บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การวิจัยเป็นการศึกษาค้นคว้าหาความจริงในสาขาวิชาต่าง ๆ อย่างมีเป้าหมายหรือวัตถุประสงค์ที่แน่นอนและดำเนินการอย่างมีระเบียบโดยถูกต้องตามหลักวิชาเพื่อให้ได้ข้อความรู้ที่เชื่อถือได้ การทำวิจัยจึงจำเป็นที่จะต้องวางแผนและตรวจสอบการดำเนินงานทุกขั้นตอน และขั้นตอนสำคัญขั้นตอนหนึ่ง คือ การวิเคราะห์ข้อมูล ซึ่งงานวิจัยส่วนใหญ่จำเป็นต้องอาศัยระเบียบวิธีการทางสถิติช่วยในการสรุปผลและตอบคำถามตามที่ตั้งจุดมุ่งหมายไว้

งานวิจัยเป็นจำนวนมากที่การวิเคราะห์ข้อมูล มีวัตถุประสงค์เพื่อต้องการทราบความสัมพันธ์ระหว่างตัวแปรสองตัว ซึ่งรวบรวมมาจากประชากรเดียวกัน โดยมักจะตั้งคำถามว่ามีความสัมพันธ์กันในลักษณะใด งานวิเคราะห์ดังกล่าวสามารถดำเนินการได้ด้วยวิธีการวิเคราะห์ทางสถิติซึ่งสามารถส่งผลสรุปได้ทั้งทิศทางและระดับความสัมพันธ์ระหว่างตัวแปร การหาความสัมพันธ์ โดยวิธีการทางสถิติมีอยู่หลายวิธีขึ้นอยู่กับข้อตกลงของแต่ละวิธี ลักษณะตัวแปร และมาตราวัดค่าของตัวแปรว่าเป็นแบบใด ดังนั้นการเลือกใช้สถิติวิเคราะห์ความสัมพันธ์จึงจำเป็นต้องทราบลักษณะสำคัญหลายประการโดยเฉพาะอย่างยิ่ง ผู้วิจัยจำเป็นต้องทราบถึงมาตราการวัดของข้อมูลต่าง ๆ ซึ่ง Steven (1946) จำแนกไว้เป็น 4 ระดับ คือ มาตรานามบัญญัติ (Nominal scales) มาตราจัดอันดับ (Ordinal scales) มาตราอันตรภาค (Interval scales) และ มาตราอัตราส่วน (Ratio scales)

ความสัมพันธ์ระหว่างตัวแปรสองตัวจะมีค่ามากหรือน้อยนั้น งานการวิจัยโดยทั่วไปจะสรุปได้จากการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ซึ่งคำนวณได้จากข้อมูลของตัวแปรทั้งสอง สัญลักษณ์ที่ใช้แทนค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร คือ ρ (rho) ถ้า ρ มีค่าใกล้ ± 1 แสดงว่าตัวแปรที่ศึกษาทั้งสองตัวนั้นมีความสัมพันธ์กันมาก ถ้า ρ มีค่าใกล้ 0 แสดงว่าตัวแปรที่ศึกษาทั้งสองตัวนั้นมีความสัมพันธ์กันน้อย เครื่องหมายของ ρ นั้นเป็นการแสดงถึง ทิศทางของความสัมพันธ์ กล่าวคือ ถ้า ρ มีเครื่องหมายบวก

แสดงว่า ตัวแปรทั้งสองตัวนั้นมีความสัมพันธ์กันไปในทางเดียวกัน แต่ถ้า ρ มีเครื่องหมายลบ แสดงว่า ตัวแปรทั้งสองตัวนั้นมีความสัมพันธ์ในทางตรงข้าม ค่าสัมประสิทธิ์สหสัมพันธ์นอกจากจะเป็นค่าสถิติที่ให้ผลสรุปว่า ตัวแปรคู่ใดคู่หนึ่งมีความสัมพันธ์กันหรือไม่และในทิศทางใดแล้ว ค่าสัมประสิทธิ์สหสัมพันธ์เมื่อนำมายกกำลังสอง ยังเป็นค่าที่บอกถึงสัดส่วนความแปรปรวนของตัวแปรหนึ่งที่สามารถอธิบายจากตัวแปรอีกตัวแปรหนึ่ง

สัมประสิทธิ์สหสัมพันธ์ระหว่างสองตัวแปรเป็นที่รู้จักกันมากที่สุด คือ สัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product Moment Correlation Coefficient : r_{xy}) ซึ่งการที่จะตัดสินใจเลือกใช้วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันจะต้องคำนึงถึงข้อตกลงเบื้องต้น (Assumptions) ดังนี้

