รายงานการวิจัย
เรื่อง
โปรแกรมช่วยในการตัดสินใจเลือกวิธีวิเคราะห์ทางสถิติสำหรับข้อมูลจากตารางการณ์จร 2 ทาง
Method Selection Program for Statistical Analysis of Two-way Contingency Tables

โดย
รศ.อุมาพร จันทศร

ได้รับทุนสนับสนุนงานวิจัยจากเงินงบประมาณแผ่นดินประจำปีงบประมาณ 2553
คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ส่วนที่ 1 รายละเอียดเกี่ยวกับโครงการ

ชื่อโครงการ (ภาษาไทย) โปรแกรมช่วยในการตัดสินใจเลือกวิธีวิเคราะห์ทางสถิติสำหรับข้อมูลจากตารางการณ์จร 2 ทาง

(ภาษาอังกฤษ) Method Selection Program for Statistical Analysis of Two-way Contingency Tables

แหล่งเงิน งบประมาณแผ่นดิน

ประจำปีงบประมาณ 2553 จำนวนเงินที่ได้รับการสนับสนุน 132,400 บาท

ระยะเวลาทำการวิจัย 1 ปี ตั้งแต่ 1 ต.ค. 2552 ถึง ก.ย. 2553

รศ.อุมาพร จันทศร โทร. 02-3264111 ต่อ 6278

สาขาวิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

kcumapor@kmitl.ac.th

คำสำคัญ: โปรแกรมช่วยตัดสินใจเลือกสถิติทดสอบ , ตารางการณ์จร 2 ทาง , ตารางชนิด 2×2 หรือ r×r หรือ r×c หรือ r×c แบบมีลำดับที่

Keywords: Program for Supporting the Decision in Statistical Tests , Two way contingency Table , 2×2 or r×r or r×c or Ordered Table r×c

บทคัดย่อ

การศึกษาครั้งนี้ เป็นการสร้างโปรแกรม เพื่อช่วยให้ผู้ใช้ที่ต้องการวิเคราะห์ข้อมูลของ 2 ตัวแปรที่บันทึกเป็นความถี่ และจัดลงตารางแจกแจงสองทาง สามารถเลือกสถิติทดสอบได้ถูกต้องตามหลักทฤษฎีสถิติ คือคำนึงถึง ขนาดของตาราง(2×2 หรือ r×r หรือ r×c หรือ r×c แบบมีลำดับที่) มาตราวัดของข้อมูล (นามบัญญัติ หรือ เรียงลำดับ) ลักษณะของประชากร (หนึ่ง หรือ สอง ประชากร) ขนาดตัวอย่าง (ขนาดเล็ก หรือ ขนาดใหญ่) และประเภทของการทดสอบ(การวัดความสอดคล้อง หรือ วัดความสัมพันธ์ หรือ ทดสอบผลต่างของสัดส่วน) การใช้งานจะอยู่ในลักษณะถามตอบระหว่างผู้ใช้กับคอมพิวเตอร์ทางหน้าจอ และจะได้คำตอบในขั้นตอนสุดท้าย ซึ่งอาจจะเป็นสถิติใด ๆ จากทั้งหมด 18 แบบ ซึ่งส่วนใหญ่เป็นหัวข้อของสถิติที่ไม่ใช้พารามิเตอร์(Nonparametric Statistics) โดยโปรแกรมนี้จะนำเสนอในเว็บไซต์ชื่อ www.kmitl.ac.th/stat นอกจากนี้โปรแกรมยังได้อธิบายเนื้อหาและตัวอย่างที่เกี่ยวข้อง รวมทั้งจะแนะนำวิธีวิเคราะห์ด้วยคำสั่งจากโปรแกรม SPSS

Abstract

The objective of this study is to develop program to assist those who want to analyse data of 2 variables which collected by counting and presented in 2 way contingency table. The program would allow them to select theoretically proper test statistic regarding type of table (2×2 or r×r or r×c or Ordered Table r×c), measurement scale, population characteristics, sample size, type of hypothesis and The program will be so designed in such a way that user can communicate with computer by means of question and answer through the monitor. The final answer will end up with any kind of statistical test from 18 alternatives. All of them are nonparametric tests. For those who are interesting can find the program at website www.kmitl.ac.th/stat Moreover, the program is capable of telling the contents, demonstrating example concerned and advise the way to apply SPSS program for the analysis.

สารบัญ

บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

ทิศทางของการวิจัยของประเทศไทยอยู่ที่การวิจัยเพื่อนำไปสู่การพึ่งตนเองในระดับประเทศ โดยเริ่มจากการสร้างปัญญาให้แก่ประชาชน การพัฒนาพลังงานที่เหมาะสม พัฒนาเทคโนโลยีและซอฟแวร์ ตลอดจนผลิตยาและเวชภัณฑ์จากผลผลิตท้องถิ่น ซึ่งผลจากการพึ่งพาตนเองได้ จะทำให้ลดการนำเข้า เกิดเทคโนโลยีใหม่ ๆ ที่จะช่วยเพิ่มประสิทธิภาพการผลิตให้มีผลเพิ่มขึ้นแต่ลดต้นทุนลง (1) และเครื่องมือที่สำคัญชนิดหนึ่งของงานวิจัยคือ วิธีการทางสถิติ (Statistical Methods) ที่มีส่วนเกี่ยวข้องกับงานวิจัยตั้งแต่การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล ไปจนถึงการวิเคราะห์ข้อมูลเพื่อสรุปผลถึงกลุ่มใหญ่ หรือที่เรียกว่ากลุ่มประชากร โดยมีวิธีการวิเคราะห์ทางสถิติมากมายที่สามารถเลือกใช้ได้ โดยต้องคำนึงถึงวัตถุประสงค์ของงานวิจัย ลักษณะข้อมูลที่บันทึกมา ตลอดจนข้อกำหนดเบื้องต้นของสถิติเหล่านั้น ถ้าเลือกวิธีการวิเคราะห์ได้ถูกต้อง ผลสรุปเกี่ยวกับประชากรก็จะน่าเชื่อถือและถูกต้อง

ข้อมูลประเภทหนึ่งที่บันทึกได้ง่าย เสียค่าใช้จ่ายน้อย และใช้ทรัพยากรในการบันทึกน้อย คือ ข้อมูลจำแนกประเภท (Categorical Data) เช่น ข้อมูลเกี่ยวกับอาการบาดเจ็บของผู้ป่วยจากอุบัติเหตุทางรถยนต์ คือบาดเจ็บเล็กน้อย บาดเจ็บ และบาดเจ็บสาหัส หรือข้อมูลของผลการรักษาผู้ป่วยด้วยยาชนิดใหม่ที่ผู้ป่วยสามารถประเมินด้วยตนเองว่า มีอาการดีขึ้น เหมือนเดิม หรือแย่ลง ถ้าทำการบันทึก 2 ตัวแปรพร้อมกัน เช่น เพศของผู้ป่วยกับอาการหลังการรักษา นิยมที่จะนำเสนอเป็นความถี่ในเซลต่าง ๆ ของตารางแจกแจงความถี่แบบสองทาง ซึ่งตารางในลักษณะนี้ นิยมเรียกชื่อว่า “ตารางการณ์จร” (The Contingency Table หรือ The Cross Classification Table) ตามคำนิยามที่ Karl Pearson ได้กำหนดไว้ในปี 1904 (2)

นักวิจัยส่วนใหญ่นิยมใช้การทดสอบไคสแควร์กับตารางการณ์จร ถึงความเป็นอิสระกันระหว่าง 2 ตัวแปรที่กำหนดในแนวนอนและแถวตั้งของตารางการณ์จรนั้น ๆ เพราะการใช้การทดสอบไคสแควร์นั้นเข้าใจได้ง่าย ไม่มีข้อกำหนดเบื้องต้นมาก และสามารถใช้โปรแกรมสำเร็จรูปทั่ว ๆ ไป วิเคราะห์แทนมือได้ง่าย แต่ถ้าตัวแปรใดตัวแปรหนึ่ง หรือตัวแปรที่สองของตารางการณ์จรนั้นสามารถบอกถึงลำดับความสำคัญ เช่น ตัวอย่างข้างต้น คือ อาการบาดเจ็บที่มี 3 ระดับ คือ เล็กน้อย บาดเจ็บ และ

สาหัส ควรเลือกใช้สถิติอื่น ๆ ที่สามารถสรุปผลได้ถึงลำดับเหล่านั้นด้วย ก็จะทำให้ผลการวิเคราะห์ได้รายละเอียดมากขึ้น (3) โดยจะเรียกตารางการณ์จรเช่นนี้ว่า “Contingency Tables with ordered Categories” ซึ่งอาจจะเป็นชนิด “Single Ordered Tables หรือ Doubly Ordered Tables ก็ได้ สถิติที่ใช้วิเคราะห์ควรจะใช้ค่าสัมประสิทธิ์แกมม่า (Gamma) หรือสัมประสิทธิ์ซอมเมอร์ดี (Somer’s D) หรือใช้สถิติทดสอบของวิลคอกซันแมนวิทนีย์ชนิดที่มีค่าซ้ำมาก หรือ ฯลฯ

นอกจากประเด็นข้างต้นแล้ว ยังมีประเด็นเกี่ยวกับตัวอย่างที่เก็บรวบรวมข้อมูลมาเป็นตัวอย่าง 2 ชุดอิสระกันหรือไม่ หรือเป็นตัวอย่างชุดเดิมที่เก็บข้อมูล 2 ครั้ง เช่น ก่อนให้ยา และหลังให้ยา ถ้าจัดตารางเป็นชนิด 2x2 ก็ควรใช้สถิติทดสอบของแมคนีมาร์ หรือตารางการณ์จรที่มีตัวแปรทางแถวนอนและแถวตั้ง มีลักษณะย่อยเหมือนกัน เช่น วิธีการรักษาผู้ป่วยด้วยยา 3 ชนิด และแพทย์ 2 ท่านเป็นผู้ประเมินว่าผู้ป่วยตัวอย่าง n คน ควรจะใช้ยาชนิดใดรักษา แพทย์ทั้ง 2 ท่านให้ผลสอดคล้องกันหรือไม่ ควรใช้สถิติแคปปา (Kappa Statistic) วัดความสอดคล้องกันของผู้ประเมิน 2 คน เช่น ตัวอย่างที่ผลการวัดความสอดคล้องเป็นแบบสมบูรณ์จากผู้ป่วย 30 คนเป็นดังตารางต่อไปนี้

		แพทย์คนที่ 1		
		ยา 1	ยา 2	ยา 3
แพทย์คนที่ 2	ยา 1	10	-	-
	ยา 2	-	10	-
	ยา 3	-	-	10

จากที่กล่าวมาเบื้องต้น จะพบว่าการเลือกใช้สถิติให้เหมาะสมกับข้อมูลที่มีอยู่ หรือวัตถุประสงค์ของงานวิจัย อาจเป็นเรื่องยากสำหรับนักวิจัยที่มีความรู้ทางสถิติไม่มาก ทำให้เลือกใช้สถิติที่ไม่เหมาะสม อาจเนื่องมาจากไม่ทราบว่ามีสถิติอื่น ๆ ที่เหมาะสมสำหรับข้อมูลชุดนั้น ๆ จึงทำให้ผลการวิเคราะห์ขาดความน่าเชื่อถือ

ระบบที่พัฒนาขึ้นนี้จำกัดขอบเขตอยู่ที่ว่า เป็นระบบที่จะช่วยแนะแนวให้กับผู้ใช้ตัดสินใจเลือกวิธีวิเคราะห์ทางสถิติ โดยจะนำเสนอในลักษณะอธิบายเนื้อหาพร้อมตัวอย่างและการวิเคราะห์ และมุ่งเน้นในแง่ของนำกระบวนการที่จะทำให้ผู้ใช้สามารถทำความเข้าใจวิธีการวิเคราะห์ทางสถิติได้ดีขึ้นและน่าสนใจมากขึ้น การใช้งานจะเป็นไปในลักษณะถามตอบระหว่างคอมพิวเตอร์และผู้ใช้ทางหน้าจอ คอมพิวเตอร์จะถามรายละเอียดต่าง ๆ เกี่ยวกับข้อมูลที่บันทึกมาว่าถูกสุ่มมาจากประชากรลักษณะใด มีมาตรวัดของข้อมูลแบบใด เป็นต้น และจะนำคำตอบที่ได้ไปค้นหาวิธีการวิเคราะห์ที่เหมาะสมแบบใดแบบหนึ่งจากทั้งหมด 14 แบบ ที่เก็บไว้เป็นฐานความรู้ในโปรแกรม ส่วนวิธีการวิเคราะห์จะแนะนำให้ผู้ใช้เลือกใช้คำสั่งต่าง ๆ จากโปรแกรมสำเร็จรูป SPSS และ MINITAB โดยจะแสดงขั้นตอนการใช้คำสั่งจากโปรแกรมทั้งสองตามลำดับ ซึ่งจะทำให้ผู้ใช้ใช้งานได้ง่ายขึ้น ดังนั้นระบบที่พัฒนาขึ้นมาจึงไม่ใช่ระบบที่สร้างขึ้นมาเพื่อแข่งขันกับโปรแกรมสำเร็จรูปในด้านการวิเคราะห์ข้อมูล แต่มุ่งเน้นที่จะสนองความต้องการในส่วนของการเลือกวิธีวิเคราะห์ ซึ่งโปรแกรมสำเร็จรูปทั้งสองไม่ได้เน้นหรือจัดทำไว้

ในส่วนของสถิติวิเคราะห์ที่สามารถเลือกใช้ได้จะมีมากมายขึ้นอยู่กับลักษณะของประชากรที่สนใจศึกษา คือประชากรเดียว หรือ 2 ประชากร แบบสัมพันธ์กันหรืออิสระกัน หรือ k ประชากร ตารางที่บันทึกข้อมูลจะจัดได้ในรูปแบบใด คือ 2x2 หรือ 2xC หรือ rxC หรือ rxr ขนาดตัวอย่างมีขนาดเล็กหรือใหญ่ เป็นต้น

1. 2 วัตถุประสงค์ของโครงการวิจัย

พัฒนาระบบเพื่อช่วยตัดสินใจเลือกวิธีการวิเคราะห์ทางสถิติ เพื่อสรุปผลกับข้อมูลจากตารางการณ์จร 2 ทาง ซึ่งจะรวมทุกกรณีตั้งแต่สุ่มตัวอย่างจากประชากรเดียว สองประชากรแบบสัมพันธ์กัน สองประชากรแบบอิสระกัน หรือ k ประชากรซึ่งจะประกอบด้วยการทดสอบด้วยสถิติต่อไปนี้

1. ค่า Odds Ratio
2. ค่า Relative Risk
3. ค่าผลต่างของสัดส่วนที่สนใจ
4. การทดสอบไคสแควร์สำหรับความเป็นอิสระ

5. การทดสอบของฟิสเชอร์
6. การทดสอบของไคสแควร์สำหรับ 2 และ k ตัวอย่างที่เป็นอิสระ
7. การทดสอบของแมคนีมาร์
8. การทดสอบของโบวเคอร์ (Bowker Test)
9. การทดสอบของแมนวิทนีย์ชนิดที่มีซ้ำมาก
10. การทดสอบครัสคาล-วอลลิส ชนิดที่มีซ้ำมาก
11. การทดสอบ Jonckheere-Terpstra ชนิดที่มีซ้ำมาก
12. การทดสอบ Kappa สำหรับการวัดความสอดคล้องของผู้ประเมิน 2 คน
13. สัมประสิทธิ์แกมมา (Gamma Coefficient)
14. สัมประสิทธิ์ซอมเมอร์ (Somer's Coefficient)

1.3 ขอบเขตของโครงการวิจัย

จะเป็นการพัฒนาระบบที่มีขอบเขตดังนี้

1. เป็นการเลือกวิธีการทดสอบสมมติฐานทางสถิติเกี่ยวกับความแตกต่างกัน ของค่าสัดส่วนที่สนใจหรือความสัมพันธ์กันของสองตัวแปรในประชากรเท่านั้น ไม่รวมถึงวิธีการประมาณค่าพารามิเตอร์ (Estimation of Parameter)
2. เป็นการเลือกวิธีการทดสอบสมมติฐานทางสถิติเกี่ยวกับตารางการณ์จรชนิด 2 ทาง ซึ่งอาจจะเป็นแบบ 2x2, rxk หรือ kxk เท่านั้น ไม่รวมถึงตารางชนิดสามทางขึ้นไป (Three-Way Contingency Table)
3. เป็นการเลือกสถิติวิเคราะห์สำหรับตารางการณ์จรซึ่งมีความหมายว่า ค่าในเซลต่าง ๆ เป็นความถี่เท่านั้น ไม่รวมถึงตารางแบบอื่น ๆ ที่อาจบันทึกค่าเป็นค่าสังเกต

แบบต่าง ๆ เช่น เป็นตารางแจกแจงความถี่ หรือตารางที่บันทึกค่าสังเกตเป็นค่า 0, 1 เป็นต้น

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ระบบสามารถช่วยเสนอแนะให้นักวิจัยเลือกประเภทของการวิเคราะห์ทางสถิติได้ถูกต้องยิ่งขึ้น

2. ระบบสามารถช่วยผู้วิจัยที่ทำการเก็บรวบรวมข้อมูลแล้ว สามารถเลือกวิธีการวิเคราะห์ที่เหมาะสมให้กับข้อมูลที่มีอยู่

3. ระบบช่วยนักวิจัยในด้านการทดสอบสมมติฐานเมื่อมีข้อมูลในมาตรวัดแบบนามบัญญัติหรือเรียงลำดับ หรือแบบอันตรภาคชั้นขึ้นไป

4. เป็นการส่งเสริมให้งานวิจัยเห็นความสำคัญของการเลือกใช้วิธีการทางสถิติ ให้เหมาะสมกับงานวิจัยเพื่อเพิ่มคุณภาพงานวิจัย

5. เป็นแนวทางให้ผู้ใช้ระบบที่มีความรู้ทางสถิติไม่มากเลือกใช้วิธีการทางสถิติ หรือกำหนดเรื่องที่เกี่ยวข้องกับวิธีการทางสถิตินั้น ๆ เพื่อใช้เป็นแนวทางในการศึกษาเพิ่มเติมต่อไป

6. ระบบนี้จะสามารถช่วยผู้วิจัยที่มีความรู้ทางสถิติไม่มากนักเกิดให้ความมั่นใจในการเลือกใช้วิธีการทางสถิติ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 บทนำและทฤษฎี

ข้อมูลที่เก็บบันทึกจากเรื่องใดเรื่องหนึ่ง ค่าของตัวแปรอาจจะมีลักษณะเป็นกลุ่ม (category) เช่น ชนิดของโรคที่ผู้ป่วยเป็น หรือเกรดที่นักศึกษาได้รับ แทนที่จะเป็นค่าปริมาณ แบบต่อเนื่อง **ถ้ามีการศึกษาจากข้อมูลที่มี 2 ตัวแปรพร้อมกัน มักจะนำเสนอในรูปตารางที่เรียกว่า ตารางการณ์จร ซึ่งจะบันทึกความถี่ (frequency) ของการเกิดตัวแปรทั้งสองในกลุ่มย่อยต่างๆ**

การวิเคราะห์ตารางการณ์จร จะเป็นการศึกษาว่าตัวแปรทั้งสองที่มีลักษณะเป็นกลุ่มย่อยต่างๆที่แยกออกจากกันโดยเด็ดขาด(ซึ่งอาจจะเรียงลำดับได้) **มีความสัมพันธ์กันในเชิงเส้นตรงหรือไม่** โดยเป็นการหาสัมประสิทธิ์สหสัมพันธ์ และการทดสอบนัยสำคัญ รวมทั้งขบวนการทดสอบสมมติฐานทางสถิติเกี่ยวกับ**ความเป็นอิสระของตัวแปรทางแถวนอนและแถวตั้ง**ในตารางการณ์จร

ตัวแปรทั้งสองอาจจะมีการกำหนดล่วงหน้าได้ว่า ตัวแปรใดตัวแปรหนึ่งเป็นตัวแปรต้นหรืออิสระ(เช่นการสุ่มตัวอย่างจากสอง(หรือมากกว่า)ประชากรที่เป็นอิสระกัน)**และอีกตัวแปรเป็นตัวแปรตาม**

หรือไม่สามารถกำหนดล่วงหน้า ได้คือเป็นตัวแปรตามทั้งคู่

จะยกตัวอย่าง ตารางที่มีรูปแบบต่างๆดังนี้

ตารางที่ 1 ผลการรักษาโรคชนิดหนึ่งด้วยยาในปริมาณต่างกัน หลังการรักษาในระยะเวลาหนึ่งผู้ป่วยให้ความเห็นเกี่ยวกับยาชนิดนี้ว่า พอใจหรือไม่พอใจ เช่นได้ข้อมูลดังนี้

ผลการรักษาผู้ป่วย

ปริมาณยาที่ใช้	พอใจ	ไม่พอใจ	รวม
5 มก.	16	48	64
10 มก.	40	20	60

ตัวแปรตาม คือผลการรักษาซึ่งมี 2 คำตอบ ส่วนตัวแปรต้นคือปริมาณของยาที่ใช้ ซึ่งทั้งสองตัวแปรสามารถเรียงลำดับได้ จึงเรียกตารางเช่นนี้ว่า **Doubly Ordered Table**

ตารางที่ 2 ผลการรักษาผู้ป่วยจากยา A-D และการเปลี่ยนแปลงที่ผู้ป่วยรับรู้ในลำดับต่างๆ

ยา	การเปลี่ยนแปลงที่ผู้ป่วยรับรู้			
	มาก	ปานกลาง	น้อย	รวม
ยา A	16	5	6	27
ยาB	6	7	19	32
ยาC	5	2	7	14
ยาD	1	0	10	11

ตัวอย่างนี้ ตัวแปรตามคือการปลี่ยนแปลงที่ผู้ป่วยรับรู้ที่เรียงลำดับได้ ตัวแปรต้นคือชนิดของยาที่ใช้ จะเรียกตารางเช่นนี้ว่า **Singly Ordered Table** เนื่องจากตัวแปรทางแถวตั้งสามารถเรียงลำดับได้

และตัวอย่างเกี่ยวกับ**ตัวแปรตามแบบนามบัญญัติ** เป็นดังนี้

ตารางที่ 3 จำนวนสมาชิกพรรคการเมืองต่างๆ จากเขตท้องที่ต่างๆ

พรรคการเมืองที่สังกัด	เขตท้องที่			
	เหนือ	ใต้	กลาง	ตะวันออกเฉียงเหนือ
พรรค ก	221	160	360	140
พรรค ข	200	291	160	311
พรรค ค	208	106	316	97

ตัวอย่างนี้ตัวแปรตาม คือ พรรคการเมืองที่สังกัด เมื่อเขตท้องที่เป็นตัวแปรอิสระ จะเรียกตารางเช่นนี้ว่า **Nominal Attribute Categories Table**

ในการทดสอบสมมติฐาน เกี่ยวกับความสัมพันธ์ของตัวแปรทางแนวนอนและแนวตั้ง จำเป็นต้องมีข้อกำหนดเกี่ยวกับการได้มาซึ่งข้อมูล ว่าต้องเป็น ลักษณะสุ่ม ซึ่งอาจได้จากการวางแผนการทดลอง หรือการสำรวจตัวอย่าง ในขณะที่การใช้ข้อมูลในอดีต (Historical data) จำเป็นต้องระมัดระวัง อาจจำเป็นต้องกำหนดตัวแปรอิสระอื่นๆ ให้คงที่ เช่น ข้อมูลในตัวอย่างที่ผ่านมา ในตัวอย่างที่ 2 และข้อควรระวังอีกประการ คือ ขนาดตัวอย่าง เพียงพอที่จะใช้สถิติทดสอบได้หรือไม่ ควรให้ในแต่ละเซล มีความถี่ ที่ไม่เป็น 0 หรือน้อยมาก เป็นจำนวนมาก

การเลือกใช้การวิเคราะห์ทางสถิติที่เหมาะสม **ต้องคำนึงถึงมาตราวัดข้อมูล** ของตัวแปรเป็นสำคัญ ฉะนั้น การเข้าใจถึงมาตราวัดข้อมูลจึงเป็นหัวใจที่สำคัญ ในการเลือกใช้สถิติที่เหมาะสม (appropriate statistics) ซึ่งจะนำไปสู่การสรุปผลที่น่าเชื่อถือ จะกล่าวถึงมาตราวัดข้อมูล ดังนี้

การวัด หมายถึง การกำหนดค่าตัวเลขให้กับสิ่งของหรือเหตุการณ์ต่างๆ โดยใช้กฎ อย่างใดอย่างหนึ่ง กฎเหล่านี้มีหลายแบบจึงเป็นที่มาของมาตราวัด 4 แบบ คือ

1.มาตรานามบัญญัติ (The Nominal or Classificatory Scale)

เป็นมาตราวัดขั้นต่ำสุดซึ่งนิยมใช้ตัวเลขหรือสัญลักษณ์ใดๆ จัดข้อมูลเป็นกลุ่มต่างที่แยกออกจากกันโดยเด็ดขาด เช่น ใช้ตัวเลข 1, 2 และ 3 แทนคนไข้ที่ป่วยด้วยโรคต่างๆ ซึ่งมารับการรักษาที่โรงพยาบาลแห่งหนึ่ง ตัวเลขหรือสัญลักษณ์เหล่านี้ไม่อาจบอกความแตกต่างในเทอมของการดีกว่า สูงกว่าหรือเลวกว่า คือมีคุณค่าเท่ากันหรือเหมือนกัน และไม่มีความหมายของตัวเลขจะนำมาบวก ลบ คูณ หรือหาร ตามวิธีพีชคณิต ไม่ได้ รวมทั้งอาจสลับเปลี่ยนค่ากันได้ เช่น ให้ 0 แทนเพศหญิง 1 แทนเพศชาย หรือ 0 แทนเพศชาย และ 1 แทนเพศหญิง

ตัวอย่างอื่นๆ เช่น การบันทึกข้อมูลตามศาสนาที่นับถือ
การบันทึกข้อมูลตามสถานภาพสมรส เช่น โสด สมรส หม้าย
การบันทึกข้อมูลตามระดับการศึกษา เช่น ประถม มัธยม ปริญญาตรี

สิ่งที่นำมาวิเคราะห์ คือ ความถี่ในแต่ละกลุ่ม

2.มาตราเรียงลำดับ (The Ordinal or Ranking Scale)

เป็นการวัดที่แสดงความแตกต่างของคุณภาพเช่นกัน แต่สามารถบอกความแตกต่างของแต่ละกลุ่มได้ในเทอม มากกว่า ดีกว่า เลวกว่าได้ แต่ไม่สามารถวัดเป็นตัวเลขได้ เช่น แบ่งคนไข้ออกเป็นกลุ่มต่างๆ ดังนี้ รักษาไม่ได้ผล รักษาแล้วได้ผลดี หรือแบ่งกลุ่มออกตามความเห็นต่อเรื่องๆ หนึ่ง ดังนี้ กลุ่มไม่เห็นด้วย กลุ่มไม่มีความเห็น กลุ่มเห็นด้วย หรือกลุ่มนักศึกษาที่ได้ผลการสอบตามเกรดต่างๆ คือ A B C D F ความแตกต่างระหว่างกลุ่มนั้นสามารถบอกได้เพียงว่า ดีกว่า เลวกว่าอย่างไร แต่ไม่สามารถบอกระยะห่างของความแตกต่างได้ รวมทั้งระยะห่างนั้น ไม่จำเป็นต้องเท่ากันด้วย เช่น นักศึกษาที่สอบได้เกรด A เก่งกว่า นักศึกษาที่ได้เกรดอื่นๆ แต่ไม่ได้เก่งเป็น 4 เท่าของคนได้ F หรือ 2 เท่าของคนที่ได้เกรด B

การบันทึกข้อมูลบางอย่างจำเป็นต้องใช้มาตราวัดเช่นนี้ เช่น การชิมอาหาร 3 ชนิด ผู้ชิมจะบอกเพียงว่าชนิดใดอร่อยที่สุดและอันดับ 2, 3 หรือ ดมกลิ่นน้ำหอม ก็บันทึกได้เพียงว่า ชอบชนิด

ใดมากที่สุด และอันดับรองลงไป เราอาจใช้ตัวเลข 1, 2 และ 3 หรืออาจใช้เลข อื่นๆ 8 25 35 (แต่นิยมใช้ 1, 2 และ 3 เพราะเข้าใจได้ง่ายกว่า)

การวิเคราะห์ข้อมูลในมาตราวัดนี้ มักนำค่าลำดับที่ (คือ 1, 2 ....) มาวิเคราะห์ เช่น หาเครื่องหมายของผลต่างของลำดับที่ (sign difference)

ทั้ง 2 มาตราข้างต้นจัดได้ว่าข้อมูลที่บันทึกมาเป็นข้อมูลเชิงคุณภาพ ดังนั้นการวิเคราะห์ข้อมูลจะใช้ได้เฉพาะสถิติที่ไม่ใช้พารามิเตอร์เท่านั้น

3.มาตราอันตรภาค (The Interval Scale)

เป็นมาตราวัดที่มีคุณสมบัติเพิ่มขึ้นจากมาตราวัดแบบที่ 2 คือ ทราบระยะห่างของความแตกต่างระหว่างกลุ่ม 2 กลุ่ม อย่างชัดเจน เช่น ความแตกต่างระหว่างค่า 20 และ 30 เท่ากับ 10 และเท่ากับ ความแตกต่างระหว่างค่า 30 กับ 40 มาตราวัดนี้มักวัดค่าเป็นเชิงปริมาณ แต่ไม่มีจุดศูนย์ หรือจุดเริ่มต้นที่แท้จริง ตัวอย่างที่นิยมใช้กับมาตราวัดนี้คือ อุณหภูมิ, ค่า I.Q., ความดัน, คะแนนให้แก่ความสามารถต่างๆ (ส่วนมากเป็นคะแนนด้านจิตวิทยา) ข้อมูลเกี่ยวกับอุณหภูมิ เช่น องศาเซลเซียส 0° C คือจุดเยือกแข็งในขณะที่ 32° F เป็นจุดเยือกแข็งของฟาเรนไฮต์ ความแตกต่างระหว่าง 30° C กับ 10° C นั้นทราบเพียงว่าห่างกัน = 20 ° C แต่ไม่มีความหมายว่าเป็น 3 เท่าระหว่างกัน ดังนั้นมาตราวัดนี้ใช้ได้เพียงเครื่องหมายบวกและลบในเชิงพีชคณิตเท่านั้นยังไม่รวมกับการใช้เครื่องหมายคูณและหาร อีกตัวอย่างหนึ่งสำหรับข้อมูลที่มีมาตราวัดนี้ คือ คะแนนสอบ ถ้าได้คะแนนสอบ = 0 ก็ไม่ได้หมายความว่าไม่มีความรู้เลยหรือถ้าได้คะแนนสอบ 50 ก็ไม่ได้หมายความว่ามีความรู้เป็น 2 เท่าของผู้ที่สอบได้คะแนนเท่ากับ 25 การวิเคราะห์ข้อมูลมาตราวัดนี้ อาจใช้แบบใช้พารามิเตอร์เมื่อข้อกำหนดเบื้องต้นนั้นเป็นจริง

4.มาตราอัตราส่วน (The Ratio Scale)

มารตราวัดนี้มีความสมบูรณ์ที่สุด คือ มีค่าในเชิงตัวเลขที่แท้จริง มักเป็นข้อมูลทางด้านวิทยาศาสตร์ เช่น ทางฟิสิกส์ ซึ่งวัดความยาว น้ำหนัก ความหนาแน่น ความต้านทานไฟฟ้า เป็นต้น ซึ่งมีจุดศูนย์ที่แท้จริง เช่น น้ำหนัก 0 ก็คือไม่มีน้ำหนักเลย และ น้ำหนัก 5, 10 กิโลกรัม ก็หนักเป็น 5 และ 10เท่าตามลำดับของน้ำหนัก 1 กิโลกรัม จึงเป็นที่มาของของชื่อมาตราวัด เนื่องจากมีความหมายแท้จริงของอัตราส่วน การวิเคราะห์ข้อมูลมาตราวัดนี้ซึ่งเป็นข้อมูลเชิงปริมาณก็ควรใช้สถิติที่ใช้พารามิเตอร์ถ้าข้อกำหนดเบื้องต้นเป็นจริง แต่ถ้าไม่เป็นจริงก็ควรใช้สถิติที่ไม่ใช้พารามิเตอร์

เนื่องจากข้อมูลที่มีมาตราวัดสูงสามารถจัดใหม่ให้กลายเป็นข้อมูลมาตราวัดแบบต่ำกว่าได้ ตัวอย่างเช่น ข้อมูลเกี่ยวกับรายได้ถ้าสามารถบันทึกรายได้ของแต่ละคนได้ ก็จัดเป็นข้อมูลที่มีมาตราวัดแบบอัตราส่วน แต่เนื่องจากรายได้มักเป็นข้อมูลปกปิดไม่นิยมเปิดเผย วิธีการเก็บข้อมูลอาจทำได้อีกวิธีหนึ่ง คือจัดรายได้เป็นช่วงต่างๆ เช่น 6-10, 11-15, 16-20 ... (หน่วยพันบาท) ข้อมูลลักษณะนี้จัดเป็นมาตราวัดแบบนามบัญญัติ การวิเคราะห์ด้วยสถิติที่ไม่ใช้พารามิเตอร์จะทำได้เพียงวิธีเดียว

ขอบเขตการศึกษา

การศึกษาครั้งนี้จะศึกษาเฉพาะการหาค่าสัมประสิทธิ์สหสัมพันธ์ และ การทดสอบสมมติฐานทางสถิติ ที่วัดความสัมพันธ์ของตัวแปรทางแถวนอนและแถวตั้ง ซึ่งอาจจะเป็นกรณีที่กำหนดว่า ตัวแปรตัวแปรหนึ่งเป็นตัวแปรตามไว้ล่วงหน้า หรือบางครั้งไม่จำเป็นต้องกำหนดไว้ล่วงหน้าก็ได้ และจัดลงเป็นตัวแปรทางแถวนอนและแถวตั้งของตารางการณ์จร โดยไม่รวมถึงชุดของตารางการณ์จรแบบต่างๆ เช่น k ชุดของตารางชนิด 2 x 2 หรือ k ชุดของตารางชนิด r x c และการสร้างโมเดลความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ และแม้จะมีวิธีการแบบได้ค่าที่แท้จริง (Exact Method) ในโปรแกรมสำเร็จรูปบางชนิดแล้วก็ตาม แต่ขอบเขตการศึกษาครั้งนี้จะใช้วิธีการประมาณ (Asymptotic assumption)

ตารางชนิด r x r

การวัดความสอดคล้องกัน (Measure of agreement)

การวัดความสอดคล้องกันจะแตกต่างจากการวัดความสัมพันธ์ (Measure of association) โดยจะเป็นการศึกษาว่าผู้ประเมิน 2 คน (หรือ k คน) มีความเห็นในทิศทางเดียวกันหรือไม่ ในการจัดกลุ่มตัวอย่างลงในกลุ่มต่างๆ(หรือลำดับที่ต่างๆ) **ความสอดคล้องกันสูงจะหมายถึงมีความสัมพันธ์กันสูงด้วย แต่ความสัมพันธ์ที่สูงอาจจะไม่มีความสอดคล้องกัน** เช่น ถ้าผู้ประเมิน X ประเมินในหน่วยตัวอย่างในระดับที่สูงกว่าผู้ประเมิน Y ในทุกหน่วยตัวอย่าง นั่นคือ ความสอดคล้องกันต่ำ แต่มีความสัมพันธ์กันสูง จะกล่าวถึงสถิติที่ใช้วัดความสอดคล้องกันที่ได้จากข้อมูลแบบตารางการณ์จร ดังต่อไปนี้

Cohen's Kappa Statistic

ถ้าเป็นสถานการณ์ที่ หน่วยทดลองถูกจัดลงกลุ่มต่างๆ (categories) เช่น ในกลุ่มนักจิตวิทยา k คนจัดผู้ป่วย N คน ในกลุ่มการรักษาต่างๆ (treatment) หรือในกลุ่มของชนิดของโรคจิต โดยนักจิตวิทยาแต่ละคนจะเป็นอิสระกัน และจัดผู้ป่วยแต่ละคนเป็นอิสระต่อกัน ในการจัดกลุ่มการ

รักษา ต้องการวัดว่านักจิตวิทยาทั้ง k คน มีความเห็นสอดคล้องกันหรือไม่ ในการจัดผู้ป่วยทั้ง N คน ในการรักษาแบบต่างๆ (กลุ่มต่างๆ)

จะใช้สถิติที่เรียกว่า “Kappa statistic” หรือ Cohen’s kappa statistic

วิธีการ ถ้ามี N หน่วยตัวอย่าง ซึ่งจะถูกกำหนดโดยผู้ประเมิน k คนให้อยู่ในกลุ่มใดกลุ่มหนึ่ง จาก m กลุ่มโดย m กลุ่มนี้จะมีลักษณะเป็น นามบัญญัติ ข้อมูลที่ได้สามารถจัดลงตารางดังนี้

	กลุ่ม(categories)				
หน่วยตัวอย่าง	1	2	………….j	………………..m	ค่า S_i
1	n_{11}	n_{12}	n_{1j}	n_{1m}	S_1
2	n_{21}				S_2
3	.				
.					
.					
i	n_{i1}		n_{ij}	n_{im}	S_i
.	.		.	.	.
N	n_{N1}		n_{Nj}	n_{Nm}	S_N
ผลรวม	C_1	C_2	C_j	C_m	N_k

เมื่อ n_{ij} = จำนวนผู้ประเมิน ที่กำหนดให้ หน่วยตัวอย่าง ที่ i ลงในกลุ่มที่ j เนื่องจากผู้ประเมินแต่ละคนกำหนดให้แต่ละ หน่วยตัวอย่าง ลงในกลุ่มใดๆ ดังนั้นผลรวมทาง แถวนอน ต้องเท่ากับ k ในทุกแถวนอน ส่วนผลรวมทางคอลัมน์จะเท่ากับ $C_j = \Sigma n_{ij}$

ถ้าผู้ประเมิน มีความเห็นสอดคล้องกันอย่างสมบูรณ์ ค่าความถี่ n_{ij} ในบางเซล ควรมีค่า = k และในเซลอื่นๆ (ในแถวนอนนั้น) จะมีค่าความถี่เท่ากับ 0 แต่ถ้าไม่มีความเห็นสอดคล้องกัน ค่าความถี่ในเซลต่างๆ (ในแถวนอนหนึ่งๆ) ควรมีค่าเป็นอย่างสุ่ม (random)

ค่าสถิติ Kappa จะเป็นค่าอัตราส่วน ของความน่าจะเป็นที่คาดว่าจะเป็น เมื่อ H_0 เป็นจริง (มีความเป็นอิสระกัน) กับ ความน่าจะเป็นที่สูงสุด

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

เมื่อ P(A) = ค่าสัดส่วนที่ k Raters ที่มีความเห็นสอดคล้องกัน

P(E) = ค่าส่วนที่ k Raters ที่มีความเห็นสอดคล้องกันโดยบังเอิญ(by chance)

K คือ ค่าสถิติ Kappa

k คือ จำนวนผู้ประเมิน

K = 1 ถ้ามีความเห็นสอดคล้องอย่างสมบูรณ์

K = 0 ถ้ามีความเห็นไม่สอดคล้องกัน (No agreement among the raters)

ค่า $P(E) = \sum_{j=1}^{m} p_j^2$ เมื่อ $p_j = C_j/Nk$

$$P(A) = \frac{1}{N}\sum_{i=1}^{N} S_i = \left[\frac{1}{Nk(k-1)}\sum_{i=1}^{N}\sum_{j=1}^{m} n_{ij}^2\right] - \frac{1}{k-1}$$

= ค่าเฉลี่ยของ S_i

เมื่อ $$S_i = \frac{\sum_{j=1}^{m}\begin{pmatrix} n_{ij} \\ 2 \end{pmatrix}}{\begin{pmatrix} k \\ 2 \end{pmatrix}} = \frac{1}{k(k-1)}\sum_{j=1}^{m} n_{ij}(n_{ij}-1)$$

ตัวอย่างที่1 การศึกษาเกี่ยวกับปลาตัวผู้ ในระหว่างการวางไข่ จะพบว่าปลาตัวผู้จะเปลี่ยนสีและสร้างอาณาเขต สร้างรัง และมีความก้าวร้าว ให้นักวิจัย 4 คน สังเกตปลา 29 ตัว ว่าเปลี่ยนสีไปในกลุ่มต่างๆ 5 ระดับ น้อยมาก ถึง มากที่สุด (กลุ่ม 1 – 5) ได้ข้อมูลดังนี้

	Coloration Category					
ปลาตัวที่	1	2	3	4	5	S_i
1	-	-	-	-	4	12/12 = 1
2	2	-	2	-	-	4/12 = .333
3	-	-	-	-	4	12/12= 1
4	2	-	2	-	-	4/12 = .333
5	-	-	-	1	3	6/12 = .50
6	1	1	2	-	-	2/12 = .167
7	3	-	1	-	-	6/12 = .50

8	3	-	1	-	-	6/12 = .50
9	-	-	2	2	-	4/12 = .333
10	3	-	1	-	-	6/12= .333
11	-	-	-	-	4	12/12= 1
12	4	-	-	-	-	12/12 = 1
13	4	-	-	-	-	12/12 = 1
14	4	-	-	-	-	12/12 = 1
15	-	-	3	1	-	6/12 = .50
16	1	-	2	1	-	2/12 = .333
17	-	-	-	2	2	4/12 = .333
18	-	-	-	-	4	12/12 = 1
19	-	-	3	-	1	6/12 = .50
20	-	1	3	-	-	6/12 = .50
21	-	-	1	-	3	6/12 = .50
22	-	-	3	1	-	6/12 = .50
23	4	-	-	-	-	12/12 = 1
24	4	-	-	-	-	12/12 = 1
25	2	-	2	-	-	4/12 = .333
26	1	-	3	-	-	6/12 = .50
27	2	-	2	-	-	4/12 = .333
28	2	-	2	-	-	4/12 = .333
29	-	1	2	-	1	2/12 = .167
C_j	42	3	37	8	26	
p_j	.362	.026	.319	.069	.224	

$$= \frac{42}{4 \times 29} \qquad = \frac{3}{4 \times 29}$$

วิธีทำ

จากค่า p_i ที่ได้จาก C_j/Nk ในแต่ละคอลัมภ์

จะได้ $P(E) = \Sigma p_j^2$

$= .362^2 + .026^2 + \ldots + .224^2 = .2884$

P(A) หาจาก S_i ดังนี้

$$S_1 = \frac{1}{k(k-1)} \Sigma\, n_{1j}(n_{1j} - 1)$$

$$= \frac{1}{4(4-1)} [0+0+0+0+4(3)]$$

$= 12/12$

$= 1$

$$S_2 = \frac{1}{4(4-1)} [2(1) + 0 + 2(1) + 0 + 0]$$

$= 4/12$

$= .333$

.

.

.

