บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในส่วนของวรรณคดีที่เกี่ยวข้อง นำเสนอเป็น 5 ตอน คือ

ตอนที่ 1 มาตราการวัด

ตอนที่ 2 สถิติที่ใช้ในการวิจัย

ตอนที่ 3 การแจกแจงที่เกี่ยวข้องกับการวิจัย

ตอนที่ 4 วิธีการรวมเซลล์

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

<u>ตอนที่ 1</u> มาตราการวัด (Measurement scales)

ในการใช้สถิติเพื่อวิเคราะห์ข้อมูลในการวิจัยนั้น ผู้วิจัยจำเป็นต้องทราบลักษณะของข้อมูลที่ใช้ในการวิจัยเสียก่อน เพื่อเป็นเครื่องวินิจฉัยว่า จะใช้วิธีการทางสถิติแบบใดในการวิเคราะห์จึงจะถูกต้องเหมาะสม ดังนั้นผู้วิจัยจึงควรทราบถึงมาตราการวัดของข้อมูลก่อน โดย Stevens (1946) ได้จำแนกมาตราการวัดออกเป็น 4 ระดับ ดังนี้

1. มาตรานามบัญญัติ (Nominal scales) เป็นมาตราการวัดในระดับต่ำสุด เป็นการวัดแบบง่าย ๆ โดยการจำแนกหรือแยกประเภท (classification) ตามคุณลักษณะที่ไม่เหมือนกัน อาจเป็นเพียงการเรียกชื่อ ดังนั้นการวัดในระดับนี้บางทีจึงไม่เป็นที่ยอมรับว่าเป็น " การวัด " เพราะไม่สามารถบอกปริมาณมากน้อยได้ เป็นแต่เพียงแสดงให้เห็นความแตกต่างของสิ่งต่าง ๆ เท่านั้น ตัวอย่างการวัดในระดับนี้ เช่น การจำแนกคนตามเพศ เป็นเพศชาย - หญิง หรือจำแนกคนตามศาสนาที่นับถือ เป็นคนที่นับถือศาสนาพุทธ คริสต์ อิสลาม เป็นต้น

การกำหนดตัวเลขให้กับสิ่งต่าง ๆ ในมาตรานามบัญญัติจะเป็นเพียงตัวแทนประเภทของสิ่งที่ถูกวัด หรือเพื่อใช้ในการสื่อความหมายให้จำได้ง่ายหรือสะดวกในการนับ โดยที่ตัวเลขดังกล่าวไม่มีความหมายในเชิงปริมาณแต่ประการใด ดังนั้นจึงไม่สามารถนำตัวเลขเหล่านั้นมา

บวก ลบ คูณ หรือ หารกันได้ นอกจากกระทำได้แต่เพียงการนับจำนวนหรือหาความถี่ของลักษณะที่มีเหมือนกันเท่านั้น เช่น กำหนดตัวเลข 1 แทนเพศชาย และเลข 2 แทนเพศหญิง เป็นต้น

2. มาตราจัดอันดับ (Ordinal scales) การวัดตามมาตรานี้จะมีระดับการวัดสูงกว่ามาตรานามบัญญัติ ลักษณะการวัดเป็นการจัดอันดับ (rank order) ให้กับคุณสมบัติบางอย่างว่ามากกว่าหรือน้อยกว่าข้อมูลอื่น ๆ ซึ่งสามารถจัดอันดับข้อมูลตามตำแหน่งได้จากมากที่สุดไปหาน้อยที่สุด หรือจากน้อยที่สุดไปหามากที่สุดได้ เช่น ผลการเรียนของนักเรียนตามลำดับ ดีมาก ดี ปานกลาง พอใช้ ต้องปรับปรุง หรือ นักเรียนที่สอบได้ที่ 1, 2, 3,...เป็นต้น ผลการวัดในมาตรานี้ช่วงห่างระหว่างอันดับจะไม่เท่ากัน และไม่ทราบว่าห่างกันเป็นปริมาณเท่าใด เช่น จะบอกว่านักเรียนที่สอบได้ที่ 1 เก่งกว่านักเรียนที่สอบได้ที่ 2 เท่ากับ นักเรียนที่สอบได้ที่ 3 เก่งกว่านักเรียนที่สอบได้ที่ 4 ไม่ได้ เป็นต้น ดังนั้นข้อมูลที่ปรากฏออกมาเป็นตัวเลขในระดับนี้ไม่สามารถนำมา บวก ลบ คูณ หรือหารกันได้ จะบอกได้แต่เพียงว่าต่างกันไปในทางไหน หรือทำให้ทราบทิศทางเท่านั้น

3. มาตราอันตรภาค (Interval scales) มาตราการวัดในมาตรานี้มีลักษณะเหมือนมาตราจัดอันดับทุกอย่าง แต่มีคุณสมบัติดีกว่าตรงที่แต่ละหน่วยของการวัดมีระยะห่างเท่าๆ กัน จึงสามารถเปรียบเทียบกันได้ว่ามากน้อยกว่ากันเท่าใด แต่ไม่สามารถบอกได้ว่าเป็นกี่เท่าของกันและกันเพราะมาตรานี้ไม่มีศูนย์แท้หรือศูนย์สมบูรณ์(true zero or absolute zero) เช่น ผู้ที่สอบได้ศูนย์คะแนน ไม่ได้หมายความว่าไม่มีความรู้ในวิชานั้น เป็นแต่เพียงผู้สอบคนนั้นทำข้อสอบฉบับนั้นไม่ได้เลย เป็นต้น ด้วยเหตุนี้ข้อมูลที่วัดได้ในมาตรานี้จึงไม่สามารถนำมาคูณ หรือหารกันได้ จะใช้ได้เฉพาะบวกหรือลบกัน เท่านั้น