1. ตัวแปรทั้งสองเป็นตัวแปรต่อเนื่อง และมีการแจกแจงของประชากรเป็นแบบปกติสองตัวแปร (Bivariate normal distribution)

2. ความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นความสัมพันธ์เชิงเส้นตรง (Linear relationship)

3. มีความเป็นอิสระของข้อมูล (Independence between pairs)

การหาค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันนั้น Lindeman, Merenda and Gold (1980:63) กล่าวว่า ต้องเป็นไปตามข้อตกลงเบื้องต้นซึ่งไม่ควรที่จะฝ่าฝืน ถ้าฝ่าฝืนข้อตกลงเบื้องต้นข้อใดข้อหนึ่งก็ไม่ควรใช้วิธีหาค่าสัมประสิทธิ์สหสัมพันธ์วิธีนี้ควรเลือกวิธีหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างสองตัวแปรวิธีอื่นแทน

งานวิจัยทางสังคมศาสตร์มีแนวโน้มที่จะศึกษาตัวแปรครั้งละหลายๆ ตัว ซึ่งมักจะมีลักษณะแตกต่างกัน ข้อมูลจากตัวแปรบางตัวอาจวัดค่าได้ในมาตราอัตราส่วน หรืออันตรภาคแต่ข้อมูลจากตัวแปรบางตัวจะวัดค่าได้เพียงมาตราอันดับ หรือมาตรานามบัญญัติเท่านั้น ในบางกรณีผู้วิจัยจะเก็บรวบรวมข้อมูลที่มีลักษณะของตัวแปรอยู่ในมาตรานามบัญญัติ ซึ่งตัวแปรที่ศึกษาจะเป็นตัวแปรที่แบ่งเป็นระดับย่อย ๆ หลายระดับ เช่น เพศ วุฒิการศึกษา อาชีพ ภูมิลำเนา ฯลฯ

ในกรณีที่ผู้วิจัยต้องการศึกษาความสัมพันธ์ระหว่างตัวแปร โดยตัวแปรทั้งสองอยู่ในมาตรานามบัญญัติจำแนกได้หลายระดับต่างๆกัน สามารถนำมาจำแนกให้อยู่ในรูปตารางการณ์จร (Contingency table) ค่าที่ปรากฏในตารางการณ์จร จะเป็นความถี่ของค่าสังเกตที่รวบรวมมาได้ ซึ่งเรียกว่า ข้อมูลจำนวนนับ (Counted data) หรือข้อมูลจำแนกประเภท

(Categorical data) ตัวแปรแต่ละตัวแบ่งเป็นระดับย่อย ๆ ตั้งแต่ 2 ระดับขึ้นไป อยู่ในรูปตาราง RxC อาจเป็นแบบ 2x2 2x3 (หรือ 3x2) 3x3 3x4 (หรือ 4x3) เป็นต้น การทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว ในลักษณะนี้จะใช้การทดสอบไคสแควร์สำหรับการทดสอบความเป็นอิสระ (Chi-Square test of independence) การสรุปผลของการทดสอบด้วยไคสแควร์นี้ ถ้าผลการทดสอบปรากฏว่าปฏิเสธสมมติฐานศูนย์แสดงว่า ตัวแปรทั้งสองมีความสัมพันธ์กันอย่างมีนัยสำคัญเท่านั้นยังไม่สามารถบอกได้ว่ามีความสัมพันธ์กันมากหรือน้อยในระดับใด

การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรที่อยู่ในมาตรานามบัญญัตินั้น โดยปกติแล้วผู้วิจัยต้องการผลสรุปของการวิเคราะห์ที่สามารถให้รายละเอียดได้ชัดเจน กล่าวคือ ต้องการหาผลสรุปของการวิเคราะห์ที่สามารถบอกระดับความสัมพันธ์ของตัวแปรทั้งสองได้ ผู้วิจัยจำเป็นต้องวิเคราะห์ความสัมพันธ์ โดยใช้สถิติการวิเคราะห์ที่มีพื้นฐานมาจากค่าไคสแควร์ (Measures based on ChiSquare) ซึ่งมีสูตรที่ได้พยายามปรับค่าไคสแควร์ให้มีค่าใกล้เคียงกับค่า Random Variable ของไคสแควร์ และยังเป็นค่าที่ตอบคำถามได้ดีกว่าไคสแควร์ คือ สามารถบอกได้ว่าขนาดของความสัมพันธ์เป็นเท่าไร (Jacobson,1976: 436-438,Liebetran 1983:13-14, Blalock 1981:303-305,สมเพลิน เกษมรัตนสันติ 2532:38) สถิติเหล่านี้เรียกว่า สถิติวัดขนาดความสัมพันธ์ที่มีพื้นฐานมาจากค่าไคสแควร์ ได้แก่ ค่าสัมประสิทธิ์ฟี(ϕ) ค่าสัมประสิทธิ์ทีของชูโพร(Tschuprow's Contigency Coefficient :T) ค่าสัมประสิทธิ์ของเพียร์สัน (Pearson's Contigency Coefficient:C)ค่าสัมประสิทธิ์ วี ของแครมเมอร์ (Cramer's Contigency Coefficient:V) โดยปกติแล้วค่าสัมประสิทธิ์สหสัมพันธ์จะมีค่าสูงสุดเท่ากับ 1 แต่ค่าสัมประสิทธิ์ความสัมพันธ์ ที่มีพื้นฐานมาจากไคสแควร์ในแต่ละวิธีมีข้อจำกัด และมีความเหมาะสมกับสถานการณ์แตกต่างกันดังนี้