S_{29}

จากค่าเหล่านั้น $P(A) = \frac{1+.333+\ldots+.169}{29}$

$= .5804$

หรือจะคำนวณ P(A)จากสูตรคือ

$$\left[\frac{1}{Nk(k-1)}\sum_i \sum_j n_{ij}^2\right] - \frac{1}{k-1}$$

$$P(A) = \frac{1}{29(4)(3)}\left[4^2 + 2^2 + 2^2 + \ldots\ldots\ldots\ldots +\right] - \frac{1}{4-1}$$

$$= \frac{318}{348} - \frac{1}{3}$$

$= .5804$

ดังนั้น $$K = \frac{P(A) - P(E)}{1 - P(E)}$$

$$= \frac{.5804 - .288}{1 - .288}$$

$$= .41$$

มีความสอดคล้องกันในระดับปานกลางของนักวิจัย 4 คน

การทดสอบนัยสำคัญของค่า K

การหาการแจกแจงค่า K กรณี N เล็ก จะยุ่งยาก ดังนั้นจะใช้การประมาณด้วยการแจกแจงปกติ เมื่อ N ใหญ่ โดย

$$E(K) = 0$$

$$V(K) \approx \frac{2}{Nk(k-1)} \times \frac{P(E) - (2k-3)[P(E)]^2 + 2(k-2)\sum p_j^3}{[1-P(E)]^2}$$

$$Z = \frac{K}{\sqrt{\text{var}K}} \sim N(0,1)$$

เมื่อ $H_0 : K = 0$ ไม่มีความสอดคล้อง

$H_1 :\ K > 0$ มีความสอดคล้อง

จากตัวอย่างที่ผ่านมา จงทดสอบค่านัยสำคัญของ K ที่ได้ที่ $\alpha = .01$

$H_0 : K = 0$

$H_1 :\ K > 0$

หาเทอม $\sum p_j^3 = .362^3 + .026^3 + .319^3 + .069^3 + .224^3$

$$= .092$$

$$V(K) = \frac{2}{29(4)(3)} \times \frac{.288\ [(2)(4) - 3][.288]^2 + 2(4-2)(.092)}{(1 - .288)^2}$$

$$= \frac{2}{348} \times \frac{.2413}{.5069}$$

$$= .002736$$

ดังนั้น

$$Z = \frac{0.41}{\sqrt{0.002736}} = 7.84$$

Zวิกฤต = 2.32 ตกใน CR. ปฏิเสธ H_0 ยอมรับ H_1 นั่นคือ นักวิจัย 4 คนมีความเห็นสอดคล้องกัน

นอกจากนี้ **ยังอาจวัดค่า Kappa ได้อีก แบบหนึ่ง จาก ผู้ประเมิน2 คน** ดังนี้

ตารางข้อมูลตัวอย่าง จะเป็นดังนี้

		ผู้ประเมินคนที่ 2				
		กลุ่มย่อย 1	2	.	....	r
ผู้ประเมินคนที่ 1	กลุ่มย่อย 1					
	กลุ่มย่อย 2					
	.					
	.					
	.					
	.					
	r					

ความถี่ในเซลต่างๆ คือ จำนวนครั้งที่ ผู้ประเมินคนที่ 1 และ2 มีความเห็นเหมือนกันในการจัดหน่วยตัวอย่างลงในกลุ่มย่อย หนึ่งๆ เช่น ความเห็นของจิตแพทย์2 คน (Psychiatrist) ที่จำนวนคนไข้กลุ่มหนึ่งลงใน 3 กลุ่มของชนิดโรคจิต คือ

Endogenous depression; ED

Reactive depression; RD

Obsessional neurosis; ON

		จิตแพทย์ B			
		ED	RD	ON	t_i = ผลรวม
จิตแพทย์A	ED	10	-	-	10
	RD	-	10	-	10
	ON	-	-	10	10
	U_i=ผลรวม	10	10	10	n=30

หมายความ**ว่าจิตแพทย์ทั้ง 2 มีความเห็นสอดคล้องกันอย่างสมบูรณ์**

การจัดกลุ่มทาง แถวนอน และ แถวตั้ง ควรจัดให้เหมือนกันทั้งทาง แถวนอน และ แถวตั้ง (คือถ้าจัดจาก ED RD และ ON ทาง แถวนอน ก็ควรจะจัดเช่นนี้ทาง แถวตั้ง (และอาจสลับที่กลุ่มย่อยแบบใดๆ ก็ได้ แต่ทาง แถวนอน และ แถวตั้ง ควรจัดเหมือนกัน)

$$\text{สถิติ} \quad K = \frac{nD-\sum t_i U_i}{n^2 - \sum t_i U_i}$$

เมื่อ n = จำนวนความถี่ทั้งหมดในตาราง

D = ผลรวมความถี่ในเส้นทะแยงมุมของตาราง

t_i = ผลรวมความถี่ใน แถวนอนที่ i

U_i = ผลรวมความถี่ใน แถวตั้ง ที่ i

จากตัวอย่างจิตแพทย์

$$K = \frac{30(30) - (10\times10+10\times10+10\times10)}{900 - 300} = 1$$

ซึ่งแสดงถึง ความสอดคล้องกันอย่างสมบูรณ์

หลักการนี้ นอกจากจะใช้วัดความสอดคล้องของ 2 ผู้ประเมิน **แล้วยังอาจใช้วัด Reliability ของสองการทดสอบ** ซึ่งวัดจากหน่วยทดลอง n สิ่ง โดยจัดลงในกลุ่มย่อยต่างๆ

ตัวอย่างที่ 2 Pathologist 2 ท่าน คือ X และ Y ได้จำแนก 118 slides เพื่อแสดงถึงการมี carcinoma of the uterine cervise อย่างเป็นอิสระโดยผลจำแนกเป็น 4 ระดับ คือ

1. negative
2. atypical aquamous hyperplasia
3. carcinoma in situ
4. aquamous or invasive carcinoma ได้ข้อมูลดังตารางต่อไปนี้

Pathologist X	Pathologist Y				รวม
	1	2	3	4	
1	22	2	2	0	26
2	5	7	14	0	26
3	0	2	36	0	38
4	0	1	17	10	28
รวม	27	12	69	10	118

จงวัดความสอดคล้องของแพทย์ทั้ง 2 ท่าน

วิธีทำ

$$\text{จาก}\quad K = \frac{nD - \sum t_i U_i}{n^2 - \sum t_i U_i}$$

$$= \frac{118(22+7+36+10)-(26\times 27+26\times 12+38\times 69+28\times 10)}{118^2-(26\times 27+26\times 12+38\times 69+28\times 10)}$$

$$= .493$$

หมายความว่าแพทย์ทั้ง 2 ท่าน มีความเห็นสอดคล้องกันในระดับปานกลาง

สรุป ค่า Kappa ต้องใช้กับข้อมูลที่เป็น **Nominal** แต่ค่า Kappa ยังมีข้อด้อยคือ จากตารางบางชนิด อาจคำนวณได้ค่า Kappa เป็นลบได้ ในกรณีนี้การหาค่า Kappa เพียงอย่างเดียว อาจจะไม่เพียงพอที่จะอธิบายถึงความสอดคล้อง จำเป็นต้องศึกษาในขั้นสูงต่อไป เช่น การสร้างโมเดลแสดงความสอดคล้องกัน

การทดสอบของโบว์เคอร์ (The Bowker test (1984))

เป็นการทดสอบที่ขยายมาจากการทดสอบของ McNemar ซึ่งใช้กับตารางชนิด 2 x 2 โดยขยายเป็นการทดสอบที่ใช้กับตารางชนิด r x r (r > 2) และใช้ในกรณีที่เปรียบเทียบประชากรกลุ่มเดียวที่ถูกกระทำแบบก่อนและหลังการทดลอง ที่มีคำตอบเป็นกลุ่มย่อยที่มากกว่า 2 โดยจะได้ตารางชนิด 3 x 3 ขึ้นไป (เช่น 4 x 4 , 5 x 5)ดังเช่นตัวอย่างต่อไปนี้

ตัวอย่าง ในการเปรียบเทียบอาการข้างเคียงหลังจากการใช้ยา(side – effect) ชนิดเดิมกับยาชนิดใหม่กับตัวอย่างคนป่วย 1 กลุ่ม โดยคาดหวังว่ายาชนิดใหม่จะช่วยลดผลอาการข้างเคียงหลังการใช้ยาให้น้อยลง (เมื่ออาการข้างเคียงหลังการใช้ยาแยกเป็น 3 กลุ่มย่อยคือ ไม่มี มีเล็กน้อย มีรุนแรง) สมมุติว่าได้ตารางแสดงความถี่ในกลุ่มย่อยต่างๆจากตัวอย่างคนป่วย 158 คน ที่ใช้ยาชนิดเก่า และยาชนิดใหม่ ดังนี้

		อาการข้างเคียงหลังการใช้ยาชนิดใหม่			
		ไม่มี	**มีเล็กน้อย**	**มีรุนแรง**	**รวม**
อาการข้างเคียง หลังการใช้ยาชนิดเดิม	**ไม่มี**	**83**	**4**	**3**	**90**
	มีเล็กน้อย	**17**	**22**	**5**	**44**
	มีรุนแรง	**4**	**9**	**11**	**24**
	รวม	**104**	**35**	**19**	**158**

จากตาราง ถ้าพิจารณาค่าเหนือและใต้เส้นทะแยงมุม แต่ละตำแหน่ง (n_{ij} และ n_{ji} เช่น n_{13} และ n_{31}) จะพบว่าค่าใต้เส้นทะแยงมุมจะมีค่ามากกว่าค่าเหนือเส้นทะแยงมุม คือ $n_{21} = 17 > n_{12} = 4$ หรือ $n_{31} = 4 > n_{13} = 3$ หรือ $n_{32} = 9 > n_{23} = 5$ นั่นแสดงว่ามีแนวโน้มที่จะพบว่ายาชนิดใหม่จะช่วยลดอาการข้างเคียงให้น้อยลงกว่ายาชนิดเดิม แม้ว่าจะพบว่ามีคนป่วยที่ไม่มีอาการข้างเคียงจากยาทั้ง 2 ชนิด ในจำนวนหนึ่งก็ตาม

Bowker ได้เสนอสถิติทดสอบเพื่อทดสอบว่าโอกาสที่จะพบค่าความถี่ในตารางเช่นนี้ มีค่าต่างกันในตำแหน่งเซลเดียวกันที่อยู่เหนือและใต้เส้นทะแยงมุมหรือไม่ ด้วยสูตรดังนี้

$$\chi^2 = \sum \frac{\left(n_{ij} - n_{ji}\right)^2}{n_{ij} + n_{ji}} \qquad \text{เมื่อ } i = 1 \text{ ถึง } r-1\ , j > i$$

โดยสถิติทดสอบนี้จะมีการแจกแจงเมื่อตัวอย่างมีขนาดใหญ่ด้วยสถิติไคสแควร์ที่ d.f. = $\frac{r\,(r-1)}{2}$ เมื่อ H_0 เป็นจริง

จากตัวอย่างข้างต้น จงทดสอบด้วยสถิติ Bowker ว่า ยาชนิดใหม่ทำให้เกิดอาการข้างเคียงแตกต่างจากยาชนิดเดิมหรือไม่ ที่ $\alpha = 0.05$

แทนค่า $\chi^2 = \frac{(4-17)^2}{21} + \frac{(3-4)^2}{7} + \frac{(5-9)^2}{14}$

$= 9.33$

ค่าวิกฤตเมื่อกำหนด $\alpha = 0.05$ ตารางไคสแควร์ที่ d.f. = $\frac{3\,(3-1)}{2}$ = 3 มีค่าเท่ากับ 7.815

ดังนั้น ค่า χ^2 คำนวณมีค่าตกในอาณาเขตวิกฤต จึงปฏิเสธ H_0 หรือการสรุปผลจากค่า p-value จะพบว่า $P(\chi^2_3 > 9.33) \simeq .026$ ซึ่งได้ผลสรุปว่าปฏิเสธ H_0 เช่นกัน (คือ .026 < .05) นั่นคือสามารถสรุปในกลุ่มคนป่วยทั่วๆไป (หรือในกลุ่มประชากร) ได้ว่ายาชนิดใหม่ให้อัตราส่วนของการเกิดอาการข้างเคียงหลังการใช้ยาแตกต่างจากยาชนิดเดิม ซึ่งจากข้อมูลตัวอย่างจะพบว่า มีอัตราส่วนที่ลดลงกว่ายาชนิดเดิมด้วยด้วยระดับนัยสำคัญ $\alpha = 0.026$

ตารางชนิด 2X 2

การวัดความสัมพันธ์ (Measure of association)

จะกล่าวถึงการวัดความสัมพันธ์ของตัวแปรทางแถวนอนและแถวตั้ง โดยใช้ขบวนการทดสอบนัยสำคัญของสถิติทดสอบสำหรับความเป็นอิสระของตัวแปรทางแถวนอนและแถวตั้ง

เมื่อตัวแปรทางแถวนอน(หรือทางแถวตั้ง)คือ**ตัวอย่างสองกลุ่มที่สุ่มจากประชากรสองกลุ่มที่เป็นอิสระกัน**

การประยุกต์ใช้คือ การเปรียบเทียบประชากร 2 กลุ่มนั้น ในสัดส่วนของเหตุการณ์ที่สนใจ

เช่น มีข้อมูลเกี่ยวกับความเชื่อเกี่ยวกับชีวิตหลังการตายของกลุ่มตัวอย่างชายและหญิง

	เชื่อ	**ไม่เชื่อ**
หญิง	435	147
ชาย	375	135

อยากเปรียบเทียบว่า เพศชายและหญิงเชื่อด้วยสัดส่วนเท่ากันหรือไม่ หรือมีความสัมพันธ์ระหว่างเพศกับความเชื่อหรือไม่ (เช่นหญิงเชื่อมากกว่าชาย) หรือความเชื่อเป็นอิสระต่อเพศหรือไม่

สรุป จะเป็นการวิเคราะห์เพื่อหาความสัมพันธ์ (Association) ระหว่าง 2 ตัวแปร ค่าสถิติที่ใช้จะมีดังนี้

1. ผลต่างของค่าสัดส่วน (Difference in proportions)

สัญลักษณ์ ให้อักษรกรีก แทนค่าพารามิเตอร์ เช่น $\prod_{ij}$, $\prod_{1+}$, $\prod_{+1}$ และอักษรโรมันแทนค่าสถิติจากตัวอย่าง เช่น ρ_{ij}

ถ้าในกรณีที่ตัวแปรทาง แถวตั้ง Y เป็นตัวแปรตาม ในขณะที่ตัวแปรทางแถวนอน X เป็นตัวแปรอิสระ จะมีการแจกแจงของค่า Y ในแต่ละระดับของ X จะเรียกว่า conditional probability for Y given the level of X หรือเรียกสั้นๆ ว่า conditional distribution, $\rho_{i/j} = n_{ij}/n_{i+}$

จากตารางข้างต้น แสดงสัญลักษณ์ ได้ดังนี้

ความเชื่อ

		เชื่อ	ไม่เชื่อ	
เพศ	หญิง	$n_{11} = 435$	$n_{12} = 147$	$n_{1+} = 582$
	ชาย	$n_{21} = 375$	$n_{22} = 135$	$n_{2+} = 509$
		$n_{+1} = 810$	$n_{+2} = 281$	$n = 1091$

จะได้ค่าสถิติต่างๆ ดังนี้

$$\rho_{11} = \frac{435}{1091} = 0.399 \quad = \text{joint probability.}$$

และจากข้อมูลดังกล่าวนี้ ตัวแปร x = เพศเป็นตัวแปรอิสระหรืออธิบาย (Explanatory variable) ดังนั้นจึงมีค่า conditional probability (conditional prob.) ดังนี้

$$p_{1/1} = p_{\text{เชื่อ/หญิง}} = \frac{435}{582} = 0.747$$

และ

$$p_{2/1} = p_{\text{ไม่เชื่อ/หญิง}} = \frac{147}{582} = 0.253$$

เพราะฉะนั้นการแจกแจง Conditional prob. ของเพศหญิง = (0.747, 0.253)

ในขณะที่การแจกแจง Conditional prob. ของเพศชาย = (0.737, 0.263)

ความเป็นอิสระ (Independence)

สองตัวแปร จะเรียกว่า เป็นอิสระกันก็ต่อเมื่อ ค่า conditional prob.ของY ทุกระดับของ x มีค่าเท่ากัน

เช่น ตัวอย่างข้างต้น สมมติว่าค่า prob. ของความเชื่อ = 0.74เท่ากันในทั้งเพศชายและหญิง นั่นคือ เพศและความเชื่อเป็นอิสระกัน

$\prod_{ij} = \prod_{i+} \times \prod_{+j}$ for i = 1…..I , j = 1……J เมื่อ $\prod_{ij}$ = ค่าพารามิเตอร์ของ ρ_{ij}

การเปรียบเทียบค่าสัดส่วนในตาราง 2 x 2

เมื่อ แถวนอน = 2 กลุ่มย่อยอิสระกัน และแถวตั้ง = กลุ่มของตัวแปรตาม

ในแถวนอน ที่ 1

ให้ $\prod_1$= prob. (success) (หรือค่า$\prod_{j/1}$ ซึ่งคือ พารามิเตอร์ของ $\rho_{j/i}$)

1- $\prod_1$= prob. (failure)

ดังนั้น ($\prod_1$, 1- $\prod_1$) คือ Conditional prob. Dist^{n} of Y ในแถวนอน 1

และ ($\prod_2$, 1- $\prod_2$) คือ Conditional prob. Dist^{n} of Y ในแถวนอน 2

ผลต่างของค่าสัดส่วนคือ $\prod_1$- $\prod_2$

จะเปรียบเทียบ prob.(success) ใน 2 แถวนอน ซึ่งจะมีค่าในช่วง (-1 และ +1)

และจะมีค่า = 0 เมื่อ $\prod_1 = \prod_2$ นั่นคือ**ตัวแปรตามเป็นอิสระกับตัวแปรอิสระ**

ค่าสถิติจากตัวอย่าง

ถ้าให้ ρ_1 = Conditional prob.ของตัวอย่างชุดที่ 1 หรือคือ $p_{1/1}$ ในตอนต้น และ
ρ_2 = Conditional prob.ของตัวอย่างชุดที่ 2 หรือคือ $p_{2/1}$ ในตอนต้น
ดังนั้นค่า $\rho_1 - \rho_2$ จะเป็นค่าประมาณของ$\prod_1$- $\prod_2$
เช่น ตัวอย่างข้างต้น ρ_1 = 435/582 = 0.747 = ค่าสัดส่วนของผู้หญิงที่เชื่อ
ρ_2 = 375/509 = 0.737 = ค่าสัดส่วนของผู้ชายที่เชื่อ
ดังนั้น ผลต่าง = 0.747 - 0.737 = 0.010

ถ้าให้สัญลักษณ์ n_1 และ n_2 = ขนาดตัวอย่างจาก 2 กลุ่ม (2แถวนอน)
ดังนั้น Standard error ของ $\rho_1 - \rho_2$ คือ

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\rho_1(1-\rho_1)}{n_1}+\frac{\rho_2(1-\rho_2)}{n_2}}$$

ซึ่งจะมีค่าลดลง เมื่อ Sample size เพิ่มขึ้น

สำหรับ ตัวอย่างขนาดใหญ่ $(1 - \infty)$ 100% Confident Interval (CI) for $\prod_1 - \prod_2$

คือ $\rho_1 - \rho_2 \pm z_{\alpha/2}\hat{\sigma}_{\hat{p}_1-\hat{p}_2}$

ตัวอย่างที่3 ข้อมูลจากการศึกษาเรื่องการใช้ Aspirin จะช่วยลดการเป็นโรคหัวใจได้อย่างไร การทดลองในระยะเวลา 5 ปี โดยผู้เข้าทดลองไม่ทราบว่าตนจะได้รับยา Aspirin หรือไม่ ได้ข้อมูลดังนี้

	เป็นโรคหัวใจ		
กลุ่มตัวอย่างที่ได้รับยา	เป็น	ไม่เป็น	รวม
placebo	189	10845	11034
Aspirin	104	10933	11037

จะถือว่า แถวนอน เป็น 2 กลุ่ม ดังนั้นจึง เป็น Independent Binomial Sample จะได้

ρ_1 = P (ป่วยเป็นโรคหัวใจ/มาจากกลุ่มที่ใช้ placebo) = $\frac{189}{11034}$

= 0.0171

ρ_2 = P (ป่วยเป็นโรคหัวใจ/มาจากกลุ่มที่ใช้ Aspirin) = $\frac{104}{11037}$

= 0.0094

ผลต่างของสัดส่วน $\rho_1 - \rho_2$ = 0.0171 – 0.0094

= 0.0077

ด้วยค่า $\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{(0.0171)(0.9829)}{11034} + \frac{(0.0094)(0.0006)}{11037}}$

= 0.0015

ดังนั้นค่า 95% CI. for $\prod_1 - \prod_2$ = 0.0077± 1.96(0.0015)

= 0.008 ± 0.003

= (0.005, 0.11)

ช่วงประมาณนี้ มีแต่ค่าบวก ดังนั้น $\Pi_1 - \Pi_2 > 0$ หรือ $\Pi_1 > \Pi_2$ หรือการใช้ยา Aspirin ช่วยลดสัดส่วน ของการเป็นโรคหัวใจลง

2. Relative risk

การใช้ผลต่างของค่าสัดส่วน **จะให้ความหมายเมื่อค่าสัดส่วนทั้งสองมีค่าใกล้ 0 หรือ 1** มากกว่ามีค่ากลางๆ เช่น ความแตกต่างระหว่างค่าสัดส่วน 2 คู่ ดังนี้

0.010 และ 0.001 ผลต่างคือ 0.009

กับ 0.410 และ 0.401 ผลต่างคือ 0.009

แต่คู่แรกจะให้ความหมายที่ดีกว่าเพราะเป็น 10 เท่าระหว่างกันและคู่ที่สอง = 1 เท่ากัน ดังนั้นค่าอัตราส่วนของสัดส่วนทั้ง 2 จะให้การอธิบายที่ดีขึ้น

ในตาราง 2 x 2 $\text{Relative risk} = \dfrac{\Pi_1}{\Pi_2}$

จากค่าสัดส่วนคู่แรกจะได้ $\text{Relative risk} = \dfrac{0.010}{0.001} = 10.0$

จากค่าสัดส่วนคู่สองจะได้ $\text{Relative risk} = \dfrac{0.410}{0.401} = 1.02$

ถ้า Relative risk = 1 นั่นคือ $\Pi_1 = \Pi_2$ หมายความว่าตัวแปรตามเป็นอิสระจากกลุ่ม (หรือตัวแปรอิสระ)ค่าสถิติจากตัวอย่าง

$$\text{Sample Relative risk} = \frac{\rho_1}{\rho_2}$$

จากตัวอย่างที่ 3 $\text{Sample Relative risk} = \dfrac{0.017}{0.0094} = 1.82$

หมายความว่า กลุ่มที่ใช้ placebo จะมีสัดส่วนการเป็นโรคหัวใจสูงกว่ากลุ่มที่ใช้ Aspirin = 82%

ส่วนการหา (1 - ∞) 100% CI. for relative risk นั้นสูตรค่อนข้างยุ่งยาก เนื่องจากการแจกแจงของ Sample Relative risk จะเบ้แม้ตัวอย่างจะใหญ่

ได้ผลคือ (1.43, 2.30) แปลผลได้ดังนี้ เราจะมั่นใจได้ 95% ว่าหลังจาก 5ปี ค่าสัดส่วนของผู้ป่วยโรคหัวใจจากกลุ่มที่ใช้ยา placebo จะมีค่าระหว่าง 1.43 ถึง 2.30 เท่าของค่าสัดส่วนของผู้ป่วยโรคหัวใจจากกลุ่มที่ใช้ยาAspirin

ถ้าเทียบกับผลค่าผลต่างสัดส่วน ($\prod_1$- $\prod_2$) จากตัวอย่างเดียวกันได้ค่า 95% CI. for $\prod_1$- $\prod_2$= 0.005,0.011 ซึ่งดูเหมือนต่างกันเพียงเล็กน้อย

ซึ่งการใช้ค่าผลต่างสัดส่วนเพียงอย่างเดียวจะทำให้สรุปผลผิดโดยเฉพาะเมื่อค่าสัดส่วนทั้งสอง มีค่าเข้าใกล้ 0

นอกจากนี้ อาจจะหาค่า Relative risk จากเหตุการณ์ Failure คือ

$$\text{Relative risk} = \frac{(1-\prod_1)}{(1-\prod_2)}$$ ซึ่งจะได้ค่าต่างจากเหตุการณ์ Success

โดยทั่วไปมักจะหา Relative riskของเหตุการณ์ที่มีค่า ความน่าจะเป็นต่ำกว่า

3. The odds ratio

เป็นค่าที่ใช้วัดความสัมพันธ์อีกค่าหนึ่ง ของตาราง 2 x 2 เมื่อให้แถวนอนเป็น 2 กลุ่มอิสระ เช่น เพศ หรือ การใช้ยากับไม่ใช้ยา ส่วนแถวตั้งหมายถึงเหตุการณ์ที่สนใจ (Success) และเหตุการณ์ที่ไม่สนใจ (Failure) ดังนี้

	Success	Failure
กลุ่ม 1		
กลุ่ม 2		

ใน แถวนอน ที่ 1 odd ของ success = $\text{odd}_1 = \frac{\prod_1}{1 - \prod_1}$

ใน แถวนอน ที่ 2 odd ของsuccess = $\text{odd}_2 = \frac{\prod_2}{1 - \prod_2}$

ค่า odd จะเป็น nonnegative และ > 1 เมื่อการเกิด success มากกว่า failure

เช่น

$\prod_1 = 0.75$ ดังนั้น ค่า odd = 0.75/0.25 = 3

แปลว่า การเกิด success จะมีโอกาสเกิดขึ้นมากเป็น 3 เท่าของ failure หรือ อาจหาค่า odd ของ failure = 0.25/0.75 = 1/3 แปลว่า failure จะเป็น 1 ใน 3 ครั้งของเกิด success และจากแถวนอนใดๆ จากค่า odd จะหาค่าความน่าจะเป็น ($\prod$) ได้จาก $\prod = \frac{\text{odd}}{\text{odd} + 1}$

ค่า odds ratio , $\theta = \frac{\text{odd}_1}{\text{odd}_2}$

คุณสมบัติของ odds ratio

1. ค่าodds ratio จะเป็น nonnegative คือ $0 < \theta < \infty$
2. ค่า odds ratio = 1 หมายความว่า ตัวแปร x และ y อิสระกัน
3. ถ้าค่า odds ratio > 1 ($1 < \theta < \infty$) เช่น $\theta = 4$ คือ odd ของการเกิดความสำเร็จใน แถวนอนที่1 เป็น 4 เท่าของ แถวนอนที่ 2 (แต่ไม่ใช่ $\prod_1$ เป็น 4 เท่าของ $\prod_2$ นั่นเป็นการแปลผลของ relative risk) นั่นคือ $\prod_1 > \prod_2$
4. ถ้าค่า odds ratio $\theta < 1(0 < \theta < 1)$ คือการเกิด success ใน แถวนอน ที่ 1 มีน้อยกว่าแถวนอนที่ 2 ($\prod_1 < \prod_2$)
5. ค่า θ ที่ห่างจาก 1 ในทิศทางหนึ่ง จะแสดงถึงความสัมพันธ์ที่สูงขึ้น เช่น $\theta = 4$ หมายความว่ามีความสัมพันธ์มากกว่า $\theta = 2$
   $\theta = 0.25$ หมายความว่ามีความสัมพันธ์มากกว่า $\theta = 0.5$
6. ค่า θ สองค่าจะแสดงถึงระดับความสัมพันธ์เดียวกัน แต่เป็นค่าส่วนกลับกัน
   เช่น $\theta = 4 = 1/0.25$ มีความสัมพันธ์เดียวกับ $0.25 = 0.25/1$
   คือ odd ในกลุ่ม 1 เป็น 0.25 เท่า ของ odd ในกลุ่ม 2
7. ถ้าการเรียงลำดับใน แถวนอน หรือแถวตั้ง สลับที่กัน ค่า θ จะเป็นส่วนกลับของค่า θ เดิม เช่นอาจจะได้ $\theta = 4$ หรือ 0.25
8. ถ้ามีการสลับที่จาก แถวนอน ⟶ แถวตั้งหรือจาก แถวตั้ง ⟶ แถวนอน จะไม่เปลี่ยนค่า odds ratio
9. ค่า odds ratio จะไม่เปลี่ยนแปลงเมื่อทั้ง 2 เซลล์ในแถวนอนใดๆ คูณด้วยค่าคงที่ที่ไม่เท่ากับ 0 เช่นเดียวกับกรณีแถวตั้ง

ถ้าตัวแปร x และ y ต่างเป็น ตัวแปรตาม อาจกำหนด ค่า θ ดังนี้

$$\theta = \frac{\prod_{11} / \prod_{12}}{\prod_{21} / \prod_{22}} = \frac{\prod_{11} \prod_{22}}{\prod_{12} \prod_{21}}$$

ค่าสถิติจากตัวอย่าง

$$\hat{\theta} = \frac{\rho_1/(1-\rho_1)}{\rho_2/(1-\rho_2} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

ซึ่งได้จากวิธี Maximum Likelihood สรุปจะคือค่าความถี่ในแนวทะแยงมุมคูณกัน

ตัวอย่าง จากตัวอย่าง 3

ค่า $\hat{\theta} = \frac{189\times10933}{104\times10845} = 1.832$

หมายความว่า มี 83% ของกลุ่มใช้ placebo จะเป็นโรคหัวใจด้วยค่า odd สูงกว่ากลุ่มที่ใช้ Aspirin

ค่า Log odds

เนื่องจากค่า odds ratio ไม่สมมาตร ($0 < \theta < \infty$) เพื่อทำให้สมมาตรจึงเปลี่ยนมาเป็นค่า log

$$\textbf{Log odds} = \textbf{log}\left(\frac{n_{11}n_{22}}{n_{12}n_{21}}\right)$$

$$= \ln n_{11} + \ln n_{22} - \ln n_{12} - \ln n_{21}$$

ซึ่งจะมีค่าในช่วง $-\infty$ ถึง $+\infty$ และ Log odds = 0 หมายถึง ตัวแปรเป็นอิสระกัน(มาจาก log1 = 0) ค่า θ และ log θ จะไม่ถูกกระทบด้วยขนาดตัวอย่าง เช่น ตัวอย่างดังนี้ ถ้าสุ่มตัวอย่างจากประชากรเดียวกัน โดยนักวิจัย 2 คน ได้ขนาดตัวอย่างต่างกันดังนี้

	75	15
	10	100
รวม	85	115

ได้ค่า $\hat{\theta} = 50$

log $\hat{\theta} = 3.91$

	750	15
	100	100
รวม	850	115

ได้ค่า $\hat{\theta} = 50$

log $\hat{\theta} = 3.91$

แต่ค่าสถิติอื่นๆ จะมีค่าต่างกัน เช่น $\chi^2 = 111.65$ และ 348.57 ตามลำดับ

อย่างไรก็ตามยังมีจุดอ่อน ของค่า odds ratio และ log odds คือ

1. การที่ค่า log odds มีค่าในช่วง - ∞ ถึง +∞ ตีความยาก ว่าตัวเลขใดจะแสดงถึงความสัมพันธ์น้อย หรือค่าใดจะแสดงความสัมพันธ์มาก
2. ในกรณีที่ n_{12} หรือ n_{21} มีค่าเป็น 0 ทั้งคู่ ค่า odds ratio และ log odds จะเป็น +∞ แต่ถ้า n_{11} และ n_{22} เป็น 0 ทั้งคู่ ค่าทั้ง 2 จะเป็น 0 ทำให้การแปลผลจะผิดไปจากความจริงเช่นข้อมูล

200	0
0	200

ตารางที่ 1

200	0
200	200

ตารางที่ 2

ถ้าดูตัวเลขในตารางจะเห็นว่าตารางที่ 1 มีความสัมพันธ์อย่างสมบูรณ์มากกว่าตารางที่ 2 แต่ถ้าคำนวณค่า odds ratio และ log odds จะได้เท่ากัน คือ + ∞ ซึ่งทำให้แปลผลผิด และถ้าเปลี่ยนเลข 0 ในตารางที่ 2 เป็นเลข 1 จะได้ว่าค่า odds ratio = 200

สรุป ควรใช้ค่า odds ratio กับตารางที่ไม่เป็น 0 เลย แต่ถ้ามีค่า 0 ให้ปรับสูตร

$$\hat{\theta} = \frac{(n_{11} +0.5)\ (n_{22} +0.5)}{(n_{12} +0.5)\ (n_{21} +0.5)}$$

และค่า log odds จะเปลี่ยนเครื่องหมายเท่านั้น เมื่อสลับแถวนอนหรือแถวตั้ง เช่น

log (4) = 1.39

ส่วน log (0.25) = - 1.39

สรุป ความเป็นอิสระระหว่าง 2 ตัวแปร

1. Relative risk ต้อง = 1
2. odd ใน row ที่ 1 = odd ใน row ที่ 2
3. odds ratio = 1
4. log odds = 0

การทดสอบนัยสำคัญ สำหรับค่า odds ratio และ log odds ratio

ในกรณีตัวอย่างเล็กหรือขนาดกลาง การแจกแจงของค่า odds ratio จะเบ้มาก ดังนั้นการอนุมาน เกี่ยวกับค่า odds ratio จะใช้ค่า log odds ratio แทน

โดยพบว่าถ้าตัวอย่างใหญ่ การแจกแจงจะประมาณด้วยการแจกแจงปกติ โดย

ค่ากลางของ $\log \hat{\theta} = \log \theta$

ค่า $SD.(\log \hat{\theta}) = \sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}$

หรือเรียก Asymptotic Standard error (ASE)

ดังนั้นที่ $(1 - \infty)$ 100% CI. for $\log \theta$ คือ

$$(\log \hat{\theta}) \pm Z_{\alpha/2}\sqrt{\frac{1}{n_{11}}+\frac{1}{n_{12}}+\frac{1}{n_{21}}+\frac{1}{n_{22}}}$$

แล้วใช้ Anti log จะได้ค่าช่วงความเชื่อมั่นของค่า odds ratio

ตัวอย่าง จาก ตัวอย่างที่ 3

จากค่า $\hat{\theta} = 1.832$

$(\log \hat{\theta}) = \log (1.832) = 0.605$

ที่ 95% CI. for $\log \theta$ คือ

$$0.605 \pm 1.96\sqrt{\frac{1}{189}+\frac{1}{10933}+\frac{1}{10845}+\frac{1}{104}}$$

$= 0.605 \pm 1.96(0.123)$

$= (0.365,0.846)$

และ 95% CI. for θ คือ

$(\exp(0.365), \exp(0.846))$

$= (e^{0.365}, e^{0.846})$

$= (1.44, 2.33)$

ช่วงนี้ไม่คลุม 1 แสดงว่ามีความสัมพันธ์กัน

คือกลุ่มที่ใช้ placebo และ ยา Aspirin มีสัดส่วนเป็นโรคหัวใจต่างกัน

หรือช่วงนี้ทำนายได้ว่า odd ของการเป็นโรคหัวใจมีอย่างน้อย 44% ที่สูงกว่าสำหรับกลุ่มที่ใช้ยา placebo เมื่อเทียบกับกลุ่มที่ใช้ aspirin

สำหรับช่วงนี้ จุดปลายมีระยะห่างไม่เท่ากันจากจุด 1.832 เนื่องจากการแจกแจงของ $\hat{\theta}$ จะเบ้ทางขวา

ในกรณีที่มีเซลมีความถี่เป็น 0 ค่า SD (log $\hat{\theta}$) ให้ใช้สูตร

$$\sqrt{\frac{1}{n_{11}+0.5}+\frac{1}{n_{12}+0.5}+\frac{1}{n_{21}+0.5}+\frac{1}{n_{22}+0.5}}$$

เช่น ตัวอย่างที่ 3 (aspirin) ถ้าใช้สูตรปรับ 0.5 จะได้ค่า odds ratio เท่ากับ

$$\frac{(189.5)(10933.5)}{(10845.5)(104.5)}$$

$= 1.828$ ซึ่งใกล้เคียงกับค่าเดิม 1.83

ความสัมพันธ์ระหว่างค่า odds ratio และค่า relative risk

$$\text{ค่า odds ratio} = \text{relative risk} \times \frac{(1-\rho_2)}{(1-\rho_1)}$$

การตีความค่า odds ratio จะไม่เหมือน relative risk

เช่น จาก ตัวอย่าง ที่ผ่านมาได้ค่า $\hat{\theta} = 1.83$ ไม่ได้หมายความว่า ρ_1 มีค่าเป็น 1.83 เท่าของ ρ_2 (ซึ่งนั่นจะแปลผลจาก relative risk =1.83) แต่หมายความว่า ค่า odd ρ_1 /(1 - ρ_1) มีค่าเป็น 1.83 เท่า ของค่า odd ρ_2 /(1 – ρ_2)

จากความสัมพันธ์ข้างต้น จะพบว่าเทอมหลังสุดจะมีค่าใกล้ 1 เมื่อโอกาสที่จะเกิด success ในกลุ่ม 2 มีค่าใกล้ 0 (คือ $\rho_2 \longrightarrow 0$)

ดังนั้น ค่า odds ratio จะเหมือน relative risk

เช่นตัวอย่าง aspirin $\hat{\theta} = 1.83$ ในขณะที่ relative risk=1.82

ในกรณีเช่นนี้ odds ratio = 1.83 จะหมายถึง ρ_1 จะมากเป็น 1.83 เท่าของ ρ_2

ความสัมพันธ์นี้มีประโยชน์ ในกรณีที่ relative risk หาค่าไม่ได้ (เช่นเก็บข้อมูลแบบ cross sectional)

4. การทดสอบของฟิชเชอร์ (The Fisher exact probability test)

การทดสอบนี้เหมาะที่จะใช้วิเคราะห์**ข้อมูลที่มีขนาดเล็กมาก** เพื่อทดสอบว่า**ประชากรทั้ง 2 ที่เป็นอิสระกัน** แตกต่างกันอย่างมีนัยสำคัญหรือไม่ในค่าสัดส่วนของเหตุการณ์ที่สนใจ เมื่อข้อมูลที่นำมาวิเคราะห์เป็นแบบนามบัญญัติหรือแบบเรียงลำดับ**ที่มีเพียงสองคำตอบ**

ข้อมูล จะประกอบด้วยตัวอย่าง 2 ชุดที่เป็นอิสระต่อกัน และตัวอย่างทั้ง 2 นั้นจะประกอบด้วยลักษณะใดลักษณะหนึ่ง ซึ่งแยกจากกันเด็ดขาด สามารถจัดข้อมูลลงในตาราง 2 x 2 ดังนี้

	กลุ่มที่ 1	กลุ่มที่ 2	ผลรวม
ลักษณะ +	A	B	A+B
ลักษณะ -	C	D	C+D
ผลรวม	A+C = n_1	B+D = n_2	N

กลุ่มที่ 1 และ 2 หมายถึงตัวอย่าง 2 กลุ่มที่เป็นอิสระกัน ซึ่งอาจจะเป็นกลุ่มทดลองกับกลุ่มควบคุม, เพศชายกับเพศหญิง, กลุ่มคนที่มีงานทำกับกลุ่มคนที่ว่างงาน หรือกลุ่มคนที่นิยมพรรคการเมือง ก กับกลุ่มที่นิยมพรรคการเมือง ข

ส่วนทางแถวนอนที่มีเครื่องหมาย – และ + นั้นหมายถึง การแบ่งกลุ่มออกเป็น 2 ลักษณะใดๆเช่น การสอบตกและสอบผ่าน, การเห็นด้วยและไม่เห็นด้วย, มีค่าสูงกว่ามัธยฐานและค่าต่ำกว่ามัธยฐาน

วิธีการรวบรวมข้อมูล คือการสำรวจตัวอย่างขนาด n ซึ่งจะเป็นกลุ่มที่ 1 ขนาด n_1 (หรือ = A+C) และกลุ่มที่ 2 ขนาด n_2 (หรือ = B+D) หาจำนวนที่ตกในแต่ละลักษณะในแต่ละกลุ่ม เช่น ในกลุ่มที่ 2 มีจำนวนตัวอย่าง = B ที่มีลักษณะ + เป็นต้น

การทดสอบด้วยวิธีนี้ สามารถใช้ได้ดีกับ**ตัวอย่างขนาดเล็กคือ $n < 20$**

ข้อกำหนดเบื้องต้น

1. ข้อมูลประกอบด้วยตัวอย่างสุ่มขนาด n_1 จากประชากรที่ 1 และตัวอย่างสุ่มขนาด n_2 จากประชากรที่ 2 และให้ $N = n_1 + n_2$
2. ตัวอย่างทั้ง 2 เป็นอิสระต่อกันทั้งภายในกลุ่มและระหว่างกลุ่ม
3. แต่ละตัวอย่างจะถูกจัดให้เป็นลักษณะ + หรือ – ลักษณะใดลักษณะหนึ่งเท่านั้น ดังนั้นมาตราวัดที่จะใช้ได้สำหรับการทดสอบนี้คือ นามบัญญัติหรือเรียงลำดับ

สมมติฐาน อาจจะเป็นการทดสอบสองหางหรือหางเดียวก็ได้ ดังนี้

ถ้าให้ P_1 และ P_2 คือค่าสัดส่วนของลักษณะ +(หรือ–) ของประชากรที่ 1 และ 2 ตามลำดับ

H_0 : ค่าสัดส่วนของลักษณะ +(หรือ–) มีค่าเท่ากันในประชากรทั้ง 2

H_1 : ค่าสัดส่วนของลักษณะ +(หรือ–) มีค่าแตกต่างในประชากรทั้ง 2

หรือ $H_0 : P_1 = P_2$ $H_1 : P_1 \neq P_2$

H_0 : ค่าสัดส่วนของลักษณะ +(หรือ–) ในประชากรที่ 1 $\leq$ ค่าสัดส่วนของลักษณะ +(หรือ–) ในประชากรที่ 2

H_1 : ค่าสัดส่วนของลักษณะ +(หรือ–) ในประชากรที่ 1 > ค่าสัดส่วนของลักษณะ +(หรือ–) ในประชากรที่ 2

หรือ $H_0 : P_1 \leq P_2$ $H_1 : P_1 > P_2$

วิธีการ สามารถหาค่าความน่าจะเป็น (exact probability) ที่จะเกิดความถี่ที่ได้จากการสังเกต (observed frequency) และมีค่าใดๆ ตกอยู่ในตาราง 2 x 2 เมื่อกำหนดให้ผลรวมมาร์จินัลมีค่าคงที่ (fixed marginal totals) ได้ โดยหาจากฟังก์ชันการแจกแจงแบบไฮเปอร์จิออเมตริก (hyper geometric distribution) ดังนี้

$$P = \frac{\begin{pmatrix} A+C \\ A \end{pmatrix}\begin{pmatrix} B+D \\ B \end{pmatrix}}{\begin{pmatrix} N \\ A+B \end{pmatrix}} = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

จะพบว่าสูตรนี้คือ **ค่าอัตราส่วนของผลคูณของแฟคตอเรียลของผลรวมทางแถวนอนและแถวตั้งต่อความถี่ของแต่ละเซลและความถี่ทั้งหมด**

ตัวอย่างเช่น สมมติได้ข้อมูลขนาด N = 19 และจัดลงในตาราง 2 x 2 ได้ดังนี้

ตารางที่ 1

	กลุ่มที่ 1	กลุ่มที่ 2	
+	0	5	5
–	10	4	14
	10	9	19

จะได้ความน่าจะเป็นที่ 19 หน่วยตัวอย่างนี้ ตกอยู่ใน 4 เซลของตาราง 2 x 2 ดังที่เป็นอยู่นี้คือ

$$P = \frac{5!14!10!9!}{19!0!5!10!4!} = 0.0108$$

ถ้าสังเกตจะพบว่าตัวอย่างข้างต้น เป็นกรณีที่มีความถี่ = 0 ในเซลใดเซลหนึ่ง **แต่เมื่อกำหนดให้ผลรวมมาร์จินัลมีค่าคงที่** ดังนั้น จึงสามารถจัดความถี่ลงในตาราง 2 x 2 ได้หลายแบบ เช่นตัวอย่างต่อไปนี้

ตารางที่ 2

	กลุ่มที่ 1	กลุ่มที่ 2	
+	6	1	7
–	1	4	5
	7	5	12

ซึ่งจะได้ค่า

$$P_1 = \frac{7!\ 5!\ 7!\ 5!}{12!\ 6!\ 1!\ 1!\ 4!} = 0.04399$$

หรือสามารถจัดความถี่ลงในตาราง 2 x 2 ได้อีกแบบหนึ่ง **โดยที่ผลรวมมาร์จินัลคงที่**ดังตาราง 2.1

ตาราง 2.1

	กลุ่มที่ 1	กลุ่มที่ 2	
+	7	0	7
–	0	5	5
	7	5	12

$$P_2 = \frac{7!5!7!5!}{12\ !6!0!0!5!} = 0.00126$$

ดังนั้นในตัวอย่างนี้ ความน่าจะเป็นที่จะได้ 12 ตัวอย่าง ตกอยู่ในเซลต่างๆ ของตาราง 2 x 2 เมื่อผลรวมมาร์จินัลคงที่

$$= 0.04399+0.00126 = 0.04529$$

ซึ่งคือค่า P – value ใช้ช่วยตัดสินใจว่าข้อมูลจากตารางที่ 2 จะทำให้สามารถปฏิเสธ H_0 เพื่อยอมรับ H_1 (แบบทางเดียว) ได้หรือไม่

การปรับความถี่ในเซลต่างๆ ทั้ง 4 ในตาราง 2 x 2นั้น มีเกณฑ์อย่างไร ให้ลองพิจารณาตัวอย่างต่อไปนี้ จะเข้าใจง่ายขึ้น