4. มาตราอัตราส่วน (Ratioe scales) เป็นมาตราการวัดในระดับสูงที่สุด กล่าวคือ ครอบคลุมคุณสมบัติทุกอย่างของมาตราอันตรภาค แต่ดีกว่ามาตราอันตรภาคตรงที่มีศูนย์แท้ ข้อมูลในระดับนี้สามารถ บวก ลบ คูณ หรือหารกันได้ มาตราการวัดที่จัดอยู่ในระดับนี้มักจะเป็นการวัดทางฟิสิกส์ เช่น ความสูง ความยาว น้ำหนัก เวลา เป็นต้น

<u>ตอนที่ 2</u> สถิติที่ใช้ในการวิจัย

สถิติวิเคราะห์สัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ (CRAMER'S V COEFFICIENT)

ในการหาความสัมพันธ์ระหว่างตัวแปรโดยตัวแปรสองตัวที่อยู่ในมาตรานามบัญญัติ จำแนกได้หลายระดับต่างๆ กัน สามารถนำมาจำแนกให้อยู่ในรูปตารางการณ์จร (Contingency table) สามารถนำสถิติการวิเคราะห์ที่มีพื้นฐานมาจากค่าไคสแควร์ (Measures based on Chi-Square) มาหาค่าสัมประสิทธิ์ความสัมพันธ์ ได้แก่ ค่าสัมประสิทธิ์ฟี (ϕ) ค่าสัมประสิทธิ์ของชูโพร (Tschuprow's Contigency Coefficient:T) ค่าสัมประสิทธิ์ของเพียร์สัน (Pearson's Contigency Coefficient:C) ค่าสัมประสิทธิ์ วี ของแครมเมอร์ (Cramer's Contigency Coefficient:V) โดยปกติแล้วค่าสัมประสิทธิ์สหสัมพันธ์จะมีค่าสูงสุดเท่ากับ 1 แต่ค่าสัมประสิทธิ์ความสัมพันธ์ที่มีพื้นฐานมาจากไคสแควร์ในแต่ละวิธีมีข้อจำกัดและมีความเหมาะสมกับสถานการณ์แตกต่างกัน ดังนี้

สัมประสิทธิ์ฟี จะมีค่าเท่ากับ 0 เมื่อตัวแปรไม่มีความสัมพันธ์กันเลยและจะมีค่าเท่ากับ 1 เมื่อตัวแปรทั้งสองมีความสัมพันธ์กันอย่างสมบูรณ์ แต่จะเท่ากับ 1 เฉพาะกรณีที่เป็นตารางการณ์จร 2x2 เท่านั้น ถ้าจำนวนแถวและจำนวนสดมภ์มากกว่านี้ ค่าสัมประสิทธิ์ฟีจะมีค่ามากกว่า 1 ได้ (Jacobson 1976: 437)

สัมประสิทธิ์ T ของชูโพร ระหว่างตัวแปรสองตัวจะมีค่าเท่ากับ 1เสมอ หากตัวแปรทั้งสองมีความสัมพันธ์กันอย่างสมบูรณ์ แต่ยังมีข้อแม้ว่า จำนวนแถวและจำนวนสดมภ์จะต้องเท่ากัน ถ้าหากจำนวนแถวและจำนวนสดมภ์ไม่เท่ากัน ค่าสัมประสิทธิ์ T จะมีค่าต่ำกว่า 1 (Blalock 1981:304)

สำหรับสัมประสิทธิ์ C ของเพียรสัน มีข้อจำกัดดังที่ Yule and Kendall (1950:53-54) ได้กล่าวสอดคล้องกับ Siegel (1956: 201) ว่าค่าสัมประสิทธิ์ C จะมีค่าเท่ากับ 0 เมื่อตัวแปรทั้งสองไม่มีความสัมพันธ์กัน แต่จะไม่มีโอกาสที่จะสูงถึง 1 ได้เลย เมื่อจำนวนแถวและสดมภ์มีค่าเท่ากับค่าสูงสุดของสัมประสิทธิ์ C จะมีค่าเท่ากับ $\sqrt{(r-1)/r}$ หรืออาจเขียนได้ว่า $0 \leqslant C \leqslant \sqrt{(r-1)/r} < 1$ เมื่อ r แทนจำนวนแถว ดังเช่นตารางการณ์จร 4x4 ค่าสัมประสิทธิ์ C จะมีค่าสูงสุด = 0.866 ถ้าตารางการณ์จร 5x5 ค่า

สัมประสิทธิ์ C จะมีค่าสูงสุด = 0.894 เป็นต้น สัมประสิทธิ์ C เป็นที่รู้จักโดยทั่วไปแต่ด้วยข้อจำกัดดังกล่าวจึงได้มีการศึกษา และพยายามสร้างดัชนีความสัมพันธ์ที่แก้ความจำกัดได้ สามารถใช้ได้กับตารางการณ์จรขนาดต่าง ๆ สามารถมีค่าสูงสุดได้เท่ากับ 1 ได้แก่ สัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ดังที่ Hays(1973:745) Blalock(1981:305) Jacobson (1976:438) แนะนำให้ใช้แทนสัมประสิทธิ์ C