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

สัมประสิทธิ์ T ของชูโพรระหว่างตัวแปรสองตัวจะมีค่าเท่ากับ 1 เสมอหากตัวแปรทั้งสองมีความสัมพันธ์กันอย่างสมบูรณ์ แต่ยังมีข้อแม้ว่าจำนวนแถว และจำนวนสดมภ์จะต้องเท่า

กัน ถ้าหากจำนวนแถวและจำนวนสดมภ์ไม่เท่ากัน ค่าสัมประสิทธิ์ T จะมีค่าต่ำกว่า 1 (Blalock 1981:304)

สำหรับสัมประสิทธิ์ C มีข้อจำกัดดังที่ Yule and Kendall (1950:53-54) ได้กล่าวสอดคล้องกับ Siegel (1956: 201) ว่าค่าสัมประสิทธิ์ C จะมีค่าเท่ากับ 0 เมื่อตัวแปรทั้งสองไม่มีความสัมพันธ์กัน แต่จะไม่มีโอกาสที่จะสูงถึง 1 ได้เลย เมื่อจำนวนแถวและสดมภ์มีค่าเท่ากัน ค่าสูงสุดของสัมประสิทธิ์ 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 ค่าสัมประสิทธิ์ C จะมีค่าสูงสุด = 0.894 เป็นต้น สัมประสิทธิ์ C เป็นที่รู้จักโดยทั่วไปแต่ด้วยข้อจำกัด ดังกล่าวจึงได้มีการศึกษา และพยายามสร้างดัชนีความสัมพันธ์ที่แก้ความจำกัดได้ แครมเมอร์ (Harald Crammer) 1946 ได้ศึกษาสัมประสิทธิ์ในรูปตารางการณ์จร RxC โดยศึกษาจากสูตรการคำนวณสัมประสิทธิ์ฟี กล่าวคือ เป็นการสร้างดัชนีความสัมพันธ์ที่ขยายขนาดของตารางการณ์จร 2x2 เป็น RxC และเสนอการใช้ค่าสัมประสิทธิ์แครมเมอร์ (V) โดยทำให้สัมประสิทธิ์ความสัมพันธ์ที่ได้สามารถมีค่าสูงสุดได้เท่ากับ 1 และสามารถใช้ได้กับตารางการณ์จรขนาดต่าง ๆ ดังที่ Hays (1973:745) Blalock (1981:305) Jacobson (1976:438) แนะนำให้ใช้แทนสัมประสิทธิ์ C

การคำนวณหาค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ มีสูตรที่ใช้ในการคำนวณดังนี้

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

โดยที่ χ^2 คือ $\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$ มีชั้นแห่งความเป็นอิสระเท่ากับ $(R-1)(C-1)$