ตัวอย่างที่4 ถ้าสุ่มตัวอย่างขนาด 7 และ 8 จากประชากรที่ 1 และ 2 ตามลำดับ โดยที่ประชากรทั้ง 2 เป็นอิสระต่อกัน เมื่อทำการจำแนกว่าแต่ละตัวอย่างมีคุณลักษณะ + หรือ – อย่างไร ได้ข้อมูลตัวอย่างดังนี้

		ตัวอย่างที่		
		1	2	
คุณลักษณะ	+	5	1	6
	–	2	7	9
		7	8	15

ถ้ามีสมมติฐานการวิจัยว่า สัดส่วนของคุณลักษณะ + (หรือ –) ในประชากรที่ 1 จะสูงกว่าในกลุ่มประชากรที่ 2 หรือไม่

ถ้าให้ P_1^+ และ P_2^+ แทนค่าสัดส่วนของคุณลักษณะ + ในประชากรที่ 1 และ 2 ตามลำดับ ดังนั้นสมมติฐานทางสถิติ คือ

$$H_0 : P_1^+ = P_2^+$$

$$H_1 : P_1^+ > P_2^+$$

ดังนั้นค่า P – value ที่จะหาจากฟังก์ชันการแจกแจงไฮเปอร์จิออเมตริกนั้นต้องหาในกรณีที่เป็นตัวอย่างจริงและกรณีที่มีค่ามากกว่าตัวอย่างจริง (extreme or more extreme) ดังนี้

จากตัวอย่าง

	1	2
+	5	1
–	2	7

ค่า $\hat{P}_1^+ = 5/7$ และค่า $\hat{P}_2^+ = 1/8$

ดังนั้น $\hat{P}_1^+ - \hat{P}_2^+ = (5/7) - (1/8) = 0.589$

ดังนั้น **ต้องหาตาราง2 x 2ที่มีความถี่ในเซล ต่างไปจากตัวอย่างที่ได้จริงนี้และมีค่า** $\hat{P}_1^+ - \hat{P}_2^+ >$ **0.589** ซึ่งจะเรียกว่า more extreme ถ้าจัดความถี่ลงใน 4 เซลของตาราง 2 x 2 ในทุกวิธีแล้วจะได้ดังนี้

ตารางที่ 3 ตัวอย่างการคำนวณหาค่าความน่าจะเป็นในการทดสอบ 2 หาง และหางเดียวของการทดสอบฟิชเชอร์

ตาราง		$\hat{P}_1^+$	$\hat{P}_2^+$	$\hat{P}_1^+ - \hat{P}_2^+$	P(table)	Obs.	Other	Total
I	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 6 & 0 & 6 \\ \hline - & 1 & 8 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.857	0	.857	.0014	.001	.000	.001
II	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 5 & 1 & 6 \\ \hline - & 2 & 7 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.714	.125	.589	.0336	.035	.006	.041
III	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 4 & 2 & 6 \\ \hline - & 3 & 6 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.571	.250	.321	.1958	.231	.084	.315
IV	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 3 & 3 & 6 \\ \hline - & 4 & 5 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.429	.375	.054	.3916	.622	.378	1.000
V	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 2 & 4 & 6 \\ \hline - & 5 & 4 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.286	.500	-.214	.2937	.378	.231	.608
VI	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 1 & 5 & 6 \\ \hline - & 6 & 3 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	.143	.625	-.482	.0783	.084	.035	.119
VII	$\begin{array}{c\|c\|c\|c} & 1 & 2 & \\ \hline + & 0 & 6 & 6 \\ \hline - & 7 & 2 & 9 \\ \hline & 7 & 8 & 15 \end{array}$	0	.750	-.750	.0056	.006	.001	.007

จากตารางที่ 3 ตัวอย่างที่สังเกตได้จริงคือ ข้อมูลในตารางที่ II ซึ่งมีค่า $\hat{P}_1^+ - \hat{P}_2^+ = .589$ **เมื่อ H_1 ระบุการมากกว่าจึงมีเพียงกรณีเดียว คือ ตาราง I ที่ $\hat{P}_1^+ - \hat{P}_2^+$ = .857 ซึ่งมากกว่า .589 เรียกกรณีนี้ว่า**

เป็น more extreme ดังนั้นค่า p – value ของกรณีนี้จึงคือผลบวกของค่าความน่าจะเป็นจากตาราง II และตาราง I ดังนี้

$$\frac{6!\ 9!\ 7!\ 8!}{15!\ 5!\ 1!\ 2!\ 7!} + \frac{6!\ 9!\ 7!\ 8!}{15!\ 6!\ 0!\ 1!\ 8!}$$

= 0.0336 + 0.0014 (ดังแสดงในค่า p(table)ในแถวตั้งที่ 4 ของตารางที่3)

= 0.035

ซึ่งมีค่าไม่มากกว่า $\alpha = 0.05$ ดังนั้นน่าจะเชื่อว่า $H_1 : P_1^+ > P_2^+$ เป็นจริงในตัวอย่างเดียวกันนี้

ถ้าสมมติฐานวิจัยเปลี่ยนไป คือเป็นการทดสอบสองหาง ดังนั้นค่า p – value จะหาจากกรณี more extreme ทั้ง 2 ด้าน พิจารณาตารางที่ 3 พบว่า

ตาราง I เป็น more extreme ทางมากกว่า (เพราะค่า $\hat{P}_1^+ - \hat{P}_2^+ > .589$) และ

ตารางVII เป็น more extreme ทางน้อยกว่า (เพราะค่า $\hat{P}_1^+ - \hat{P}_2^+$ มีค่าน้อยที่สุด)

ดังนั้น ค่า p – value = ค่าความน่าจะเป็นจากตาราง II+ตาราง I + ตารางVII

= 0.0366+0.0014+0.0056 (จากแถวตั้ง P(table))

= 0.041

มีค่าน้อยกว่า $\alpha = 0.05$ จึงน่าปฏิเสธ H_0 นั่นคือ ค่าสัดส่วนต่างกันใน 2ประชากร

ตัวอย่างที่5 เพื่อให้เข้าใจมากขึ้น จะลองเปลี่ยนตัวอย่างที่สังเกตได้จริงว่าคือข้อมูลในตาราง III (จากตารางที่ 3)

ค่า $\hat{P}_1^+ = 4/7 = 0.571$

ค่า $\hat{P}_2^+ = 2/8 = 0.250$

ดังนั้น $\hat{P}_1^+ - \hat{P}_2^+ = 0.571 - 0.250 = 0.321$

ดังนั้น กรณีที่เรียกว่าเป็น more extreme (ในทิศทางเดียวกัน) คือ ข้อมูลในตาราง I และ II ซึ่งมีค่า $\hat{P}_1^+ - \hat{P}_2^+ > 0.321$

ค่า p – value ของการทดสอบหางเดียว $H_1 : P_1^+ > P_2^+$ จึงคือ ความน่าจะเป็นจากตารางที่ III+II+I

= 0.1958+0.0014+0.0336

= 0.231 ฉะนั้น การสรุปผลคือ ยอมรับ H_0 (หรือปฏิเสธ H_1)

และค่า p – value ของการทดสอบสองหาง จึงคือ ความน่าจะเป็นจากตารางที่ III+II+I +VII+VI

= 0.1958+0.0014 +0.0336 +0.0056+0.0783

= 0.315 เช่นเดิมการสรุปผลคือ ปฏิเสธ H_1

หมายเหตุ ในการทดสอบสองหาง จำนวน more extreme ทางมากกว่าและน้อยกว่ามักจะเท่ากัน เพื่อให้เป็นการทดสอบสมมติฐานแบบที่มีอาณาเขตวิกฤตปลายหางซ้ายขวาใกล้เคียงกัน หรือช่วงการยอมรับจะอยู่กลางของโค้งการแจกแจง

ค่าในแถวตั้งที่ 5, 6 และ 7 จะได้กล่าวถึงในการใช้ตารางสำเร็จรูปดังต่อไปนี้ วิธีการดังกล่าวข้างต้นแม้จะมีผลดีในแง่ที่สามารถคำนวณหาค่าความน่าจะเป็นที่แท้จริงได้ถูกต้อง แต่จะพบว่าในทางปฏิบัติจะเสียเวลา และน่าเบื่อมาก จึงได้มีผู้สร้างตารางสำเร็จรูป เมื่อมีข้อจำกัดว่าขนาดตัวอย่าง $n_1+n_2 = N \leq 15$ วิธีการใช้ตารางอธิบายได้ดังนี้

1. ให้พิจารณาค่าผลรวมทางแถวนอนและแถวตั้งของตาราง 2 x 2 จากตัวอย่างที่สังเกตได้จริง ให้

S_1 = ค่าผลรวมมาร์จินัลที่มีค่าน้อยที่สุด (อาจจะเป็นแถวนอนหรือแถวตั้งก็ได้)

S_2 = ค่าผลรวมมาร์จินัลที่มีค่าน้อยอันดับที่ 2 จากแถวที่ตรงกันข้ามกับ S_1เสมอ เช่น

S_1 = ผลรวมทางแนวนอน S_2 จะต้องเป็นผลรวมแนวตั้ง เช่น

X	

S_1 = ผลรวมมาร์จินัลที่น้อยที่สุด

S_2 = ผลรวมมาร์จินัลที่น้อยที่สุดอันดับที่ 2

2. ให้ X = ความถี่ที่สังเกตได้จริงจากตัวอย่างขนาด N ซึ่งอยู่ในเซลที่ S_1และ S_2 ตัดกัน (ดังรูปตารางในข้อ 1)

3. ในแต่ละค่าของ N, S_1, S_2, X ในตารางการแจกแจงของสถิติฟิชเชอร์ จะมีแถวตั้ง 3 แถวใหญ่

แถวที่ 1 obs. คือค่า p ที่ได้จาก H_1 : แบบหางเดียว

แถวที่ 2 other คือค่า p ที่ได้จากการจัดตาราง 2 x 2 ในทิศทางตรงข้ามกับ H_1: แบบหางเดียว

แถวที่ 3 Total คือค่า p ที่ได้จาก H_1 แบบสองหาง

ตัวอย่างที่6 จากอาสาสมัคร 21 คน เลือกมา 11 คนอย่างสุ่มแล้วให้ทานยาซึ่งสามารถดูดซึมแอลกอฮอล์ และ 10 คน ที่เหลือไม่ให้ทานยาใดๆ หลังจากนั้นแต่ละคนจะได้ดื่มเหล้า หนึ่งชั่วโมงหลังจากนั้นได้ทดสอบลมหายใจ (Breathalyser test) ซึ่งจะแสดงว่ามีแอลกอฮอล์ในกระแสเลือดหรือไม่ ได้ผลดังตาราง

	ผล +	ผล –	
ทานยา	2	9	11
ไม่ทานยา	7	3	10
	9	12	

จงทดสอบสมมติฐานว่า สัดส่วนของผู้ที่ได้ผลการทดสอบเป็น + ในกลุ่มที่ทานยาและไม่ทานยาแตกต่างกันหรือไม่

วิธีทำ $H_0 : P_1^+ = P_2^+$ โดย 1 หมายถึงกลุ่มที่ทานยา

$H_1 : P_1^+ \neq P_2^+$ และ 2 หมายถึงกลุ่มที่ไม่ทานยา

เนื่องจาก $n_1+n_2 = 21 = N$ ใหญ่กว่า 15 จึงไม่สามารถใช้ตาราง ได้จึงใช้วิธีหาความน่าจะเป็นจากฟังก์ชันไฮเปอร์จิออเมตริกโดยตรง ในลำดับแรก จัดความถี่ลงในตาราง2 x 2 เมื่อให้ผลรวมมาร์จินัลคงที่ ในทุกวิธีที่เป็นไปได้ก่อน

	+ (1)	-	(2)		3)= ตัวอย่างจริง		(4)		(5)	
1	0	11	1	10	2	9	3	8	4	7
2	9	1	8	2	7	3	6	4	5	5

$P_1 = .000034$ $P_2 = .001684$ $P_3 = .022454$ $P_4 = .117885$ $P_5 = .282924$

$\hat{P}_2^+ - \hat{P}_1^+ = .9$ $\hat{P}_2^+ - \hat{P}_1^+ = .71$ $\hat{P}_2^+ - \hat{P}_1^+ = .52$ $\hat{P}_2^+ - \hat{P}_1^+ = .33$ $\hat{P}_2^+ - \hat{P}_1^+ = .14$

	+ (6)	-	(7)		(8)		(9)		(10)	
1	5	6	6	5	7	4	8	3	9	2
2	4	6	3	7	2	8	1	9	0	10

$P_6 = .330078$ $P_7 = .188616$ $P_8 = .050522$ $P_9 = .005613$ $P_{10} = .000187$

$\hat{P}_2^+ - \hat{P}_1^+ =-.05$ $\hat{P}_2^+ - \hat{P}_1^+ = -.25$ $\hat{P}_2^+ - \hat{P}_1^+ = -.44$ $\hat{P}_2^+ - \hat{P}_1^+ = -.63$ $\hat{P}_2^+ - \hat{P}_1^+ = -.82$

เมื่อค่า $P_1,\ldots, P_{10}$ เป็นค่าตามฟังก์ชันไฮเปอร์จิออเมตริกของตารางที่ 1,…., 10 ตามลำดับ ถ้าพิจารณาค่า $\hat{P}_2^+ - \hat{P}_1^+$ (เพื่อให้เป็นค่าบวก แทนที่จะใช้ค่า $\hat{P}_1^+ - \hat{P}_2^+$) จากทุกตารางทั้ง 10 ตาราง จะพบว่า เมื่อตัวอย่างจริงคือตารางที่ (3) จะมีตารางที่ (1) และ (2) เป็น more extreme ทางมากกว่า และจะมีตารางที่ (9) และ (10) เป็น more extreme ทางน้อยกว่า

ดังนั้นค่า p – value ของการทดสอบสองทางคือค่าความน่าจะเป็นจากตารางที่(3)+(2)+(1)+(9)+(10)

= .022454 + .001684 + .000034 + .005613 + .000187

= .02997

มีค่าน้อยกว่า $\alpha = .05$ ดังนั้น น่าจะปฏิเสธ H_0 เพื่อยอมรับ H_1 นั้นคือ ยาชนิดนี้ทำให้สัดส่วนของผู้ทานยาแล้วได้ผลทดสอบเป็น + ต่างจากกลุ่มที่ไม่ทานยา

หมายเหตุ ถ้าเราพอจะมีหลักฐานทางด้านเคมีหรือชีววิทยาเกี่ยวกับยาชนิดนี้โดยคาดเดาว่า ยาชนิดนี้อาจจะมีผลดี คือช่วยลดสัดส่วนของผู้ที่ได้ผลการทดสอบเป็น + ลงได้ ดังนั้นสมมติฐานทางสถิติคือ

$H_0 : P_1^+ = P_2^+$ หรือ $H_0 : P_2^+ = P_1^+$

$H_1 : P_1^+ < P_2^+$ $\quad$ $H_1 : P_2^+ > P_1^+$

จากตาราง 2 x 2 ทุกวิธีที่เป็นไปได้ 10 ตารางที่ผ่านมานั้นหาค่า $P_2^+ - P_1^+$ ของข้อมูลตัวอย่างจริงและกรณีที่เป็น more extreme ซึ่งคือตาราง (1) และ (2) ดังนั้น

p – value = ค่า P จากตาราง (3) + (2)+ (1)

= .022454 + .001684 + .000034

= .0224172

p – value มีค่าน้อยกว่า $\alpha = .05$ ดังนั้น ปฏิเสธ H_0 เพื่อยอมรับ H_1 ยาชนิดนี้มีผลดีจริงตามที่โฆษณา คือ ลดสัดส่วนของผู้ที่ได้ผลการทดสอบเป็น + ลง ที่ระดับนัยสำคัญ 0.05

5. การทดสอบของแมคนีมาร์ (McNemar test for significance of change)

เป็นการทดสอบเพื่อดูความเปลี่ยนแปลงในรูป “ก่อนและหลัง” (before and after) **โดยใช้กับกลุ่มตัวอย่างชุดเดียว (คือสุ่มตัวอย่างจากประชากรกลุ่มเดียว) แต่ทำการทดลอง 2 ครั้ง ก่อนและหลังจากการใช้ทรีทเมนต์หนึ่ง**เพื่อดูว่ามีการเปลี่ยนแปลงระหว่างก่อนและหลังการทดลองหรือไม่ เช่น

การทดสอบความคิดเห็นของคนกลุ่มหนึ่งก่อนและหลังการฟังอภิปราย

การทดสอบอาการเปลี่ยนแปลงของคนกลุ่มหนึ่งก่อนและหลังการใช้ยาชนิดหนึ่ง

การทดสอบความคิดเห็นของคนกลุ่มหนึ่งก่อนและหลังการอ่านบทความในหนังสือพิมพ์

ดังนั้น สามารถวัดประสิทธิภาพของทรีทเมนต์หนึ่งๆ (การอภิปราย, ยาชนิดหนึ่ง, บทความในหนังสือพิมพ์) ว่ามีผลทำให้เกิดการเปลี่ยนแปลงหรือไม่

ข้อกำหนดเบื้องต้น

1. ข้อมูลประกอบด้วยข้อมูลคู่ n คู่ และสามารถจัดลงตาราง 2 x 2 ได้
2. มาตรวัดของข้อมูลเป็นแบบนามบัญญัติหรือแบบเรียงลำดับที่มีเพียง 2 กลุ่มย่อยเท่านั้น เช่น ใช่ กับไม่ใช่ หรือ + กับ –
3. ข้อมูลคู่แต่ละคู่เป็นอิสระกัน แต่ภายในคู่สัมพันธ์กัน

ข้อมูล

เช่น ต้องการทดสอบทรีทเมนต์ คือ การเต้นรำประกอบเพลงว่าจะมีผลต่อการลดน้ำหนักหรือไม่ สุ่มตัวอย่างคนที่ต้องการลดน้ำหนัก ด้วยวิธีนี้มา n คน ในขั้นตอนแรกถามความคิดเห็นเกี่ยวกับผลที่คาดว่าจะได้รับก่อนการเต้นรำ ว่าจะได้ผลหรือไม่ได้ผล(+,-) และหลังจากนั้นทดลองให้เต้นรำประกอบเพลงไประยะเวลาหนึ่ง แล้วถามความคิดเห็นเช่นเดิมอีกครั้ง ดังนี้

คนทดลองที่	ความเห็นก่อนการเต้นรำ	ความเห็นหลังการเต้นรำ
1	+	-
2	+	+
3	+	+
.	.	.
.	.	.
.	.	.
n	-	-

ซึ่งสามารถนำมาจัดลงในตารางแจกแจง 2 x 2 ได้ดังนี้

		หลังให้ทรีทเมนต์	
		-	+
	+	A	B
ก่อนให้ทรีทเมนต์	-	C	D

รวม = n

เมื่อ A, B, C, D = จำนวนความถี่ที่มีลักษณะหนึ่งๆ จากจำนวนตัวอย่างขนาด n

คือ A = จำนวนคนที่มีความคิดเห็นก่อนการให้ทรีทเมนต์ว่าจะได้ผลดี(+) แต่เมื่อได้รับทรีทเมนต์แล้วมีความเห็นว่าไม่ได้ผล (-) หรือ + เปลี่ยนเป็น –

D = จำนวนคนที่มีความคิดเห็นก่อนการให้ทรีทเมนต์ว่าไม่ได้ผลดี (-) แต่หลังจาก

ส่วน B และ C = จำนวนคนที่มีความเห็นเหมือนเดิมหลังจากให้ทรีทเมนต์แล้ว คือ จาก + เปลี่ยน + หรือ จาก – เปลี่ยนเป็น –

ดังนั้น จะเห็นว่า**จำนวนความถี่ A และ D จะแสดงถึงการเปลี่ยนแปลง**

สมมติฐาน สามารถทดสอบได้ทั้งหางเดียวและสองหาง ดังนี้

H_0 : P(A) = P(D) = 0.5 หรือมีการเปลี่ยนแปลงก่อนและหลังการให้ทรีทเมนต์ในจำนวนเท่าๆ กัน หรือ $P(+\longrightarrow -) = P(- \longrightarrow +)$

H_1 : $P(A) \neq P(D)$ หรือมีการเปลี่ยนแปลงก่อนและหลังการให้ทรีทเมนต์ในจำนวนที่แตกต่างกัน หรือ $P(+\longrightarrow -) \neq P(- \longrightarrow +)$

หรือ $H_0 : P(A) \leq P(D)$ $H_1 : P(A) > P(D)$

นั่นคือ สงสัยว่ามีการเปลี่ยนแปลงจาก + เป็น – มีมากกว่าการเปลี่ยนแปลงจาก – เป็น + หรือไม่

$H_0 : P(A) \geq P(D)$ $H_1 : P(A) < P(D)$

นั่นคือ สงสัยว่ามีการเปลี่ยนแปลงจาก + เป็น – มีน้อยกว่าการเปลี่ยนแปลงจาก – เป็น + หรือไม่

วิธีการ

การหาผลสรุปของสมมติข้างต้นนี้ อาจดูได้ในลักษณะเปรียบเทียบความถี่ที่ได้จากการทดลองกับความถี่คาดหวัง (ภายใต้ H_0) ว่าแตกต่างกันหรือไม่

โดยความถี่ คาดหวังตาม H_0 จะเป็นดังนี้ จำนวนความถี่ที่คาดหวังในช่อง A และ D ควรมีค่าเท่ากัน = $\frac{A+D}{2}$ และใช้การทดสอบของไคสแควร์ ได้ดังสูตร

$$\chi^2 = \sum \frac{(O_i - E_j)^2}{E_i}$$ และในกรณีนี้สนใจเฉพาะ A และ D

ฉะนั้น

$$\chi^2 = \frac{\left(A - \frac{(A+D)}{2}\right)^2}{\frac{(A+D)}{2}} + \frac{\left(D - \frac{(A+D)}{2}\right)^2}{\frac{(A+D)}{2}}$$

จะได้

$$\chi^2 = \frac{(A-D)^2}{(A+D)}$$

แต่เนื่องจาก χ^2 เป็นการแจกแจงแบบต่อเนื่อง จึงจำเป็นต้องปรับค่าต่อเนื่อง จะได้

$$\chi^2 = \frac{(|A-D|-1)^2}{(A+D)} \qquad \text{ด้วย d.f.} = 1$$

การหาอาณาเขตวิกฤต หาได้จากตาราง χ^2 ที่ d.f. = 1 คือ จะปฏิเสธสมมติฐานเบื้องต้น ถ้า ค่า χ^2 จากตัวอย่างมีค่ามากกว่าค่าวิกฤติ

หมายเหตุ ถ้าใช้ตัวสถิติทดสอบในเทอมของ χ^2 จะใช้ได้เฉพาะการทดสอบ 2 หางเท่านั้น

ตัวอย่างที่ 7 ก่อนที่จะมีการถ่ายทอดทางโทรทัศน์ถึงการอภิปรายโต้ตอบกันของผู้สมัครชิงตำแหน่งประธานาธิบดี 2 คน คือ Regan และ Carter จาก 2 พรรคคือ รีพับรีกัน และ เดโมแครท ตามลำดับได้สุ่มตัวอย่างผู้มีมีสิทธิ์ออกเสียงมา 75 คนเพื่อสอบถามว่าจะเลือกผู้สมัครจากพรรคใด พบว่ามี 41 คน จะลงเสียงให้ Carter และอีก 34 คนจะลงให้ Regan แต่หลังจากได้ชมการอภิปรายของผู้สมัครทั้ง 2 ทางโทรทัศน์แล้ว ปรากฏว่ากลุ่มตัวอย่าง 75 คนชุดเดิมนี้ เปลี่ยนความตั้งใจที่จะเลือกดังนี้ มี 13คน จากผู้ที่เดิมจะลงเสียงให้ Carter เปลี่ยนใจจะลงเสียงให้ Regan และมี 7 คนของ ผู้ที่เดิมจะลงเสียงให้ Regan เปลี่ยนใจมาลงเสียงให้ Carter สรุปได้ดังตารางนี้

		หลังจากอภิปราย		
		Regan	Carter	
ก่อนการอภิปราย	Carter	13	28	41
	Regan	27	7	34
				รวม = 75

ผู้วิจัยต้องการทราบว่าการอภิปรายทางโทรทัศน์มีผลทำให้มีการเปลี่ยนความชอบ ของผู้มีสิทธิ์ ลงคะแนนเลือกประธานาธิบดีในสัดส่วนเท่ากันหรือไม่

วิธีทำ

1. $H_0 : P(\text{Regan} \longrightarrow \text{Carter}) = P(\text{Carter} \longrightarrow \text{Regan})$

   $H_1 : P(\text{Regan} \longrightarrow \text{Carter}) \neq P(\text{Carter} \longrightarrow \text{Regan})$

2.ใช้การทดสอบของ McNemar เพราะเป็นการเปลี่ยนแปลง“ก่อนและหลัง”

$$\chi^2 = \frac{(|A-D|-1)^2}{(A+D)} = \frac{(|13-7|-1)^2}{20} = 1.25$$

3.อาณาเขตวิกฤตหาจากตาราง χ^2 คือ $\chi^2 > \chi^2$ ตาราง

กำหนด $\alpha = 0.05, \chi^2_{.05,1} = 3.84$

และ 1.25 <3.84 ไม่ตกในอาณาเขตวิกฤตยอมรับ H_0

สรุปได้ว่า มีการเปลี่ยนแปลงในการลงเสียงให้พรรคใดหลังจากได้ฟังการอภิปรายทางโทรทัศน์ในสัดส่วนเท่าๆ กัน

ตัวอย่างขนาดเล็ก ในกรณีที่ค่า (A+D)/2 มีค่าน้อยกว่า 5 การประมาณตัวสถิติทดสอบ χ^2 ด้วยการแจกแจงแบบไคสแควร์ที่ d.f. = 1 ดังกล่าวข้างต้นแล้วนี้ไม่เหมาะสม

แนะนำให้ใช้การทดสอบแบบทวินามโดย n = A+D และค่า x = ค่าที่น้อยที่สุดจากค่า (AหรือD) และ p = 0.5 **การทดสอบทวินามนี้สามารถใช้ทดสอบทางเดียวที่บ่งบอกทิศทางได้ด้วย**

จะแสดงวิธีการทดสอบแบบทวินาม ในตัวอย่างที่ผ่านมา ซึ่ง A+D = 20 ดังนั้น n = 20 ค่าที่เล็กที่สุดระหว่าง (AหรือD) =7=x จากตารางทวินามที่ n =20, p=0.5 จะพบว่า P (x $\leq$7)=0.132 ดังนั้นค่า P – value = 2(0.132)=0.264 ซึ่งมากกว่า 0.05 ยอมรับH_0 ได้ผลเช่นเดียวกับการใช้ χ^2

สรุป ในหัวข้อที่ผ่านมาเป็นการกล่าวถึงการวัดความสัมพันธ์ของตัวแปรทางแถวนอนและแถวตั้ง ของ**ตารางชนิด 2 x 2** ด้วยสถิติแบบต่างๆ คือ

1. ค่าผลต่างของค่าสัดส่วน
2. ค่า Relative risk
3. ค่า Odds ratio
4. The Fisher Exact Probability Test
5. The McNemar Test

ซึ่งการแปลผลว่าตัวแปรทั้งสองสัมพันธ์กันหรือไม่ โดยพิจารณาจาก ค่าสถิติเหล่านั้น และทำการทดสอบนัยสำคัญที่ระดับนัยสำคัญหนึ่งๆ (หรือ หาช่วงเชื่อมั่นของค่าพารามิเตอร์ ที่ระดับความเชื่อมั่นหนึ่ง)

ตารางชนิด r x c

กรณีแรกจะกล่าวถึงตารางแบบที่ข้อมูลมีมาตราวัดแบบนามบัญญัติ(Nominal Attribute Categories)

จะกล่าวถึงค่าสถิติที่ใช้วัดความสัมพันธ์ หรือเรียกว่า สัมประสิทธ์สหสัมพันธ์ (Correlation coefficient) แบบต่างๆ ก่อน และตอนท้ายจะกล่าวถึงสถิติทดสอบความเป็นอิสระดังนี้

1. ค่าสัมประสิทธิ์คาร์เมอร์ ; C^2 (The Cramer coefficient)

ค่าสถิติคาร์เมอร์**จะใช้วัดความสัมพันธ์ของคุณลักษณะหรือตัวแปรที่มีลักษณะเป็นกลุ่ม (Nominal scale) สองตัวแปรที่ถูกสุ่มมาจากประชากรหนึ่ง** โดยจะมีค่าคงเดิมไม่ว่าการจัดตารางจะใช้แถวนอนและแถวตั้งเป็นคุณลักษณะใด และอาศัยการวัดจากพื้นฐานของสถิติทดสอบความเป็นอิสระของ χ^2 ดังนี้

$$\text{ค่าสัมประสิทธ์ของคาร์เมอร์ ; } C^2 = \sqrt{\frac{\chi^2}{N(t-1)}}$$

เมื่อ $t = \min(r, k)$ และ $\chi^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$

ดังนั้นค่า C^2 จะมีค่า $0 \le C^2 \le 1$ โดยไม่มีค่าเป็นลบ

การแปลความหมายทำได้ดังนี้

ค่า C^2	การแปลผล
0 – 0.25	สัมพันธ์น้อย
0.26 – 0.50	สัมพันธ์ปานกลาง
0.51 – 0.75	สัมพันธ์ค่อนข้างมาก
0.76 – 1.0	สัมพันธ์มาก

เนื่องจากค่า $(C^2)^2$ เป็นลิเนียร์ฟังก์ชัน (linear function) ของค่า χ^2 ดังนั้นค่า P – value ของค่าχ^2 จะมีค่าเท่ากับP – value ของค่า$(C^2)^2$ จากข้อมูลเดียวกันและสามารถสรุปถึงค่า C^2 ได้เช่นกัน

และ**ถ้าตารางการจรณ์เป็นชนิด 2 x 2** ค่าสัมประสิทธิ์นี้จะเรียกชื่อว่า**ค่าสัมประสิทธิ์ฟาย**

$$\boldsymbol{\emptyset^2} = \sqrt{\frac{\chi^2}{N}}$$

ตัวอย่างที่8 สุ่มตัวอย่างจากรายงานทางเศรษฐกิจของรัฐเซาท์คาโรไรนา ปี 1972 ซึ่งแสดงข้อมูลจากการสำมะโม(census) ในปี 1970 เกี่ยวกับผู้อาศัยอยู่ในรัฐซึ่งมีอายุอย่างน้อย 25 ปีขึ้นไปในขณะนั้นโดยสนใจตัวแปรเพศ และจำนวนปีที่เข้ารับการศึกษา ได้ข้อมูลตัวอย่างดังตารางต่อไปนี้ จงวัดความสัมพันธ์ พร้อมทั้งทดสอบนัยสำคัญ ที่ระดับนัยสำคัญ .05

จำนวนผู้ที่อาศัยอยู่ในรัฐ (พันคน)

จำนวนปีที่เข้ารับการศึกษา	เพศ	
	ชาย	หญิง
0	18	16
ประถม 1 — 4	67	55
5 และ 6	61	67
7	44	49
8	52	59
มัธยม 1 — 3	130	179
4	120	144
วิทยาลัย 1 — 3	47	59
4	38	42
ปริญญา	23	12
รวม	600	682

วิธีทำ จากตารางการจรณ์ ชนิด10 x 2 ดังนั้น t = min (10,2) = 2

$$\text{ค่าสัมประสิทธิ์ของคาร์เมอร์} = \sqrt{\frac{\chi^2}{\mathrm{N}\,(2-1)}}$$

$$= \sqrt{\frac{12.06}{1282(1)}}$$

$$= 0.097$$

เมื่อต้องการวัดนัยสำคัญ หาค่า P – value จากค่า $\chi^2 = 12.06$ ที่ d.f.= 9 พบว่า P – value มากกว่า 0.05 ดังนั้นตัวแปรทั้งสองเป็นอิสระกัน

ข้อจำกัดของค่าสัมประสิทธิ์คาร์เมอร์

1. ค่าสัมประสิทธิ์คาร์เมอร์ จะมีค่าเป็น 0 เมื่อตัวแปรทั้ง 2 ไม่มีความสัมพันธ์กันซึ่งคุณสมบัติข้อนี้ คล้ายสัมประสิทธิ์สหสัมพันธ์อื่นๆ เมื่อไม่มีความสัมพันธ์กันอย่างสมบูรณ์ ค่าสัมประสิทธิ์ควรจะมีค่าเป็น 0
2. ค่าสัมประสิทธิ์คาร์เมอร์ ที่มีค่าเท่ากับ 1 ไม่ได้หมายความว่าตัวแปรทั้ง 2มีความสัมพันธ์กันอย่างสมบูรณ์ ซึ่งต่างจากค่าสัมประสิทธิ์สหสัมพันธ์ทั่วๆ ไป ที่มีความหมายว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันอย่างสมบูรณ์ **ค่าสัมประสิทธิ์คาร์เมอร์เท่ากับ 1 จะหมายถึงตัวแปรทั้ง 2 มีความสัมพันธ์กันอย่างสมบูรณ์ เฉพาะกรณีที่ตารางการจรณ์เป็นชนิด r = k และตาราง การณ์จรนี้ จะต้องมีแต่ละแถวนอน และแต่ละแถวตั้ง มีเพียงเซลเดียว (single cell) ที่มีค่าความถี่ไม่เป็น 0** ดังนั้นจะแสดงตัวอย่างเมื่อตารางการณ์จรเป็นชนิด r = k และ $c^2 = 1$ ดังนี้

จากการสุ่มตัวอย่างสุนัขมา 50 ตัวจาก 3 พันธุ์ และแยกตามขนาดจะได้ข้อมูลดังตารางต่อไปนี้

พันธุ์	ขนาดสุนัข			รวม
	ใหญ่	กลาง	เล็ก	
อัลเซเชียน	20	0	0	20
บลูด๊อก	0	10	0	10
เทอร์เรีย	0	0	20	20
รวม	20	10	20	50

คำนวณได้ค่า $\chi^2 = 100$ และ $C^2 = 1$ หมายความว่าตัวแปรพันธุ์ และขนาดสุนัขมีความสัมพันธ์กันอย่างสมบูรณ์ **ในสองทิศทาง คือถ้าทราบตัวแปรหนึ่งจะทำนายหรือคาดคะเนค่าตัวแปรที่เหลือได้** เช่น ทราบว่าเป็นสุนัขพันธุ์เทอร์เรีย ก็บอกได้ว่ามีขนาดเล็ก หรือแปลความกลับกัน

ในกรณีตารางการจรณ์ชนิดใดๆ ที่ $r \neq k$ ค่าสัมประสิทธิ์คาร์เมอร์อาจจะมีค่า= 1โดยมีความหมายว่ามีความสัมพันธ์ระหว่าง 2 ตัวแปรอย่างสมบูรณ์ **แต่ในทิศทางเดียวกันเท่านั้น (only one direction)** เช่น ในกรณีที่ r < k และสัมประสิทธ์มีค่า = 1 จะมีเพียงหนึ่งเซลทีมีควาถี่ไม่เป็น 0 ในแต่ละแถวตั้ง แต่จะต้องมีบางแถวนอน (some rows) ที่มีมากกว่าหนึ่งเซลที่มีความถี่ไม่เป็นศูนย์ (จะมี k – r เซลที่มีความถี่ไม่เป็นศูนย์) ดังนี้ **ในกรณีนี้จึงมีความสัมพันธ์กันอย่างสมบูรณ์ จากตัว**

แปรทางแถวตั้ง ไปสู่ตัวแปรทางแถวนอนแต่จะไม่มีความสัมพันธ์อย่างสมบูรณ์จากตัวแปรทางแถวนอนไปสู่ตัวแปรทางแถวตั้ง และความสัมพันธ์จะมีทิศทางตรงข้ามกับที่กล่าวข้างต้น เมื่อค่าสัมประสิทธิ์ = 1 และ $r > k$ ดังจะแสดงตารางชนิด $r \times k$ ที่คำนวณค่า $C^2 = 1$ ดังนี้

อาชีพ

เพศ	พยาบาล	นักบิน	ประชาสัมพันธ์	รวม
หญิง	50	0	30	80
ชาย	0	20	0	20
รวม	50	20	30	100

แปลผลได้ว่า เมื่อทราบตัวแปรทางแถวตั้ง คือ อาชีพ เช่น นักบิน ก็พอจะทำนายได้ว่า ต้องเป็นเพศชาย (แถวนอน)

หรือ

อาชีพ	เพศ		รวม
	หญิง	ชาย	
พยาบาล	50	0	50
นักบิน	0	20	20
ประชาสัมพันธ์	30	0	30
รวม	80	20	100

สรุปได้ว่าในกรณีที่ค่าสัมประสิทธิ์ = 1 เมื่อ $r \neq k$ ความสัมพันธ์จะมีลักษณะเป็นแบบสมบูรณ์ชนิดไม่สมมาตร (asymmetrical perfect relation) กล่าวคือ มีความสัมพันธ์อย่างสมบูรณ์เพียงทิศทางเดียวแต่ไม่เป็นทั้งสองทิศทาง

3. การหาค่าสัมประสิทธิ์คาร์เมอร์ได้ ต้องมาจากการใช้การทดสอบของ χ^2 ดังนั้นข้อจำกัดของ χ^2 ควรจะเป็นจริง คือ ขนาดตัวอย่างต้องมีขนาดใหญ่ (เพื่อไม่ให้เกิดกรณีที่จะมีมากกว่า 20 เปอร์เซนต์ของจำนวนเซลทั้งหมด มีความถี่คาดหวังน้อยกว่า 5)
4. **ค่าสัมประสิทธิ์คาร์เมอร์ ไม่สามารถนำไปเปรียบเทียบกับค่าสัมประสิทธิ์อื่นๆ ได้** เช่น สัมประสิทธิ์ของเพียร์สัน (ยกเว้นกรณีเป็นตารางชนิด 2 x 2) หรือสัมประสิทธิ์ของสเปียร์แมนและเคนดอลล์

แม้จะมีข้อจำกัดดังกล่าวของค่าสัมประสิทธิ์คาร์เมอร์ แต่จะพบว่าค่าสัมประสิทธิ์นี้มีข้อดีหลายประการคือ

1. มีค่าสูงสุด = 1 ในขณะที่ค่า C มีค่าสูงสุดไม่เท่ากับ 1 แม้ว่าความหมายของการที่ค่าสัมประสิทธิ์ = 1 จะไม่มีความหมายถึงความสัมพันธ์อย่างสมบูรณ์เหมือนกันกับค่าสัมประสิทธิ์อื่นๆ ดังกล่าวในข้อ 2 ของข้อจำกัดแล้ว
2. สามารถเปรียบเทียบค่าสัมประสิทธิ์คาร์เมอร์ ในกรณีตารางแจกแจงสองทาง r x k ใดๆ ได้

ความคิดพื้นฐานเกี่ยวกับสถิติประเภท PRE

โดยทั่วๆ ไปข้อมูลในตาราง ซึ่งจะก่อให้เกิดความสัมพันธ์ (มีความถี่เฉพาะช่องที่อยู่บนเส้นทแยง นอกเส้นทแยงเป็น 0 หมด) ไม่ค่อยปรากฏบ่อย ทั้งนี้เพราะมีปัจจัยอื่นๆนอกเหนือไปจากตัวแปรอิสระ ซึ่งไม่ได้บรรจุเอาไว้ในตารางมากระทบต่อตัวแปรตาม จึงมักพบแต่เพียงตารางข้อมูลที่มีความสัมพันธ์ระหว่างตัวแปรเป็นขนาดกลางๆ หรือขนาดมาก แต่ไม่ถึงขั้นสมบูรณ์ ด้วยเหตุนี้การตีความจึงมีความสำคัญ

สมมติว่ามีตารางข้อมูล ซึ่งถูกดัดแปลงให้เป็นค่า Probability เราเรียก joint probability distribution (A_j, B_k) ดังนี้

ข้อมูลที่ถูกดัดแปลงเป็นค่า probability แล้ว

	A1	A2	รวม
B1	.20	.15	.35
B2	.10	.30	.40
B3	.10	.15	.25
	.40	.60	1.0

นั่นคือถ้า P (A_j, B_k) คือ P(A_1, B_1) ก็จะมีค่า probability = 0.20 หรือ P(A_2, B_1) = 0.15 เป็นต้น

สมมติว่ามีการสุ่มเอาหน่วยใดหน่วยหนึ่งออกมาจากตารางนี้ โดยที่เราไม่ทราบเลยว่ามาจากช่วงใดของ A แล้วถามว่า หน่วยนี้มาจากช่วงใดของ B คำตอบก็ควรจะเป็น B_2 ทั้งนี้เพราะ B_2 มีผลรวมที่ใหญ่ที่สุด คือ 0.40 โอกาสที่จะเกิดช่วง B_2 จึงมากที่สุด แต่ก็เป็นไปได้ที่หน่วยที่สุ่มออกมานั้น ตามความเป็นจริงแล้วอาจมาจากช่วง B_1 หรือB_3 ก็ได้ ดังนั้นจึงเกิดความผิดพลาดในการทำนาย = 0.35+0.25 = 0.60

การทำนายเช่นนี้และวิธีการหาค่าความผิดพลาดดังข้างต้น เป็นการใช้ประโยชน์จากค่า probability ที่ผลรวมของตัวแปรตามแต่เพียงอย่างเดียว

คราวนี้สมมติใหม่ว่าหน่วยที่ถูกสุ่มมานั้นทราบว่ามาจากช่วง A_1 ดังนั้นจึงควรทำนายว่าหน่วยนี้ควรมาจาก B_1 เพราะช่อง A_1B_1 มีค่า probability ที่ใหญ่ที่สุดคือ 0.20

นั่นคือ probability ในการเดาถูกคิดจากเฉพาะช่วง A_1 เท่านั้นจะเท่ากับ 0.20/0.40 = 50% หรือคิดได้อีกวิธีหนึ่งว่าถ้า probability การเดาถูก = 0.20 probability ของการเดาผิดจะเท่ากับ 0.10+0.10 =0.20 เช่นกัน นั่นคือต่างก็เท่ากับ 50%

ดังนั้นถ้าเทียบกับการดู probability เฉพาะแต่ที่ผลรวมของ B อย่างเดียวแล้ว การทราบข่าวสารเพิ่มขึ้น (เช่นมาจากช่วง A_1 เป็นต้น) ย่อมทำให้การทำนายผิดพลาดลงได้ นั่นคือลดลงไปเท่ากับ 60 – 50 =10 %

ความผิดพลาดที่ลดลงไปอันเนื่องมาจากการทราบข่าวสารเพิ่มขึ้น (ข่าวสารซึ่งมาจากตัวแปรอิสระ) เปรียบเทียบกับความผิดพลาดที่เกิดจากการทำนายจากผลรวมอย่างเดียวโดยไม่ทราบข่าวสารอะไรเลย เราเรียกอัตราส่วนนี้ว่า Proportional Reduction in Error (PRE) ซึ่งเป็นรูปหนึ่งของการวัดความสัมพันธ์ระหว่างตัวแปร การที่ความผิดพลาดในการทำนายลดลง แสดงว่าตัวแปรตัวหนึ่งอธิบายตัวแปรอีกตัวแปรอีกตัวหนึ่งได้ ตัวแปรทั้งสองจึงมีความสัมพันธ์กัน

แบบฟอร์ม (สูตร) ในการหาค่า PRE โดยทั่วๆ ไปคือ $(E_1 – E_2)/ E_1$

E_1 คือ ความผิดพลาดเมื่อไม่ทราบข่าวสารจากตัวแปรอิสระ คือ ดูจากผลรวมตัวแปรตาม แต่เพียงอย่างเดียว

E_2 คือ ความผิดพลาดที่ยังไม่เกิดขึ้น แม้มีข่าวสารเพิ่มขึ้นมาจากตัวแปรอิสระ

โดยทั่วๆ ไปสถิติแบบ PRE มักจะหาค่าโดยอาศัยแบบฟอร์มของรูปสูตรดังข้างบน เพียงแต่ว่าการหาค่าความผิดพลาดทั้ง E_1 และ E_2 แตกต่างกันไปตามแต่นักสถิติผู้คิดค้นนั้นๆ จะคิดออกมาได้

2.สถิติ PRE ที่ใช้กับตารางซึ่งตัวแปรเป็นนามบัญญัติ

2.1 Goodman and Kruskall's Lambda

สัมประสิทธิ์สหสัมพันธ์นี้ **จะใช้วัดความสัมพันธ์ของสองตัวแปรที่ถูกสุ่มมาจากประชากรหนึ่ง** ซึ่งมีมาตราวัดแบบ**นามบัญญัติ**เท่านั้น