แครมเมอร์ (Harald Crammer) 1946 ได้ศึกษาสัมประสิทธิ์ในรูปตารางการณ์จร RxC โดยศึกษาจากสูตรการคำนวณสัมประสิทธิ์ฟี $\phi = \sqrt{\chi^2/n}$ กล่าวคือ เป็นการสร้างดัชนีความสัมพันธ์ที่ขยายขนาดของตารางการณ์จร 2x2 เป็น RxC แก้ข้อจำกัดสัมประสิทธิ์ T ของชูโพร ว่าหากจำนวนแถวและจำนวนสดมภ์ไม่เท่ากันแล้ว ค่าสัมประสิทธิ์ T จะมีค่าต่ำกว่า 1 ซึ่ง สูตรการคำนวณ คือ $T = \sqrt{\chi^2/n\sqrt{(r-1)(c-1)}}$ และแก้ข้อจำกัดสัมประสิทธิ์ C ของเพียร์สัน ดังที่ได้กล่าวแล้ว สูตรการคำนวณ คือ $C = \sqrt{\chi^2/(\chi^2+n)}$ สถิติวิเคราะห์ความสัมพันธ์เหล่านี้ มีพื้นฐานมาจากค่าไคสแควร์ (Measures base on Chi-Square) พยายามปรับค่าไคสแควร์ให้อยู่ในรูป $(k\chi^2)/n$ รวมทั้งสูตรการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ด้วย

การคำนวณหาค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์มีสูตรในการคำนวณดังนี้

$$V = \sqrt{\chi^2/n(m-1)} \quad \text{หรือ} \quad V = \phi/\sqrt{(m-1)}$$

$$\text{โดยที่} \quad \phi = \sqrt{\chi^2/n}$$

$$\text{โดยที่} \quad \chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} \quad ; \nu = (r-1)(c-1)$$

เมื่อ V หมายถึง ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์

χ^2 หมายถึง ค่าไคสแควร์ (Pearson's χ^2 Statistic)

n หมายถึง ขนาดของกลุ่มตัวอย่าง

m หมายถึง ค่าต่ำสุดของจำนวนแถวหรือจำนวนสดมภ์

O_{ij} หมายถึง ค่าความถี่ที่สังเกตได้ (Observed Frequency) ในแถวที่ i สดมภ์ที่ j

E_{ij} หมายถึง ค่าความถี่ที่คาดหวัง (Expected Frequency) ในแถวที่ i สดมภ์ที่ j

โดยที่ค่า E_{ij} สามารถคำนวณที่ได้จากสูตร

$$E_{ij} = \frac{R_i C_j}{n}$$

เมื่อ R_i หมายถึง ความถี่รวมในแถวที่ i

C_j หมายถึง ความถี่รวมในสดมภ์ที่ j

ข้อตกลงเบื้องต้นของสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์

ในการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์นั้น จะทำการคำนวณค่าไคสแควร์ก่อน ด้วยเหตุนี้ Marascuilo (1971:410) จึงได้ให้ข้อตกลงเบื้องต้นของสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ไว้ดังนี้

1. มีความเป็นอิสระในการรวบรวมข้อมูล
2. ตัวแปรทั้งสองเป็น Multinomial
3. ค่าความถี่ที่คาดหวังภายในเซลล์ของตารางต้องมีค่ามากกว่า 5 หรือ $E_{ij} > 5$
4. มีกลุ่มตัวอย่างเพียงกลุ่มเดียว

การทดสอบนัยสำคัญของสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์

ทดสอบนัยสำคัญของสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ได้โดยสถิติทดสอบไคสแควร์สำหรับทดสอบการเป็นอิสระ จากสูตร ดังนี้

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad ; \nu = (r-1)(c-1)$$

ในการคำนวณหาค่าสัมประสิทธิ์ความสัมพันธ์แบบอื่นๆ เมื่อคำนวณค่าสัมประสิทธิ์ความสัมพันธ์แล้ว จึงนำไปทดสอบนัยสำคัญเพื่อต้องการทราบค่าความสัมพันธ์นั้นมีความเชื่อมั่นอย่างมีนัยสำคัญทางสถิติหรือไม่ โดยนำค่าสัมประสิทธิ์ความสัมพันธ์ที่ได้ไปทดสอบด้วยวิธีการทางสถิติที่เหมาะสม แต่สำหรับค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ต้องคำนวณค่าไคสแควร์มาก่อน ซึ่งค่าไคสแควร์ที่ได้นี้ ความจริงเป็นการทดสอบความเป็นอิสระของตัวแปรทั้งสองว่ามีความสัมพันธ์กันหรือไม่ โดยเมื่อ คำนวณค่าไคสแควร์ได้แล้วก็นำค่าที่ได้ไปเปรียบเทียบกับค่าไคสแควร์มาตรฐาน ที่ชั้นแห่งความอิสระเท่ากับ (r-1)(c-1) ปกติเมื่อค่าไคสแควร์ที่คำนวณได้มีนัยสำคัญทางสถิติ แล้วจะเป็นการบอก แต่เพียงว่าตัวแปรทั้งสองมีความสัมพันธ์กัน เท่านั้น ยังสรุปไม่ได้ว่าตัวแปรทั้งสองมีความสัมพันธ์กันในระดับใด ซึ่งสามารถรู้ได้จากการคำนวณหาค่าสัมประสิทธิ์ต่อไปในภายหลัง