n คือ ขนาดของกลุ่มตัวอย่าง

m คือ ค่าต่ำสุดของจำนวนแถวหรือจำนวนสดมภ์

ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะสูงหรือต่ำนั้น ขึ้นอยู่กับลักษณะความสัมพันธ์ระหว่างตัวแปร และจากสูตรการคำนวณจะเห็นว่าขนาดตัวอย่าง ขนาดของตารางการณ์จร จำนวนแถว และสดมภ์ มีความเกี่ยวข้องกับค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ กล่าวคือ ถ้าขนาดความสัมพันธ์มีค่าคงที่ขนาดตัวอย่างสูง ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะต่ำ และจำนวนของแถวและสดมภ์ในตารางการณ์จรสูง ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ก็จะต่ำเนื่องจากได้นำค่าต่ำสุดของจำนวนแถวหรือสดมภ์ (m-1) มาใช้ในการคำนวณซึ่งถ้าจำนวนแถว(R)หรือจำนวนสดมภ์ (C) ค่าใดค่าหนึ่งต่ำแล้วจะให้ค่า m-1 ต่ำส่งผลให้ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์สูง แต่การคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์นั้นยังมีความเกี่ยวข้องกับค่าไคสแควร์ กล่าวคือ ข้อมูลที่อยู่ในรูปตารางการณ์จรตารางเดียวกันนั้น จะต้องคำนวณค่าไคสแควร์มาก่อน แล้วจึงนำค่าไคสแควร์ที่ได้มาแทนค่าในการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ซึ่งจากสูตรการคำนวณค่าไคสแควร์ ถ้าจำนวนขนาดของตารางการณ์จร (RxC) มาก จำนวนทั้งแถวและสดมภ์มากขึ้นค่าไคสแควร์จะมากขึ้น และจำนวนตัวอย่างมากขึ้นค่าไคสแควร์ก็จะมากทั้งที่ความสัมพันธ์ยังคงเหมือนเดิม ค่าไคสแควร์ที่สูงมากไม่สามารถใช้เป็นเครื่องบ่งชี้ได้ว่า ตัวแปรทั้งสองมีความสัมพันธ์กันมาก (สุชาติ ประสิทธิ์รัฐสินธุ์ และคณะ,2523:33) การคำนวณไคสแควร์นั้นจะเห็นว่าหลักใหญ่ คือ เป็นการเปรียบเทียบความแตกต่างของความถี่ที่คาดหวังกับความถี่ที่ได้จากการสังเกต ซึ่งในที่นี้ความถี่ที่คาดหวังจะเป็นตัวพารามิเตอร์และความถี่ที่สังเกตได้ จะเป็นค่าประมาณจากพารามิเตอร์ ถ้าทั้งสองค่านี้ใกล้เคียงกันการทดสอบจะไม่มีนัยสำคัญ ส่งผลให้ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ไม่มีนัยสำคัญด้วย

ค่าความถี่ที่คาดหวังในเซลล์ (E_{ij}) ของตารางการณ์จรนั้นไม่ควรต่ำ เพราะถ้ามีค่าต่ำแล้ว อาจทำให้การตัดสินใจสรุปผลที่ได้คลาดเคลื่อนจากที่ควรเป็น นักสถิติหลายท่านได้เสนอแนะเกี่ยวกับค่าความถี่ที่คาดหวังขั้นต่ำ (Minimum Expected Frequency) ว่าควรจะมีขนาดเท่าใดซึ่งได้มีคำแนะนำแตกต่างกันไป และในหนังสือสถิติมาตรฐานที่ใช้กันอยู่ทั่วไปนั้นมักจะแนะนำว่าค่าความถี่ที่คาดหวังควรเป็น 5 หรือมากกว่า (Marascuilo and Mcsweeney 1977, Yamane 1973) ในทางปฏิบัติถ้าความถี่ที่คาดหวังน้อยกว่า 5 ให้รวมเซลล์ที่มีความถี่ที่คาดหวังต่ำกับเซลล์ที่อยู่ถัดไปเพื่อให้ได้ความถี่ที่คาดหวังเป็น 5 หรือมากกว่า แล้วจึงจะทำการคำนวณค่าไคสแควร์ (Cochran,1952:314-345) และในการคำนวณค่าสัมประสิทธิ์ความ