แนวความคิดพื้นฐานที่ได้อธิบายมาดังกล่าวแล้วนั้นเป็นการวัดความสัมพันธ์ของตัวแปรที่รู้จักกันในชื่อ Lambda ($\hat{\lambda}$) ซึ่งมักจะมี subscript เป็น r ($\hat{\lambda}_r$) ถ้าเราหาค่า $\hat{\lambda}$ โดยกำหนดตัวแปร

ทางแถวนอนให้เป็นตัวแปรตามและ subscript เป็น c ($\hat{\lambda}_c$) ถ้ากำหนดให้ตัวแปรทางด้านแถวตั้งเป็นตัวแปรตาม ตัวอย่างการหาค่า $\hat{\lambda}_r$ เป็นดังนี

ตารางที่ 4 ข้อมูลเกี่ยวกับคำสั่งศาลในการลงโทษและประเภทของความผิดต่างๆ

ประเภทความผิด	ประเภทของคำพิพากษา			
	ปรับ	รอลงโทษ	จำคุก	รวม
ขโมย	5	30	5	40
ปล้น	0	30	20	50
ปลอมแปลง	5	0	5	10
รวม	10	60	30	100

ถ้าให้ประเภทของความผิดเป็นตัวแปรตาม สูตรสำหรับหาค่า $\hat{\lambda}$ ซึ่งได้ดัดแปลงให้คิดง่ายขึ้นได้แก่

$$\hat{\lambda}_r = \frac{\sum(\text{เซลที่ใหญ่ที่สุดในแต่ละแถวตั้ง}) - \text{ค่าMax. ของผลรวมแถวนอน}}{N - \text{ค่าMax.ของผลรวมแถวนอน}}$$

$$\hat{\lambda}_r = \frac{(5+30+20)-50}{100-50} = \frac{5}{50}$$

$$= 0.10$$

ซึ่งหมายความว่า ถ้าเราทราบประเภทของคำพิพากษา เราสามารถจะทำนายประเภทของความผิดได้โดยสามารถลดความผิดพลาดของคำทำนายได้ 10%

ค่า Lambda จะมีค่าอยู่ระหว่าง 0 ถึง 1 ในกรณีที่การทราบข่าวสารของตัวแปรอิสระไม่ช่วยลดความผิดพลาดในการทำนายตัวแปรตามเลย ค่า $\hat{\lambda}_r$ จะเป็น0 เรียกว่าไม่มี predictive association เลย ในทำนองเดียวกันถ้าค่า $\hat{\lambda}_r$ มีค่าเป็น 1 แสดงว่าการทราบข่าวสารของตัวแปรอิสระทำให้ไม่มีการทำนายผิดพลาดเลย เรียกว่า Complete predictive association

แนวคิดนี้ไม่เหมือนกับการทดสอบความอิสระของ $\chi 2$ และไม่เหมือนกับการหาค่าความสัมพันธ์ของ Cramer's C^2 นั่นคือ แม้ว่า $\hat{\lambda}_r = 0$ แต่ตัวแปรทั้ง 2 ก็อาจมีความสัมพันธ์กันเพียงแต่ความสัมพันธ์นี้ไม่อยู่ในลักษณะเป็น predictive association กล่าวคือการทราบตัวแปรอิสระไม่ช่วยลดความผิดพลาดในการทำนายตัวแปรตาม สำหรับการทดสอบนัยสำคัญของ χ^2 แม้ว่าตัว

แปรทั้งสองมีความสัมพันธ์ทางสถิติเล็กมากจนเกือบจะไม่เหลืออยู่แล้ว แต่ก็อาจสรุปได้ว่ามีนัยสำคัญทางสถิติที่สูงมาก ซึ่งมักจะเกิดเมื่อขนาดตัวอย่างใหญ่มากซึ่งทำให้ผู้วิจัยเข้าใจว่าเขาพบความสัมพันธ์แล้ว และนำไปประยุกต์กับประชากรในสังคมซึ่งที่จริงแล้วอาจจะไม่เป็นความจริง

สำหรับค่า Cramer's C^2 นั้น ถ้าเมื่อใดที่ตัวแปรไม่มีสัดส่วนแสดงความสัมพันธ์กันเลยจนได้ค่า $C^2 = 0$ แล้ว $\hat{\lambda}_r$ ก็จะเท่ากับ 0 ด้วย (แต่ถ้าค่า $\hat{\lambda}_r = 0$, C^2 ไม่จำเป็นต้องเท่ากับ 0) และเมื่อใดที่เกิดความสัมพันธ์อย่างสมบูรณ์จนกระทั่งเกิด prefect predictive association นั่นคือ $\hat{\lambda}_r = 1$ แล้ว C^2 จะเท่ากับ 1 ด้วย (แต่ถ้า $C^2 = 1$ ค่า $\hat{\lambda}_r$ ไม่จำเป็นต้องเท่ากับ 1)

ค่า Lambda จึงมีค่าเหนือกว่า χ^2 และ C^2 เพราะสามารถบอกpredictive association ได้ คือบอก "ค่าที่อธิบายได้" นั่นเอง กล่าวคือตัวแปรหนึ่งซึ่งเป็นตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ สำหรับค่า $\hat{\lambda}_c$ ก็สามารถคำนวณหาได้เช่นกัน

$$\hat{\lambda}_c = \frac{\sum(\text{เซลที่ใหญ่ที่สุดในแต่ละแถวนอน}) - \text{ค่าMax. ของผลรวมแถวตั้ง}}{\text{N} - \text{ค่าMax.ของผลรวมแถวตั้ง}}$$

สำหรับข้อมูลในตารางที่ 4 ได้ค่า $\hat{\lambda}_c$ ดังนี้

$$\hat{\lambda}_c = \frac{(30+30+5) - 60}{100 - 60}$$

$$= 0.125$$

นั่นคือถ้าเราทราบประเภทของความผิดเราสามารถทำนายคำพิพากษาได้โดยลดความผิดพลาดในการทำนาย 12.5%

ถ้าเปรียบเทียบค่า $\hat{\lambda}_c$ และ $\hat{\lambda}_r$ สำหรับข้อมูลชุดนี้จะพบว่า $\hat{\lambda}_c$ เป็นค่าที่ดีกว่าเพราะมีค่าสูงกว่า $\hat{\lambda}_r$ จึงลดความผิดพลาดได้มาก อย่างไรก็ตามผู้วิจัยคงต้องตัดสินใจว่าจะใช้ $\hat{\lambda}_c$ หรือ $\hat{\lambda}_r$ โดยดูที่ตัวแปรเป็นสำคัญ หรือไม่อาจจะตัดสินใจโดยดูค่าที่สูงสุดระหว่าง $\hat{\lambda}_c$ กับ $\hat{\lambda}_r$ ตัวแปรบางตัวสามารถมีสถานะภาพที่บางครั้งก็เป็นตัวแปรอิสระได้ บางครั้งก็เป็นตัวแปรตามได้ แต่ตัวแปรบางตัวถ้าพิจารณาดูอย่างรอบคอบแล้วอาจเป็นได้เฉพาะตัวแปรอิสระเท่านั้น ผู้วิจัยจึงต้องเลือกใช้ให้ถูกต้อง

เนื่องจากค่า Lambda เป็นค่า กำหนดทิศทางของตัวแปรได้ ว่าตัวแปรใดเป็นตัวแปรอิสระ ตัวแปรใดเป็นตัวแปรตาม จึงเป็นสถิติประเภทกำหนดทิศทาง (Directional หรือ Asymmetric Statistic) อย่างไรก็ตาม**ค่า$\hat{\lambda}$ ประเภทไม่กำหนดทิศทาง (Indirectional หรือ Symmetric Statistic) ก็สามารถหาได้เช่นกัน** ตัวอย่างเช่น

$$\hat{\lambda}_{rc} = \frac{(5+30+20) + (30+30+5) - 50 - 60}{2(100) - 50 - 60}$$
$$= 0.111$$

จะสังเกตได้ว่า เนื่องจากสูตรการหาค่า $\hat{\lambda}_{rc}$ เป็นการรวมเอาทั้งสูตร $\hat{\lambda}_c$และ $\hat{\lambda}_r$เข้าด้วยกัน ฉะนั้นค่า $\hat{\lambda}_{rc}$ จึงเป็นค่าที่อยู่ระหว่าง $\hat{\lambda}_c$และ $\hat{\lambda}_r$เสมอ มีความหมายว่าถ้าเราทราบตัวแปรหนึ่ง สามารถทำนายตัวแปรอีกตัวหนึ่งได้ ก็จะลดความผิดพลาดในการทำนาย 11.1 % นั่นก็คือ ถ้าเราไม่ระบุตัวแปรว่าตัวแปรใดเป็นตัวแปรอิสระ ตัวแปรใดเป็นตัวแปรตาม อย่างน้อยที่สุดก็สามารถลดความผิดพลาดในการทำนายเฉลี่ยแล้ว = 11.1% กล่าวคือในสภาพแวดล้อมบางอย่าง เราปรารถนาที่จะใช้ค่าที่ไม่บอกทิศทาง ($\hat{\lambda}_{rc}$) มากกว่า $\hat{\lambda}_r$ หรือ $\hat{\lambda}_c$

อย่างไรก็ตามค่า **Lambda เป็นค่าที่ไม่ไว (insensitive) ต่อความสัมพันธ์ของตัวแปรทุกๆ รูปแบบจึงเป็นที่นิยมใช้กันน้อยกว่า Goodman and Kruskall's Tau ซึ่งเป็นค่าที่ไวต่อความสัมพันธ์หลายๆ รูปแบบ**

ค่า $\hat{\lambda}$ ที่ขาดความไว ก็คือถ้าเผอิญตัวเลขในตารางแสดงว่ามีความสัมพันธ์กันบ้างไม่มากก็น้อย แต่ต้องไม่ใช่ 0 ตัวอย่างดังปรากฏในข้อมูลตารางที่ 5

ตารางที่5 ข้อมูลแสดงความสัมพันธ์ระหว่างเพศและการสูบบุหรี่

		การสูบบุหรี่		
		สูบ	ไม่สูบ	รวม
เพศ	ชาย	101	10	111
	หญิง	45	44	89
	รวม	146	54	200

ถ้าพิจารณาจากค่าตัวเลขในตาราง ตัวเลขเหล่านี้ยังแสดงถึงการมีความสัมพันธ์อยู่บ้าง โดยเฉพาะเพศชาย ซึ่งมีทั้งหมด 111 คน เป็นผู้สูบบุหรี่เสีย 101 คนแม้ว่าเพศหญิง จะมีผู้สูบบุหรี่พอๆ กับไม่สูบบุหรี่ก็ตาม ตารางนี้ถ้าคิดค่าสถิติฟาย แล้วปรากฏว่า = .204 ซึ่งหมายความว่าตัวแปรทั้งคู่มีความสัมพันธ์กัน 20.4%

แต่ถ้าพิจารณา $\hat{\lambda}$ โดยให้การสูบบุหรี่เป็นตัวแปรตาม ดังนั้นหากจะพิจารณาแต่ตัวแปรตามโดยไม่มีข่าวสารจากตัวแปรอิสระเลยในจำนวนทั้งสิ้น 200 คน เราย่อมทำนายว่าสูบบุหรี่ทั้งหมดนั่นคือมีการทำนายผิด 54 คน 54 นี้ย่อมเป็นค่า E1 นั่นเอง แต่ถ้าเรามีข่าวสารเกี่ยวกับเพศเข้า

มาให้ทายโดยแยกเพศชาย 111คน เพศหญิง 89คนในบรรดาเพศชายถ้าเราทำนายว่าสูบบุหรี่ ก็ย่อมมีการทำนายผิด
10 คน ในทำนองเดียวกัน เพศหญิงก็ทำนายผิด 44 คน รวมความผิดพลาดทั้งสิ้นเป็น 10+44 = 54 คนซึ่งก็คือ E2 นั่นเอง

$$\hat{\lambda}_c = \frac{E1 - E2}{E2} = \frac{54 - 54}{54} = 0$$

หรือมิฉะนั้นจะคิดตามสูตรก็ได้ตัวเลขเช่นเดียวกันคือ

$$\hat{\lambda}_c = \frac{(101+45)-146}{200-146} = 0$$

ซึ่งจริงๆ แล้วตัวเลขจากตารางยังปรากฏมีความสัมพันธ์กันอยู่ถึง 20.4% แสดงว่า $\hat{\lambda}$ ไม่ไวต่อตัวเลขในตารางบางชนิดที่ทำให้ค่า E1 และ E2 เกิดเท่ากัน ทั้งนี้เพราะค่า $\hat{\lambda}$ ไม่ได้เอาข้อมูลจากทุกๆ เซลมาร่วมคิดด้วย โดยเฉพาะเซลที่มีความถี่ รองๆ ลงไปจากเซลที่มีความถี่สูงสุด **ถ้าพิจารณาถึงจุดอ่อนในเรื่องนี้ Goodman and Kruskall's Tau จะเป็นค่าสถิติที่เหนือกว่าค่า Lambda** เพราะนำค่าของทุกๆ เซลมาหาความผิดพลาด และยังได้คำนึงถึงแต่ละหน่วยในความถี่รวมในแต่ละเซลอย่าง Lambda

การทดสอบนัยสำคัญของค่า $\hat{\lambda}$

ถ้าเรากำหนดขนาดตัวอย่างให้ใหญ่แล้ว Sampling distribution ของค่า $\hat{\lambda}$ จะกระจายเป็นโค้งปกติและมีค่า Z เป็น

$$Z = \frac{\hat{\lambda}_r - \lambda_r}{\hat{\sigma}_{(\hat{\lambda}_r)}}$$

โดยมีเงื่อนไขว่าค่าλ_r (ของประชากร) ต้องไม่เป็น 0 และมีค่า $\hat{\sigma}_{(\hat{\lambda}_r)}$ มีสูตรดังนี้

$$\hat{\sigma}_{(\hat{\lambda}_r)} = \sqrt{\frac{\left(n - \sum_j f_{ij}\right)\left(\sum_j f_{ij} + f_{+j}(max.) - 2\sum_j {}^{*} f_{ij}\right)}{\left(N - f_{+j}(max.)\right)^3}}$$

เมื่อ $\sum_{j} f_{ij}$ คือ ผลรวมความถี่ของเซลที่ใหญ่ที่สุดในแต่ละแถวตั้ง

$f_{+j}(\max.)$ คือ ค่ามากสุดของผลรวมแถวนอน

$\sum_{j}{}^{*} f_{ij}$ คือ ผลรวมของความถี่ที่ใหญ่ที่สุดของแต่ละแถวตั้ง แต่ความถี่ที่ใหญ่ที่สุดนี้ต้องอยู่ ตรงกับแถวนอนที่มี $f_{+j}(\max.)$ (ถ้ามีหลายจำนวนก็เอาจำนวนเหล่านั้นรวมกัน ถ้ามีเทอมเดียวก็ใช้เพียงค่านั้น)

จากตัวอย่างเรื่องประเภทคำพิพากษาและความผิดที่ผ่านมา ได้ค่า Standard error ดังนี้

$$\hat{\sigma}(\hat{\lambda}_{r}) = \sqrt{\frac{[100-(5+30+20)][(5+30+20)+50-2(30+20)]}{(100-50)^{3}}}$$

$$= 0.042$$

ซึ่งสามารถนำค่านี้ไปทดสอบนัยสำคัญและหาค่าประมาณของ λ_{r} (ของประชากร) แบบช่วงได้ สำหรับค่า λ_{c} ก็ทำได้เช่นกัน

หมายเหตุ จากข้อจำกัดของ $\hat{\sigma}\,\hat{\lambda}_{r}$ ข้างต้น ต้องใช้ทดสอบสมมติฐานที่ว่า ค่าพารามิเตอร์ เป็นค่าเฉพาะค่าใดค่าหนึ่ง (ในช่วง 0 – 1) ที่ไม่ใช่ค่า 0, 1

คือ $H_0: \lambda_{r} = C$ เท่านั้น โดย C ต้องไม่เท่ากับ 0 หรือ1

2.2 Goodman and Kruskall's Tau

ใช้วัดความสัมพันธ์ของสองตัวแปรที่ถูกสุ่มมาจากประชากรหนึ่ง และมีมาตราวัดแบบนามบัญญัติ

จากข้อมูลชุดเดียวกับที่ใช้หาค่า Lambda คือตารางที่ 4 สมมติว่าเราจะทำนายคำตัดสินการลงโทษของศาล โดยที่ไม่มีข่าวสารเกี่ยวกับตัวแปรทางด้านประเภทของความผิดเลย ข้อมูลให้แต่รายละเอียดว่า 10รายถูกปรับ 60รายถูกภาคทัณฑ์ และ 30ราย ถูกจำคุก รวมทั้งสิ้น 100คนแต่เราไม่ทราบว่า 10รายไหนจาก 100รายจะถูกปรับหรือ 60รายไหนจะถูกภาคทัณฑ์ หรือ 30รายไหนจะถูกจำคุก นั่นก็เหมือนกับว่าเราได้รับกล่องซองจดหมายมากล่องใหญ่บรรจุคำสั่งเป็นซองๆมา100 ซอง โดยปิดผนึก และเราถูกสั่งให้แยกซองใส่กล่องย่อย 3 กล่องที่เขียนกำกับว่า เป็นกล่องสำหรับพวกถูกปรับ ถูกภาคทัณฑ์ และถูกจำคุก โดยผู้สั่งให้ข่าวสารแต่เพียงว่า 10 % เป็นปรับ 60% เป็นภาคทัณฑ์ และ 30% เป็นจำคุก เมื่อไม่มีข้อมูลอื่นๆ อีกสิ่งที่ทำได้ก็คือใช้วิธีเสี่ยงทายคือหยิบซองโดยการสุ่มนับจำนวนให้ถูกต้อง และแยกใส่ลงไปในกล่องย่อยทั้ง 3 ใบ

โดยวิธีนี้เราคาดว่าจะเกิดความผิดพลาดสักเท่าไรในการแยกซองเหล่านั้น

สำหรับในกล่องย่อยที่เป็นการปรับนั้น คือ 10 ซอง ความน่าจะเป็นของความผิดพลาดก็คือ ความน่าจะเป็นที่ดึงซองจดหมายออกมา 1 ฉบับ แล้วพบว่าเป็นการภาคทัณฑ์หรือจำคุกแทนที่จะเป็นการปรับ ความน่าจะเป็นของความผิดพลาดคือ

$$60/100 + 30/100 = 90/100$$

เนื่องจากเรามีคน 10 รายที่ต้องถูกปรับโดยมีความน่าจะเป็นของความผิดพลาดเป็น90/100 หรือ 0.9 ดังนั้นถ้าคิดเป็นจำนวนความผิดพลาดจะได้

$$= 0.9 \times 10 = 9$$

สำหรับคำพิพากษาอื่นๆ ก็คิดจำนวนความผิดพลาดได้ในทำนองเดียวกันกล่าวคือถ้าเราเลือกซองใส่ในกล่องย่อยภาคทัณฑ์จำนวนความผิดพลาดทั้งหมดจะได้

$$(10/100+30/100) \times 60 = 24$$

และถ้าเราเลือกซองใส่ในกล่องย่อยการจำคุก จำนวนความผิดพลาดทั้งหมดจะได้

$$(10/100+60/100) \times 30 = 21$$

ดังนั้นรวมทั้งหมดจะได้ความผิดพลาดคือ 9+24+21 = 54 ซึ่งก็คือค่า E1 นั่นเอง (ความผิดพลาดที่ดูจากผลรวมแต่เพียงอย่างเดียวโดยไม่มีข่าวสารจากตัวแปรอิสระ)

ในการหาค่า E2 คือการทำนายคำตัดสินของศาล โดยได้รับความรู้เพิ่มเติมขึ้นว่าบุคคลเหล่านั้นประพฤติผิดอย่างไร (ซึ่งเป็นตัวแปรอิสระ) เช่นทราบมาว่า 40 คนเป็นขโมยและถูกศาลสั่งปรับเสีย 5 คน ภาคทัณฑ์เสีย 30 คน และจำคุก 5 คน แต่เราก็ไม่ทราบว่า 5 คนไหนถูกปรับ 30 คน

ไหนถูกภาคทัณฑ์ และ 5 คนไหนถูกจำคุกดังนั้นในการจัดซอง 40 ซองให้อยู่ในกล่องย่อยของการปรับ การภาคทัณฑ์ และจำคุกจึงทำความผิดพลาดได้ดังนี้

ก. การจัดซอง 5 ซองให้อยู่ในกล่องปรับจะเกิดความผิดพลาดเป็นจำนวน

$$\frac{35}{40} \text{ x } 5 = 4.375$$

ข. การจัดซอง 30 ซองให้อยู่ในกล่องภาคทัณฑ์จะเกิดความผิดพลาดเป็นจำนวน

$$\frac{10}{40} \text{ x } 30 = 7.5$$

ค.การจัดซอง 5 ซองให้อยู่ในกล่องจำคุกจะเกิดความผิดพลาดเป็นจำนวน

$$\frac{35}{40} \text{ x } 5 = 4.375$$

ดังนั้นสำหรับแถวนอนที่ หนึ่ง จะเกิดความผิดพลาดรวม = 4.375+7.5+4.375 =16.25จำนวน ฉะนั้น ในแถวนอนที่สองและสามก็หาความผิดพลาดรวมได้เช่นกัน ดังนั้น E2 จะมีค่าเป็น

$$\begin{aligned} E2 = & \left(\frac{35}{40}\times 5\right)+\left(\frac{10}{40}\times 30\right)+\left(\frac{35}{40}\times 5\right)+ \\ & \left(\frac{50}{50}\times 0\right)+\left(\frac{20}{50}\times 30\right)+\left(\frac{30}{50}\times 20\right)+ \\ & \left(\frac{5}{10}\times 5\right)+\left(\frac{10}{10}\times 0\right)+\left(\frac{5}{10}\times 5\right) \end{aligned}$$

$$= 45.25$$

ดังนั้น T_c (เมื่อให้แถวตั้งเป็นตัวแปรตาม) ซึ่งให้ค่า proportional reduction in errors ของการทำนายคำตัดสินในการลงโทษโดยทราบประเภทของอาชญากรรมจึงมีค่าดังนี้

$$T_c = \frac{E1 - E2}{E1}$$

$$= \frac{54 - 45.25}{54} = 0.16$$

ซึ่งตีความหมายได้ว่าการมีความสัมพันธ์กันระหว่างประเภทของอาชญากรรม และการตัดสินลงโทษ โดยทราบวิธีต่างๆ ทำให้สามารถลดความผิดพลาดในการทำนาย (PRE) ได้ถึง 16%

ถ้าได้เปลี่ยนการทำนายประเภทของการลงโทษ โดยพิจารณาจากผลรวมแต่เพียงอย่างเดียว ไปเป็นการทำนายประเภทของการลงโทษ โดยพิจารณาจากประเภทของอาชญากรรมที่ก่อขึ้นด้วย

และเพื่อความสะดวกในการคำนวณหาค่า Tau c จึงมีสูตรที่ดัดแปลงแล้วดังนี้

$$T_c = \frac{\sum_1^r \sum_1^k \frac{N_{ij}^2}{N_{i.}} - \sum^k (\frac{N_{.j}}{N})^2}{N - \frac{\sum^k (N_{.j})^2}{N}}$$

$$\sum_1^r \sum_1^k \frac{N_{ij}^2}{N_{i.}} = \frac{25}{40} + \frac{900}{40} + \frac{25}{40} + \frac{0}{50} + \frac{900}{50} + \frac{400}{50} + \frac{25}{10} + \frac{0}{10} + \frac{25}{10}$$

$$= 54.75$$

$$\sum^k (\frac{N_{.j}}{N})^2 = \frac{(10)^2 + (60)^2 + (30)^2}{100}$$

$$= 46$$

$$T_c = \frac{54.75 - 46}{100 - 46}$$

$$= 0.16$$

ซึ่งมีค่าเท่ากับข้างบนดังได้คำนวณไว้แล้ว

ค่า T_c นั้น เมื่อเปลี่ยนตัวแปรตามใหม่ ก็สามารถคำนวณหาได้คล้ายคลึงกับที่กล่าวมาแล้ว สำหรับตารางขนาด 2 x 2 จะได้ค่า $T_c = T_r$ = ค่าสถิติฟาย ($Ø^2$) จึงไม่นิยมใช้หาค่า Goodman and Kruskall's Tau ในตาราง 2 x 2 เพราะการหาค่า$Ø^2$ อาจสะดวกกว่า

ค่า Tau เหมือนค่าLambda ที่มีขีดจำกัดล่างอยู่ที่ 0 และขีดจำกัดบนอยู่ที่ 1

ค่า 0 หมายถึงไม่มีความสัมพันธ์ระหว่างตัวแปร

ส่วนค่า 1 หมายถึงมีความสัมพันธ์อย่างสมบูรณ์

และค่า Tau นี้มีลักษณะเหมือนLambda อีกประการหนึ่งคือเป็นค่าสถิติที่มีทิศทาง แต่ถ้าปรากฏว่าตัวแปรทั้ง 2 ไม่สามารถบอกทิศทางกันได้ (ไม่สามารถบอกหรือกำหนดได้ว่าตัวใดเป็นตัวแปรอิสระ ตัวใดเป็นตัวแปรตาม) ก็สามารถคำนวณหาค่า T_{rc} ได้เช่นกัน

3. การทดสอบไคสแควร์สำหรับกลุ่มตัวอย่าง r กลุ่มที่เป็นอิสระกันหรือการทดสอบของ χ^2 เกี่ยวกับความเป็นเอกภาพของสัดส่วน

(χ^2 test for r Independent samples หรือ The χ^2 test for Homogeneity of Proportions)

การทดสอบนี้ขยายมาจากการทดสอบของ χ^2 สำหรับ 2 กลุ่มตัวอย่างที่เป็นอิสระกัน ดังนั้นเหมาะที่จะใช้เมื่อตัวอย่างมีขนาดใหญ่

ข้อมูล ใช้เมื่อข้อมูลคือความถี่ (frequency) ที่ได้จากมาตรวัดแบบนามบัญญัติหรือเรียงลำดับ การทดสอบยังคงคล้ายกับการทดสอบของ χ^2 สำหรับกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระต่อกัน คือ ทำการเปรียบเทียบความถี่ที่สังเกตได้ (observed frequency) กับความถี่คาดหวัง (Expected frequency) โดยข้อมูลจะจัดลงตารางแจกแจง 2ทางชนิด r x c เมื่อมี c แถวตั้ง และ r แถวนอน โดยที่แบบทดสอบนี้จะสุ่มตัวอย่าง r ประชากร ด้วยขนาดตัวอย่าง $n_1, n_2, \ldots, n_r$ ตามลำดับ ข้อมูลที่ได้จากตัวอย่างจะสรุปๆ ได้ดังตารางต่อไปนี้

ตัวอย่างที่	ลักษณะย่อย						ขนาดตัวอย่าง
	1	2	3	…	j	c	
1	O_{11}	O_{12}	O_{13}	…	…	O_{1c}	n_1
2							n_2
3							.
.							.
i	…	…	…	…	O_{ij}	…	.
r	O_{r1}	O_{r2}	O_{r3}	…	…	O_{rc}	n_r
รวม	n_1	n_2	n_3	…	…	n_c	

สมมติฐาน คือ สัดส่วนในแต่ละประชากรจะเท่ากันหรือเป็นเอกภาพ นั่นคือ

H_0 : ตัวอย่าง r กลุ่มนี้มาจากประชากรเดียวกัน หรือสัดส่วนของแต่ละลักษณะย่อยในประชากร r กลุ่ม ไม่แตกต่างกัน

H_1 : ตัวอย่าง r กลุ่มนี้มาจากประชากรที่ต่างกัน

หรือ

$$H_0 : \pi_{11} = \pi_{21} = \pi_{31} = ... = \pi_{r1} = \pi_1$$

$$\pi_{12} = \pi_{22} = \pi_{32} = ... = \pi_{r2} = \pi_2$$

.

.

.

$$\pi_{1c} = \pi_{2c} = \pi_{3c} = ... = \pi_{rc} = \pi_c$$

หรือ $\pi_{1j} = \pi_{2j} = \pi_{3j} = ... = \pi_{rj} = \pi_j$ (j =1, 2,..c)

หรือ H_0 : ตัวแปรทางแถวนอนและแถวตั้งเป็นอิสระต่อกัน

สถิติที่ใช้ทดสอบ

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

เมื่อ O_{ij} = ความถี่ที่สังเกตได้จากตารางแจกแจง 2 ทางที่แถวนอน i และแถวตั้ง j

E_{ij} = ความถี่คาดหวังที่ได้จากตารงแจกแจง 2 ทางที่แถวนอน i และแถวนอนตั้ง j

ซึ่งการหาค่า E_{ij} ภายใต้ H_0 ยังคงคล้ายกับการทดสอบ χ^2 ในตัวอย่าง 2 กลุ่ม

คือ $E_{ij} = \dfrac{\text{(ผลรวมของแถวนอนที่ i)(ผลรวมของแถวตั้งที่ j)}}{\text{จำนวนความถี่ทั้งหมด}}$

ค่า E_{ij} ไม่ควรมีค่า < 5 เป็นจำนวนมากกว่า 20% ของจำนวนเซลทั้งหมด ถ้าเกิด เหตุการณ์เช่นนี้ อาจแก้ไขโดยรวมกลุ่มที่ใกล้กันเข้าด้วยกันซึ่งจะทำให้ d.f. ลดลง

ค่าสถิตินี้สามารถประมาณการแจกแจงได้ด้วย χ^2 ที่ d.f. = (k – 1)(r – 1)

ดังนั้น การหาอาณาเขตวิกฤต ก็ทำการเปรียบเทียบค่า χ^2 ที่ได้จากสูตรนี้กับค่า χ^2 จากตารางที่ระดับนัยสำคัญหนึ่ง

ตัวอย่างที่ 1 ในการศึกษาถึงความชอบของแม่บ้านต่อวิธีการบรรจุมันฝรั่งทอดกรอบแบบต่างๆ จากแม่บ้านที่มีรายได้ต่างๆ กัน 3 กลุ่ม จึงสุ่มตัวอย่างแม่บ้านจาก 3กลุ่มรายได้มาด้วยขนาด 80 100 และ 120 คน แล้วทดสอบถึงความชอบของวิธีการบรรจุ 3 แบบ คือ บรรจุในกล่องแบบโลหะ (Metallic) หรือบรรจุในกระดาษที่ทำด้วยขี้ผึ้ง (Waxed paper) หรือแบบกระดาษแก้ว (Cellophane) ได้ข้อมูลดังนี้

แม่บ้านที่มีรายได้	ชนิดการบรรจุที่ชอบ			รวม
	Metalic	Waxed Paper	Cellophane	
รายได้สูง	15	25	40	80
รายได้ปานกลาง	27	33	40	100
รายได้ต่ำ	55	40	25	120
รวม	97	98	105	300

จงหาผลสรุปว่า กลุ่มแม่บ้านที่มีรายได้แตกต่างกัน 3 กลุ่ม ดังกล่าวมีค่าสัดส่วนของความชอบต่อการบรรจุมันฝรั่งทอดด้วยวิธีต่างๆ แตกต่างกันหรือไม่ ที่ α = 0.01 หรือรายได้และชนิดการบรรจุที่ชอบ มีความสัมพันธ์กันหรือไม่

วิธีทำ เนื่องจากข้อมูลมีมาตราวัดแบบนามบัญญัติ แบบ 3 กลุ่มตัวอย่างที่เป็นอิสระกันและสนใจเกี่ยวกับค่าสัดส่วนของเหตุการณ์ย่อย ดังนั้น สถิติที่เหมาะสมจึงคือ χ^2

จาก $$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{k} \frac{(\mathrm{O}_{ij} - \mathrm{E}_{ij})^2}{\mathrm{E}_{ij}}$$

คำนวณหาค่าความถี่คาดหวังของเซลต่างๆ ดังนี้

เช่นหาค่า $\mathrm{E}_{22} = \frac{98 \times 100}{300} = 32.67$

และไม่มีเซลใดที่มีค่า E_{ij} น้อยกว่า 5 จึงไม่จำเป็นต้องปรับตาราง

คำนวณหาค่า $$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{k} \frac{(\mathrm{O}_{ij} - \mathrm{E}_{ij})^2}{\mathrm{E}_{ij}} \quad \text{for } i = 1, 2, 3$$

$j = 1, 2, 3$

$= 25.70$

ที่ระดับนัยสำคัญ 0.01 d.f. = (3 –1)(3 – 1) = 4

ค่า χ^2 วิกฤต = 13.277

ค่า χ^2 คำนวณ ตกอยู่ในอาณาเขตวิกฤต จึงปฏิเสธ H_0 และยอมรับ H_1 นั่นคือ สัดส่วนของความชอบวิธีการบรรจุมันฝรั่งทอดแบบต่างๆ มีค่าแตกต่างกันในประชากร 3 กลุ่ม หรือมีความสัมพันธ์กันระหว่าง รายได้กับชนิดของการบรรจุที่ชอบ

4. การทดสอบแบบ Fisher's exact เมื่อตารางเป็นชนิด r x c หรือคือการทดสอบของ Freeman-Halton

จากวิธีการทดสอบของ Fisher จากตารางแบบ 2 x 2 สามารถขยายไปใช้กับตารางชนิด r x c ได้ ด้วยการนำเสนอของ Freeman และ Halton (1951) (ซึ่งยังคงใช้หลักการเดียวกันคือขนาดตัวอย่างเล็กๆ ซึ่งไม่ควรใช้ χ^2 – test เพราะจะเกิดค่าความถี่คาดหวังที่น้อยกว่า 5 เกิน 20 % ของจำนวนเซลทั้งหมด นั่นคือ ตัวสถิติทดสอบจะไม่มีการแจกแจงประมาณได้ด้วยการแจกแจงแบบไคสแควร์ที่ d.f. (r-1)(c-1))

วิธีการคำนวณจะยังคงใช้หลักการคิดค่าความน่าจะเป็นของเหตุการณ์จากการแจกแจงแบบไฮเปอร์จีออเมตริกจากตาราง r x c หนึ่งๆ เมื่อกำหนดให้ผลรวมมาร์จินัลคงที่ และกำหนดให้

H_0 : ค่าความน่าจะเป็นของเหตุการณ์ย่อยต่างๆใน r ประชากรไม่ต่างกัน

หรือ H_0 : ไม่มีความสัมพันธ์ระหว่างตัวแปรในแถวนอนและแถวตั้ง

นั้นเป็นจริง (เมื่อให้ r แถวนอนแทน r ประชากร และ c แถวตั้งแทนลักษณะย่อยต่างๆ)

ด้วยค่าความน่าจะเป็น $$P = \frac{\Pi_i (n_{i+}!)\ \Pi_j (n_{+j}!)}{n!\ \Pi_{i,j} (n_{ij})!}$$

และเมื่อรวมค่าความน่าจะเป็น P นี้กับค่า P ที่คำนวณจากตาราง r x c แบบอื่นๆที่เป็นไปได้ในกรณีที่มีค่าสังเกตุน้อยกว่า เมื่อกำหนดให้ผลรวมมาร์จินัลคงที่ จะได้ค่า p-value และนำไปเปรียบเทียบกับค่าระดับนัยสำคัญที่กำหนด $= \alpha$ ถ้าค่า p-value น้อยกว่าค่า α ที่กำหนด จะปฏิเสธ H_0 : ไม่มีความสัมพันธ์ระหว่างตัวแปรในแถวนอนและแถวตั้ง ที่ระดับนัยสำคัญ α

<u>หมายเหตุ</u> เพื่อความเข้าใจในการคำนวณค่า p-value สามารถทำความเข้าใจจากหัวข้อของ Fisher's exact test กรณีตารางชนิด 2 x 2 ซึ่งจะใช้หลักการเดียวกัน แต่วิธีนี้คิดจากตารางชนิด r x c แทนเท่านั้น

ตัวอย่าง จากตัวอย่างผู้ป่วยที่ติดเชื้อซึ่งเกิดการระคายเคืองอย่างรุนแรง (severe irritation) ในบริเวณใดบริเวณหนึ่งของร่างกายคือ ที่จมูก ลำคอ หรือที่บริเวณหู จากผู้ป่วยที่มีถิ่นฐานต่างๆ 6 แหล่ง คือ A ถึง F สมมติว่าได้ข้อมูลจำนวนผู้ป่วยในกลุ่มย่อยต่างๆดังนี้

แหล่งถิ่นฐาน	บริเวณที่ระคายเคือง		
	จมูก	ลำคอ	บริเวณหู
A	1	0	1
B	1	1	0
C	0	1	0
D	1	1	0
E	8	0	7
F	0	1	1

ต้องการทดสอบว่าในกลุ่มประชากรจากตัวอย่างที่สุ่มมานี้ มีความสัมพันธ์ระหว่างแหล่งถิ่นฐานของผู้ป่วย กับบริเวณที่เกิดการระคายเคืองหรือไม่ (หรือสัดส่วนของผู้ป่วยที่เกิดการระคายเคืองในบริเวณต่างๆ (จมูก คอ หู) แตกต่างกันในกลุ่มผู้ป่วยที่มาจากแหล่งถิ่นฐาน A ถึง F)

วิธีทำ จากข้อมูลตัวอย่างจะพบว่าตัวอย่างจากแต่ละแหล่งถิ่นฐาน (ประชากร) มีขนาดเล็กมาก (คือ $n_A = 2$, $n_B = 2$, … , $n_F = 2$) และตารางเป็นชนิด 6 x 3 ข้อมูลชนิดนี้ควรใช้การทดสอบของ Freeman - Halton แทนที่จะใช้ χ^2 test for Homogeneity เนื่องจากจะเกิดค่าความถี่คาดหวังที่น้อยกว่า 5 เป็นจำนวนมาก(เกิน 20 % ของจำนวนเซลทั้งหมด)

คำนวณค่า P จากตารางค่าสังเกตุนี้จะได้ค่า 0.00005769

และค่า p-value ของการเกิดตารางค่าสังเกตุนี้หรือมีค่าน้อยกว่า = .0261 (จากโปรแกรมสำเร็จรูป STATXACT)

ดังนั้น สามารถสรุปได้ว่า จะปฏิเสธ H_0 : ไม่มีความสัมพันธ์ระหว่างบริเวณที่เกิดการระคายเคืองกับ แหล่งถิ่นฐานของผู้ป่วย ด้วยค่า α ที่แท้จริง = 2.61 %

นั่นคือ มีความสัมพันธ์ระหว่างตัวแปรทั้งสอง ซึ่งถ้าพิจารณาดูจำนวนในเซลต่างๆของตาราง จะพบว่าจากแหล่งถิ่นฐาน E จะเกิดการระคายเคืองที่บริเวณจมูก และบริเวณหูในสัดส่วนที่สูงมากกว่าที่อื่นๆ ในขณะที่แหล่งถิ่นฐานอื่นๆ (A-D และ F) มีสัดส่วนที่ใกล้เคียงกัน จึงไม่เป็นที่น่าสงสัยที่จะปฏิเสธ H_0

หมายเหตุ ถ้าใช้ χ^2 test for Homogeneity กับตารางค่าสังเกตุนี้จะได้ค่า $\chi^2_{cal} = 14.96$

และได้ค่า p-value = $P(\chi^2_{10} \geq 14.96) = .1335$ ซึ่งมีค่าใหญ่ ไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = 0.05$

ซึ่งจะได้ผลตรงกันข้ามกับวิธีของ Freeman - Halton ดังกล่าวข้างต้น

ตารางชนิด r x c แบบมีลำดับที่

ข้อมูลกลุ่มบางประเภท สามารถจัดลำดับความแตกต่างระหว่างกลุ่มได้ เช่น ผลการรักษาผู้ป่วย จำแนกเป็น มีอาการแย่ลง มีอาการคงเดิม มีอาการดีขึ้น ฉะนั้นถ้าสามารถเลือกใช้สถิติที่คำนึงถึงการลดหลั่นกันเหล่านั้น ก็จะเป็นการใช้ข้อมูลได้ครบถ้วนมากยิ่งขึ้น ทำให้ผลสรุปน่าเชื่อถือมากขึ้น ในหัวข้อนี้ จะกล่าวถึงตารางชนิด r x c สองแบบ

คือ **ชนิดที่ตัวแปรทางแถวนอนและแถวตั้งจัดลำดับที่ได้ทั้งสอง ซึ่งจะเรียกตารางที่ได้ว่า Doubly ordered Table**

หรือจะมีเพียงหนึ่งตัวแปร(อาจเป็นทางแถวนอนหรือแถวตั้ง)ที่จัดลำดับได้ ซึ่งจะเรียกตารางที่ได้ว่า Singly ordered Table

โดยลำดับแรกจะกล่าวถึงค่าสัมประสิทธิ์สหสัมพันธ์ และในลำดับท้าย จะกล่าวถึงสถิติที่ใช้ทดสอบความเป็นอิสระของตัวแปร ทางแถวนอนและแถวตั้ง ที่จัดลำดับได้

1.สถิติ PRE ที่ใช้กับตารางซึ่งตัวแปรเป็นเรียงลำดับ

1.1 Goodman and Kruskall's Gamma (γ)

สถิตินี้จะใช้กับตารางแบบ Doubly ordered Table

ตารางที่ข้อมูลมีมาตราวัดแบบเรียงลำดับ (Ordinal scale) นั้น แม้ข้อมูล ที่บรรจุอยู่ในตารางจะเป็นความถี่เช่นเดียวกับนามมาตรา แต่เนื่องจากช่วงต่างๆ มีลำดับความมากน้อยต่างกัน สถิติที่ใช้กับตารางที่ข้อมูลมีมาตราวัดแบบเรียงลำดับ จึงพยายามใช้ประโยชน์จากการเรียงลำดับมากน้อยของช่วงมาเป็นหลักในการวัด PRE คือดูว่าความรู้หรือข่าวสารในเรื่องลำดับของข้อมูลทางตัวแปรอิสระจะลดความผิดพลาดในการทำนายลำดับของตัวแปรตามได้มากน้อยเพียงไร

เพื่อที่จะเข้าใจการเปรียบเทียบลำดับของข้อมูลที่มีมาตราแบบเรียงลำดับและความเข้าใจในกฎแห่งความคิดของสถิติแบบ **Gamma และ Somer's d** จึงใช้ตัวอย่างต่อไปนี้เป็นแนวทางในการสร้างความเข้าใจ

ตารางที่ 6 ข้อมูลแสดงความสัมพันธ์ระหว่างน้ำหนักและส่วนสูง (เป็นกรณีตัวอย่างขนาดเล็กมากและไม่มีจำนวนซ้ำ เพื่อเข้าใจได้ง่าย)

		ความสูง		
		สูง	กลาง	ต่ำ
น้ำหนัก	มาก		ดาว	
	กลาง	เดือน		
	ต่ำ			เด่น

สำหรับการเปรียบเทียบลำดับกันนั้น จำเป็นต้องจับคู่ข้อมูลเป็นคู่ๆ เพื่อเทียบดูว่าใครจะสูงกว่ากันและใครจะหนักกว่ากัน

จำนวนคู่	ลำดับความสูง	ลำดับน้ำหนัก	ลักษณะลำดับทั้งสอง
เดือนกับดาว	เดือนสูงกว่าดาว	ดาวหนักกว่าเดือน	ขัดแย้งกัน
เดือนกับเด่น	เดือนสูงกว่าเด่น	เดือนหนักกว่าเด่น	คล้อยตามกัน
ดาวกับเด่น	ดาวสูงกว่าเด่น	ดาวหนักกว่าเด่น	คล้อยตามกัน

ข้อมูลจากตารางที่ 6 **ซึ่งมีเพียง 3 คน** จะได้จำนวนคู่ทั้งหมดที่เป็นไปได้ 3 คู่ สรุปได้ว่ามีลำดับที่ขัดแย้ง (discordant pair เขียนย่อแทนว่า D) หนึ่งคู่และมีลำดับคล้อยตามกัน (concordant pairs เขียนย่อว่า C) 2 คู่ ดังนั้น ถ้าให้ความสูงเป็นตัวแปรอิสระ และน้ำหนักเป็นตัวแปรตาม ถ้าเราจะทำนายน้ำหนักของคนๆ หนึ่งโดยทราบส่วนสูง ก็ควรทำนายในลักษณะคล้อยตามกันจะมีความถูกต้องมากกว่า กล่าวคือ ถ้าทราบว่าสูงมากควรทำนายว่าหนักมากด้วย ทั้งนี้เพราะการจัดลำดับนั้นมีจำนวนคู่ที่คล้อยตามกันมากกว่า จำนวนคู่ที่ขัดแย้งกัน ดังนั้นในการทำนายเช่นนี้ จะมีความผิดพลาดไป 1 คู่ นั่นคือ $E_2 = 1$ นั่นเอง