ตอนที่ 3 การแจกแจงที่เกี่ยวข้องกับงานวิจัย

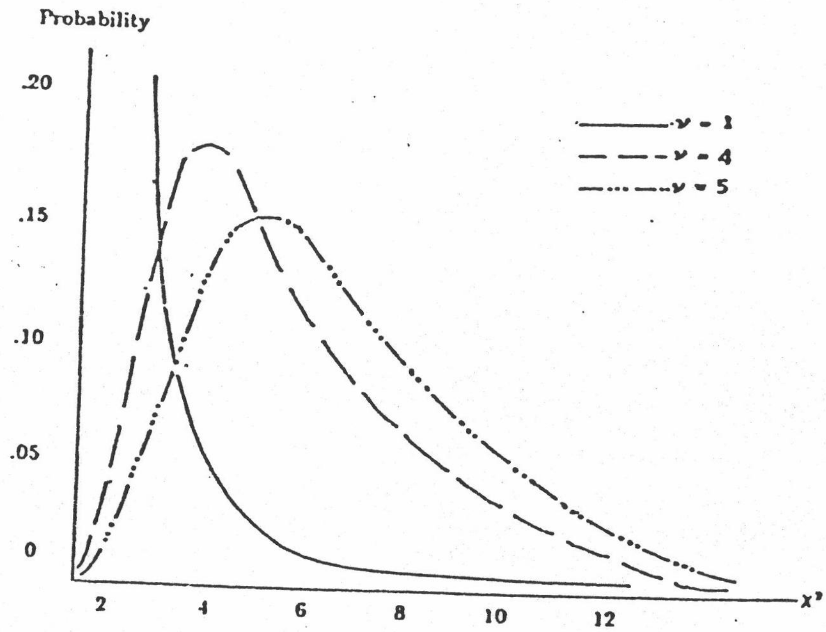
การแจกแจงไคสแควร์ เป็นการแจกแจงอีกแบบหนึ่งที่มีความสำคัญมาก ซึ่ง F.R. Helmert นักฟิสิกส์ชาวเยอรมัน ได้เป็นผู้ค้นคิดเป็นคนแรกในปี ค.ศ.1876 ต่อมาในปี ค.ศ. 1900 Karl Pearson ได้พัฒนาและเผยแพร่จนเป็นที่รู้จักกันอย่างแพร่หลาย

ฟังชั่นการแจกแจงไคสแควร์

การแจกแจงไคสแควร์เป็นการแจกแจงอีกประเภทหนึ่งของการแจกแจงแกมม่า (gamma distribution) ฟังชั่นการแจกแจงของไคสแควร์สามารถกำหนดในรูปของฟังก์ชั่นแกมม่าได้ดังนี้

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\nu/2)2^{\nu/2}}\, x^{\nu/2-1}\, e^{-x/2} & ;\ 0 < x < \infty \\ 0 & ;\ x \text{ มีค่าอื่น ๆ} \end{cases}$$

เมื่อ Γ คือฟังก์ชั่นแกมม่ามีค่าตามสูตร $= \int_0^{\infty} x^{p-1} e^{-x} dx$;p>0 เรียกตัวแปรสุ่ม x ว่ามีการแจกแจงไคสแควร์ ที่ชั้นแห่งความเป็นอิสระเท่ากับ ν การแจกแจงไคสแควร์จะกำหนดด้วยสัญญลักษณ์ χ^2 ซึ่งกราฟของการแจกแจงไคสแควร์ มีลักษณะดังแผนภาพที่ 1



แผนภาพที่ 1 การแจกแจงไคสแควร์ ที่มีชั้นแห่งความเป็นอิสระต่างกัน

การประยุกต์ใช้การแจกแจงแบบไคสแควร์ที่รู้จักกันทั่วไป คือ การอธิบายลักษณะการแจกแจงของ s^2 ของกลุ่มตัวอย่าง ซึ่ง ถ้า s^2 เป็นความแปรปรวนของตัวอย่างเชิงสุ่มขนาด n ที่สุ่มจากประชากรที่มีการแจกแจงปกติ ที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 แล้ว

$$\chi^2_{n-1} = \frac{(n-1)s^2}{\sigma^2}$$

จะมีค่าแจกแจงแบบไคสแควร์ ที่ชั้นแห่งความเป็นอิสระเท่ากับ n-1

<u>คุณสมบัติของการแจกแจงไคสแควร์</u>

1. ลักษณะการแจกแจงไม่สมมาตร โดย χ^2 จะมีค่าตั้งแต่ 0 ถึง ∞

2. การแจกแจงจะมีลักษณะแตกต่างกันตามค่าของชั้นแห่งความเป็นอิสระถ้า $\nu \geqslant 3$ การแจกแจงจะเบ้บวกหรือเบ้ขวา (positively skewed distribution) และเมื่อชั้นแห่งความเป็นอิสระมีค่าเพิ่มขึ้น การแจกแจงไคสแควร์จะมีลักษณะใกล้เป็นการแจกแจงแบบปกติมาตรฐาน

3. ค่าออร์ดิเนตสูงสุด จะอยู่ที่ χ^2

4. จุดสูงสุดของโค้งมีเพียงจุดเดียว (unimodal)

5. การแจกแจงเป็นไปตามกฎของการแจกแจงแบบมัลติโนเมียล (Multinomial Distribution)