สัมพันธ์ 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

ดังนั้นถ้าความถี่ที่คาดหวังในเซลล์ของตารางการณ์จร ในการวิเคราะห์มีค่าน้อยกว่าหรือเท่ากับ 5 หรือ $E_{ij}\leq 5$ ซึ่งในทางปฏิบัติมักจะพบลักษณะเช่นนี้เสมอจึงทำให้ผู้ใช้สถิติมีปัญหาในการใช้สถิติดังกล่าว เนื่องจากเกิดความลังเลว่าจะใช้วิธีการวิเคราะห์ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์อย่างไร ระหว่างวิธีการวิเคราะห์ในลักษณะแตกต่างกันใน 2 กรณีที่มีการรวมเซลล์และไม่รวมเซลล์ ถ้าเลือกใช้วิธีการวิเคราะห์โดยไม่รวมเซลล์แล้วจะเป็นวิธีการวิเคราะห์ที่ผู้ใช้ได้ฝืนคำแนะนำหรือข้อตกลงเบื้องต้น และในแง่การวิเคราะห์อีกลักษณะหนึ่งถ้าผู้ใช้สถิติวิเคราะห์สัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์เลือกใช้วิธีการวิเคราะห์โดยการรวมเซลล์แล้ว ก็ไม่ทราบว่าค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะแตกต่างไปจากที่ควรจะเป็น หรือไม่ เพราะการวิเคราะห์ลักษณะนี้จะต้องนำข้อมูลจากตารางเดิมที่ได้มาทำการรวมเซลล์ของข้อมูล เนื่องจากได้คำนวณค่าความถี่ที่คาดหวังในเซลล์นั้นแล้วมีค่าน้อยกว่าหรือเท่ากับ ซลล์แล้วจะได้ลักษณะตารางใหม่ที่มีขนาด RxC น้อยกว่าขนาดตาราง RxC เดิม และขนาดตาราง RxC นั้นมีผลต่อค่าไคสแควร์ดังที่กล่าวมาแล้วว่า ค่าไคสแควร์จะมีค่าต่ำเมื่อจำนวนของตารางการณ์จรลดลง เพราะทั้งจำนวนแถว(R) และสดมภ์(C) มีความเกี่ยวข้องนั่นเองแต่ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์นั้นจะนำเฉพาะค่าต่ำสุดของจำนวนแถว หรือสดมภ์เพียงค่าใดค่าหนึ่งเท่านั้นมาเกี่ยวข้อง (m-1) ซึ่งถ้าการวิเคราะห์โดยการรวมเซลล์ค่าต่ำสุดของจำนวนแถวหรือสดมภ์(m-1)ที่นำมาคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์โดยบางกรณีของลักษณะตารางยังเป็นค่าเดิม แต่อย่างไรก็ตามสูตรการคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์จะต้องคำนวณค่าไคสแควร์มาก่อนและการวิเคราะห์ที่ใช้วิธีการรวมเซลล์กับไม่รวมเซลล์เมื่อความถี่ที่คาดหวังมีขนาดเล็กนั้นจะเห็นว่าการวิเคราะห์ที่ไม่รวมเซลล์จะมีจำนวนของเซลล์ในตารางการณ์จรมากกว่าย่อมส่งผลให้ค่าไคสแควร์แตกต่างจากการวิเคราะห์ที่ใช้วิธีการรวมเซลล์ แต่สำหรับค่าต่ำสุดของจำนวนแถวหรือจำนวนสดมภ์ของ

การคำนวณค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์นั้นอาจจะมีค่าแตกต่างหรือไม่แตกต่าง ขึ้นอยู่กับลักษณะตารางผลการวิเคราะห์ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์เมื่อความถี่ที่คาดหวังมีขนาดเล็ก ในลักษณะการเปรียบเทียบระหว่างการวิเคราะห์ที่มีการรวมเซลล์กับไม่รวมเซลล์ ขณะนี้ยังไม่มีหลักฐานและข้อสรุปว่าควรจะเลือกใช้การวิเคราะห์วิธีใด กรณีไหนสามารถที่จะใช้ได้ในลักษณะของการวิเคราะห์ที่ไม่มีการรวมเซลล์ เมื่อความถี่ที่คาดหวังมีขนาดเล็ก และจะมีความคลาดเคลื่อนกับลักษณะการวิเคราะห์ที่มีการรวมเซลล์เป็นเช่นไรนั้นยังไม่มีผู้ใดศึกษา ผู้นำสถิติวิเคราะห์ดังกล่าวไปใช้ ยังเป็นปัญหาในการตัดสินใจที่จะเลือกใช้ ระหว่างการวิเคราะห์ที่มีการรวมเซลล์กับไม่รวมเซลล์ว่าจะใช้วิธีการวิเคราะห์ใด กรณีไหน ซึ่งถ้าหากว่าผลการวิเคราะห์ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ เมื่อความถี่ที่คาดหวังมีขนาดเล็ก ได้ผลสรุปที่ไม่แตกต่างกันแล้วผู้ใช้สถิติน่าจะเลือกใช้ วิธีการวิเคราะห์ กรณีที่ไม่มีการรวมเซลล์เพราะไม่ต้องยุ่งยากเพิ่มขั้นตอนในการคำนวณแต่ขณะนี้ยังไม่มีหลักฐานและข้อสรุปที่ชัดเจนได้

ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ระหว่างการวิเคราะห์ โดยวิธีการรวมเซลล์กับไม่รวมเซลล์วิธีการวิเคราะห์ใดมีความเหมาะสมในการเลือกใช้ เพื่อทำให้ผลของการวิเคราะห์มีประสิทธิภาพนั้น สามารถพิจารณาได้จากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE) โดยถ้าพบว่าวิธีการวิเคราะห์ใดให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำแล้ว ควรจะเลือกใช้วิธีการวิเคราะห์นั้น และค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์เป็นค่าที่สามารถบอกความมากน้อยของความสัมพันธ์แต่ไม่สามารถบอกเป็นระดับช่วงได้ ดังนั้น ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ที่ได้ของวิธีการวิเคราะห์ทั้งสองวิธีดังกล่าวจะมีความแตกต่างกันหรือไม่สามารถพิจารณาได้จากค่ามัธยฐาน ซึ่งเป็นค่าที่อยู่ตำแหน่งกลางของ ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ที่เรียงลำดับความมากน้อยของแต่ละวิธีการวิเคราะห์จากการทดลอง การวิจัยนี้นอกจากจะเสนอค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ค่ามัธยฐานแล้ว ยังได้ทดสอบความแตกต่าง ลักษณะการแจกแจงของ ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ที่ได้จากการทดลองทุกๆ ค่าระหว่างการวิเคราะห์ โดยการรวมเซลล์กับไม่รวมเซลล์ในตารางการณ์จร เมื่อความถี่ที่คาดหวังมีขนาดเล็ก และทดสอบความแตกต่างค่ามัธยฐานของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ระหว่างวิธีการวิเคราะห์ที่แตกต่างกันของทั้งสองกรณีดังกล่าวจากผลการทดลองโดยใช้วิธีซิมูเลท(Simulate) ซึ่งจะทำให้ผลสรุปที่เด่นชัดขึ้นภายใต้ภาวะที่คล้าย

กับการทดลอง คือสามารถกำหนดลักษณะการแจกแจงของประชากร ค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร สามารถระบุ หรือจำกัดขนาดตัวอย่างขนาดตาราง ได้ด้วยเครื่องคอมพิวเตอร์ ซึ่งเป็นวิธีที่สามารถดำเนินการได้ในปัจจุบัน

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบผลของการวิเคราะห์โดยการรวมเซลล์ กับไม่รวมเซลล์ที่มีต่อค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ โดยมีวัตถุประสงค์เฉพาะ ดังนี้

1.เพื่อเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ระหว่างการวิเคราะห์โดยการรวมเซลล์กับไม่รวมเซลล์ในตารางการณ์จร เมื่อความถี่ที่คาดหวังมีขนาดเล็ก โดยที่สัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่าเท่ากับ 0.0 , 0.1 ,..., 0.9

2.เพื่อเปรียบเทียบลักษณะการแจกแจง ของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ระหว่างการวิเคราะห์โดยการรวมเซลล์กับไม่รวมเซลล์ในตางการณ์จร เมื่อความถี่ที่คาดหวังมีขนาดเล็กโดยที่สัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่าเท่ากับ 0.0 , 0.1 ,..., 0.9

3.เพื่อเปรียบเทียบค่ามัธยฐาน ของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ ระหว่างการวิเคราะห์โดยการรวมเซลล์ กับไม่รวมเซลล์ในตารางการณ์จร เมื่อความถี่ที่คาดหวังในเซลล์มีขนาดเล็ก โดยที่สัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่าเท่ากับ 0.0 , 0.1 ,..., 0.9

ข้อตกลงเบื้องต้น

การวิจัยครั้งนี้เลือกใช้โปรแกรมย่อยสับรูทีน (subprogram subroutine) ซึ่งเป็นคำสั่ง หรือกลุ่มคำสั่งที่เขียนขึ้นมาเพื่อให้เครื่องคอมพิวเตอร์ ทำการคำนวณปัญหาเฉพาะอย่าง หรือปัญหาเฉพาะเรื่อง ตามกรณีที่ต้องการศึกษา โปรแกรมย่อยสับรูทีนนี้ มีหลักฐานและ

การศึกษาในรูปแบบการแปลงข้อมูลซึ่งมีอยู่แล้ว สร้าง Random number เพื่อสร้างลักษณะตัวแปรของประชากรที่ต้องการศึกษา

ขอบเขตของการวิจัย

1. ตัวแปรที่ศึกษาในการวิจัยครั้งนี้ มีดังต่อไปนี้

   1.1 ตัวแปรตาม (Dependent variables) ได้แก่

      1. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ กรณีการวิเคราะห์ที่มีการรวมเซลล์กับไม่รวมเซลล์ จากผลการทดลอง

      2. ลักษณะการแจกแจง ของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ กรณีการวิเคราะห์ที่มีการรวมเซลล์กับไม่รวมเซลล์จากผลการทดลอง