สำหรับการหาค่า E_1 คือการหาความผิดพลาดในการทำนายโดยดูจากอันดับของตัวแปรตามแต่เพียงอย่างเดียว และไม่ได้ใช้ความรู้เกี่ยวกับอันดับทางตัวแปรอิสระเข้ามาช่วยเลยนั้น วิธีทำก็คือจับเอาคู่ใดคู่หนึ่งออกมาเปรียบเทียบแล้วทำนายทางด้านน้ำหนัก (ตัวแปรตาม) อย่างเดียว สมมุติว่าเป็นคู่ของดาวกับเดือน การทำนายคงจะต้องใช้หลักการสุ่มเป็นสำคัญ เช่น ถ้าสุ่มหยิบได้ชื่อ ดาวขึ้นมาก็บอกว่า ดาวหนักมากกว่าเดือน (ซึ่งถูกต้องตามความเป็นจริง) หรือถ้าสุ่มได้เดือนก็บอกว่าเดือนหนักกว่าดาว (ซึ่งไม่ถูกต้องตามความเป็นจริง) นั่นคือถ้าสุ่มแบบนี้หลายๆ ครั้งเป็นระยะเวลายาวนาน จะพบว่าโอกาสของการทำผิดเท่ากับครึ่งหนึ่งของจำนวนการสุ่มครั้งหนึ่งของจำนวนการสุ่มทั้งหมด นั่นคือเกิดความผิดพลาดเป็น 0.50 สำหรับในกรณีนี้มีคู่ที่เป็นไปได้ทั้งหมด 3 คู่ จึงเกิดความผิดพลาดทั้งหมด = 0.5 x 3 = 1.5 ครั้ง ซึ่งเลขจำนวนนี้ก็คือค่า E_1 นั่นเอง

ดังนั้น Gamma (γ) $= \dfrac{E_1 - E_2}{E_1}$

$$= \frac{1.5 - 1}{1.5}$$

$$= 0.33$$

นั่นคือเราสามารถลดความผิดพลาดในการทำนายได้ 33% จากการทำนายอันดับของน้ำหนักแต่เพียงอย่างเดียวไปเป็นการทำนายอันดับของน้ำหนักโดยดูอันดับของความสูงเป็นตัวชี้แนะ ทั้งนี้เนื่องจากตัวแปรอิสระและตัวแปรตามมีความสัมพันธ์กัน (อธิบายกันได้) จึงสามารถทำเช่นนี้ได้

ตัวอย่างที่ยกมาข้างต้นมีข้อมูลเพียง 3 หน่วย เท่านั้น สำหรับตารางข้อมูลโดยทั่วไปข้อมูลจะมีมากกว่านี้มากมาย การจะเทียบลำดับลำดับช่วงของตัวแปรทั้งสองจึงต้องใช้วิธีอื่นซึ่งง่ายกว่าแต่ยังคงไว้ซึ่งพื้นฐานของสูตร $\dfrac{E_1 - E_2}{E_1}$ ตามเดิม

โดยทั่วๆ ไปสูตรของการหาค่า Gamma คือ Gamma $(\gamma) = \frac{C - D}{\text{untied pairs}}$

C คือจำนวนคู่ที่มีลำดับคล้อยตามกันทั้งหมด

D คือจำนวนคู่ที่มีลำดับขัดแย้งกันทั้งหมด

สำหรับค่า untied pairs นั้นคือจำนวน C+D นั้นเอง แต่เพื่อความเข้าใจในการจัดคู่ทั้งหมดของข้อมูลแบบตาราง หัวข้อแทรกต่อไปนี้จะสร้างความเข้าใจในเรื่องการจัดคู่และจะได้ใช้ต่อไปในเรื่องการคำนวณหาค่าสถิติแบบ Somer 's d, Kendall's Tau ต่างๆ

Tied pairs และ Untied pairs (จำนวนคู่แบบผูกและแบบไม่ผูก) ถ้าสมมุติมีข้อมูลจำนวนทั้งหมด 100 หน่วย จะคำนวณหาจำนวนคู่ทั้ง tied pairs และ untied pairs ได้ดังนี้

$$\text{จำนวนคู่ทั้งหมด} = \frac{N(N-1)}{2}$$

$$= \frac{100(99)}{2} = 4950 \text{ คู่}$$

ถ้าสมมุติให้ T เป็นจำนวนคู่ที่เป็นไปได้ทั้งหมด T จะประกอบด้วย

$$T = C + D + T_r + T_c + T_{rc}$$

หรือ $= T_u + T_r + T_c + T_{rc}$

T_u = เป็นการจับคู่กับหน่วยต่างๆ ที่ไม่อยู่ในแถวนอนและแถวตั้งเดียวกับเซลของตนเอง ปกติจะ มีค่าเท่ากับผลบวกของคู่ที่คล้อยตามกันทั้งหมด และคู่ที่ขัดแย้งกันทั้งหมด

T_r = การจับคู่กับเซลที่อยู่ตรงกับแถวนอนของตนเอง (pairs tied on row)

T_c = การจับคู่กับเซลต่างๆ ที่อยู่ตรงกับแถวตั้งของตนเอง (pairs tied on column)

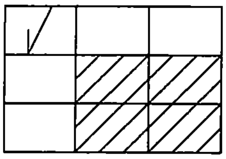
T_{rc} = การจับคู่กันในเซลเดียวกัน

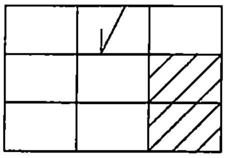
ตัวอย่างการจับคู่ต่างๆ ทำได้ดังนี้

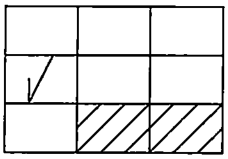
ตารางที่ 7 ข้อมูลไขว้ระหว่างการศึกษาและรายได้มีดังนี้

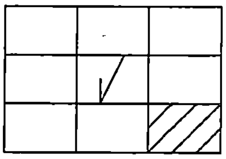
		การศึกษา		
		สูง	กลาง	ต่ำ
รายได้	สูง	20	10	5
	กลาง	10	10	10
	ต่ำ	5	10	20

ก. การหาคู่ที่คล้อยตามกันทั้งสองตัวแปรใช้ผังตารางดังข้างล่างนี้







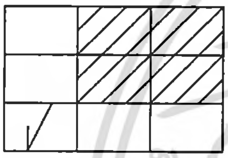


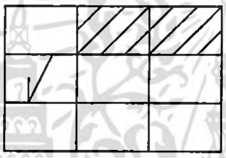
ให้จับคู่กันระหว่างเซลที่ทแยงจากมุมซ้ายบนลงมามุมขวาล่าง โดยเซลบนและกลุ่มเซลล่างต้องไม่อยู่ในแถวนอนหรือสดมภ์เดียวกัน (ดูรูปประกอบ)

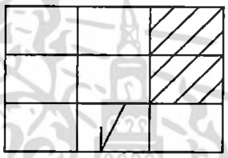
ค่า C = 20(10+10+10+20) + 10(10+20) + 10(10+20)+10(20)

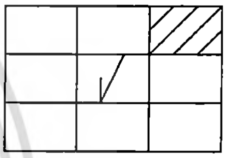
= 1800 คู่

ข. การหาคู่ที่ขัดแย้งกันระหว่างตัวแปรทั้งสองใช้ผังดังข้างล่างนี้









ให้จับคู่ในเซลที่ทแยงมุมล่างซ้ายขึ้นไปบนมุมขวาบน โดยเซลล่างและกลุ่มเซลบนต้องไม่อยู่ในแถวนอนและแถวตั้งเดียวกัน (ดูรูปประกอบ)

ค่า D = 5(10+10+10+5) + 10(10+5) + 10(10+5)+10(5)

= 525 คู่

ดังนั้น untied pairs = 1800 + 525

= 2325 คู่

เมื่อคู่ที่เป็นคู่คล้อยตามกัน และคู่ที่ขัดแย้งกัน ไม่ผูกอยู่ในสดมภ์หรือแถวนอนเดียวกัน ก็สามารถทำให้เปรียบเทียบอันดับกันได้ทั้งสองตัวแปร

ค. การหาจำนวนคู่ที่อยู่ตรงกันกับแถวนอนของตนเอง (pairs tied on row) ทำได้ดังนี้คือ

T_r = 20(10+5) +10(5)+.........แถวนอนที่ 1

10(10+10) + 10(10)+....... แถวนอนที่ 2

5(10+20) + 10(20)............แถวนอนที่3

= 1000 คู่

ง. การหาจำนวนคู่ที่ตรงกับแถวตั้งของตนเอง (pairs tied on column) หาได้ดังนี้

T_c = 20(10+5) +10(5) +…………………แถวตั้งที่ 1

10(10+10) +10(10)+……………… แถวตั้งที่ 2

5(10+20) +10(20) ……………… แถวตั้งที่ 3

= 1000

จ. การจับคู่ภายในเซลเดียวกันเอง (pairs tied on row and column) หาได้ดังนี้ เช่น ในช่อง a_{11} ซึ่งมีค่าเท่ากับ 20 จะมีจำนวนคู่ที่เป็นไปได้ทั้งหมด

a_{11} = 19+18+17+16+15+…………..1 190 คู่

a_{12} = 9+8+7+6+5+……………….1 45 คู่

ทำดังนี้ครบทุกเซล ผลรวมของจำนวนคู่ทั้งหมดจะเป็น 625 คู่ หรือหาด้วยอีกวิธีหนึ่ง คือ

T_{rc} = 4950 – 1800 – 525 – 1000 – 1000

= 625 คู่

สำหรับข้อมูลจากตารางที่ 7 ได้ค่า Gamma ดังนี้

$$\gamma = \frac{1800 - 525}{1800 + 525}$$

$$= + 0.55$$

ตีความหมายได้ดังนี้คือ **ให้ดูเครื่องหมายแยกจากตัวเลข ถ้าเครื่องหมายเป็นบวกให้ทำนายตัวแปรตัวหนึ่งในลักษณะที่คล้อยตามกับตัวแปรอีกหนึ่ง** เช่นในเรื่องนี้ ถ้าพบผู้ที่มีการศึกษาสูงก็ควรทำนายว่ามีรายได้สูงด้วย เพราะมีคู่ที่มีลักษณะคล้อยตามกันมากกว่าคู่ที่ขัดแย้งกัน เครื่องหมายจึงได้บวก และการทำนายเช่นนี้จะลดความผิดพลาดลงได้ 55%

เนื่องจากสูตรของ Gamma ไม่ได้ใช้คู่ที่ผูกอยู่กับแถวนอนหรือแถวตั้งเข้ามาเกี่ยวข้องเลย การหาจำนวนคู่ที่คล้อยตามกันและขัดแย้งกันไม่ได้บ่งบอกว่าตัวแปรใดเป็นตัวแปรอิสระ ตัวแปรใดเป็นตัวแปรตาม **ค่าGamma จึงเป็น Symmetric หรือ indirectional statistic (ไม่บอกทิศทาง)**

ตัวอย่างข้อมูลที่มีคู่ที่ขัดแย้งกันมากกว่าคู่ที่คล้อยตามกัน

ตารางที่ 8 ข้อมูลไขว้ระหว่างการศึกษาและรายได้

		การศึกษา			
		สูง	กลาง	ต่ำ	รวม
รายได้	สูง	4	6	25	35
	กลาง	15	5	10	30
	ต่ำ	24	8	3	35
	รวม	43	19	38	100

ค่า $C = 4(5+10+8+3) + 15(8+3) + 6(10+3) + 5(3)$

$= 352$

ค่า $D = 24(6+5+25+10) + 8(25+10) + 15(6+25) + 5(25)$

$= 1974$

ดังนั้น $\gamma = \dfrac{352-1974}{352+1974}$

$= -0.697$

ซึ่งตีความหมายว่า ถ้าจะให้ทำนายรายได้โดยดูคุณสมบัติจากการศึกษา **ต้องกล่าวว่าผู้มีการศึกษาสูงจะมีรายได้ต่ำ** หรือถ้าทำนายการศึกษาโดยดูคุณสมบัติจากรายได้ ต้องกล่าวว่าผู้มีรายได้สูงจะมีการศึกษาต่ำ**(ทั้งนี้เพราะจำนวนคู่ที่ขัดแย้งกันมีมากกว่าคู่ที่คล้อยตามกัน)** **และการทำนายนี้จะลดความผิดพลาดได้ 69.70%**

Gamma มีค่าผันแปรอยู่ระหว่าง –1 ถึง +1 แต่ไม่เป็นปัญหาในการตีความ เพราะต้องแยกเครื่องหมายออกจากตัวเลข ความหมายของตัวเลขยังอิงอยู่กับหลัก probability คือดูค่า 0 ถึง 1 หรือเทียบเป็นเปอร์เซนต์คือ จาก 0 ถึง 100 เปอร์เซ็นต์

ค่า – 1 หมายถึง คู่แบบไม่ผูก (untied pairs) ทุกคู่เป็นคู่ที่ขัดแย้งกัน ส่วนค่า + 1 คือคู่แบบไม่ผูกทุกคู่เป็นคู่ที่คล้อยตามกัน

โดยสรุป ค่า Gamma เป็นวิธีที่เปลี่ยนจากการดูข้อมูลเป็นกลุ่มในแต่ละเซล (เช่น แบบของ Lambda) มาเป็นการศึกษาเป็นรายกรณี คือเปรียบเทียบกันเป็นคู่ๆ ซึ่งมีความละเอียดกว่าแบบ Lambda มาก อย่างไรก็ตาม Gamma ก็ไม่ได้เอาคู่ที่เป็นแบบผูก (tied pairs) มาศึกษาด้วย ซึ่งถ้าจำนวนคู่แบบผูกมีเป็นจำนวนมากการตีความหมายของ Gamma ก็จะมีการบิดพลิ้วไปจากข้อเท็จจริงมาก อันเป็นจุดอ่อนของ Gamma

ผลกระทบของผลรวมของแถว (นอนหรือตั้ง) ต่อค่า Gamma ถ้าผลรวมมีความเบ้จะทำให้เกิดจำนวนคู่แบบผูก มีมากกว่าจำนวนคู่แบบไม่ผูก, ค่า Gamma เป็นค่าที่หาจากคู่ที่ไม่ผูกโดยตรง ซึ่งได้ลดจำนวนลงถ้าจำนวนคู่ที่ผูกเพิ่มขึ้น ดังนั้น การสรุปความสัมพันธ์ของตารางทั้งตารางเกิดจากการดูจำนวนคู่ซึ่งเป็นส่วนน้อย จึงเป็นข้อสรุปที่ไม่แม่นยำและดูไม่สมเหตุสมผลนัก **ดังนั้น ตารางที่จะใช้หาค่า Gamma จึงควรเป็นตารางที่มีผลรวมกระจายสม่ำเสมอกัน** และข้อควรระมัดระวังที่ควรสังเกต คือ การจัดลำดับตัวแปรทางแถวนอนและแถวตั้ง ต้องอยู่ในลักษณะมาก ⟶ น้อย หรือ น้อย⟶ มาก เหมือนกัน

1.2 Somer's d Statistics

เป็นค่าสถิติที่พยายามเอาคู่ซึ่งผูกกับตัวแปรตามเข้ามาเป็นตัวหารในสูตรด้วยเช่น ข้อมูลจากตารางที่ 8 ในเรื่องการศึกษาและรายได้ถ้าสมมติให้การศึกษาเป็นตัวแปรตามค่า d จะได้จากสูตร

$$d_c = \frac{C-D}{C+D+T_c}$$

$$= \frac{1800-525}{1800+525+100}$$

$$= \frac{1275}{3325} = 0.383$$

เนื่องจากตัวหารมีค่ามากกว่าของ Gamma จึงทำให้ค่า d น้อยกว่า Gamma เสมอ แต่ก็เป็นการแก้จุดอ่อนของ Gamma ที่ไม่ได้ใช้จำนวนคู่ซึ่งผูกอยู่กับตัวแปรใดตัวแปรหนึ่งเลย จึงทำให้ค่า Gamma สูงเกินเหตุเพราะดูแต่จำนวนคู่ที่ไม่ผูกกับตัวแปรใดๆ จำนวนคู่ที่ไม่ผูกนี้ถ้ามีน้อย ก็ทำให้เกิดความผิดพลาดในการทำนายมาก

ดังนั้น ค่า somer's d จึงเป็นสถิติที่ดีกว่าค่า Gamma แต่ก็มีข้อจำกัดว่าเหมาะสมเฉพาะกับตารางที่จำนวนเซลของตัวแปรตามที่จะต้องมีขนาดอย่างน้อยที่สุด เท่ากับจำนวนช่วงของตัวแปรอิสระเท่านั้น จึงจะได้ค่า Somer's d ที่สูงสุด

ค่า Somer's d จะอยู่ระหว่าง ± 1 การตีความหมายก็ทำเช่นเดียวกับค่า Gammaและเนื่องจากการนำเอาคู่ที่ผูกอยู่กับตัวแปรตามเข้ามาเกี่ยวข้องด้วย **ค่า Somer's d จึงเป็นสถิติแบบมีทิศทางเช่นเดียวกับค่า Lambda**

นั่นคือ Somer's d สามารถหาค่าที่เอาตัวแปรตัวหนึ่งเป็นตัวแปรตามได้ จากตารางที่ 8 เดิม ถ้าให้รายได้เป็นตัวแปรตาม ค่า d หาได้จากสูตร

$$d_r = \frac{C-D}{C+D+T_r}$$

$$= \frac{1800 - 525}{1800+525+1000}$$

$$= 0.383$$

ในขณะเดียวกัน Somer's d แบบไม่มีทิศทางก็หาค่าได้เช่นกัน

$$d_{rc} = \frac{(C+D+T_r)d_r + (C+D+T_c)d_c}{(C+D+T_r)+(C+D+T_c)}$$

ค่า d_{rc} จะเหมาะสมมาก ถ้าตารางไม่มีค่า 0 ในเซลใดโดยอย่างน้อยที่สุดเป็นจำนวนสองแถวตั้งและสองแถวนอน และค่า d_{rc} จะมีค่าอยู่ระหว่าง d_r และ d_c

2. การทดสอบของวิลคอกซัน แมนท์วิทนีย์

(The Wilcoxon-Mann-Whitney Test or The Wilcoxon Rank Sum Test)

เป็นสถิติที่ใช้ทดสอบค่ากลาง(นิยมใช้ค่ามัธยฐาน)ของสองประชากรที่เป็นอิสระกัน

บางครั้งเรียกว่า Mann-Whitney U test หรือ Mann-Whitney- Wilcoxon test โดย Wilcoxon ได้ศึกษากรณีใช้ผลรวมลำดับที่ (rank sum) เป็นตัวสถิติทดสอบโดยที่ Mann และ Whitney ได้ชี้ถึงความสัมพันธ์ระหว่างตัวสถิติทดสอบที่เขาตั้งขึ้นกับของ Wilcoxon การทดสอบนี้นับได้ว่าเป็นการทดสอบที่มีประสิทธิภาพมากที่สุด มักนิยมใช้เพื่อเลี่ยงการใช้การทดสอบแบบที (The t-test) ในสถิติที่ใช้พารามิเตอร์ หรือเมื่อข้อมูลมีมาตราวัดต่ำกว่าแบบอันตรภาค

ข้อกำหนดเบื้องต้น

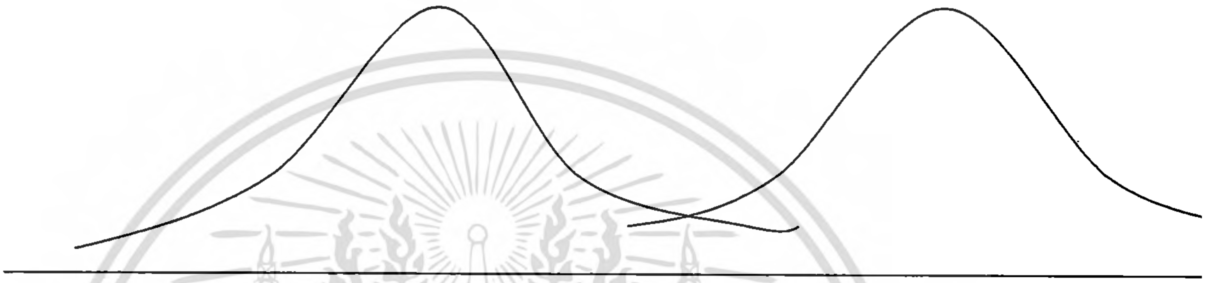
1. ข้อมูลประกอบด้วยตัวอย่างสุ่มด้วยค่า $X_1, X_2 \ldots X_{n1}$ จากประชากรที่ 1 และตัวอย่างสุ่มอีก 1 ชุด ด้วยค่าสังเกต $Y_1, Y_2 \ldots Y_{n2}$ จากประชากรที่ 2 ซึ่งเป็นอิสระกัน

2. ตัวอย่าง 2 ชุดนี้เป็นอิสระกัน

3.ค่าตัวแปรสุ่มมีค่าต่อเนื่อง (continuous)

4.มาตรวัดอย่างน้อยเป็นแบบเรียงลำดับ (ordinal scale)

5. ฟังก์ชันการแจกแจง ของ 2ประชากร ต่างกันเฉพาะค่ากลาง (ซึ่งนิยมวัดด้วยมัธยฐาน, M_x, M_y) นั่นคือ ประชากรทั้ง 2 ต้องมีการแจกแจงที่เหมือนกัน ต่างกันเฉพาะค่ากลางเท่านั้น เช่น สมมติว่าประชากรทั้ง 2 มีการแจกแจงปกติ ดังนั้น อาจเขียนรูปโค้งการแจกแจง ดังนี้



หมายความว่ารูปโค้งการแจกแจงของ 2 ประชากรต้องเหมือนกัน (ในแง่การกระจาย) ต่างกันเฉพาะค่ากลางเท่านั้น

หมายเหตุ ในทางปฏิบัติไม่จำเป็นต้องทราบว่ามีการแจกแจงแบบใด

สมมติฐาน ถ้าให้ M_xและ M_y แทนค่ามัธยฐานของประชากรที่ 1 และ 2 ตามลำดับ อาจทำการทดสอบสองหางหรือหางเดียว ได้ดังนี้

$$H_0 : M_x = M_y$$
$$H_1 : M_x \neq M_y$$

หรือ

$$H_0 : M_x \geq M_y$$
$$H_1 : M_x < M_y$$

หรือ

$$H_0 : M_x \leq M_y$$
$$H_1 : M_x > M_y$$

สถิติที่ใช้ทดสอบ ในที่นี้จะเสนอวิธีการของ wilcoxon(1945) และ Mann, Whitney (1947) ซึ่งต่างก็เสนอวิธีการทดสอบของตนเอง และในที่สุดสามารถหาความสัมพันธ์ของทั้ง 2วิธีดังต่อไปนี้

1.วิธีการของ wilcoxon ได้ใช้แนวคิดคล้ายการทดสอบของ wilcoxon signed Rank test คือใช้ผลรวมของลำดับที่ (sum of rank or rank sum) ของตัวอย่างชุดหนึ่งในข้อมูลรวมทั้งหมด (n_1+n_2จำนวน) ที่ได้เรียงลำดับจากน้อยไปหามาก โดยคาดว่าถ้า H_0 เป็นจริง ในข้อมูลรวมทั้งหมดนั้นค่าลำดับที่ของตัวอย่างชุดหนึ่งควรจะมีคละกันไปทั้งค่าน้อย ปานกลาง และมาก ซึ่งจะทำให้ได้ผลรวมลำดับที่ ค่าหนึ่งที่ไม่มากเกินไปหรือน้อยเกินไป แต่ถ้า H_1 เป็นจริง ค่าผลรวมของลำดับที่จากตัวอย่างชุดหนึ่งจะมีค่ามาก หรือน้อยเกินไปดังตัวอย่างต่อไปนี้

ถ้ามีตัวอย่างสุ่มขนาด 4 ด้วยค่าตัวแปรสุ่ม X และอีกชุดหนึ่งด้วยขนาด 5 ด้วยตัวแปรสุ่ม Y ปรากฏว่าเมื่อนำทั้ง 9 จำนวนรวมกัน และเรียงลำดับ

และให้ S = ผลรวมของลำดับที่ของข้อมูล Xในข้อมูลรวมทั้งหมด

$$= \sum_{i=1}^{n_1} \text{Rank}(X_i)$$

ถ้าข้อมูลรวมทั้งหมด เมื่อนำมาเรียงลำดับแล้วได้ลำดับที่ดังนี้

ชุดที่ 1 YYYYYXXXX กรณีนี้จะได้ค่า S = 6+7+8+9 = 30

หรือ ชุดที่ 2 XXXXYYYYY กรณีนี้จะได้ค่า S = 1+2+3+4 = 10

หรือ ชุดที่3 XYXYXYXYY กรณีนี้จะได้ค่า S = 1+3+5+7 = 16

จะพบว่าในตัวอย่างรวมชุดที่ 1 ตัวแปร X อยู่ในตอนท้ายได้ค่า S = 30 มีค่าใหญ่มากในตัวอย่างชุดนี้ น่าจะคาดเดาว่า ประชากรกลุ่ม X มีแนวโน้มที่จะมีค่ามากกว่ากลุ่มY

ในตัวอย่างรวมชุดที่ 2 ตัวแปร X อยู่ในตอนต้น ได้ค่า S = 10 มีค่าน้อย ดังนั้น น่าจะทำให้ยอมรับ H_1 : ประชากร X มีแนวโน้มที่จะมีค่าน้อยกว่า Y

และในตัวอย่างรวมชุดที่ 3 ตัวแปร X อยู่ในลักษณะผสม (mix) กันอย่างดีกับ Y ทำให้มีลำดับที่ ทั้งค่าน้อย ปานกลาง และมาก ได้ค่า S = 16 ซึ่งมีค่าปานกลางจากตัวอย่างนี้ น่าจะทำให้เรายอมรับ H_0 :ประชากร X และYมีค่ามัธยฐานไม่ต่างกัน

Wilcoxon ได้สร้างตารางแสดงค่าความน่าจะเป็นของค่า S ที่น้อยหรือมากเกินไปนี้ ซึ่งสามารถใช้ตารางดังกล่าวหาค่า p – value เพื่อตัดสินใจยอมรับ H_0 หรือปฏิเสธ H_0ได้ แต่เนื่องจากค่า S ที่เล็กที่สุดจะแตกต่างกันไปตามขนาดตัวอย่างที่สุ่มมา จึงทำให้การสร้างตารางยากขึ้น และค่อนข้างใหญ่ ทำให้ไม่สะดวกในการใช้ ในที่นี้จึงไม่เสนอวิธีการของ Wilcoxonโดยตรงนี้ แต่จะปรับสูตรสถิติที่ใช้ทดสอบให้สัมพันธ์กับค่า S นี้ และสอดคล้องกับวิธีของ Mann, Whitney ซึ่งจะได้เสนอในลำดับต่อไป

ในกรณีตัวอย่างขนาดใหญ่ สามารถประมาณการแจกแจงของ S ด้วยการแจกแจงปกติ ดังสูตร

$$Z = \frac{(S \pm 0.5) - n_1(n_1 + n_2 + 1)/2}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

2.**วิธีการของ Mann, Whitney** มักเรียกชื่อการทดสอบ ของเขาทั้งสองว่า Mann-Whitney U test ซึ่งกำหนดให้ตัวสถิติ U คือ การนับจำนวนค่าสังเกตในตัวอย่างชุดหนึ่งที่นำหน้า (exceeding) แต่ละค่าสังเกตในตัวอย่างอีกชุดหนึ่งในข้อมูลที่นำมารวมกันและเรียงลำดับ การคำนวณหาค่า U สามารถทำได้ง่ายไม่จำเป็นต้องใช้คอมพิวเตอร์ และวิธีการนี้ยังเป็นพื้นฐานในการหาช่วงความเชื่อมั่นของผลต่างค่ามัธยฐานใน 2 ประชากรด้วย

ให้ $U = \sum_{i=1}^{n_1} U_i$

= ผลรวม (จำนวนค่า Y ที่น้อยกว่าหรือนำหน้าX_i ในข้อมูลรวมทั้งหมดที่เรียงลำดับแล้ว) เมื่อ $i = 1, 2 \ldots n_1$

เช่น มีข้อมูลรวม YYYYYXXXX จะได้ U = 5+5+5+5 = 20

XXXXYYYYY จะได้ U = 0

XYXYXYXYY จะได้ U = 0+1+2+3 =6

จะเห็นว่า ค่าU ที่ใหญ่เกินไปหรือน้อยเกินไปทำให้น่าจะเชื่อว่า H_1เป็นจริง ในขณะที่U ที่มีค่าปานกลางจะทำให้เชื่อว่า H_0เป็นจริง ซึ่งจะสอดคล้องกับค่า S ของ Wilcoxon

นอกจากการนับจำนวนเพื่อหาค่า U แล้ว อาจใช้สูตรหาค่า Uดังนี้

$$U = n_1n_2 + \frac{n_2(n_2 +_1)}{2} - S_2$$

เมื่อ S_2 = ผลรวมลำดับที่ของตัวแปร Y จากตัวอย่างขนาด n_2 ลองพิจารณาตัวอย่างต่อไปนี้ ถ้ามีตัวอย่างสุ่มชุดที่ 1 และ 2 ด้วยค่าสังเกต ดังนี้

ข้อมูล X : 110 70 53 51

ข้อมูล Y : 78 64 75 45 82

รวมทั้งหมดเข้าด้วยกัน และเรียงลำดับจะได้

45	51	53	64	70	75	78	82	110
Y	X	X	Y	X	Y	Y	Y	X

หาค่า U = 1+1+2+5 = 9

ถ้าใช้สูตร $U = (4 \times 5) + \dfrac{5(5+1)}{2} - (1+4+6+7+8)$

$= 9$ เท่ากันกับวิธีนับข้างต้น

Mann-Whitney ได้สร้างตารางค่าความน่าจะเป็นเมื่อ u มีค่าต่างๆ ที่ค่า n_1, n_2 ต่างๆ กัน แต่การใช้ตารางจำเป็นต้องเลือกใช้ค่า U ที่มีค่าน้อยที่สุด เพราะค่าความน่าจะเป็นที่คำนวณในตารางเป็นความน่าจะเป็นด้านซ้ายของโค้งการแจกแจง

การเลือกใช้ ค่า Uที่น้อยที่สุด ให้ใช้ความสัมพันธ์

$U' = n_1 n_2 - U$ แล้วเลือกค่า U' หรือ U ที่เล็กที่สุด ในกรณีทดสอบ 2 หาง หรือ หางเดียวด้านซ้าย

เช่นตัวอย่างข้างต้น $U' = (4 \times 5) - 9 = 11$ ดังนั้นเลือกใช้ $U = 9$ เป็นสถิติที่ใช้ทดสอบแต่การสร้างตารางการแจกแจงของค่า U จะง่ายขึ้น เนื่องจากค่าต่ำสุดของ $U = 0$ เสมอ ในกรณีตัวอย่างใหญ่สามารถประมาณการแจกแจงค่า U ด้วยการแจกแจงปกติมาตรฐานได้ค่า

$$Z = \frac{U - n_1 n_2/2}{\sqrt{n_1 n_2 (n_1 + n_2 + n_1)/12}}$$

3. **วิธีการของ wilcoxon และ Mann-Whitney** Mann-Whitney ได้แสดงความสัมพันธ์ระหว่างสถิติทีใช้ทดสอบของเขากับของ wilcoxon พบว่า

ถ้าให้ $T = S - \dfrac{n_1(n_1+1)}{2}$ แล้วค่าT ที่ได้จะมีค่าเท่ากับค่าUนั่นเอง

หลักในการหาอาณาเขตวิกฤตยังคงคล้ายการพิจารณาค่า S เนื่องจากค่า T มีความสัมพันธ์ในเชิงเส้นตรงกับค่า S ดังนั้น ค่า T ที่มากเกินไปหรือน้อยเกินไปจะทำให้ปฏิเสธ H_0 เพื่อยอมรับ H_1 แต่การสร้างตารางการแจกแจงความน่าจะเป็นของค่า T จะง่ายขึ้น เนื่องจากค่าเล็กที่สุดของ $T = 0$ เสมอ

ดังนั้น สถิติที่ใช้ทดสอบ คือ $T = S - \dfrac{n_1(n_1+1)}{2}$

เมื่อ S = ผลรวมลำดับที่ของตัวอย่าง ขนาด n1 ในข้อมูลรวมที่เรียงลำดับแล้ว

การตัดสินใจ ใช้ตารางแสดงค่าวิกฤตของสถิติทีใช้ทดสอบ T

ในกรณีการทดสอบ**สองหาง**จะปฏิเสธ H_0 ถ้าพบว่าค่า T น้อยเกินไปหรือใหญ่เกินไป

อาณาเขตวิกฤต คือ $T < W_{\alpha/2}$ หรือ $T > W_{1-\alpha/2}$

เมื่อ $W_{1-\alpha/2} = n_1 n_2 - W_{\alpha/2}$

เมื่อเป็นการทดสอบ**หางเดียว ด้านน้อยกว่า** $H_1 : M_x < M_y$ จะปฏิเสธ H_0 เมื่อพบว่าค่า T น้อยเกินไป

อาณาเขตวิกฤต คือ $T < W_{\alpha}$

เมื่อเป็นการทดสอบ**หางเดียว ด้านมากกว่า** $H_1 : M_x > M_y$ จะปฏิเสธ H_0 เมื่อพบว่าค่า T ใหญ่เกินไป

อาณาเขตวิกฤต คือ $T > W_{1-\alpha}$ เมื่อ $W_{1-\alpha} = n_1 n_2 - W_{\alpha}$

ตัวอย่างที่10 เพื่อทดสอบความแข็งของเหล็กที่ได้จาก 2 แหล่ง คือ A และ B ได้สุ่มตัวอย่างเหล็กจากแหล่งA มา 4 ชิ้น และจากแหล่ง B มา 5 ชิ้น แล้วนำ 2 ชิ้นมาขัดถูกัน พิจารณาดูว่าชิ้นใดมีร่องรอยเสียหายมากกว่า ให้เป็นชิ้นที่มีความแข็งน้อยกว่า ทำเช่นนี้กับตัวอย่างทั้ง 9 ชิ้น แล้วให้ลำดับที่ 1 แก่ชิ้นที่แข็งน้อยที่สุด จนถึงอันดับที่ 9 คือชิ้นที่แข็งที่สุด

ได้ข้อมูลผลการทดลองดังนี้	A	A	A	B	A	B	B	B	B
ลำดับที่	1	2	3	4	5	6	7	8	9

จงทดสอบสมมติฐานว่า เหล็กจากแหล่งทั้ง 2 มีความแข็งไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05

วิธีทำ ใช้การทดสอบของ wilcoxon และMann Whitney เนื่องจากเป็นกรณี 2 กลุ่มตัวอย่างที่เป็นอิสระกัน และมีมาตรวัดข้อมูลแบบเรียงลำดับ พิจารณาในแง่ค่ากลาง คือ

H_0 : ค่ามัธยฐานของความแข็งของเหล็กจาก 2 แหล่งไม่แตกต่างกัน

H_1 : ค่ามัธยฐานของความแข็งของเหล็กจาก 2 แหล่งแตกต่างกัน

หรือ $H_0 : M_x = M_y$

$H_1 : M_x \neq M_y$

หาสถิติที่ใช้ทดสอบ $T = S - \frac{n_1(n_1+1)}{2}$

หา S จากแหล่ง $A = 1 + 2 + 3 + 5 = 11$

$$T = 11 - \frac{4(5)}{2}$$

$$= 1$$

หาค่าวิกฤตจากตาราง เมื่อเป็นการทดสอบสองหางที่ $\alpha = 0.05$ ดังนั้นหาค่าวิกฤตที่ค่า $P = 0.025$ $n_1 = 4$ และ $n_2 = 5$ ได้ค่าวิกฤต = 2

ดังนั้น $W_{\alpha/2} = 2$

หาค่า $W_{1-\alpha/2} = n_1n_2 - W_{\alpha/2} = (4 \times 5) - 2 = 18$

อาณาเขตวิกฤต คือ $T < 2$ หรือ $T > 18$

T ข้อมูลตัวอย่าง = 1 จึงตกในอาณาเขตวิกฤต จึงปฏิเสธ H_0

สรุปว่า เหล็กจาก 2 แหล่งมีความแข็งไม่เท่ากันหรือ ต่างกัน

กรณี ties ถ้ามีค่าสังเกตบางค่ามีค่าเท่ากัน วิธีการลำดับที่ให้ค่าเฉลี่ยของลำดับที่ควรจะเป็นจริง ดังจะแสดงต่อไป

กรณีตัวอย่างใหญ่ กรณี n_1 หรือ $n_2 > 20$ สามารถประมาณการแจกแจงของสถิติ T ได้ด้วยการแจกแจงปกติ ด้วยตัวสถิติทดสอบ Z ดังนี้

$$Z = \frac{T - n_1n_2/12}{\sqrt{n_1n_2(n_1 + n_2 + 1)/12}} \sim N(0,1)$$

การหาอาณาเขตวิกฤตให้หาจากโค้งการแจกแจงปกติมาตรฐาน

ตัวอย่างที่ 11 จากเด็กนักเรียนในชั้นมัธยมแห่งหนึ่งสุ่มตัวอย่างเด็กที่มาจากต่างจังหวัด 12 คน และเด็กในเขตกรุงเทพมหานคร 36 คน ต้องการทราบว่าเด็กจากต่างจังหวัดมีร่างกายที่แข็งแรงกว่าเด็กในเขต กทม.หรือไม่ จึงทำการทดสอบความแข็งแรงได้ข้อมูลดังนี้ ถ้าคะแนนที่ได้ต่ำแสดงถึงความแข็งแรงน้อย และคะแนนที่ได้สูงแสดงว่ามีความแข็งแรงมาก

X_i เด็กจากต่างจังหวัด		Yi เด็กในเขต กทม.					
14.8	10.6	12.7	16.9	7.6	2.4	6.2	9.9
7.3	12.5	14.2	7.9	11.3	6.4	6.1	10.6
5.6	12.9	12.6	16.0	8.3	9.1	15.3	14.8
6.3	16.1	2.1	10.6	6.7	6.7	10.6	5.0
9.0	11.4	17.7	5.6	3.6	18.6	1.8	2.6
4.2	2.7	11.8	5.6	1.0	3.2	5.9	4.0

วิธีทำ เลือกใช้ wilcoxon-Mann-Whitney test

$H_0 : M_x \leq M_y$

$H_1 : M_x > M_y$ รวมข้อมูลทั้งหมดเข้าด้วยกัน แล้วเรียงลำดับ

X	Y	ลำดับที่	X	Y	ลำดับที่	X	Y	ลำดับที่
	1.0	1		6.2	17		11.3	33
	1.8	2	6.3		18	11.4		34
	2.1	3		6.4	19		11.8	35
	2.4	4		6.7	20.5	12.5		36
	2.6	5		6.7	20.5		12.6	37
2.7		6	7.3		22		12.7	38
	3.2	7		7.6	23	12.9		39
	3.6	8		7.9	24		14.2	40
	4.0	9		8.3	25		14.8	41.5
4.2		10	9.0		26	14.8		41.5
	5.0	11		9.1	27		15.3	43
	5.6	13		9.9	28		16.0	44
	5.6	13		10.6	30.5	16.1		45
5.6		13		10.6	30.5		16.9	46
	5.9	15	10.6		30.5		17.7	47
	6.1	16		10.6	30.5		18.6	48

มี ties เกิดขึ้น ค่าลำดับทีให้ จะคือค่าเฉลี่ยของลำดับที่ควรจะเป็นเช่น ค่า 5.6 มี 3 ค่า ลำดับที่ควรจะเป็นคือ 12 13 และ 14 ดังนั้นค่าเฉลี่ยของทั้ง 3 ค่าคือ 13 จึงให้ลำดับที่ของค่า 5.6 ทั้ง 3 ค่า = 13 (ได้แสดงค่า ties ด้วยเครื่องหมายวงเล็บ)

หาค่า S = ผลรวมลำดับที่ของ X

$= 6+10+13+18+22+26+30.5+34+36+39+41.5+45$

$$= 321$$

ดังนั้น $$T = S - \frac{n_1(n_1+1)}{2}$$

$$= 321 - \frac{12(12+1)}{2}$$

$$= 243$$

ด้วยการทดสอบหางเดียวทางด้านขวา ที่ระดับนัยสำคัญ 0.05 หาค่าวิกฤตด้านขวาจึงได้อาณาเขตวิกฤต คือ $Z > 1.645$ (เมื่อเป็นกรณีตัวอย่างขนาดใหญ่) จากข้อมูลตัวอย่าง คำนวณสถิติทดสอบ

$$Z = \frac{243-(12)(36)/2}{\sqrt{(12)(36)(12+36+1)/12}}$$

$$= 27/42$$

$$= 0.64$$

Z ไม่ตกในอาณาเขตวิกฤต จึงไม่สามารถปฏิเสธ H_0

นั่นคือ เด็กจากต่างจังหวัดไม่แข็งแรงกว่าเด็กในเขต กทม.