6. สำหรับการแจกแจงของไคสแควร์ เมื่อชั้นแห่งความเป็นอิสระเท่ากับ ν จะมีลักษณะ ดังนี้

6.1 ค่าเฉลี่ยจะมีค่าเท่ากับ ν และมีค่าความแปรปรวนเท่ากับ 2

6.2 ค่ามัธยฐานเท่ากับ $\nu-2$

6.3 ค่าความเบ้ของการแจกแจงเท่ากับ $\sqrt{8/\nu}$

<u>ตอนที่ 4</u> วิธีการรวมเซลล์

หลักการรวมเซลล์นั้นไม่ได้มีกฎเกณฑ์ที่แน่นอนตายตัว ผู้ทำวิจัยจำเป็นต้องพิจารณาความเหมาะสมและความเป็นไปได้ของการรวม สำหรับการวิจัยนี้จะสนใจในวิธีการรวมเซลล์โดยกำหนดเงื่อนไข ดังนี้

1. ตัวแปรทั้งสองเป็นตัวแปรที่อยู่ในมาตรานามบัญญัติจำแนกได้หลายระดับ

2. ตัวแปรมีระดับที่ไม่เป็นลำดับและไม่เกิน 5 ระดับ

3. การรวมเซลล์ไม่ทำให้กลุ่มเซลล์ที่รวมนั้นผิดความหมาย

4. จะรวมกลุ่มเซลล์ของตัวแปรด้านแถวหรือด้านสดมภ์ด้านใดด้านหนึ่ง โดยยึดหลักในการรวมเพื่อต้องการให้ค่าความถี่ที่คาดหวังในเซลล์มีค่ามากกว่า 5

5. จะรวมกลุ่มเซลล์ในด้านสดมภ์เท่านั้น

เนื่องจากว่าการศึกษาครั้งนี้เป็นการศึกษาการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร โดยที่ ตัวแปรทั้งสองเป็นตัวแปรที่อยู่ในมาตรานามบัญญัติจำแนกได้หลายระดับ ซึ่งในการวิจัยผู้ทำวิจัย สามารถที่จะนำมาจำแนกให้อยู่ในรูปตาราง RxC และสามารถที่จะให้ตัวแปรใดเป็นตัวแปรด้านแถวหรือสดมภ์ก็ได้ ใคร่จะอธิบายให้เห็นได้ดังนี้

ถ้าให้ A เป็นตัวแปรด้านแถว(R) มี 3 ระดับ และ B เป็นตัวแปรด้านสดมภ์(C) มี 4 ระดับ เขียนตาราง RxC จะมีขนาด 3x4 (AxB) แต่ถ้าเปลี่ยนให้ B เป็นตัวแปรด้านแถว A เป็นตัวแปรด้านสดมภ์ เขียนตาราง RxC จะมีขนาด 4x3 (BxA) และถ้านำตาราง RxC ทั้ง 2 ลักษณะ มาคำนวณค่าความถี่ที่คาดหวัง ($E_{ij}=R_iC_j/n$) จากตาราง 3x4 (AxB) จะได้ $E_{34} = R_3C_4/n$ และตาราง 4x3 (BxA) จะได้ $E_{43} = R_4C_3/n$ จะเห็นว่า $E_{34} = E_{43}$ นอกจากจะสลับให้ตัวแปรใดเป็นตัวแปรด้านแถวหรือด้านสดมภ์ดังกล่าวแล้ว ผู้ทำวิจัยยังสามารถที่จะสลับตำแหน่งด้านแถวหรือตำแหน่งด้านสดมภ์ด้านใดด้านหนึ่ง และถ้าเป็นการพิจารณาทางด้านสดมภ์แล้ว ผู้ทำวิจัยสามารถที่จะเรียงลำดับสดมภ์ที่มีผลรวมของความถี่ที่มีค่าจากมากไปน้อย เนื่องจาก ความถี่ของค่าสังเกตที่อยู่ในแต่ละแถวและสดมภ์เป็นอิสระต่อกันย่อมสามารถสลับตำแหน่งจากเดิมได้ ผลรวมของสดมภ์ใดที่มีค่าต่ำย่อมเป็นผลให้ค่าความถี่ที่คาดหวังที่อยู่ในสดมภ์นั้นมีโอกาสต่ำด้วย ดังจะเห็นได้จากตัวอย่างตารางที่เสนอไว้ข้างล่างต่อจากนี้ ซึ่งมีผลรวมของความถี่ในแต่ละสดมภ์เป็น 33,30,28 และ 9 ค่าความถี่ที่คาดหวังที่มีค่าต่ำ (E_{ij}<5) จะเป็นเซลล์ที่อยู่ในสดมภ์ที่มีผลรวมต่ำที่สุด จากตารางตัวอย่างดังกล่าวนี้คือสดมภ์ที่ 4 ซึ่งมีผลรวมความถี่เท่ากับ 9 เมื่อค่าความถี่ที่คาดหวังของสดมภ์ใดมีค่าต่ำเช่นนี้แล้ว ในวิธีการรวมเซลล์ผู้ทำวิจัยจะรวมเซลล์ในสดมภ์ที่มีค่าผลรวมต่ำดังกล่าวกับสดมภ์ที่มีค่าผลรวมต่ำถัดไปเข้าด้วยกัน ด้วยหลักการพิจารณาที่กล่าวมาการวิจัยนี้จึงสนใจในวิธีการรวมกลุ่มเซลล์ และการรวมกลุ่มเซลล์นี้จะรวมเฉพาะกลุ่มเซลล์ของตัวแปรด้านสดมภ์เท่านั้น โดยจะทำการรวมกลุ่มเซลล์ในสดมภ์ปลายสุดกับสดมภ์ที่อยู่ถัดขึ้นมาตามลำดับ และในการรวมกลุ่มเซลล์ของตัวแปรด้านสดมภ์นี้จะรวมให้เหลือระดับตัวแปรด้านสดมภ์อย่างต่ำจำนวน 2 ระดับ(2 สดมภ์)