      3. ค่ามัธยฐานของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ กรณีการวิเคราะห์ที่มีการรวมเซลล์ กับไม่รวมเซลล์จากผลการทดลอง

   1.2 ตัวแปรอิสระ (Independent variables) ได้แก่

   วิธีการวิเคราะห์ ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ในลักษณะที่แตกต่างกัน 2 กรณี ได้แก่

      1. กรณีการวิเคราะห์ที่ไม่รวมเซลล์ เมื่อความถี่ที่คาดหวังบางเซลล์มีค่าน้อยกว่าหรือเท่ากับ 5

      2. กรณีการวิเคราะห์ที่มีการรวมเซลล์ เมื่อความถี่ที่คาดหวังมีค่าน้อยกว่าหรือเท่ากับ 5

2. ค่าสัมประสิทธิ์สหสัมพันธ์ของประชากรที่ศึกษา คือ ρ = 0.0,0.1,...,0.9

3. การวิจัยครั้งนี้ศึกษาจากประชากรที่มีลักษณะการแจกแจงแบบปกติสองตัวแปร (Bivariate normal distribution) เท่านั้น

4. การวิจัยครั้งนี้กำหนดกลุ่มตัวอย่างที่ศึกษามีขนาด 200

เนื่องจากว่าการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์นั้น ควรใช้กลุ่มตัวอย่างประมาณ 20 เท่า ของตัวแปร และถ้าไม่ทราบจำนวนตัวแปร ก็ควรใช้กลุ่มตัวอย่าง ไม่ต่ำกว่า 100 (Lindeman,Merenda and Gold 1980:55) และเนื่องจากการวิจัยครั้งนี้มีสถานการณ์ที่ศึกษาหลายกรณี ผู้วิจัยจึงกำหนดขนาดตัวอย่างที่ศึกษาเพียงขนาดเดียว

5. จากการสำรวจงานวิจัยที่วิเคราะห์ความสัมพันธ์ของข้อมูลนามบัญญัติ ในการทำวิทยานิพนธ์ของนิสิตบัณฑิตวิทยาลัย คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างปีการศึกษา 2526-2531 จำนวน 30 เล่ม พบว่า ในงานวิจัยนั้นมีการแบ่งระดับตัวแปรที่ศึกษาในมาตรานามบัญญัติออกเป็น 2-5 ระดับเป็นส่วนใหญ่ โดยคิดเป็นร้อยละ 80

ผู้วิจัยจึงกำหนดตารางการณ์จรที่ใช้ศึกษามีขนาดดังนี้ คือ

2x3 2x4 2x5 3x3 3x4 3x5 4x4 4x5 5x5

ขนาดตารางที่มี $R = C$ ได้แก่ตาราง 3x3 4x4 5x5

ขนาดตารางที่มี $R \neq C$ ได้แก่ตาราง 2x3 2x4 2x5 3x4 3x5 4x5

6. การวิจัยครั้งนี้กำหนดการรวมเซลล์ที่มีค่าความถี่ที่คาดหวังมีขนาดเล็ก (น้อยกว่าหรือเท่ากับ 5)ในด้านสดมภ์เท่านั้น ดังเหตุผลซึ่งได้เสนอวิธีการรวมเซลล์ไว้ในบทที่ 2โดยรวมกลุ่มเซลล์ในสดมภ์ จำนวน 1 ถึง 3 สดมภ์

7. การวิจัยครั้งนี้กำหนดจำนวนสดมภ์ในการรวม ดังต่อไปนี้

7.1 กำหนดการรวมเซลล์ที่มีค่าความถี่ที่คาดหวังมีขนาดเล็ก (น้อยกว่าหรือเท่ากับ 5)จำนวน 1 สดมภ์โดยรวมสดมภ์ดังกล่าวเข้ากับสดมภ์ที่อยู่ถัดไปเพื่อให้อยู่ในสดมภ์เดียวกันเป็น 1 สดมภ์ (ลดลง 1 สดมภ์)

ศึกษาได้กับตาราง 2x3 2x4 2x5 3x3 3x4 3x5 4x4 4x5 5x5

ทำการทดลองทั้งหมด 90 กรณี

7.2 กำหนดการรวมเซลล์ที่มีค่าความถี่ที่คาดหวังมีขนาดเล็ก (น้อยกว่าหรือเท่ากับ 5) จำนวน 2 สดมภ์ โดยรวม 2 สดมภ์ดังกล่าวเข้ากับสดมภ์ที่อยู่ถัดไปเพื่อให้อยู่ในสดมภ์เดียวกันเป็น 1 สดมภ์ (ลดลง 2 สดมภ์)