2.1 การทดสอบของ Wilcoxon- Mann- Whitney แบบมีซ้ำมาก

(The Wilcoxon-Mann-Whitney Test with Ties)

ถ้าค่าข้อมูลที่เป็นลำดับที่ในกลุ่มใดๆ 2 กลุ่มนั้นมีค่า ซ้ำกันมาก อาจจัดข้อมูลใหม่เป็นตารางแจกแจง 2 ทาง โดยแถวนอนมี 2 แถว หมายถึง 2 กลุ่มตัวอย่างที่อิสระกัน และแถวตั้งจัดเป็นค่าลำดับที่ และนับความถี่ลงเซลต่างๆ ซึ่งตารางนี้จะมีลักษณะแบบเรียงลำดับทางแถวตั้ง อาจ**เรียกชื่อว่าเป็น Singly ordered contingency table** เช่นมีข้อมูลดังต่อไปนี้

สุ่มตัวอย่างคนไข้ 4 คน ให้ใช้ยา A และคนไข้ 3 คนใช้ยา B หลังจากเวลาผ่านไประยะหนึ่งประเมินผลการรักษา โดยให้ แพทย์ประเมินในระดับความพอใจ จาก 1 – 3 (1 น้อยที่สุด) ได้ข้อมูลดังนี้

ยา A	1 1 2 2
ยา B	2 3 3

n_A = 4 จำนวน

n_B = 3 จำนวน

จะพบว่าค่าข้อมูลซึ่งมีค่า 1 – 3 มีค่าซ้ำกันมาก การให้ลำดับที่ต้องใช้ค่าเฉลี่ย ของลำดับที่

โดย ค่า 1 มีสองค่า ได้ลำดับที่เฉลี่ยของลำดับ 1 และ 2 คือ 1.5

ค่า 2 มีสามค่า ได้ลำดับที่เฉลี่ยของลำดับ 3, 4 และ 5 คือ 4

ค่า 3 มีสองค่า ได้ลำดับที่เฉลี่ยของลำดับ 6 และ 7 คือ 6.5

จะวิเคราะห์ข้อมูลแบบตารางที่เรียกว่า Singly ordered contingency table ด้วยสถิติ Wilcoxon – Mann – Whitney แบบปรับ ties ดังนี้

จัดข้อมูลลงตารางการณ์จร 2 ทาง และบันทึกความถี่ จะได้ตารางการณ์จรดังนี้

	ผลการรักษาตามลำดับความพอใจ			ผลรวม
	1	2	3	
ยา A	2	2	0	4 = n_A
ยา B	0	1	2	3 = n_B
				n = 7

ถ้าพิจารณาตารางใหม่นี้ จะพบว่าสามารถใช้สถิติ χ^2 วิเคราะห์ได้ เพราะเป็นตารางการณ์จร แต่เนื่องจากขนาดตัวอย่างเล็กมาก รวมทั้งความถี่คาดหวังจะมีค่าน้อยกว่า 5 มากเกิน 20 % ของจำนวนเซลทั้งทั้งหมด การใช้ χ^2 จึงไม่เหมาะสม (ได้ค่า p – value = 0.31 ซึ่งก็ไม่พบนัยสำคัญระหว่างยาทั้งสอง)

เราสามารถใช้สถิติทดสอบของ Wilcoxon – Mann – Whitney แบบมี ties มากๆ ดังนี้

จากสูตร $$T = S - \frac{n_1(n_1+1)}{2}$$

เมื่อ S = ผลรวมลำดับที่ของยา A = 1.5+1.5+4+4 = 11

ดังนั้น $$T = 11 - \frac{4(5)}{2} = 1$$ (หรือ T = 11 เมื่อคิดจากกลุ่ม B)

ซึ่งได้ค่า p – value = 0.0857 พบว่ามีนัยสำคัญที่ระดับ $\alpha = 0.10$ ซึ่งดีกว่าใช้ χ^2

เมื่อ H_0 : ยา A และ ยา B ให้การรักษา ไม่ต่างกัน

H_1 : ยา A และ ยา B ให้การรักษา ต่างกัน

หรือ H_0 : ไม่มีความสัมพันธ์กันระหว่างยาที่ใช้กับผลการรักษา

H_1 : มีความสัมพันธ์กันระหว่างยาที่ใช้กับผลการรักษา

ตารางที่ปรับเป็นตารางการณ์จรข้างต้น จะมีเพียงทางเดียวที่เรียงลำดับ (อาจจะเป็นแนวนอน หรือ แถวตั้งก็ได้) จึงเรียกว่า **Singly ordered contingency table** และอาจจะขยายเป็น 3 แถวนอน (หรือแถวตั้ง) ในแถวที่ไม่ได้เรียงลำดับ เช่นเป็น กรณี 3 กลุ่มตัวอย่างหรือ 4 กลุ่มตัวอย่าง และสามารถใช้สถิติทดสอบที่ใช้กับมากกว่า 2 ประชากร ได้เช่นกัน โดยใช้หลักการคำนวณเช่นเดียวกันนี้ ซึ่งจะได้กล่าวถึงในหัวข้อการทดสอบของ Kruskal – Wallis ต่อไป

3. การทดสอบของครัสคาลและวัลลิส

(The Kruskal-Wallis One-Way Analysis of Variance By Rank Test)

ใช้ทดสอบว่า **ประชากร k กลุ่มที่เป็นอิสระกันมีค่ามัธยฐานเท่ากันหรือไม่** โดยมีวิธีการที่สำคัญคือ ค่าคาดหมายของลำดับที่ของข้อมูลตัวอย่างแต่ละกลุ่ม ควรมีค่าพอๆ กัน ข้อมูลที่นำมาทดสอบประกอบด้วยข้อมูลจากตัวอย่างสุ่ม k ชุด แต่ละชุดอาจมีขนาดตัวอย่างแตกต่างกัน ข้อมูลที่จะใช้วิเคราะห์ต้องมีมาตราวัดอย่างน้อยเป็นแบบเรียงลำดับ (Ordinal Scale) และมีการแจกแจงแบบต่อเนื่อง

การทดสอบนี้นิยมใช้แทนการทดสอบแบบเอฟ (F-test) ในสถิติที่ใช้พารามิเตอร์ในกรณีที่ข้อกำหนดเบื้องต้นของการทดสอบเอฟ ไม่เป็นจริง

สมมติฐาน

H_0 : ค่ามัธยฐานของประชากร k กลุ่มไม่แตกต่างกัน

H_1 : ค่ามัธยฐานของประชากรอย่างน้อย 1 คู่ แตกต่างกัน

วิธีการ อาจสรุปขั้นตอนการทดสอบได้ดังนี้

1. จัดลำดับของข้อมูลทั้งหมดร่วมกัน จากน้อยไปหามากโดยให้คะแนนต่ำสุดมีลำดับที่ 1 และคะแนนสูงสุดเป็นลำดับที่ n เมื่อ n เป็นจำนวนข้อมูลทั้งหมด

2. หาผลรวมของลำดับที่ในข้อมูลแต่ละชุด คือ R_i i = 1, 2, …k

3.คำนวณค่าสถิติ

$$H = \left[\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1)$$

เมื่อ k = จำนวนประชากรที่เป็นอิสระกัน

R_i = ผลรวมของลำดับที่ในตัวอย่างที่ i, i = 1, …k

n_i = ขนาดตัวอย่างชุดที่ i, i = 1, …k

$$n = \sum_{i=1}^{k} n_i$$

ตามทฤษฎีจะพบว่าH_0 เป็นจริง H จะมีการแจกแจงประมาณได้ด้วย χ^2 ที่ d.f. = k – 1 ถ้า n_i **มีค่าใหญ่พอสมควร**

4.การหาอาณาเขตวิกฤตและสรุปผล สามารถแยกได้ตามขนาดตัวอย่าง คือ

4.1 เมื่อ $n_i > 5$ การแจกแจงของค่าสถิติ H ประมาณด้วย χ^2 ที่ d.f. = k – 1เมื่อกำหนดระดับนัยสำคัญ = α หาอาณาเขตวิกฤต จากตาราง χ^2 ที่ d.f. = k – 1จะปฏิเสธ H_0 เมื่อค่าของ H มากกว่าหรือเท่ากับ χ^2 จากตาราง

4.2 เมื่อ k =3 และ $n_i \leq 5$ในแต่ละ k ใช้ตารางที่ kruskal สร้างไว้ โดยแสดงค่าวิกฤตของ H พร้อมทั้งความน่าจะเป็นที่จะเกิดค่า H นั้นๆ สามารถจะทำการเปรียบเทียบค่า H หรือ p – value ก็ได้ คือจะปฏิเสธH_0 เมื่อค่าH จากตัวอย่างมากกว่าหรือเท่ากับค่า H จากตารางที่ระดับนัยสำคัญα

ตัวอย่างที่ 12 กำหนดให้มีวิธีการสอน 3 วิธีคือ ก ข และ ค นำไปใช้สอนวิชาคณิตศาสตร์ให้กับนักศึกษาจากตัวอย่างนักศึกษาที่มีพื้นฐานความรู้ใกล้เคียงกันทั้งหมด 13 คน ถูกสุ่มให้เรียนด้วยวิธีทั้ง 3 ด้วยจำนวน4, 4 และ 5 คนตามลำดับ ทำการทดสอบได้ผลดังนี้

คะแนนจากวิธี ก	คะแนนจากวิธี ข	คะแนนจากวิธี ค
82	71	91
80	79	93
81	78	84
83	74	90
		88

วิธีการสอนทั้ง 3 วิธี ว่าจะให้ผลแตกต่างกันหรือไม่ ที่ระดับนัยสำคัญ 0.05

วิธีทำ จะเลือกใช้วิธีทดสอบของ Kruskal – Wallis เนื่องจากเป็นตัวอย่าง 3 กลุ่มที่เป็นอิสระกัน และข้อมูลมีมาตราวัดเป็นแบบอันตรภาค วิธีการทดสอบทำได้ ดังนี้

1.จัดลำดับของข้อมูลที่นำมารวมเป็นชุดเดียวกัน โดยจัดลำดับจากน้อยไปหามากได้ข้อมูลลำดับ ดังนี้

วิธี ก	วิธี ข	วิธี ค
7	1	12
5	4	13
6	3	9
8	2	11
		10
$R_1 = 26$	$R_2 = 10$	$R_3 = 55$

2.หาผลรวมของลำดับที่ในข้อมูลแต่ละชุดได้ค่า $R_1 = 26, R_2 = 10, R_3 = 55$

3.คำนวณค่าสถิติจากสูตร

$$H = \left[\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1)$$

$$= \left[\frac{12}{13(13+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - \left[\frac{26^2}{4}+\frac{10^2}{4}+\frac{55^2}{5}\right] - 3(13+1)$$

$= 10.68$

4. หาอาณาเขตวิกฤตและสรุปผล

เนื่องจาก $n_i \leq 5$ ฉะนั้นจะหาค่าวิกฤตH จากตารางของ kruskal ที่ระดับนัยสำคัญ$\alpha = .05$ จะได้ $H = 5.6176$

ซึ่ง H ที่คำนวณจากตัวอย่าง = 10.68 ที่มีค่ามากกว่า H จากตาราง ดังนั้นปฏิเสธ H_0 นั่นคือ ยอมรับ H_1 สรุปได้ว่าวิธีการสอนทั้ง 3 แบบให้ผลต่างกัน

ในกรณีที่หน่วยตัวอย่างมีลำดับที่ เท่ากัน จะกำหนดลำดับที่ให้เท่ากับค่าเฉลี่ยและให้

$$H = \frac{\left[\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1)}{1-\frac{\sum(t^3-t)}{n^3-n}}$$

เมื่อ t เป็นจำนวนหน่วยตัวอย่างที่มีค่าเท่ากันในแต่ละลำดับที่
ค่า H นี้จะมีการแจกแจงแบบไคสแควร์โดยประมาณ ด้วย d.f. = k – 1

ตัวอย่างที่ 13 ในการพิจารณากลุ่มอาชีพ 4 ประเภท เพื่อศึกษาระดับการครองชีพ (ซึ่งคิดจากการมีของใช้ชนิดต่างๆ ภายในบ้าน) ว่าแตกต่างกันหรือไม่ จึงสุ่มตัวอย่างคนจากอาชีพดังกล่าวมา 9, 7, 8, และ 7 คนตามลำดับบันทึกคะแนนซึ่งคิดจากการมีของใช้ชนิดต่างๆ ภายในบ้านได้ข้อมูล X_{ij} ดังตารางต่อไปนี้

เป็นเจ้าของกิจการ		ลูกจ้างเอกชน		ข้าราชการ		เกษตรกร	
X_{1j}	R_{1j}	X_{2j}	R_{2j}	X_{3j}	R_{3j}	X_{4j}	R_{4j}
42	13	39	9.5	27	2	33	6
65	30	41	12	32	4.5	28	3
52	24	50	22	44	15.5	22	1
44	15.5	55	25	65	30	36	7
50	22	46	18	39	9.5	44	15.5
39	9.5	62	27	48	19	49	20
63	28	44	15.5	56	26	39	9.5
65	30			50	22		
32	4.5						
รวม	176.5		129.0		128.5		62.0

วิธีทำ H_0: ระดับการครองชีพของกลุ่มอาชีพทั้ง 4 ไม่แตกต่างกัน

H_1: ระดับการครองชีพของกลุ่มอาชีพทั้ง 4 แตกต่างกันอย่างน้อย 1 คู่

เมื่อแปลงคะแนน X_{ij} ให้เป็นลำดับที่ R_{ij} จะพบว่ามีค่าที่ซ้ำกัน ดังนี้

1.ที่X_{ij} = 32 มีจำนวนซ้ำ = 2 จำนวน

2.ที่X_{ij} = 39 มีจำนวนซ้ำ = 4 จำนวน

3.ที่X_{ij} = 44 มีจำนวนซ้ำ = 4 จำนวน

4.ที่X_{ij} = 50 มีจำนวนซ้ำ = 3 จำนวน

5.ที่X_{ij} = 65 มีจำนวนซ้ำ = 3 จำนวน

ดังนั้นค่า t จะมีค่าดังนี้

$$t = 2, 4, 4, 3, 3$$

$$\sum (t^3 - t) = (2^3 - 2)+(4^3 - 4)+(4^3 - 4)+(3^3 - 3)+(3^3 - 3)$$

$$= 114$$

จาก
$$H = \frac{\left[\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(n+1)}{1 - \frac{\sum(t^3 - t)}{n^3 - n}}$$

$$H = \frac{\frac{12}{31(32)}\left[\frac{(176.5)^2}{9}+\frac{(129)^2}{7}+\frac{(128.5)^2}{8}+\frac{(62)^2}{7}\right] - 3(32)}{1 - \frac{114}{31^3 - 31}}$$

$$= 6.20$$

เมื่อ $\alpha = 0.01$ จากตารางการแจกแจงแบบ χ^2 ค่าวิกฤตคือ = 11.34

ดังนั้นค่า H ไม่ตกในอาณาเขตวิกฤต นั่นคือ ยอมรับ H_0 สรุปว่าระดับการครองชีพของกลุ่มอาชีพทั้งสี่ไม่แตกต่างกัน

3.1 การทดสอบ Kruskal-Wallis แบบมีซ้ำมาก

(The Kruskal-Wallis Test with Ties)

ในกรณีที่ค่าข้อมูลใน k กลุ่มอิสระกันนั้น มีค่าซ้ำกันมาก อาจจะซ้ำกันภายในกลุ่ม หรือระหว่างกลุ่มก็ได้ เช่นมีข้อมูลดังนี้

จากวิทยาลัย 3 แห่ง สุ่มตัวอย่างคณาจารย์ เพื่อประเมินการบริหารงานของผู้บริหารใน 5 ระดับคือ แย่ที่สุด ปานกลาง ดี ดีมาก ยอดเยี่ยม โดยให้มีค่าคะแนน คือ 1 – 5 ตามลำดับ ถ้าได้ข้อมูลด้วยขนาดตัวอย่างจาก 3 วิทยาลัย คือ 15, 10, และ 12 ดังนี้

วิทยาลัย A	1 1 1 1 1 1 2 2 2 2 3 3 3 4 4	$n_A = 15$
วิทยาลัย B	1 1 2 2 3 3 3 4 4 5	$n_B = 10$
วิทยาลัย C	1 2 2 3 3 3 4 4 4 4 5 5	$n_C = 12$

จากข้อมูลข้างต้นจะเห็นว่ามีค่าซ้ำกันมากในกลุ่มค่า 1– 5 ดังนั้น อาจจัดข้อมูลลงตาราง 2 ทาง จะได้ดังนี้

วิทยาลัย	ผลการประเมิน					
	1 แย่ที่สุด	2 ปานกลาง	3 ดี	4 ดีมาก	5 ยอดเยี่ยม	
A	6	4	3	2	0	$n_A = 15$
B	2	2	3	2	1	$n_B = 10$
C	1	2	3	4	2	$n_C = 12$

จะพบว่า ตัวแปรทางแถวตั้ง มีลักษณะเรียงลำดับ

จึงเรียกว่า ตารางแบบ Singly Ordered Table

ถ้าทำการวิเคราะห์ความสัมพันธ์ด้วยสถิติทดสอบ χ^2 จะได้ค่า $\chi^2_{\text{cal.}} = 7.02$ ด้วย p-value = 0.53 ยอมรับ H_0 แต่ถ้าสังเกตจะพบว่าค่าความถี่ในเซลต่างๆ มีค่าน้อยมาก และมีค่า 0 รวมอยู่ด้วย ดังนั้นการใช้สถิติ χ^2 ไม่น่าเหมาะสมสำหรับข้อมูลชุดนี้ (เนื่องจากจะเกิดความถี่คาดหวังที่น้อยกว่า 5 มีมากเกิน 20 % ของจำนวนเซลทั้งหมด)

สามารถจะใช้สถิติของ Kruskal wallis แบบมี ties วิเคราะห์ข้อมูลชุดนี้แทน ซึ่งจะได้ H = 6.43 ด้วย p-value = 0.04 ซึ่งพบว่าปฏิเสธH_0 ยอมรับH_1 (ในขณะที่χ^2 ไม่พบความแตกต่าง)

เมื่อ H_0 : ผู้บริหารวิทยาลัยทั้ง 3 ได้รับผลการประเมินไม่ต่างกัน

H_1 : ผู้บริหารวิทยาลัยทั้ง 3 ได้รับผลการประเมินต่างกัน

หรือ H_0 : ผลการประเมินที่ได้รับ เป็นอิสระกับผู้บริหารในแต่ละวิทยาลัย

H_1 : ผลการประเมินที่ได้รับ มีความสัมพันธ์กับผู้บริหารในแต่ละวิทยาลัย

จะแสดงวิธีการคำนวนค่า H แบบปรับ ties ด้วยตัวอย่างดังต่อไปนี้

ในการเปรียบเทียบค่ากลางของผลการสอนจากอาจารย์ 3 ท่าน (A, B, C) ในวิชาหนึ่ง ได้สุ่มตัวอย่างจากนักศึกษา 3 กลุ่มนี้ เป็นจำนวน 20, 10 และ 20 คนตามลำดับ ถ้าผลการสอบให้เป็น 3 ระดับดังนี้ สอบผ่าน (p) สอบตก (F) และสอบผ่านอย่างมีเงื่อนไข (C) ได้ผลการบันทึกข้อมูลดังนี้

กลุ่ม	A	B	C
	F	F	F
	F	F	F

F	F	F
F	F	P
F	P	P
P	P	P
P	P	P
.	.	.
.	.	.
.	.	.
20 คน	10 คน	20 คน

เพื่อให้เห็นภาพของข้อมูลชัดเจนยิ่งขึ้น จึงจัดข้อมูลข้างต้นลงตาราง 2 ทาง ในรูปความถี่ ได้ดังนี้

ผลสอบ	อาจารย์		
	A	B	C
สอบตก, F	5	4	3
สอบผ่าน, P	10	5	15
มีเงื่อนไข, C	5	1	2
รวม	20	10	20

จงหาผลสรุปว่า อาจารย์ทั้ง 3 ท่าน โดยเฉลี่ยแล้วให้ผลสอบ แตกต่างกันหรือไม่ ที่ $\alpha = 0.05$

วิธีทำ ใช้ Kruskal Wallis แบบมี ties

		A	B	C	จำนวนซ้ำ
แปลงเป็นลำดับที่	1	5	4	3	12
	2	10	5	15	30
	3	5	1	2	8
		n_1=20	n_2=10	n_3=20	

เพราะฉะนั้น 12 จำนวน แรกที่สอบตก จะได้ ลำดับที่ = 1+2+…….12/12 = 6.5

30 จำนวนที่สอบผ่าน จะได้ ลำดับที่ = 13+14+……..+ 42/30 = 27.5

8 จำนวนที่มีเงื่อนไข จะได้ ลำดับที่ = 43+………+50/8 = 46.5

หาผลรวมลำดับที่ ในแต่ละอาจารย์ จะได้

	A	B	C
R_i	540	210	525
R_i^2	291600	44100	275625
$\frac{R_i^2}{n_i}$	14580	4410	13781.2

$$\sum_{i=1}^{3} \frac{R_i^{\ 2}}{n_i} = 32771.25$$

โดยค่า R_i ของกลุ่ม A คิดจาก 5×6.5+10×27.5+5×46.5 = 540

และปรับ tie

$$\sum t_i\,(t_i - 1)\,(t_i + 1) = (12\times11\times13) + (30\times29\times31) + (8\times7\times9) = 29190$$

$$n^3 - n = 50^3 - 50 = 124950$$

$$H = \frac{\frac{12}{n(n+1)}\sum\frac{R_i^{\ 2}}{n_i} - 3(n+1)}{1 - \frac{\sum\left(t^3 - t\right)}{n^3 - n}}$$

$$= \frac{\frac{12}{50\times51}\times32771.25 - 3(50+1)}{1 - \frac{29190}{124950}}$$

$$= \frac{1.218}{0.766}$$

$$= 1.59$$

$\chi^2_{2,.05} = 5.99$ ไม่ตก CR. ยอมรับ H_0

4. การทดสอบสมมติฐานแย้งแบบมีลำดับของ Jonckheere – Terptra

(The Jonckheere Terpstra Test for Order Alternative)

บางครั้งต้องการที่จะทดสอบสมติฐานดังนี้ $H_0 : M_1 = M_2 = ... = M_k$

$H_1 : M_1 \leq M_2 \leq .... \leq M_k$

เช่น ในการศึกษาถึงประสิทธภาพของยา ผู้วิจัยต้องการทราบว่า ตัวอย่างที่ถูกทดลองจะแสดงผลตอบสนองมากขึ้น เมื่อเพิ่มขนาดยา (increased dosage) หรือไม่

สมมติฐานแย้งเช่นนี้เรียกว่า สมมติฐานแย้งแบบมีลำดับ (ordered alternative)

ข้อกำหนดเบื้องต้น

1. ข้อมูลประกอบด้วย k กลุ่มตัวอย่างด้วยขนาด $n_1, n_2, ...n_k$
2. ค่าสังเกตเป็นอิสระกันทั้งภายในกลุ่มและระหว่างกลุ่ม
3. ตัวแปรสุ่มเป็นค่าต่อเนื่อง
4. มาตราวัดอย่างน้อยต้องเป็นมาตราวัดแบบเรียงลำดับ
5. ประชากรทั้ง k กลุ่มมีลักษณะเหมือนกัน ยกเว้นต่างกันที่ค่ากลาง (location parameter)

สมมติฐาน

H_0 : ตัวอย่าง k กลุ่มมาจากประชากรเดียวกัน หรือประชากรที่เหมือนกัน

H_1 : ตัวอย่าง k กลุ่มมาจากประชากรซึ่งมีค่ามัธยฐาน ดังนี้ $M_1 \leq M_2 \leq M_3 .... \leq M_k$

สถิติที่ใช้ทดสอบ

$$J = \sum_{i<j} U_{ij}$$

เมื่อ U_{ij} เป็นจำนวนคู่ของค่าสังเกต (a,b) ซึ่ง $X_{ia} < X_{jb}$ หมายความว่า เราจะเปรียบเทียบทุกคู่ของค่าสังเกตตัวอย่าง (ดังจะแสดงด้วยตัวอย่าง)

การตัดสินใจ จะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า J คำนวณค่าได้ $\geq$ ค่าวิกฤต J ที่ α, k และ $n_1,...,n_k$จากตารางวิกฤตของ Jonckheere

<u>Teis</u> ถ้าค่า $X_{ia} = X_{jb}$ ให้ค่า U_{ij}=1/2

ตัวอย่าง 14 ในการทดลองเพื่อเปรียบเทียบ ขนาดปุ๋ย 3 ประเภท คือ 10, 20, และ 30 กก. ต่อไร่ในการปลูกข้าวในท้องที่แห่งหนึ่ง จึงสุ่มสุ่มเลือกพื้นที่เพาะปลูกมาทดลองทั้งสิ้น 22 ไร่ ให้มีสภาพการปลูกคล้ายคลึง กันทุกประการ ยกเว้นการให้ปุ๋ยทดลองด้วยขนาด 8, 7, และ 7 ไร่ สำหรับการใช้ปุ๋ยจำนวน10, 20, และ 30 กก. ต่อไร่ตามลำดับ เมื่อถึงฤดูเก็บเกี่ยวบันทึกน้ำหนักของข้าวที่ได้เป็นกิโลกรัมจากแต่ละไร่ ดังตารางต่อไปนี้

จงหาผลสรุปว่าการใช้ปุ๋ยมากขึ้นจะทำให้ผลผลิตมากขึ้นตามลำดับหรือไม่

ผลผลิตจากไร่ ที่ใช้ปุ๋ย 10 กก.	ผลผลิตจากไร่ ที่ใช้ปุ๋ย 20 กก.	ผลผลิตจากไร่ ที่ใช้ปุ๋ย 30 กก.
54.0	79.8	98.6
67.0	82.0	99.5
47.2	88.8	95.8
71.1	79.6	93.3
62.7	85.7	98.9
44.8	81.7	91.1
67.4	88.5	94.5
80.2		

วิธีทำ

H_0 : ตัวอย่างทั้ง 3 มาจากประชากรที่มีค่ามัธยฐานไม่ต่างกัน

H_1 : ค่ามัธยฐาประชากรทั้ง 3 มีแนวโน้มที่จะลดลงจากการใช้ปุ๋ย 30 กก. ไปจนถึง 10 กก. ต่อไร่

สถิติที่ใช้ทดสอบ จะเปรียบเทียบค่าสังเกตจากกลุ่ม 10 กับกลุ่ม 20 เช่น 54.0 กับ 79.8 54.0 กับ82.0,….ถ้าค่าสังเกตจากกลุ่ม 10 น้อยกว่ากลุ่ม 20 ให้ค่า $U_{ij} = 1$ ถ้าค่าสังเกตจากกลุ่ม 10 มากกว่ากลุ่ม 20 ให้ค่า $U_{ij} = 0$

ดังนั้น จากทุกคู่ที่เป็นไปได้ของกลุ่ม 10+20 จะได้ค่า $U_{ij} = 54$

และ จากทุกคู่ที่เป็นไปได้ของกลุ่ม 10+30 จะได้ค่า $U_{ij} = 56$

และ จากทุกคู่ที่เป็นไปได้ของกลุ่ม 20+30 จะได้ค่า $U_{ij} = 49$

ฉะนั้น J = 54+56+49 = 159

การตัดสินใจ จากตาราง หาค่าวิกฤต J ที่ $k = 3$ ขนาดตัวอย่าง = 7, 7, 8
ได้ ความน่าจะเป็นที่ค่า $J > 159$ มีค่าน้อยกว่า .005
ดังนั้น P value < .005 ปฏิเสธ H_0 ยอมรับ H_1

หรือค่าวิกฤตที่ $\alpha = .04605$ คือ 109 ซึ่ง J ตัวอย่าง = 159 มีค่าใหญ่กว่าจึงตกในอาณาเขตวิกฤต

ตัวอย่างขนาดใหญ่ ใช้การแจกแจงปกติประมาณ โดยให้ค่า

$$Z = \frac{J - \dfrac{(n^2 - \sum^k n_i^2}{4}}{\sqrt{\dfrac{n^2(2n + 3) - \sum^k n_i^2(2n_i + 3)}{72}}} \sim N(0,1)$$

จากตัวอย่างข้างต้น $n = 8+7+7 = 22$

$$\sum_i^k n_i^2 = 8^2+7^2+7^2 = 162$$

$$\sum_i^k n_i^2\,(2n_i+3) = 8^2\,((2 \text{ x } 8)+3) + 7^2\,((2 \text{ x } 7)+3) + 7^2\,((2 \text{ x } 7)+3)$$

$$= 2882$$

$$Z = \frac{159 - \dfrac{\left[22^2 - 162\right]}{4}}{\sqrt{\dfrac{\left[22^2(2 \times 22 + 3) - 2882\right]}{72}}} = 4.73$$

p-value ของค่า $Z = 4.73$ มีค่าน้อยกว่า .001 จึงปฏิเสธ H_0

4. 1 การทดสอบของ Jonckheere – Terptra แบบมีซ้ำมาก

(The jonckheere – Terptra Test with ties)

สามารถใช้แนวคิดในหัวข้อ 2 และ 4 -มาประยุกต์ใช้ได้เช่นกัน แต่ตารางการจรณ์ที่ปรับได้ใหม่ **จะมีลักษณะเรียงลำดับทางแถวนอนและแถวตั้ง จึงเรียกว่า Doubly ordered contingency table** เช่นได้ตารางการณ์จรดังนี้

ปริมาณยาที่ใช้	ผลการรักษาที่สังเกตเห็น			
	ไม่เปลี่ยนแปลง(1)	เปลี่ยนแปลงบ้าง(2)	เปลี่ยนแปลงมาก(3)	
5 มก./วัน	10	12	17	$n_1 = 39$
10 มก./วัน	7	9	11	$n_2 = 29$
15 มก./วัน	7	8	12	$n_3 = 27$

สถิติที่ใช้ทดสอบคือ สถิติ J แบบปรับ ties ซึ่งคำนวณได้ตามสูตรของ J และมี

H_0 : ค่าเฉลี่ยของประชากรที่ k กลุ่มไม่ต่างกัน

H_1 : ค่าเฉลี่ยจะมีค่าลดหลั่นตามลำดับของการใช้ยาจากมาก⟶ น้อย

หรือ

H_0 : ปริมาณยาที่ได้เป็นอิสระกับผลการรักษา

H_1 : ปริมาณยาที่มากขึ้น จะทำให้ผลการรักษาที่สังเกตเห็นเปลี่ยนแปลงมากขึ้น

ตอนที่2 งานวิจัยที่เกี่ยวข้อง

มีการสร้างโปรแกรมเพื่อช่วยนักวิจัยเลือกสถิติวิเคราะห์ให้เหมาะสมที่สุดหลายงาน เช่น งานของ Cheng – Hong Yang และคณะ ได้สร้างโปรแกรมชื่อ SCRAR เพื่อช่วยนักวิจัยเลือกสถิติวิเคราะห์ตั้งแต่การสุ่มตัวอย่าง จนถึงวิธีการทางสถิติ ดังนำเสนอในรายงานชื่อเรื่อง “An Interactive Statistical Comparison System for Routing Problems” ที่ปรากฎที่ Web Site ชื่อ http://www.journal.au.edu/ijcem/sep97/article4.html หรืองานของ Olsen และ Bozeman ที่นำเสนอวิธีการเลือกสถิติทดสอบสำหรับการทดสอบสมมติฐานแบบต่าง ๆ สำหรับในประเทศไทยก็มีงานวิจัยในระดับวิทยานิพนธ์ ซึ่งนำเสนอการเลือกสถิติวิเคราะห์ในหัวข้อการทดสอบสมมติฐาน การวางแผนการทดลอง การวิเคราะห์สหสัมพันธ์และการถดถอย ซึ่งจะอยู่ในรูปภาษาไทย ทำให้นักวิจัยไทยได้ใช้ประโยชน์เป็นอย่างดี

บทที่ 3

วิธีการดำเนินงาน

ในการดำเนินงานสร้างเว็บไซต์และโปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ได้แบ่งการดำเนินงานออกเป็นขั้นตอนต่างๆ คือ

3.1 การศึกษาโปรแกรมคอมพิวเตอร์ที่เกี่ยวข้อง

3.2 การออกแบบและพัฒนาโปรแกรม

ดังรายละเอียดต่างๆ ดังนี้

3.1 การศึกษาโปรแกรมคอมพิวเตอร์ที่เกี่ยวข้อง

โดยศึกษาโปรแกรมเพื่อสร้างเว็บไซต์และโปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ซึ่งใช้โปรแกรมคอมพิวเตอร์ดังนี้

1. Getting started with Dreamweaver CS3. Dreamweaver Developer Center

[link : http://www.adobe.com/devnet-archive/dreamweaver/getting_started_cs3.html]

2. Learn CSS | Dreamweaver Developer Center. Adobe Developer Connection

[link : http://www.adobe.com/devnet/dreamweaver/css.html]

3. JavaScript and HTML DOM Reference. W3Schools

[link : http://www.w3schools.com/jsref/default.asp]

4. Flash Tutorial. Thailand Adobe Flash User Group. Thai Flash Dev

[link : http://www.thaiflashdev.com]

5.Flash And ActionScript 2.0. Flash Developer Center. Adobe Developer Connection

[link : http://www.adobe.com/devnet/flash.html]

3.2 การออกแบบและพัฒนาโปรแกรม

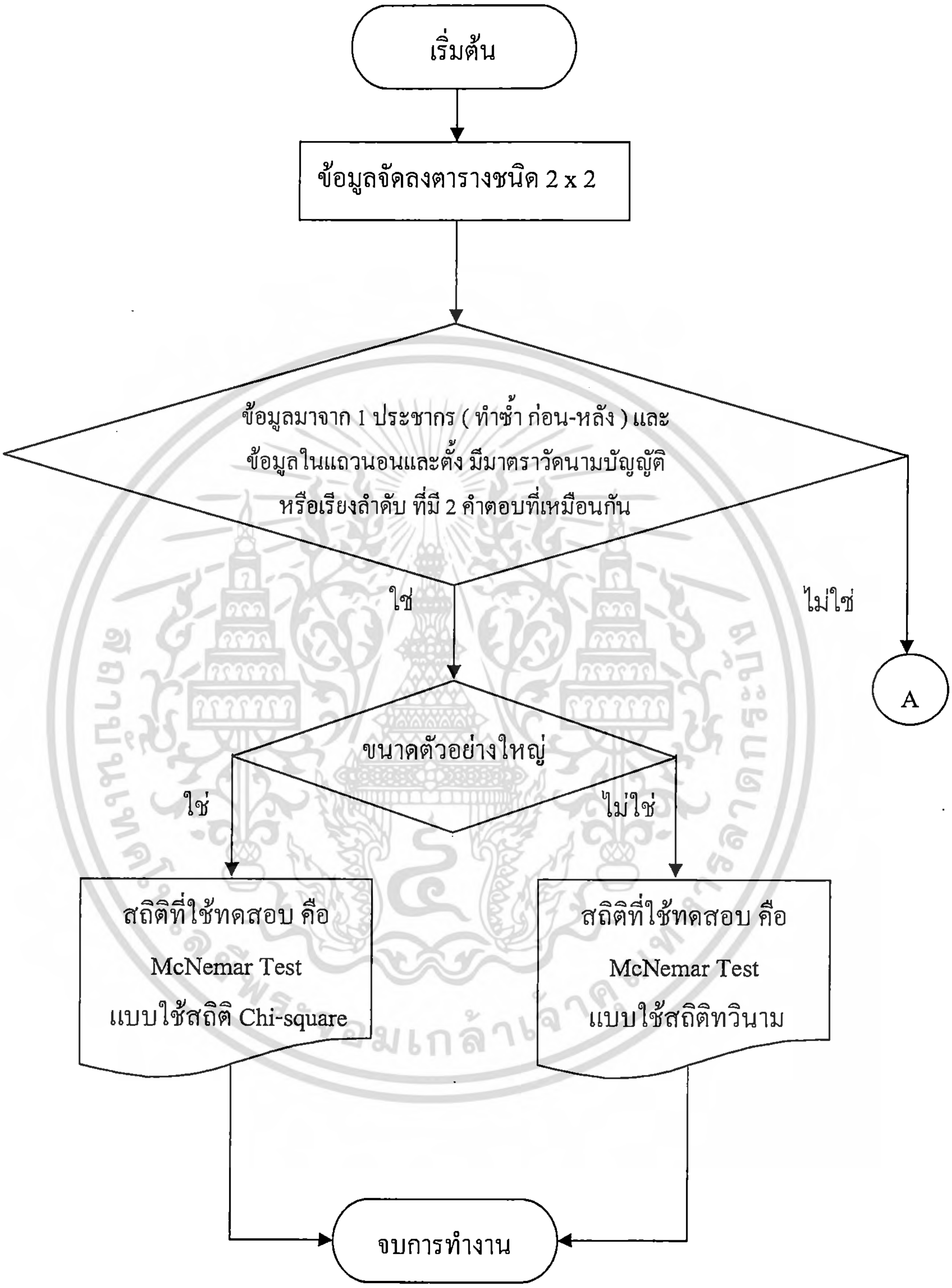
ในการออกแบบและพัฒนาโปรแกรมมีขั้นตอนดังนี้

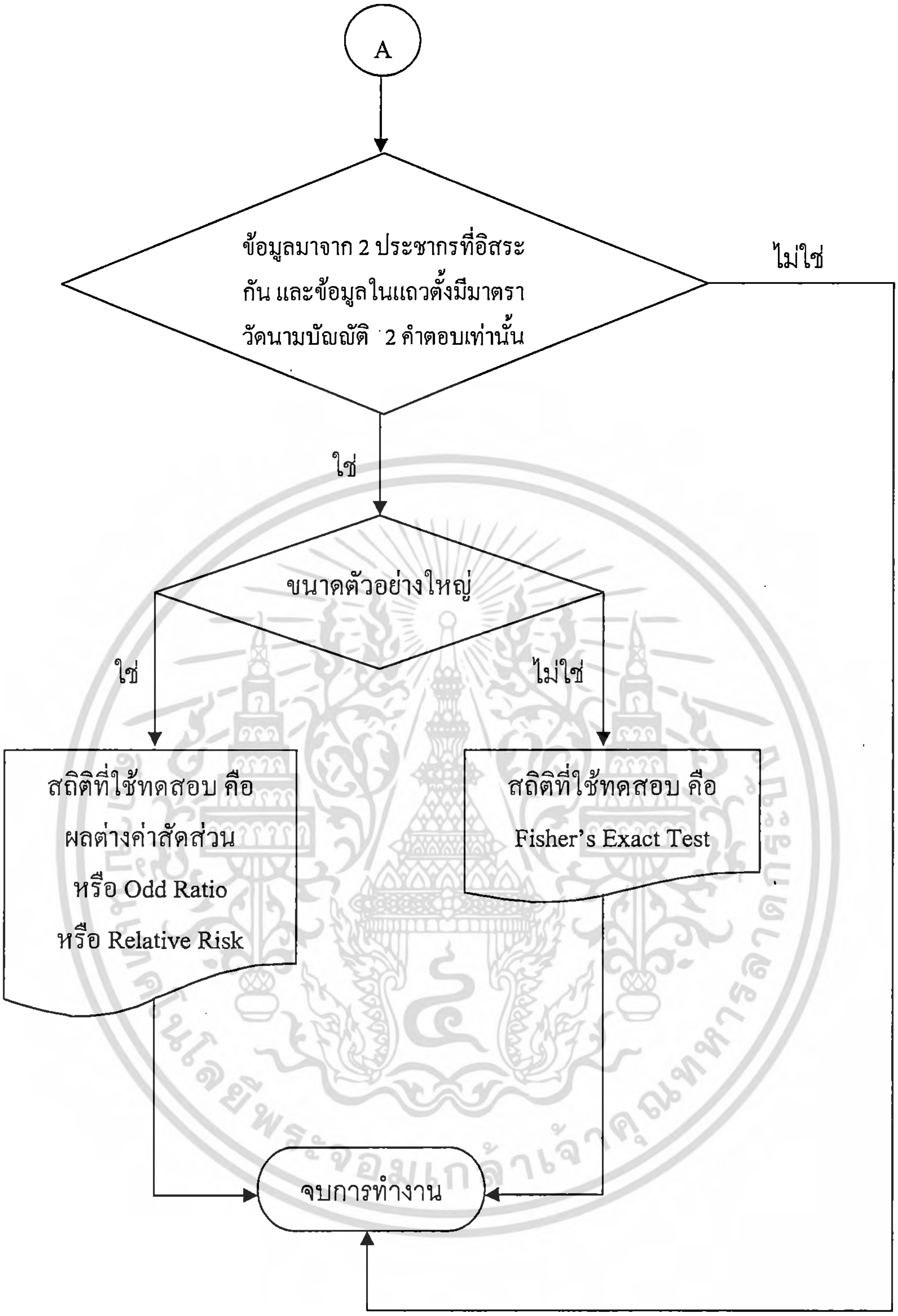
1. กำหนดจุดประสงค์และขอบเขตเพื่อให้สามารถสร้างโปรแกรมให้ได้ตามวัตถุประสงค์และมีประสิทธิภาพ

2. ออกแบบโปรแกรมให้มีความสวยงาม เป็นระเบียบ น่าสนใจ อ่านง่าย และผู้ใช้สามารถเข้าศึกษาหัวข้อต่างๆ ได้ง่าย

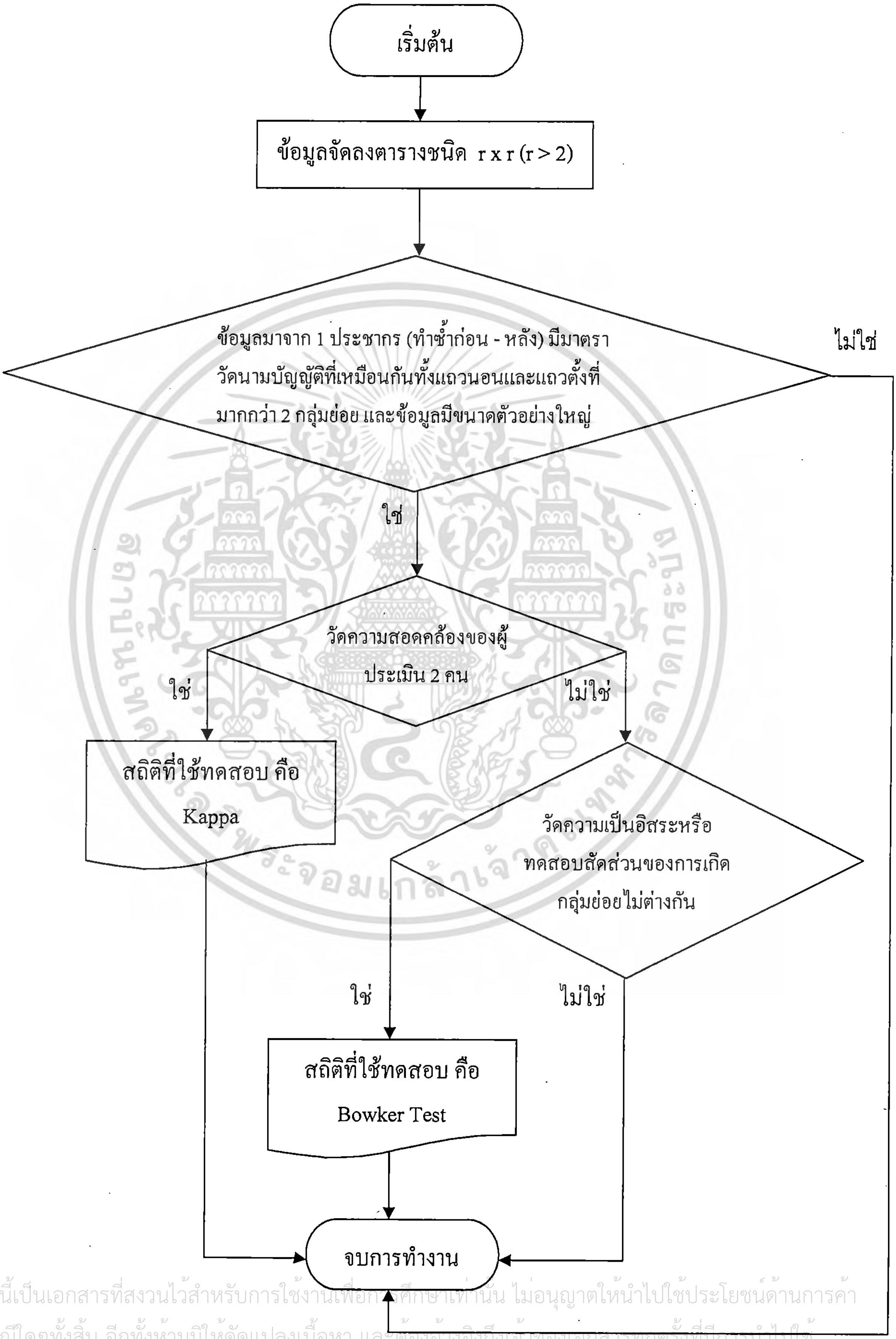
3. ออกแบบแผนผังการทำงาน (Flow Chart) ของโปรแกรมได้ดังนี้