ผู้วิจัยขอเสนอตัวอย่างลักษณะตารางที่มีข้อมูลในเซลล์บางเซลล์ เมื่อนำมาคำนวณค่าความถี่ที่คาดหวังแล้วมีค่าน้อยกว่าหรือเท่ากับ 5 ดังตารางต่อไปนี้

ตารางข้อมูลเพื่อทดสอบว่ามีความสัมพันธ์ระหว่างระดับการบริหารการศึกษา กับการอ่านหนังสือพิมพ์ประจำวัน ในเดือนมิถุนายน 2534 หรือไม่ ระดับใด จากการสอบถามผู้บริหารการศึกษา 100 คน สมมุติว่าปรากฏผล ดังนี้

ระดับการบริหารการศึกษา	พฤติกรรมการอ่านหนังสือพิมพ์ (จำนวนวัน)				รวม
	มากกว่า 25วัน	16-25วัน	5-15วัน	น้อยกว่า 5วัน	
หัวหน้าหมวด	15	10	11	4	40
ผู้ช่วยผู้บริหาร	11	12	9	3	35
ผู้บริหาร	7	8	8	2	25
รวม	33	30	28	9	100

จากข้อมูลในแต่ละเซลล์ของตารางนำมาคำนวณค่าความถี่ที่คาดหวัง ซึ่งคำนวณจากสูตร $E_{ij} = R_iC_j/n$ จะได้ค่าความถี่คาดหวัง ดังนี้

ตารางแสดงความถี่ที่คาดหวัง

ระดับการบริหารการศึกษา	พฤติกรรมการอ่านหนังสือพิมพ์ (จำนวนวัน)				รวม
	มากกว่า 25วัน	16-25วัน	5-15วัน	น้อยกว่า 5วัน	
หัวหน้าหมวด	E_{11}=13.2	E_{12}=12	E_{13}=11.2	E_{14}=3.6	40
ผู้ช่วยผู้บริหาร	E_{21}=11.55	E_{22}=10.5	E_{23}=9.8	E_{24}=3.15	35
ผู้บริหาร	E_{31}=8.25	E_{32}=7.5	E_{33}=7	E_{34}=2.25	25
รวม	33	30	28	9	100

จะเห็นว่าค่าความถี่คาดหวังที่มีขนาดเล็ก นั่นคือ มีค่าน้อยกว่าหรือเท่ากับ 5 ($E_{ij} \leq 5$) จำนวน 3 เซลล์ ได้แก่ E_{14} , E_{24} และ E_{34} ตารางลักษณะนี้ เมื่อคำนวณจำนวนของตารางการณ์จรและค่าต่ำสุดของจำนวนแถวหรือสดมภ์ จะได้ผลดังนี้

จำนวนของตารางการณ์จรของการคำนวณค่าไคสแควร์ จะเท่ากับ (R-1)(C-1) = 6 และค่าต่ำสุดของจำนวนแถวหรือสดมภ์ (m-1) ในที่นี้ คือ (R-1) = 2

จากตัวอย่างดังกล่าวนี้ผู้ใช้สถิติอาจเกิดปัญหาในการที่จะตัดสินใจเลือกใช้วิธีการวิเคราะห์ระหว่างการรวมเซลล์กับไม่รวมเซลล์ เพราะมี E_{ij} ที่น้อยกว่าหรือเท่ากับ 5 ถ้าวิเคราะห์โดยวิธีการไม่รวมเซลล์ก็จะเป็นการผิดหลักที่ฝืนคำแนะนำหรือข้อตกลงเบื้องต้นและจะได้จำนวนของตารางการณ์จรของการคำนวณไคสแควร์และค่าต่ำสุดของจำนวนแถวหรือสดมภ์ดังค่าข้างบน แต่ถ้าวิเคราะห์โดยวิธีการรวมเซลล์แล้ว ต้องทำการรวมเซลล์ที่มีความถี่คาดหวังน้อยกว่าหรือเท่ากับ 5 กับเซลล์ที่อยู่ในสดมภ์ถัดไป

จากตัวอย่างตารางเดิมจะได้ตารางใหม่ที่เกิดจากการรวมเซลล์ซึ่งมีขนาด 3x3 มีลักษณะดังนี้

ระดับการบริหารการศึกษา	พฤติกรรมการอ่านหนังสือพิมพ์ (จำนวนวัน)			รวม
	มากกว่า 25วัน	16-25วัน	น้อยกว่า 16วัน	
หัวหน้าหมวด	15	10	15	40
ผู้ช่วยผู้บริหาร	11	12	12	35
ผู้บริหาร	7	8	10	25
รวม	33	30	37	100

จากข้อมูลในตารางใหม่นี้เมื่อนำมาคำนวณค่าความถี่ที่คาดหวัง จะได้ค่าความถี่ที่คาดหวัง ดังปรากฏผลดังนี้

