparametric Statistics*. 3 rd.ed. New York : John Wiley & Sons, 1999.

## บทคัดย่อ

การ  
ซึ่งไปรวมกัน  
นำมาใช้ในกา  
วิธีการนี้ผู้ต  
Full-profile  
ทั้งหมดของ  
ชุดจะถูกเขียน  
การวิเคราะห์  
และ nonmi

## Abstract

C  
Two qualit  
evaluation  
dents eval  
Each of att  
command,

\* อาจารย์ปร