กรณีข้อมูลจัดลงตารางชนิด 2 x 2

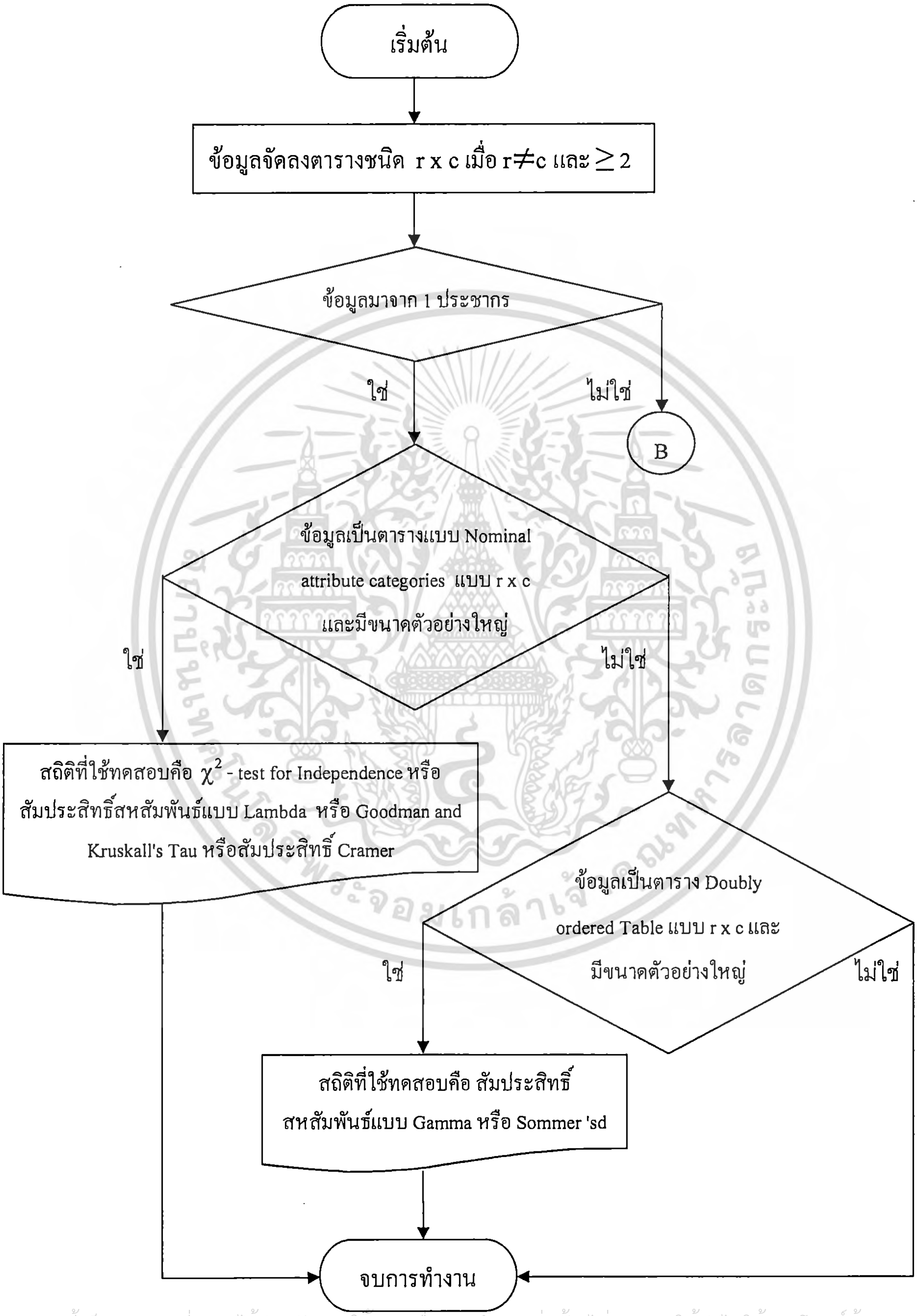


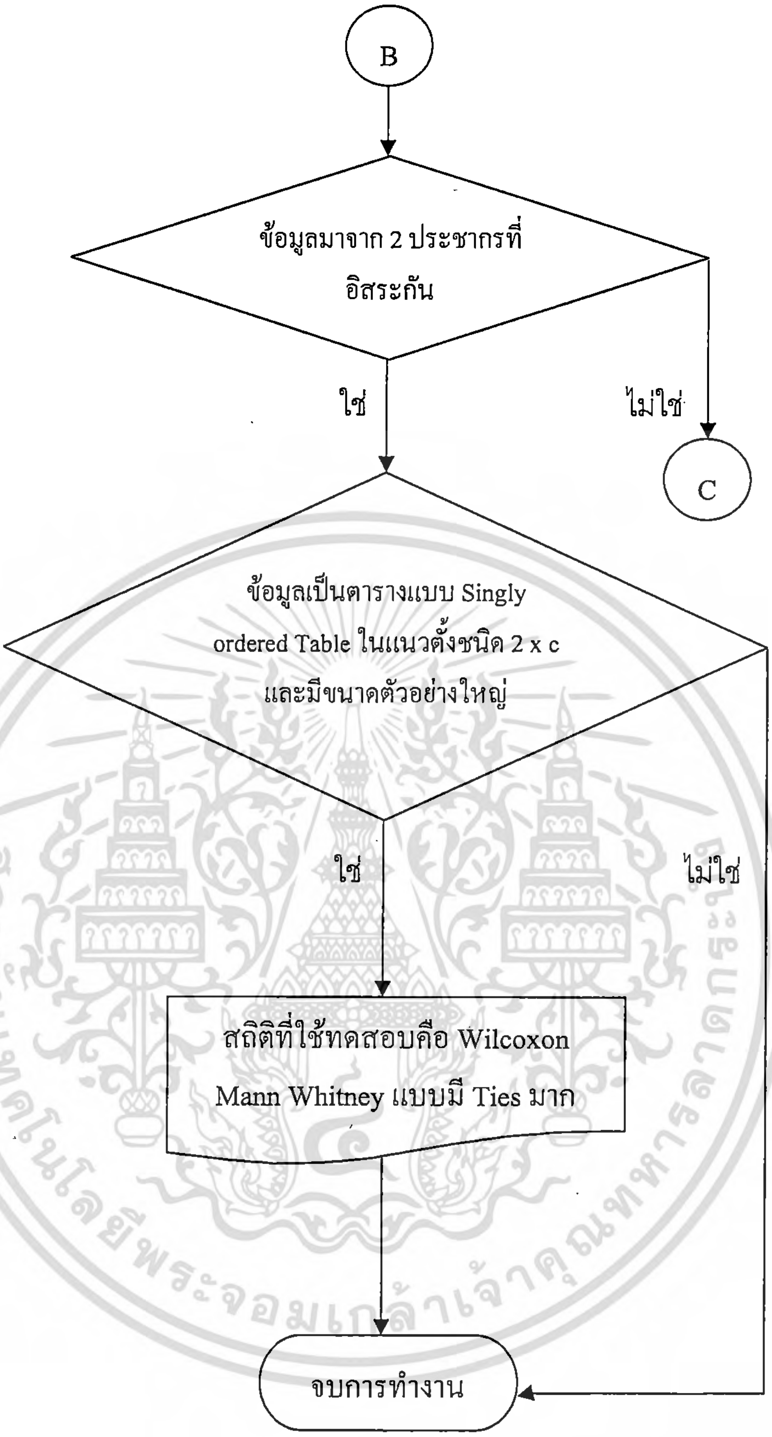

A
ข้อมูลมาจาก 2 ประชากรที่อิสระกัน และข้อมูลในแถวตั้งมีมาตราวัดนามบัญญัติ 2 คำตอบเท่านั้น
ไม่ใช่
ใช่
ขนาดตัวอย่างใหญ่
ใช่
ไม่ใช่
สถิติที่ใช้ทดสอบ คือ ผลต่างค่าสัดส่วน หรือ Odd Ratio หรือ Relative Risk
สถิติที่ใช้ทดสอบ คือ Fisher's Exact Test
จบการทำงาน

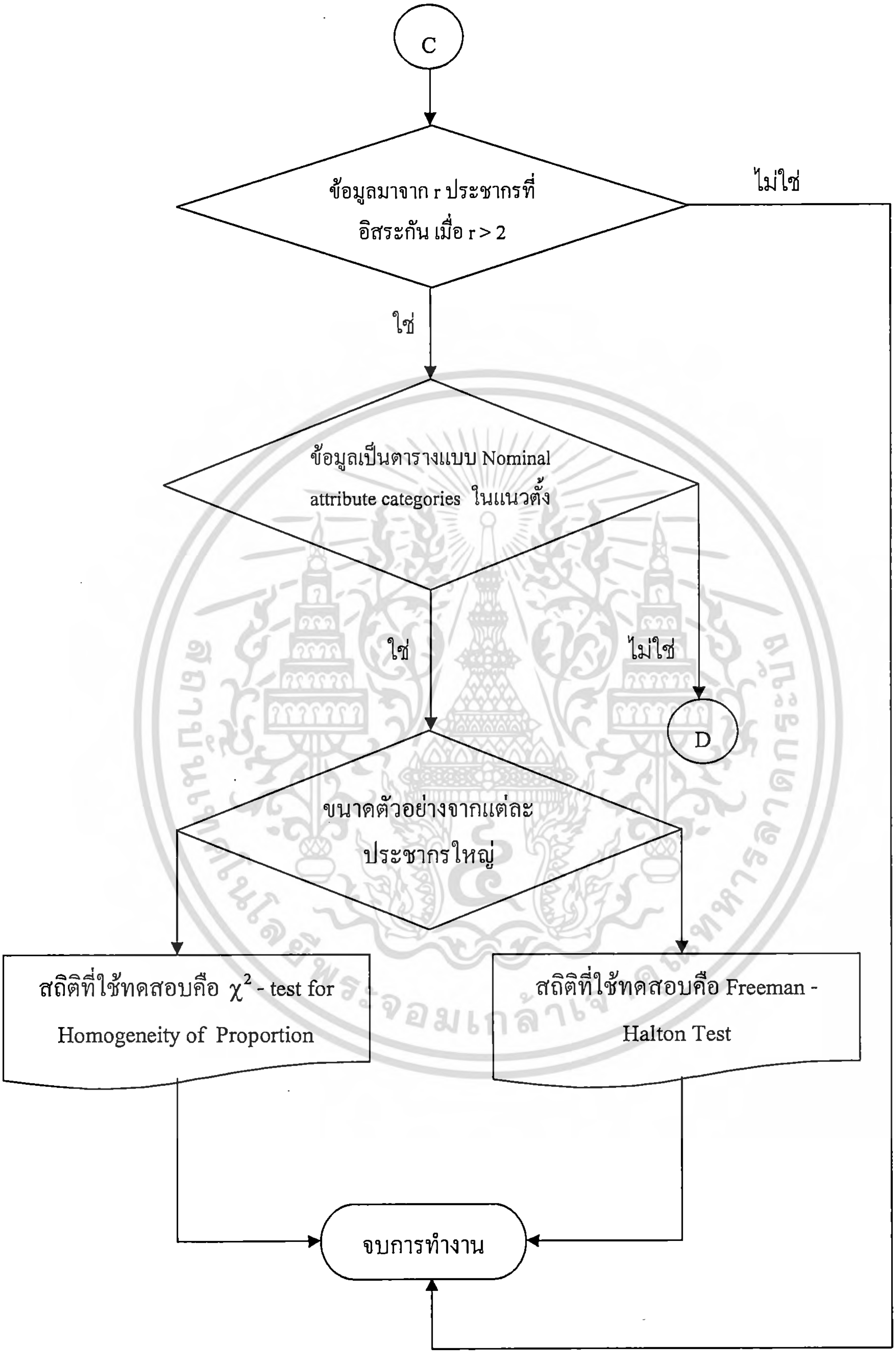
กรณีข้อมูลจัดลงตารางชนิด r x r (r > 2)

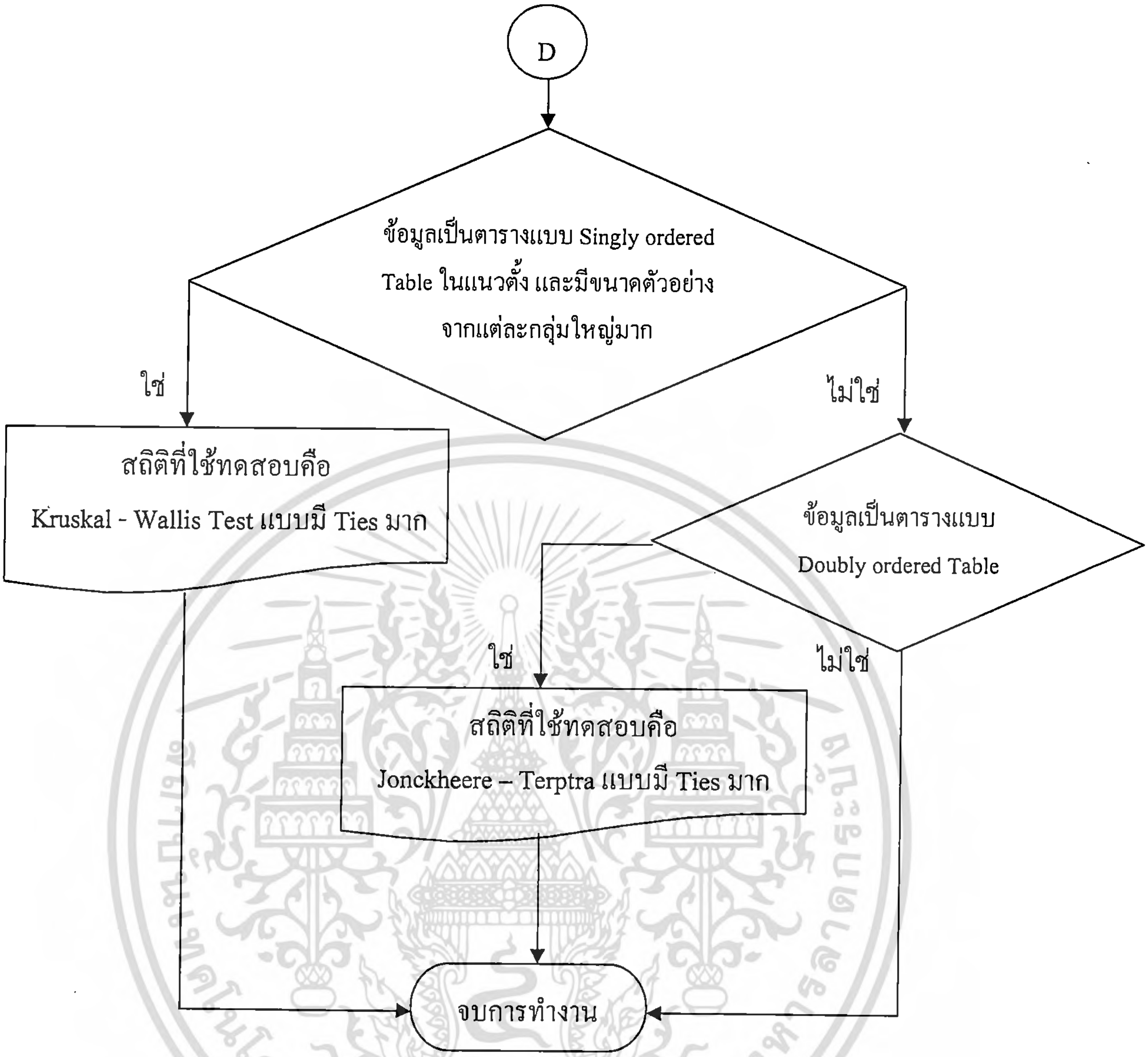


กรณีข้อมูลจัดลงตารางชนิด r x c เมื่อ r≠c และ ≥ 2




B
ข้อมูลมาจาก 2 ประชากรที่อิสระกัน
ใช่
ไม่ใช่
C
ข้อมูลเป็นตารางแบบ Singly ordered Table ในแนวตั้งชนิด 2 x c และมีขนาดตัวอย่างใหญ่
ใช่
ไม่ใช่
สถิติที่ใช้ทดสอบคือ Wilcoxon Mann Whitney แบบมี Ties มาก
จบการทำงาน


C
ข้อมูลมาจาก r ประชากรที่ อิสระกัน เมื่อ r > 2
ไม่ใช่
ใช่
ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแนวตั้ง
ใช่
ไม่ใช่
D
ขนาดตัวอย่างจากแต่ละประชากรใหญ่
สถิติที่ใช้ทดสอบคือ χ^2 - test for Homogeneity of Proportion
สถิติที่ใช้ทดสอบคือ Freeman - Halton Test
จบการทำงาน


D
ข้อมูลเป็นตารางแบบ Singly ordered Table ในแนวตั้ง และมีขนาดตัวอย่างจากแต่ละกลุ่มใหญ่มาก
ใช่
ไม่ใช่
สถิติที่ใช้ทดสอบคือ
Kruskal - Wallis Test แบบมี Ties มาก
ข้อมูลเป็นตารางแบบ Doubly ordered Table
ใช่
ไม่ใช่
สถิติที่ใช้ทดสอบคือ
Jonckheere – Terptra แบบมี Ties มาก
จบการทำงาน

บทที่ 4

ผลการศึกษา

ผลการศึกษาครั้งนี้ ได้โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ซึ่งนำเสนอในเว็บไซต์ชื่อ www.kmitl.ac.th/stat โดยมีหัวข้อต่างๆ ดังนี้

4.1 หน้าแรก

4.2 บทนำ

4.3 เนื้อหาของทฤษฎีสถิติ

4.4 วิธีวิเคราะห์ด้วยโปรแกรม SPSS

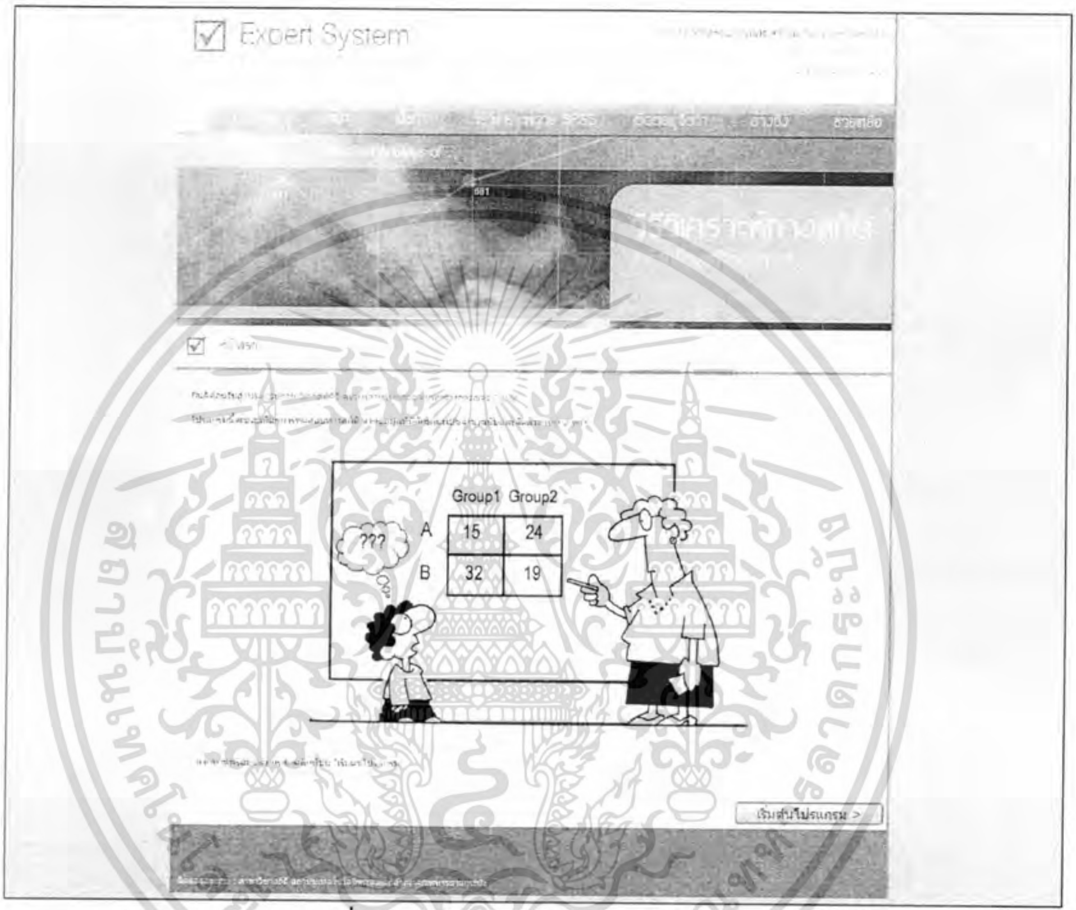
4.5 ติดต่อผู้จัดทำ

4.6 อ้างอิง

4.7 ช่วยเหลือ

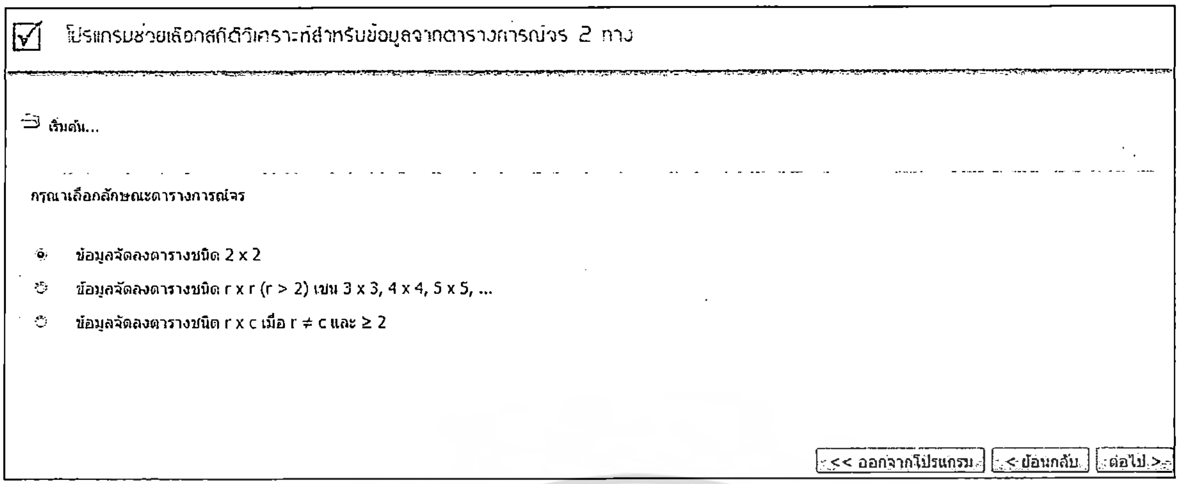
โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ประกอบไปด้วยหน้าต่างของเว็บไซต์ในหัวข้อต่างๆด้วยรายละเอียดดังนี้

4.1 หน้าแรก จะแสดงหน้าแรกของโปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ดังรูป 4-1



รูปที่ 4-1 หน้าต่างหน้าแรกของโปรแกรม

รูปที่ 4-1 เมื่อคลิกที่ **เริ่มต้นโปรแกรม** จะเข้าสู่การเริ่มต้นสอบถามผู้ใช้โปรแกรม ตั้งแต่ลักษณะของตารางการณ์จร ลักษณะของประชากร วัตถุประสงค์ของการทดสอบ ขนาดตัวอย่าง ฯลฯ และจะให้คำตอบแก่ผู้ใช้โปรแกรมว่าสถิติทดสอบใดจึงจะเหมาะสมที่สุด ดังตัวอย่างต่อไปนี้



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...

กรุณาเลือกลักษณะตารางการณ์จร

● ข้อมูลจัดลงตารางชนิด 2 x 2

○ ข้อมูลจัดลงตารางชนิด r x r (r > 2) เช่น 3 x 3, 4 x 4, 5 x 5, ...

○ ข้อมูลจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2

รูปที่ 4-2 หน้าต่างถามตอบลักษณะตารางการณ์จร

รูปที่ 4-2 แสดงหน้าต่างการถามลักษณะตารางการณ์จรของข้อมูลที่ใช้ทดสอบ เมื่อเริ่มต้นโปรแกรม โดยมีตัวเลือกของลักษณะตารางการณ์จรทั้ง 3 ลักษณะคือ ข้อมูลจัดลงตารางชนิด 2 x 2 หรือข้อมูลจัดลงตารางชนิด r x r หรือข้อมูลจัดลงตารางชนิด r x c

ส่วนของข้อมูลที่จัดลงตารางชนิด 2 x 2

☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
└ ข้อมูลจัดลงตารางชนิด 2 x 2

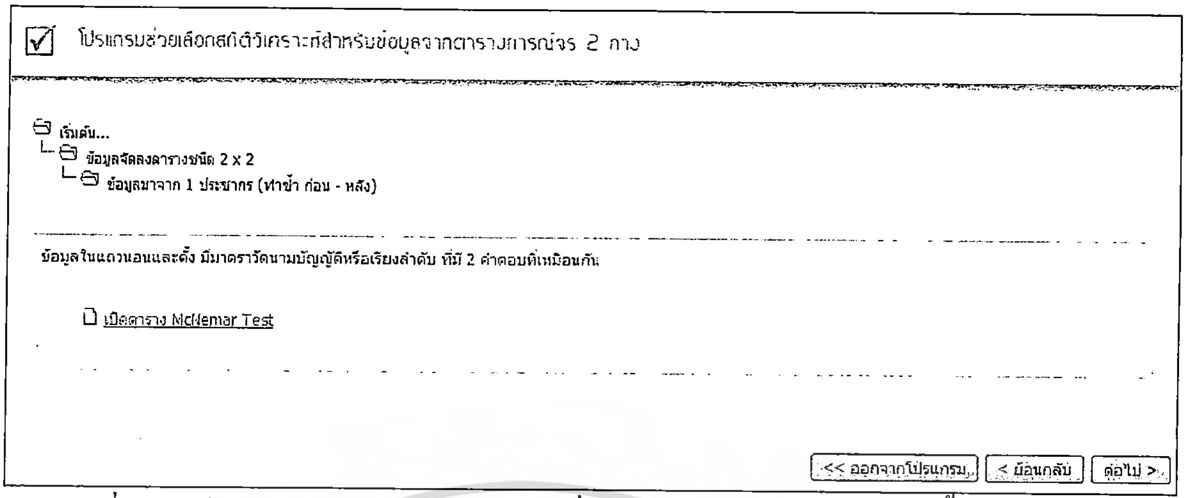
กรุณาเลือกลักษณะประชากรของข้อมูล

● ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)

○ ข้อมูลมาจาก 2 ประชากรที่อิสระกัน

รูปที่ 4-3 หน้าต่างถามลักษณะประชากรของข้อมูล

รูปที่ 4-3 แสดงการถามลักษณะประชากรของข้อมูล เมื่อข้อมูลจัดลงตารางชนิด 2 x 2 มีตัวเลือกลักษณะประชากรของข้อมูล 2 ตัวเลือกคือ ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน – หลัง) หรือข้อมูลมาจาก 2 ประชากรที่อิสระกัน



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

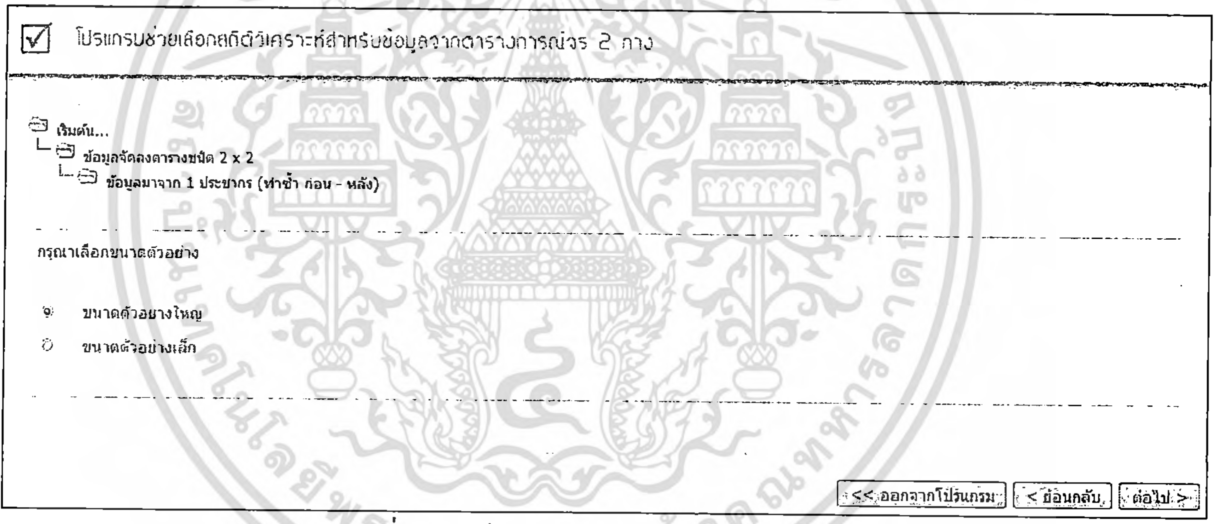
เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด 2 x 2
  - ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)

ข้อมูลในแถวนอนและตั้ง มีมาตราวัดนามบัญญัติหรือเรียงลำดับ ที่มี 2 คำตอบที่เหมือนกัน

เปิดตาราง McNemar Test

รูปที่ 4-4 หน้าต่างอธิบายลักษณะของข้อมูล เมื่อข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน – หลัง)

รูปที่ 4-4 แสดงการอธิบายลักษณะของข้อมูล เมื่อข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน – หลัง) คือ ข้อมูลในแถวนอนและตั้ง มีมาตราวัดนามบัญญัติหรือเรียงลำดับ ที่มี 2 คำตอบที่เหมือนกัน



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด 2 x 2
  - ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)

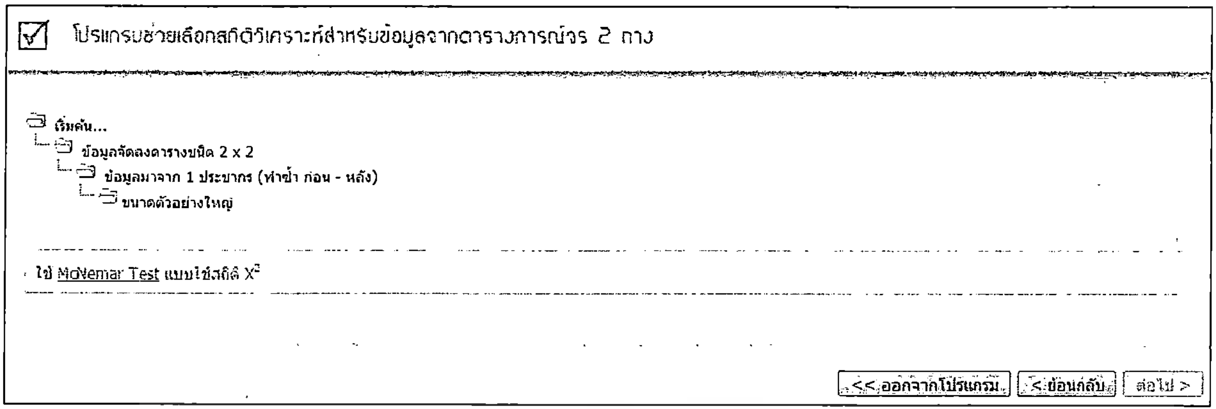
กรุณาเลือกขนาดตัวอย่าง

● ขนาดตัวอย่างใหญ่

○ ขนาดตัวอย่างเล็ก

รูปที่ 4-5 หน้าต่างถามขนาดตัวอย่างของข้อมูล

รูปที่ 4-5 แสดงการถามขนาดตัวอย่างของข้อมูล เมื่อข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน – หลัง) และข้อมูลจัดลงตารางชนิด 2 x 2 มีตัวเลือกคือ ขนาดตัวอย่างใหญ่หรือขนาดตัวอย่างเล็ก



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

- เริ่มต้น...
  - ข้อมูลจัดลงตารางชนิด 2 x 2
    - ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)
      - ขนาดตัวอย่างใหญ่

ใช้ McNemar Test แบบใช้สถิติ X^2

รูปที่ 4-6 หน้าต่างสถิติที่ใช้ทดสอบคือ McNemar Test แบบใช้สถิติ Chi-square

รูปที่ 4-6 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด 2 x 2 ข้อมูลมาจาก 1 ประชากร และขนาดตัวอย่างใหญ่คือ McNemar Test แบบใช้สถิติ Chi-square

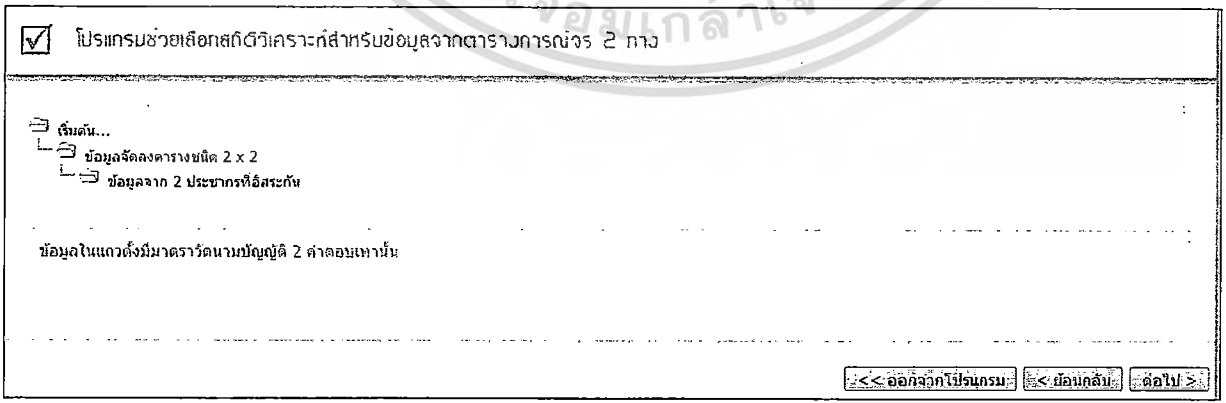
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

- เริ่มต้น...
  - ข้อมูลจัดลงตารางชนิด 2 x 2
    - ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)
      - ขนาดตัวอย่างเล็ก

ใช้ McNemar Test แบบใช้สถิติทวินาม

รูปที่ 4-7 หน้าต่างสถิติที่ใช้ทดสอบคือ McNemar Test แบบใช้สถิติทวินาม

รูปที่ 4-7 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด 2 x 2 ข้อมูลมาจาก 1 ประชากร และขนาดตัวอย่างเล็กคือ McNemar Test แบบใช้สถิติทวินาม



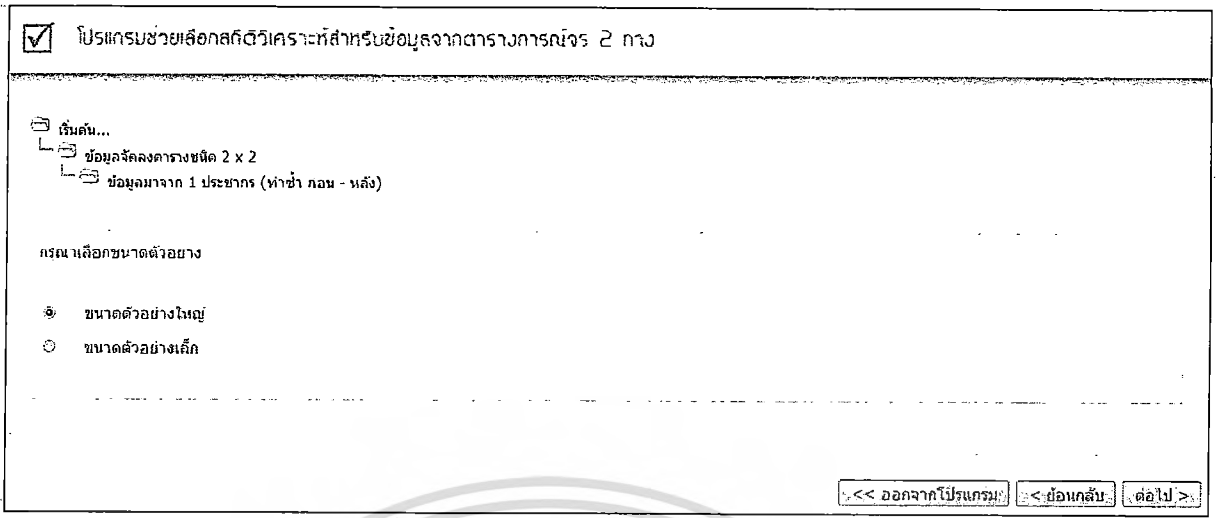
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

- เริ่มต้น...
  - ข้อมูลจัดลงตารางชนิด 2 x 2
    - ข้อมูลจาก 2 ประชากรที่อิสระกัน

ข้อมูลในแถวตั้งมีมาตรวัดนามบัญญัติ 2 คำตอบเท่านั้น

รูปที่ 4-8 หน้าต่างอธิบายลักษณะของข้อมูล เมื่อข้อมูลมาจาก 2 ประชากรที่อิสระกัน

รูปที่ 4-8 แสดงการอธิบายลักษณะของข้อมูล เมื่อข้อมูลมาจาก 2 ประชากรที่อิสระกัน คือ ข้อมูลในแถวตั้งมีมาตรวัดนามบัญญัติที่มี 2 คำตอบเท่านั้น



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
  ข้อมูลจัดลงตารางชนิด 2 x 2
    ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน - หลัง)

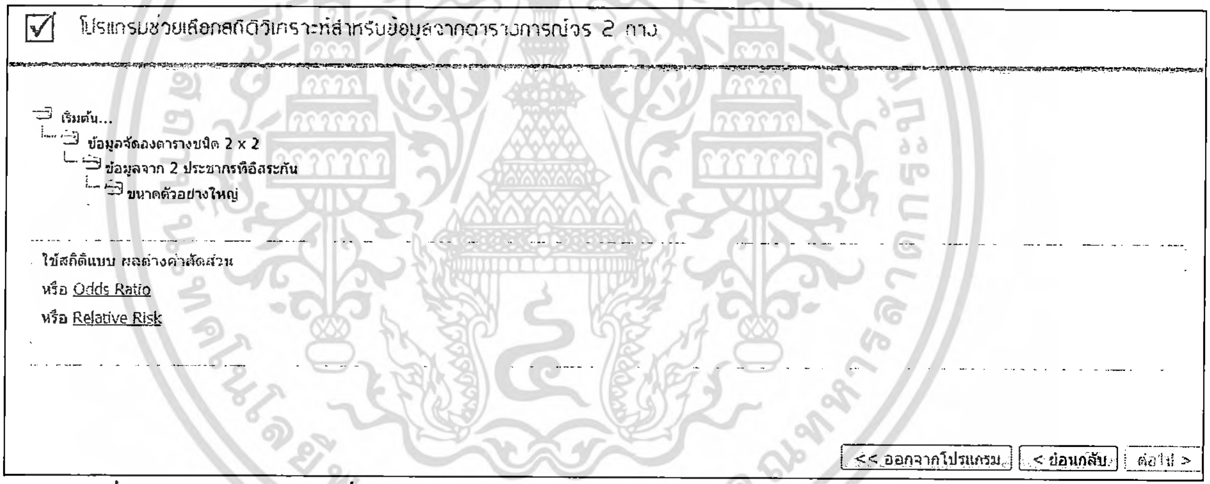
กรุณาเลือกขนาดตัวอย่าง

● ขนาดตัวอย่างใหญ่

○ ขนาดตัวอย่างเล็ก

รูปที่ 4-9 หน้าต่างถามขนาดตัวอย่างของข้อมูล

รูปที่ 4-9 แสดงการถามขนาดตัวอย่างของข้อมูล เมื่อข้อมูลมาจาก 2 ประชากรที่อิสระกันและข้อมูลจัดลงตารางชนิด 2 x 2 มีตัวเลือกคือ ขนาดตัวอย่างใหญ่หรือขนาดตัวอย่างเล็ก



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
  ข้อมูลจัดลงตารางชนิด 2 x 2
    ข้อมูลจาก 2 ประชากรที่อิสระกัน
      ขนาดตัวอย่างใหญ่

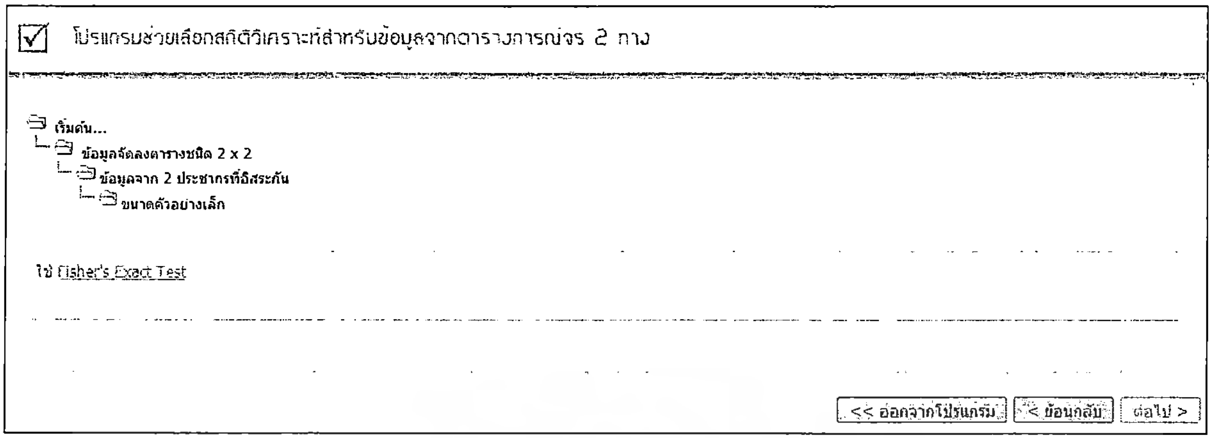
ใช้สถิติแบบ ผลต่างค่าสัดส่วน

หรือ Odds Ratio

หรือ Relative Risk

รูปที่ 4-10 หน้าต่างสถิติที่ใช้ทดสอบคือ ผลต่างค่าสัดส่วน หรือ Odds Ratio หรือ Relative Risk

รูปที่ 4-10 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด 2 x 2 ข้อมูลมาจาก 2 ประชากรที่อิสระกัน และขนาดตัวอย่างใหญ่คือ ผลต่างค่าสัดส่วน หรือ Odds Ratio หรือ Relative Risk



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

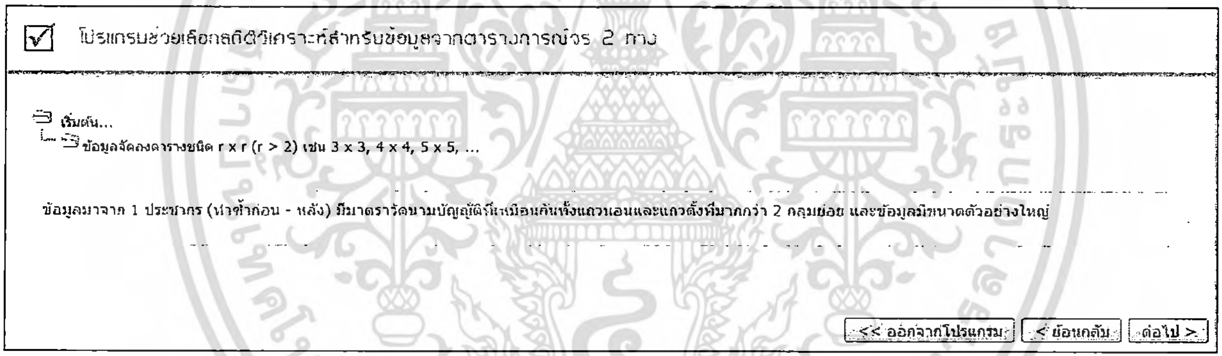
เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด 2 x 2
  - ข้อมูลจาก 2 ประชากรที่อิสระกัน
    - ขนาดตัวอย่างเล็ก

ใช้ Fisher's Exact Test

รูปที่ 4-11 หน้าต่างสถิติที่ใช้ทดสอบคือ Fisher's Exact Test

รูปที่ 4-11 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด 2 x 2 ข้อมูลมาจาก 2 ประชากรที่อิสระกัน และขนาดตัวอย่างเล็กคือ Fisher's Exact Test

ส่วนของข้อมูลที่จัดลงตารางชนิด r x r



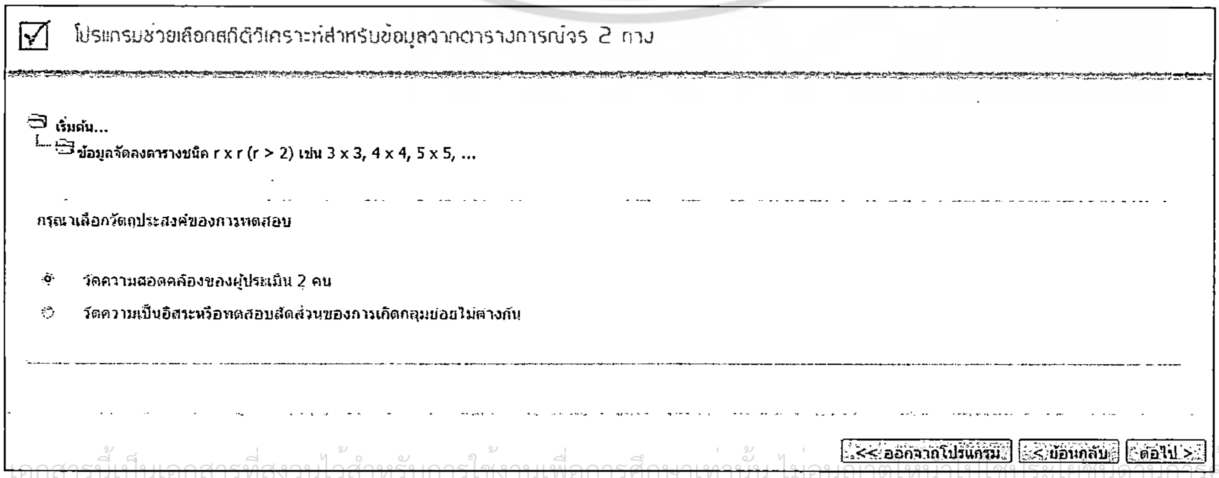
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x r (r > 2) เช่น 3 x 3, 4 x 4, 5 x 5, ...

ข้อมูลมาจาก 1 ประชากร (ทำซ้ำก่อน - หลัง) มีมาตราวัดนามบัญญัติที่เหมือนกันทั้งแถวนอนและแถวตั้งที่มากกว่า 2 กลุ่มย่อย และข้อมูลมีขนาดตัวอย่างใหญ่

รูปที่ 4-12 หน้าต่างอธิบายลักษณะของข้อมูล เมื่อข้อมูลจัดลงตารางชนิด r x r

รูปที่ 4-12 แสดงการอธิบายลักษณะของข้อมูล เมื่อข้อมูลจัดลงตารางชนิด r x r คือ ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน-หลัง) มีมาตราวัดนามบัญญัติที่เหมือนกันทั้งแถวนอนและแถวตั้งที่มากกว่า 2 กลุ่มย่อยและข้อมูลมีขนาดตัวอย่างใหญ่



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x r (r > 2) เช่น 3 x 3, 4 x 4, 5 x 5, ...