ระดับการบริหาร การศึกษา	พฤติกรรมการอ่านหนังสือพิมพ์ (จำนวนวัน)			รวม
	มากกว่า 25วัน	16-25วัน	น้อยกว่า 16วัน	
หัวหน้าหมวด	E_{11}=13.2	E_{12}=12	E_{13}=14.8	40
ผู้ช่วยผู้บริหาร	E_{21}=11.55	E_{22}=10.5	E_{23}=12.95	35
ผู้บริหาร	E_{31}=8.25	E_{32}=7.5	E_{33}=9.25	25
รวม	33	30	37	100

จากตาราง 3x4 เดิมเมื่อรวมเซลล์แล้วจะเปลี่ยนเป็นตารางขนาด 3x3 ทำการคำนวณค่าคาดหวังแล้ว จะเห็นว่าได้ค่าความถี่ที่คาดหวังมากกว่า 5 ทุกเซลล์ ดังนั้นจะได้

จำนวนของตารางการณ์จรของการคำนวณค่าไคสแควร์ จะเท่ากับ (R-1)(C-1) = 4 และค่าต่ำสุดของจำนวนแถวหรือสดมภ์(m-1) ในที่นี้จะเป็นค่า (R-1) หรือ (C-1) ก็ได้เพียงค่าใดค่าหนึ่ง จะเท่ากับ 2

จากการพิจารณาการรวมเซลล์และไม่รวมเซลล์ของลักษณะตารางดังตัวอย่างนั้น จำนวนของตารางการณ์จรของการคำนวณค่าไคสแควร์จะต่างกัน โดยวิธีการรวมเซลล์จะมีจำนวนของตารางการณ์จรน้อยกว่า สำหรับค่าต่ำสุดของจำนวนแถวหรือสดมภ์ของการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะไม่ต่างกัน

ข้อมูลที่อยู่ในรูปตารางการณ์จรนั้นมีหลายลักษณะ ถ้าลักษณะข้อมูลที่นำมาศึกษาอยู่ในลักษณะตารางขนาด 3x3 โดยที่ E_{ij}<5 แล้ว เมื่อวิเคราะห์โดยวิธีการไม่รวมเซลล์แล้วจะได้

จำนวนของตารางการณ์จรของการคำนวณค่าไคสแควร์ = (3-1)(3-1) = 4

และค่าต่ำสุดของจำนวนแถวหรือจำนวนสดมภ์ (m) = (3-1) = 2

เมื่อวิเคราะห์โดยการรวมเซลล์ จะได้ตารางใหม่ซึ่งมีขนาด 3x2 และจะได้

จำนวนของตารางการณ์จรของการคำนวณค่าไคสแควร์ = (3-1)(2-1) = 2

และค่าต่ำสุดของจำนวนแถวหรือจำนวนสดมภ์ (m) = (2-1) = 1

จากการพิจารณาการรวมเซลล์และไม่รวมเซลล์ของลักษณะตาราง 3x3 นั้น จำนวน

ของตารางการณ์จรของการคำนวณค่าไคสแควร์จะต่างกัน โดยการรวมเซลล์จะมีจำนวนของตารางการณ์จรน้อยกว่า และค่าต่ำสุดของจำนวนแถวหรือสดมภ์ของการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะต่างกันด้วย โดยที่การรวมเซลล์จะมีค่าต่ำสุดของจำนวนแถวหรือสดมภ์ต่ำกว่า

<u>ตอนที่ 5</u> งานวิจัยที่เกี่ยวข้อง

ทองดี แย้มสรวล (2530) ได้ใช้วิธีการทดลองด้วยเทคนิคมอนติคาร์โลซิมูเลชั่น ทำการศึกษาเพื่อศึกษาลักษณะการแจกแจง การควบคุมการคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบสำหรับค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน เคนดอนเทา และแครมเมอร์วี เมื่อประชากรที่ศึกษามีลักษณะการแจกแจงแบบปกติสองตัวแปร และสัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่า 0.0, 0.1,..., 0.9 ศึกษาจากข้อมูลที่มีลักษณะเป็นอันดับในรูปตารางการณ์จร 5x5 กลุ่มตัวอย่างมีขนาด 150,200 และ 250 ผลปรากฏว่าสัมประสิทธิ์ของสเปียร์แมนประมาณค่า ρ ได้ใกล้เคียงที่สุด สถิติทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน เคนดอลเทา และแครมเมอร์วี สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากัน อำนาจการทดสอบสำหรับค่าสัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน มีอำนาจในการทดสอบสูงที่สุด