ศึกษาได้กับตาราง 2x4 2x5 3x4 3x5 4x4 4x5 5x5

ทำการทดลองทั้งหมด 70 กรณี

7.3 กำหนดการรวมเซลล์ที่มีค่าความถี่ที่คาดหวังมีขนาดเล็ก (น้อยกว่าหรือเท่ากับ 5) จำนวน 3 สดมภ์ โดยรวม 3 สดมภ์ดังกล่าวเข้ากับสดมภ์ที่อยู่ถัดไปเพื่อให้อยู่ในสดมภ์เดียวกันเป็น 1 สดมภ์ (ลดลง 3 สดมภ์)

ศึกษาได้กับตาราง 2x5 3x5 4x5 5x5 ทำการทดลองทั้งหมด 40 กรณี

8. แต่ละกรณีของการวิจัยครั้งนี้จะทำการทดลองซ้ำจำนวน 4,000 ครั้ง

คำจำกัดความในการวิจัย

ค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์	หมายถึง ค่าสถิติที่บอกถึงขนาดความสัมพันธ์ระหว่างตัวแปรสองตัวที่อยู่ในมาตรานามบัญญัติ ซึ่งคำนวณได้จากกลุ่มตัวอย่างจำนวน 200 ที่สุ่มมาจากประชากรกลุ่มเดียวจากการทดลอง
การแจกแจงค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์	หมายถึง การแจกแจงสุ่มของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ที่ได้ทุก ๆ ค่า จากการทดลองจำนวน 4000 ครั้ง
ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE)	หมายถึง ค่าความคลาดเคลื่อนของค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ที่ได้จากการทดลอง 4000 ครั้งจากค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร ยกกำลังสองเฉลี่ย
ตารางการณ์จรที่ตัวแปรมีลักษณะเป็นนามบัญญัติ	หมายถึง ตารางที่แต่ละช่องสามารถบรรจุจำนวนหรือความถี่ของตัวแปรทั้งสองโดยที่ตัวแปรทั้งสองอยู่ในมาตรานามบัญญัติ ซึ่งมีการจัดประเภทอยู่ระดับ 2 - 5 รายการ

ความถี่ที่ได้จากการสังเกต หมายถึง จำนวนความถี่ของข้อมูลที่เป็นอิสระต่อกัน ประกอบเป็นกลุ่มตัวอย่าง (Observed Frequency) ซึ่งได้จากการสุ่มในเงื่อนไขใดเงื่อนไขหนึ่งจากการทดลองใช้สัญลักษณ์ O_{ij}

ความถี่ที่คาดหวัง (Expected Frequency) หมายถึง ความถี่ที่ควรจะเป็นตามสมมติฐาน ทฤษฎีหรือหลักการต่าง ๆ ใช้สัญลักษณ์ E_{ij} สำหรับการวิจัยครั้งนี้ E_{ij} คำนวณได้จาก $E_{ij} = R_i C_j / n$

ความถี่ที่คาดหวังขนาดเล็ก หมายถึง ความถี่ที่คาดหวังที่มีค่าน้อยกว่าหรือเท่ากับ 5 ทุกเซลล์ของสดมภ์นั้น

ซิมูเลชั่น (Simulation) หมายถึง เทคนิคการผลิตและสุ่ม(Technique of random number generation)เลขสุ่มที่ได้จะมีลักษณะ (Random) แต่ละหมายเลขจะมีโอกาสเกิดได้เท่าเทียมกัน

ประโยชน์ที่คาดว่าจะได้รับ

1. ผลการวิจัยครั้งนี้จะเป็นประโยชน์แก่ผู้ใช้สถิติ ได้มีผลสรุป และหลักฐานในการใช้สถิติวิเคราะห์สัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์และมีเกณฑ์ในการตัดสินใจว่า เมื่อค่าความถี่ที่คาดหวังมีขนาดเล็ก ควรจะรวมเซลล์หรือไม่รวมเซลล์อย่างไร เพื่อทำให้ผลของการตัดสินใจมีประสิทธิภาพ และเหมาะสม

2. ผลการวิจัยครั้งนี้จะทำให้ผู้ใช้สถิติ เกิดความเข้าใจในการใช้สัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ เป็นสถิติวิเคราะห์ความสัมพันธ์ และเพื่อให้ผู้ใช้สถิติได้นำค่าสัมประสิทธิ์ความสัมพันธ์ วี ของแครมเมอร์ไปขยายผลของการวิจัย ที่ต้องการวิเคราะห์ความสัมพันธ์ของข้อมูลนามบัญญัติได้กว้างขวางยิ่งขึ้น