ประวิทย์ วชิระจงกล (2529) ได้ใช้วิธีการทดลองด้วยเทคนิคมอนติคาร์โลซิมูเลชั่นทำการศึกษาตัวสถิติที่ใช้วัดความสัมพันธ์ของตัวแปรแบ่งกลุ่ม คือ สัมประสิทธิ์เงื่อนไขของเพียร์สันสัมประสิทธิ์เงื่อนไขของชูโพร สัมประสิทธิ์ วี ของแครมเมอร์ สัมประสิทธิ์การทำนายของกัทแมน และสัมประสิทธิ์แบบกูดแมนและครัสคัล เพื่อดูว่าตัวสถิติตัวใดจะใช้วัดความสัมพันธ์ของตัวแปรแบบแบ่งกลุ่มได้ดีกว่ากัน โดยพิจารณาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เมื่อค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร (ρ) ที่ระดับ 0.1, 0.3, 0.5, 0.7 และ 0.9 ข้อมูลมีขนาด 20,30,50,100,200,500 และขนาดตารางการณ์จรมีขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5 ผลปรากฏว่าที่ ρ มีค่า 0.1 และ 0.3 ตัวสถิติที่ใช้วัดความสัมพันธ์ ได้ดีพอ ๆ กัน ρ มีค่า 0.5 ขนาดตัวอย่าง 50 ลงไปแล้ว บางกรณีสัมประสิทธิ์พี สัมประสิทธิ์ของชูโพร สัมประสิทธิ์ วี ของแครมเมอร์ จะใช้วัดความสัมพันธ์ได้ดีแต่ส่วนมากแล้วที่ ρ มีค่าที่สูงกว่า 0.5 สัมประสิทธิ์เงื่อนไขของเพียร์สัน

จะให้ผลดีกว่าตัวสถิติอื่น ๆ

วีณา เตชะพนาดร (2528)ได้ศึกษาความสัมพันธ์ระหว่างสัมประสิทธิ์ C และสัมประสิทธิ์ วี ของแครมเมอร์กับค่าทดสอบไคสแควร์โดยพิจารณาถึงอิทธิพลของขนาดตัวอย่าง ขนาดตาราง และการจัดแบ่งกลุ่มข้อมูล ซึ่งศึกษาจากข้อมูลเชิงสุ่มชนิดตัวแปรปกติสองตัวแปร เมื่อกลุ่มตัวอย่างขนาด 20,30,40,50,75 และ 100 ขนาดตารางระดับ 2x2 2x3 2x4 2x5 3x3 3x4 3x5 4x4 4x5 และ 5x5 ตัวแปรปกติสองตัวแปรมีสัมประสิทธิ์สหสัมพันธ์ตั้งแต่ 0.00 ถึง 0.98 โดยแบ่งเป็นช่วง ดังนี้ 0.00 ถึง 0.40 จะมีความห่าง 0.02 0.40 ถึง 0.60 จะมีความห่าง 0.01และ 0.60 ถึง 0.98 จะมีความห่าง 0.02 ผลปรากฏว่า เมื่อพิจารณาในแง่ของการทดสอบสมมติฐานขนาดตัวอย่าง และขนาดตารางจะทำให้ระดับนัยสำคัญจากการจำลอง สูงกว่าระดับนัยสำคัญจากทฤษฎีในแต่ละขนาดตัวอย่าง จะได้ค่าวิกฤตจากการจำลองแตกต่างกัน ที่ระดับนัยสำคัญและชั้นแห่งความอิสระเดียวกัน ถ้าพิจารณาในแง่ของความสัมพันธ์ระหว่างค่าคาดหวังของสัมประสิทธิ์สหสัมพันธ์กับค่าทดสอบไคสแควร์ พบว่าเมื่อขนาดของตัวอย่างและขนาดตารางเพิ่มขึ้นค่าคาดหวังของสัมประสิทธิ์สหสัมพันธ์จะลดลง ส่วนการจัดกลุ่มข้อมูลที่แตกต่างกันจะทำให้ค่าเฉลี่ยและความแปรปรวนของค่าไคสแควร์ แตกต่างกันในแต่ละกลุ่ม

Sarndal (1974:165-187) ได้ใช้วิธีซิมูเลชั่น เพื่อทำการศึกษาเปรียบเทียบวิธีการวัดค่าความสัมพันธ์แบบต่าง ๆ จำนวน 15 วิธี ทั้งที่เป็นวิธีการวัดค่าความสัมพันธ์แบบเดิมและที่ผู้วิจัยได้พัฒนาขึ้นมาใหม่ โดยต้องการศึกษาเปรียบเทียบมโนทัศน์พื้นฐาน และคุณสมบัติของความสัมพันธ์แบบต่าง ๆ เมื่อตัวแปร X อยู่ในมาตรานามบัญญัติ ที่จำแนกประเภทได้ตั้งแต่ 2 ประเภทขึ้นไป และตัวแปร Y อยู่ในมาตรานามบัญญัติ จัดอันดับ และอันตรภาค ขนาดของตัวอย่างที่ศึกษาจำนวน 400 และ ρ =0.0 ,0.1,...,1.0 จากการศึกษาเปรียบเทียบวิธีการวัดค่าความสัมพันธ์ของเคนดอลเทา (T_c) กับของแครมเมอร์ (V) ซึ่งผู้วิจัยเปรียบเทียบเมื่อตัวแปรทั้งสองอยู่ในมาตรานามบัญญัติ และบันทึกข้อมูลลงในตารางการณ์จร ผลปรากฏว่าเมื่อ ρ มีค่าเพิ่มขึ้นจาก 0 ถึง 1 ทั้งสัมประสิทธิ์ T_c และ V จะมีค่าเพิ่มขึ้นอย่างช้า ๆ จนมีค่าถึง 1 และเมื่อใช้ค่าแก้ในการวัดค่าความสัมมพันธ์ ที่ ρ = 0 ทั้งสัมประสิทธิ์ T_c และ V มีค่าใกล้เคียงกับค่า ρ แต่โดยทั่วไปแล้วการใช้ค่าแก้ในการวัดความสัมพันธ์ระหว่างตัวแปรทั้งสอง สัมประสิทธิ์ T_c สามารถใช้ได้ผลดีกว่าสัมประสิทธิ์ V