กรุณาเลือกวัตถุประสงค์ของการทดสอบ

● วัดความสอดคล้องของผู้ประเมิน 2 คน
○ วัดความเป็นอิสระหรือทดสอบสัดส่วนของการเกิดกลุ่มย่อยไม่ต่างกัน

รูปที่ 4-13 หน้าต่างถามวัตถุประสงค์ของการทดสอบ

รูปที่ 4-13 แสดงการถามวัตถุประสงค์ของการทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x r มี 2 ตัวเลือก คือ วัดความสอดคล้องของผู้ประเมิน 2 คน หรือวัดความเป็นอิสระหรือทดสอบสัดส่วนของการเกิดกลุ่มย่อยไม่ต่างกัน

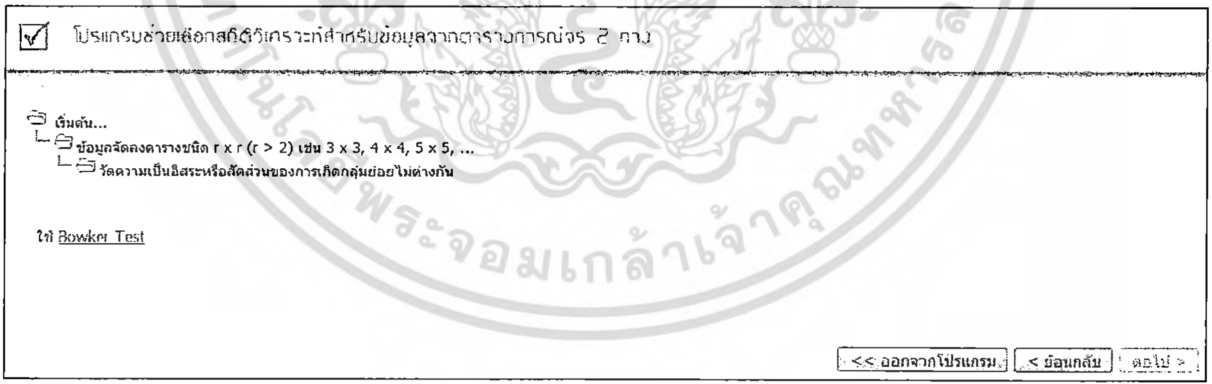
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x r (r > 2) เช่น 3 x 3, 4 x 4, 5 x 5, ...
  - วัดความสอดคล้องของผู้ประเมิน 2 คน

ใช้สถิติแบบ Kappa

รูปที่ 4-14 หน้าต่างสถิติที่ใช้ทดสอบคือ Kappa

รูปที่ 4-14 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x r เมื่อต้องการวัดความสอดคล้องของผู้ประเมิน 2 คน คือ Kappa



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

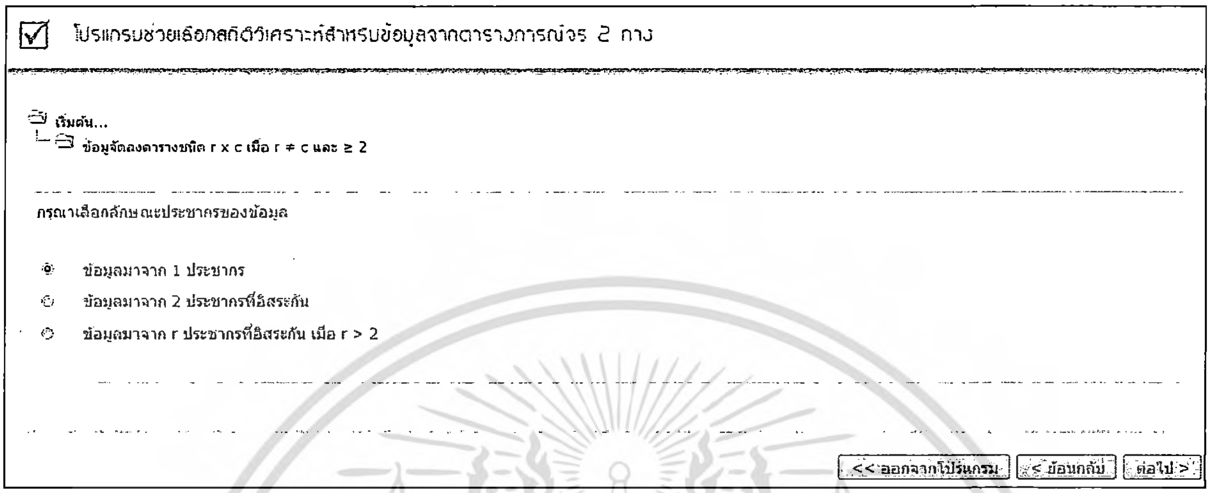
เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x r (r > 2) เช่น 3 x 3, 4 x 4, 5 x 5, ...
  - วัดความเป็นอิสระหรือสัดส่วนของการเกิดกลุ่มย่อยไม่ต่างกัน

ใช้ Bowker Test

รูปที่ 4-15 หน้าต่างสถิติที่ใช้ทดสอบคือ Bowker Test

รูปที่ 4-15 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x r เมื่อต้องการวัดความเป็นอิสระหรือทดสอบสัดส่วนของการเกิดกลุ่มย่อยไม่ต่างกันคือ Bowker Test

ส่วนของข้อมูลที่จัดลงตารางชนิด r x c



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

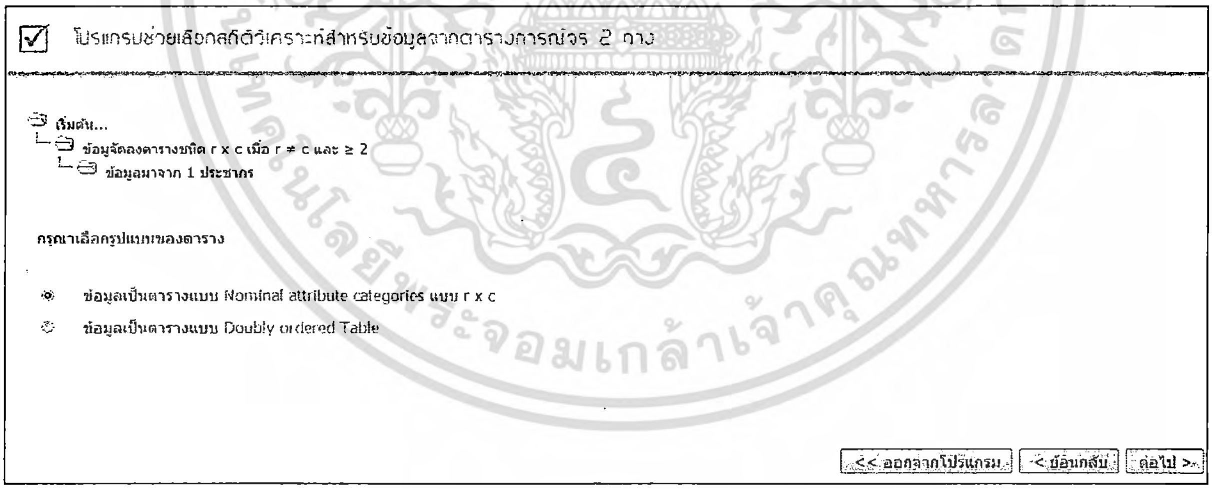
เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2

กรุณาเลือกลักษณะประชากรของข้อมูล

● ข้อมูลมาจาก 1 ประชากร
○ ข้อมูลมาจาก 2 ประชากรที่อิสระกัน
○ ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2

รูปที่ 4-16 หน้าต่างถามลักษณะประชากรของข้อมูล

รูปที่ 4-16 แสดงการถามลักษณะประชากรของข้อมูล เมื่อข้อมูลจัดลงตารางชนิด r x c มีตัวเลือกลักษณะประชากรของข้อมูล 3 ตัวเลือกคือ ข้อมูลมาจาก 1 ประชากร (ทำซ้ำ ก่อน – หลัง) หรือข้อมูลมาจาก 2 ประชากรที่อิสระกัน หรือข้อมูลมาจาก r ประชากรที่อิสระกันเมื่อ r > 2



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

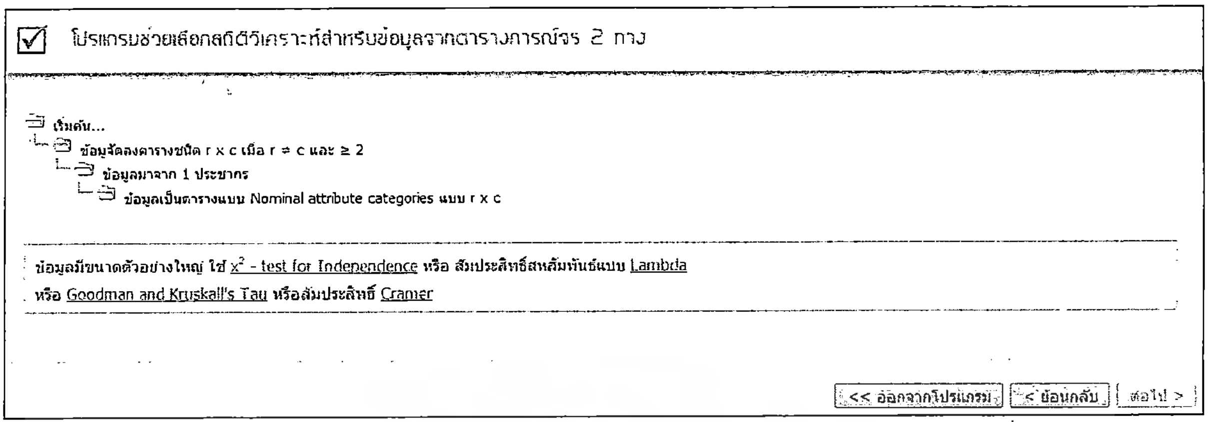
เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
  - ข้อมูลมาจาก 1 ประชากร

กรุณาเลือกรูปแบบของตาราง

● ข้อมูลเป็นตารางแบบ Nominal attribute categories แบบ r x c
○ ข้อมูลเป็นตารางแบบ Doubly ordered Table

รูปที่ 4-17 หน้าต่างถามรูปแบบของตาราง

รูปที่ 4-17 แสดงการถามรูปแบบของตาราง เมื่อข้อมูลจัดลงตารางชนิด r x c และมาจาก 1 ประชากร มี 2 ตัวเลือกคือ ข้อมูลเป็นตารางแบบ Nominal attribute categories แบบ r x c หรือข้อมูลเป็นตารางแบบ Doubly ordered Table



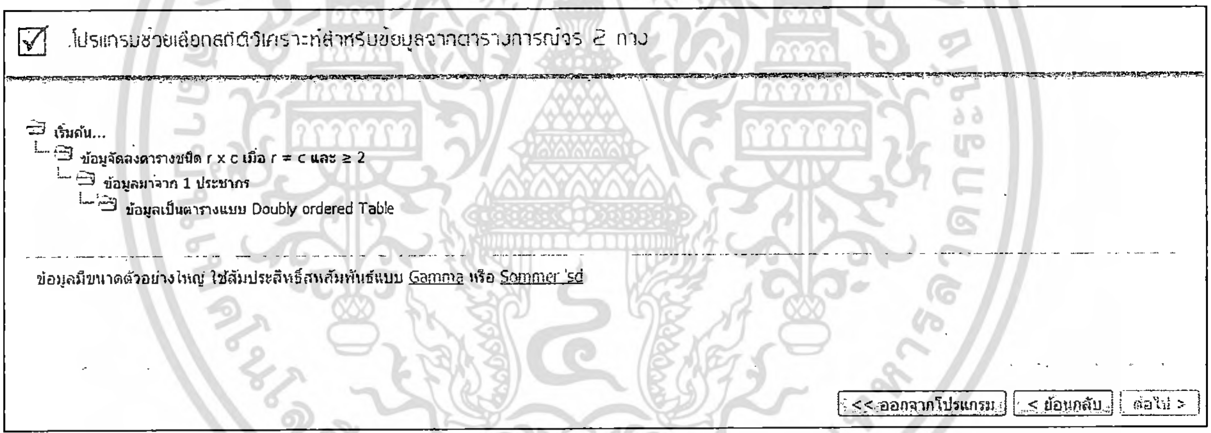
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
  - ข้อมูลมาจาก 1 ประชากร
    - ข้อมูลเป็นตารางแบบ Nominal attribute categories แบบ r x c

ข้อมูลมีขนาดตัวอย่างใหญ่ ใช้ x^2 - test for Independence หรือ สัมประสิทธิ์สหสัมพันธ์แบบ Lambda หรือ Goodman and Kruskall's Tau หรือสัมประสิทธิ์ Cramer

รูปที่ 4-18 หน้าต่างสถิติที่ใช้ทดสอบคือ Chi-square test for Independent หรือสัมประสิทธิ์สหสัมพันธ์แบบ Lambda หรือ Goodman and Kruskull's Tau หรือสัมประสิทธิ์ Cramer

รูปที่ 4-18 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก 1 ประชากร และข้อมูลเป็นตารางแบบ Nominal attribute categories คือ Chi-square test for Independent หรือสัมประสิทธิ์สหสัมพันธ์แบบ Lambda หรือ Goodman and Kruskull's Tau หรือสัมประสิทธิ์ Cramer



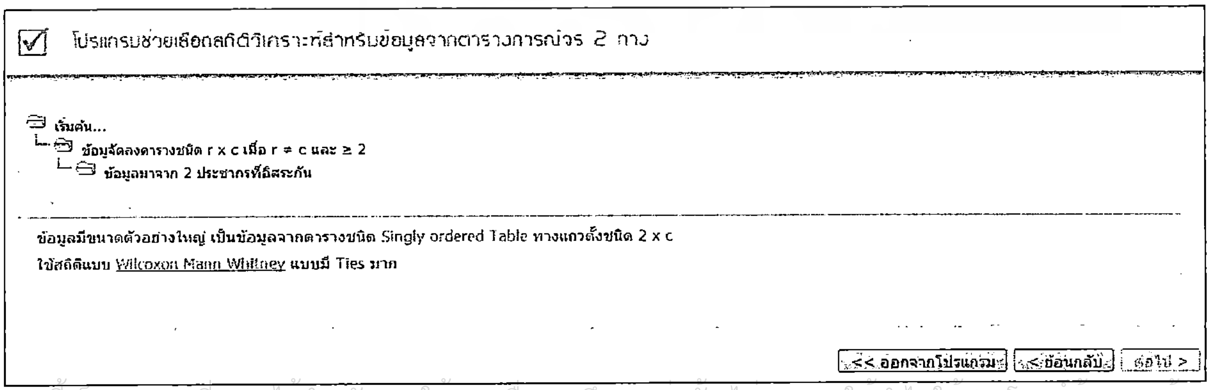
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
  - ข้อมูลมาจาก 1 ประชากร
    - ข้อมูลเป็นตารางแบบ Doubly ordered Table

ข้อมูลมีขนาดตัวอย่างใหญ่ ใช้สัมประสิทธิ์สหสัมพันธ์แบบ Gamma หรือ Sommer 'sd

รูปที่ 4-19 หน้าต่างสถิติที่ใช้ทดสอบคือ สัมประสิทธิ์สหสัมพันธ์แบบ Gamma หรือ Sommer'sd

รูปที่ 4-19 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก 1 ประชากร และข้อมูลเป็นตารางแบบ Doubly ordered Table คือ สัมประสิทธิ์สหสัมพันธ์แบบ Gamma หรือ Sommer'sd



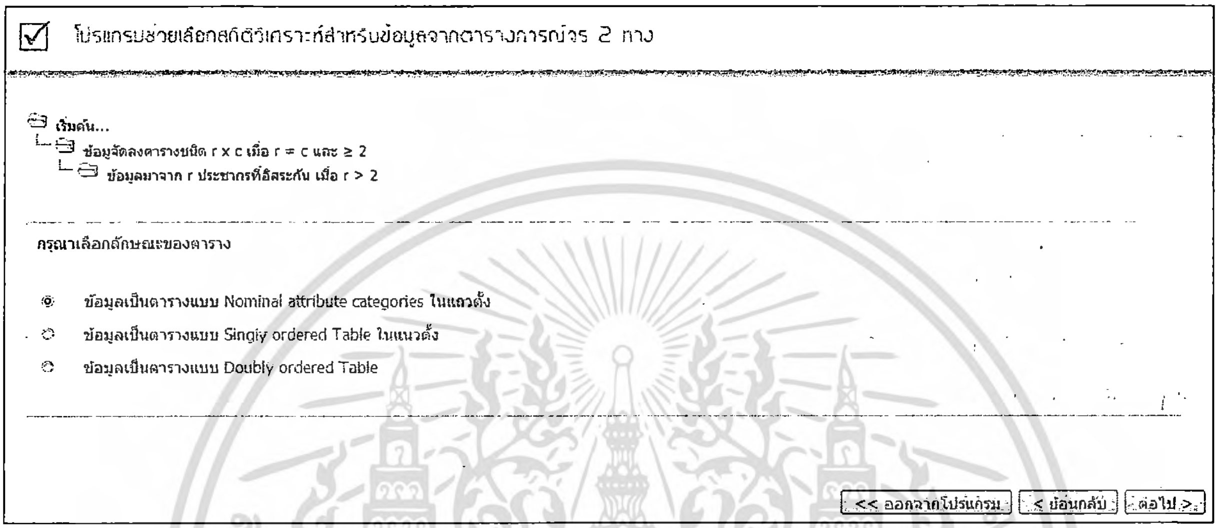
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
  - ข้อมูลมาจาก 2 ประชากรที่อิสระกัน

ข้อมูลมีขนาดตัวอย่างใหญ่ เป็นข้อมูลจากตารางชนิด Singly ordered Table ทางแถวตั้งชนิด 2 x c ใช้สถิติแบบ Wilcoxon Mann Whitney แบบมี Ties มาก

รูปที่ 4-20 หน้าต่างสถิติที่ใช้ทดสอบคือ Wilcoxon Mann Whitney แบบมี Ties มาก

รูปที่ 4-20 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก 2 ประชากรที่อิสระกัน ข้อมูลมีขนาดตัวอย่างใหญ่และจัดลงตารางชนิด Singly ordered Table คือ Wilcoxon Mann Whitney แบบมี Ties มาก



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ r = c และ ≥ 2
  - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2

กรุณาเลือกลักษณะของตาราง

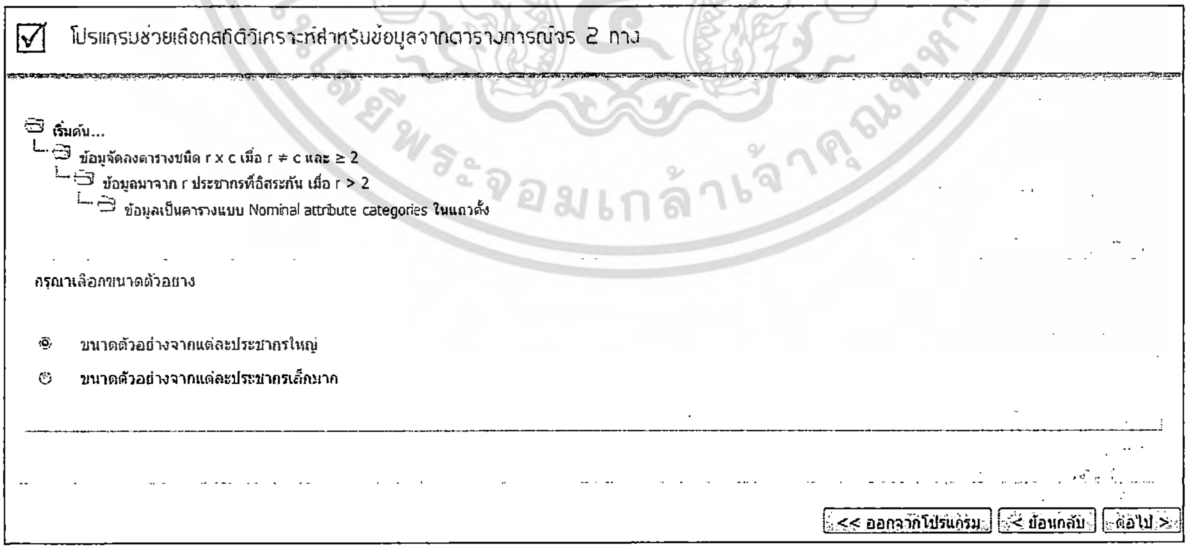
● ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง

○ ข้อมูลเป็นตารางแบบ Singiy ordered Table ในแนวตั้ง

○ ข้อมูลเป็นตารางแบบ Doubly ordered Table

รูปที่ 4-21 หน้าต่างถามรูปแบบของตาราง

รูปที่ 4-21 แสดงการถามรูปแบบของตาราง เมื่อข้อมูลจัดลงตารางชนิด r x c และมาจาก r ประชากรที่อิสระกันเมื่อ r > 2 มี 3 ตัวเลือกคือ ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง หรือข้อมูลเป็นตารางแบบ Singly ordered Table ในแถวตั้ง หรือข้อมูลเป็นตารางแบบ Doubly ordered Table



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
  - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2
    - ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง

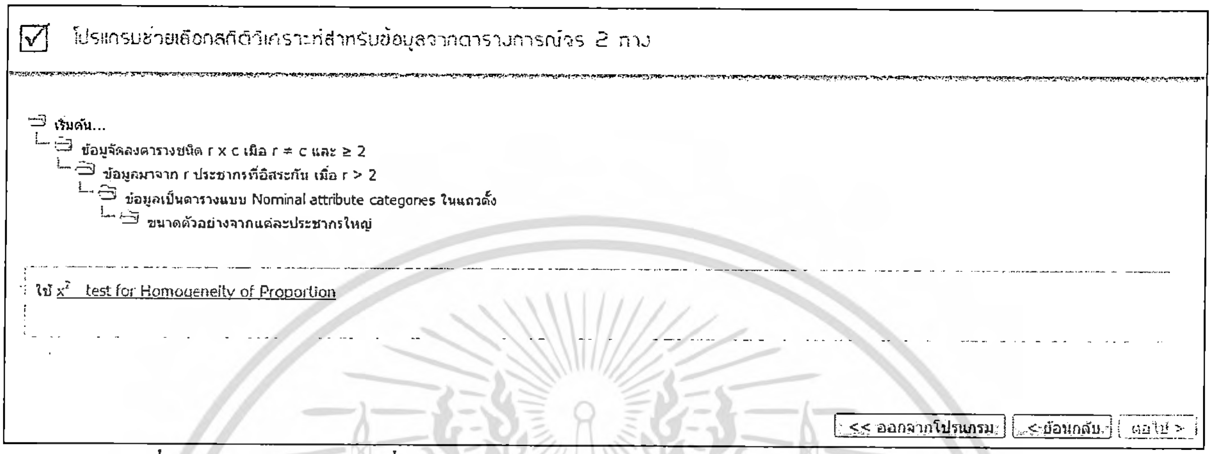
กรุณาเลือกขนาดตัวอย่าง

● ขนาดตัวอย่างจากแต่ละประชากรใหญ่

○ ขนาดตัวอย่างจากแต่ละประชากรเล็กมาก

รูปที่ 4-22 หน้าต่างถามขนาดตัวอย่างของข้อมูล

รูปที่ 4-22 แสดงการถามขนาดตัวอย่างของข้อมูล เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก r ประชากรที่อิสระกันเมื่อ r > 2 และข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง มีตัวเลือก คือ ขนาดตัวอย่างจากแต่ละประชากรใหญ่ หรือ ขนาดตัวอย่างจากแต่ละประชากรเล็กมาก



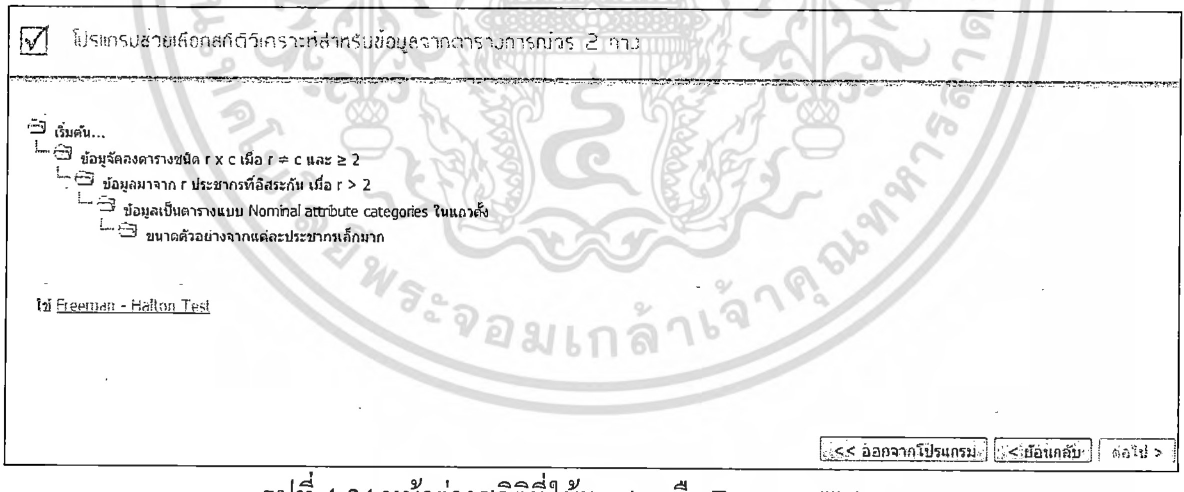
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ $r \neq c$ และ ≥ 2
  - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2
    - ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง
      - ขนาดตัวอย่างจากแต่ละประชากรใหญ่

ใช้ x^2 test for Homogeneity of Proportion

รูปที่ 4-23 หน้าต่างสถิติที่ใช้ทดสอบคือ Chi-square test for Homogeneity of Proportion

รูปที่ 4-23 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก r ประชากรที่อิสระกัน เมื่อ r > 2 และข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง มีขนาดตัวอย่างจากแต่ละประชากรใหญ่คือ Chi-square test for Homogeneity of Proportion



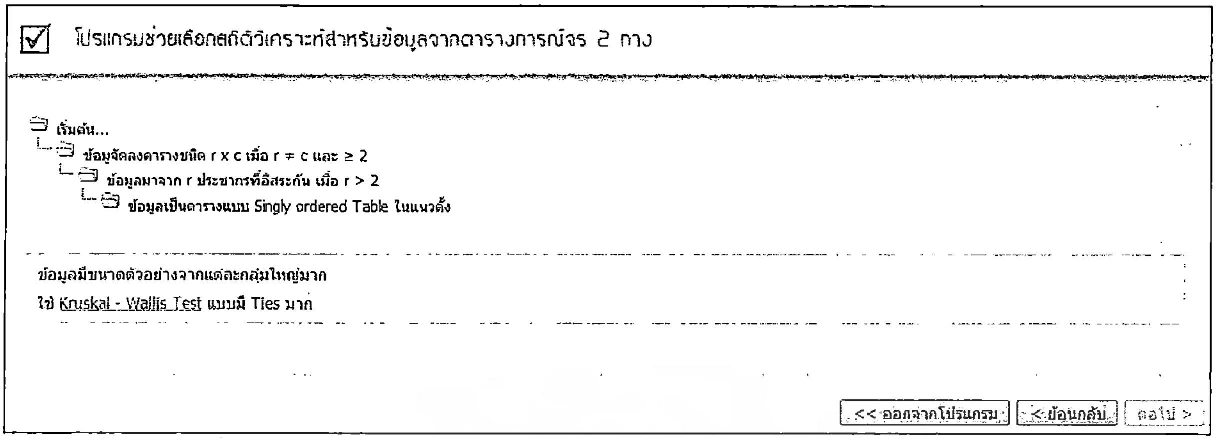
☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

เริ่มต้น...
- ข้อมูลจัดลงตารางชนิด r x c เมื่อ $r \neq c$ และ ≥ 2
  - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2
    - ข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง
      - ขนาดตัวอย่างจากแต่ละประชากรเล็กมาก

ใช้ Freeman - Halton Test

รูปที่ 4-24 หน้าต่างสถิติที่ใช้ทดสอบคือ Freeman-Halton Test

รูปที่ 4-24 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก r ประชากรที่อิสระกัน เมื่อ r > 2 และข้อมูลเป็นตารางแบบ Nominal attribute categories ในแถวตั้ง มีขนาดตัวอย่างจากแต่ละประชากรเล็กมากคือ Freeman-Halton Test



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

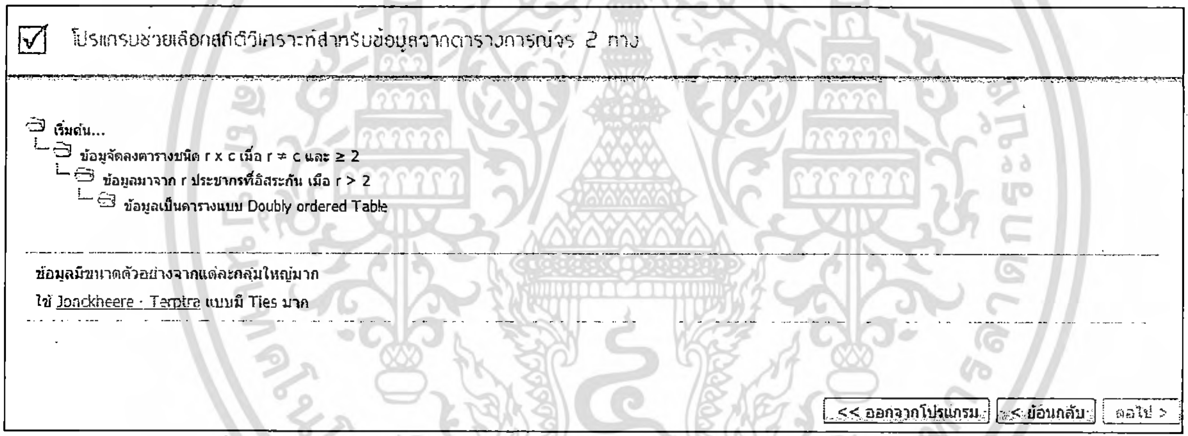
- เริ่มต้น...
  - ข้อมูจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
    - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2
      - ข้อมูลเป็นตารางแบบ Singly ordered Table ในแนวตั้ง

ข้อมูลมีขนาดตัวอย่างจากแต่ละกลุ่มใหญ่มาก

ใช้ Kruskal - Wallis Test แบบมี Ties มาก

รูปที่ 4-25 หน้าต่างสถิติที่ใช้ทดสอบคือ Kruskal-Wallis Test แบบมี Ties มาก

รูปที่ 4-25 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก r ประชากรที่อิสระกัน เมื่อ r > 2 และข้อมูลเป็นตารางแบบ Singly ordered Table ในแถวตั้ง คือ Kruskal-Wallis Test แบบมี Ties มาก



☑ โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง

- เริ่มต้น...
  - ข้อมูจัดลงตารางชนิด r x c เมื่อ r ≠ c และ ≥ 2
    - ข้อมูลมาจาก r ประชากรที่อิสระกัน เมื่อ r > 2
      - ข้อมูลเป็นตารางแบบ Doubly ordered Table

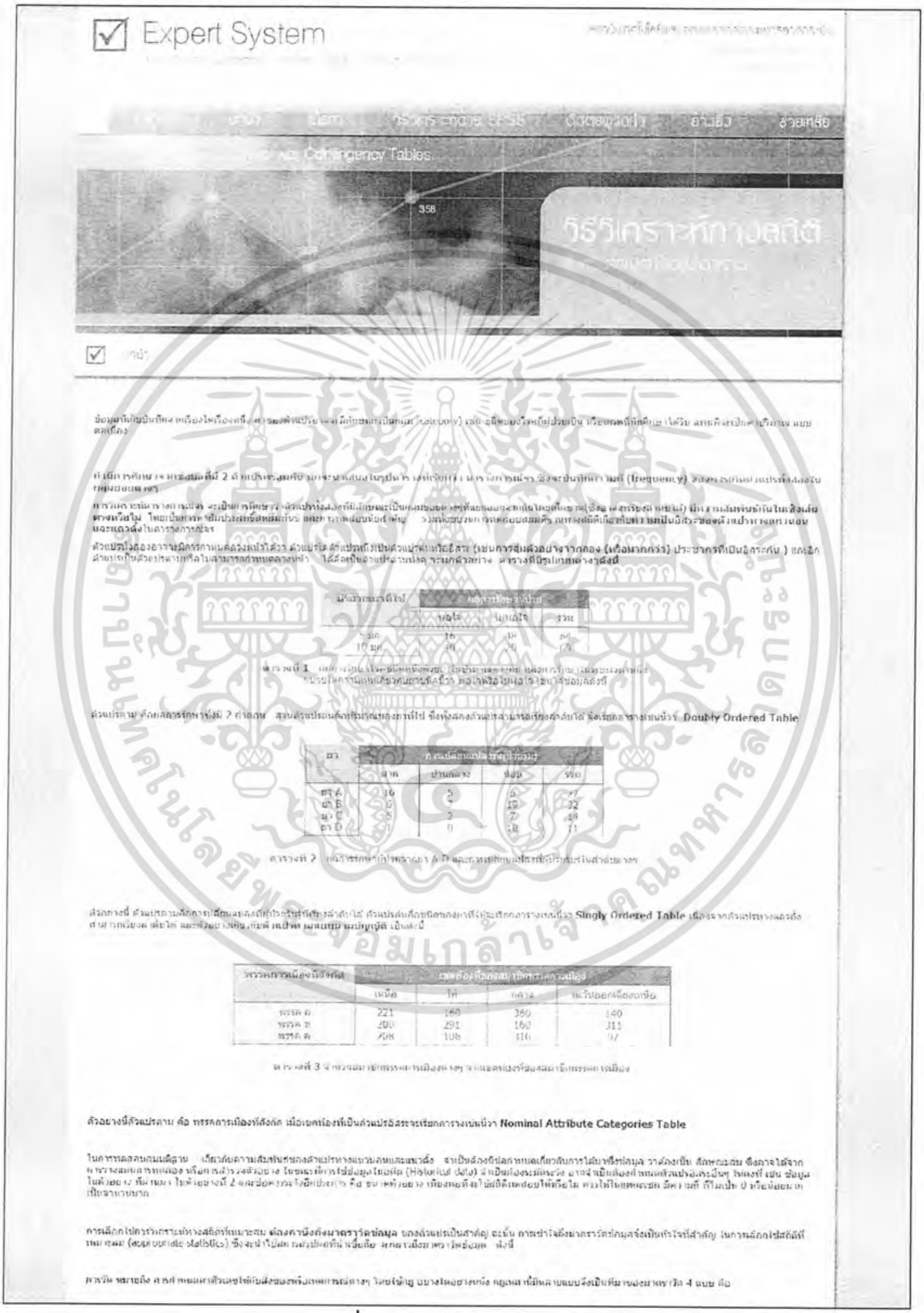
ข้อมูลมีขนาดตัวอย่างจากแต่ละกลุ่มใหญ่มาก

ใช้ Jonckheere - Terptra แบบมี Ties มาก

รูปที่ 4-26 หน้าต่างสถิติที่ใช้ทดสอบคือ Jonckheere-Terptra แบบมี Ties มาก

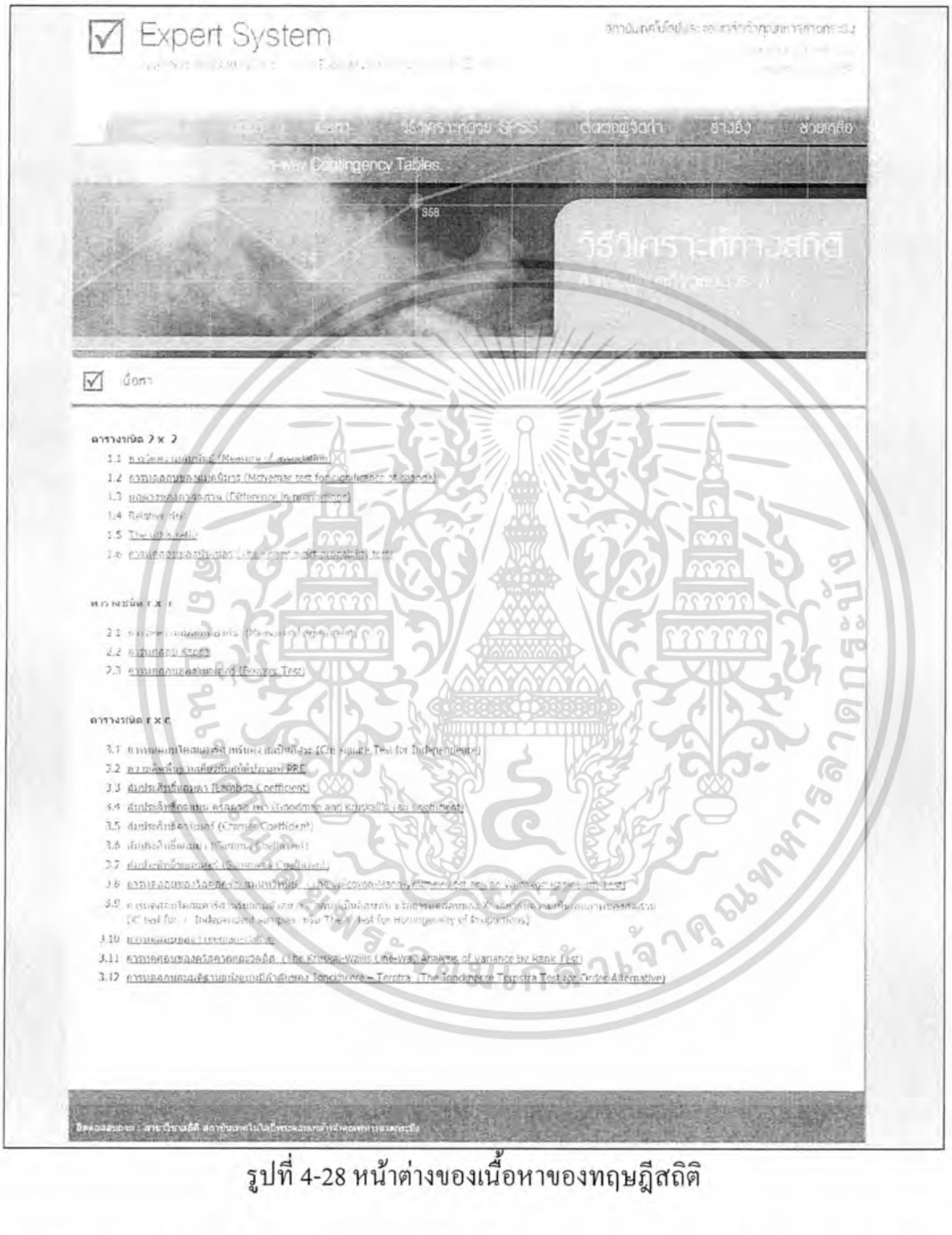
รูปที่ 4-26 แสดงสถิติที่ใช้ทดสอบ เมื่อข้อมูลจัดลงตารางชนิด r x c มาจาก r ประชากรที่อิสระกัน เมื่อ r > 2 และข้อมูลเป็นตารางแบบ Doubly ordered Table คือ Jonckheere-Terptra แบบมี Ties มาก

4.2 บทนำ เมื่อคลิกที่บทนำจะเข้าสู่การอธิบายความหมายของศัพท์ทางสถิติที่เกี่ยวข้องกับการใช้โปรแกรมนี้ เช่น ตารางการณ์จร มาตรวัดข้อมูล เป็นต้น ดังรูป 4-27



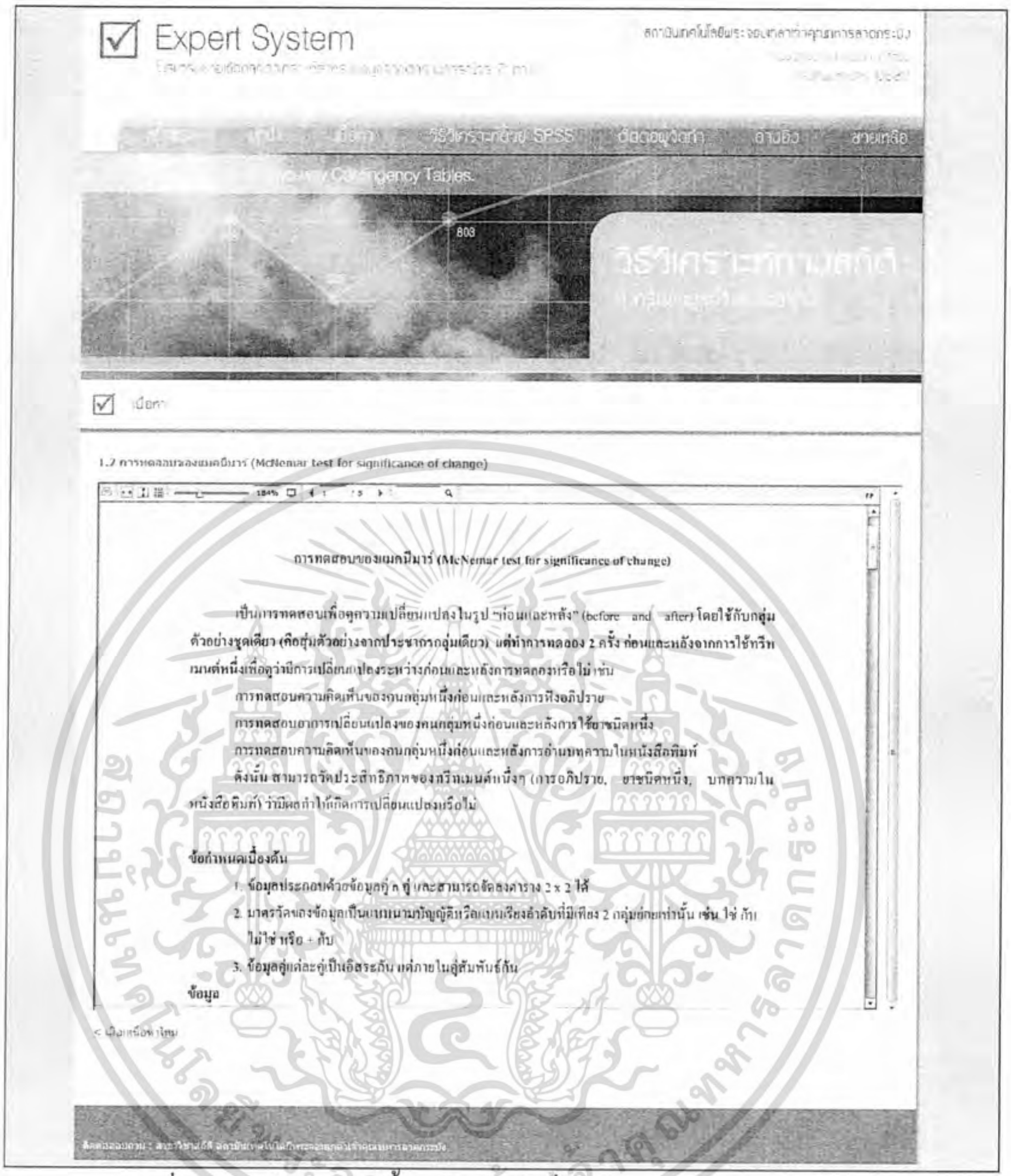
รูปที่ 4-27 หน้าต่างของบทนำ

4.3 เนื้อหาของทฤษฎีสถิติ จะแสดงหัวข้อต่างๆของเนื้อหาของการทดสอบทางสถิติ โดยแต่ละหัวข้อจะอธิบายถึงทฤษฎีทางสถิติของการทดสอบนั้นๆ ที่สามารถนำไปประยุกต์ใช้กับการทดสอบสำหรับข้อมูลจากตารางการณ์จร 2 ทาง ดังรูป 4-28



รูปที่ 4-28 หน้าต่างของเนื้อหาของทฤษฎีสถิติ

เมื่อเลือกหัวข้อเนื้อหา **การทดสอบของแมคนีมาร์ (McNemar test for significance of change)** สามารถแสดงได้ดังรูปที่ 4-29



1.2 การทดสอบของแมคนีมาร์ (McNemar test for significance of change)

การทดสอบของแมคนีมาร์ (McNemar test for significance of change)

เป็นการทดสอบเพื่อดูความเปลี่ยนแปลงในรูป "ก่อนและหลัง" (before and after) โดยใช้กับกลุ่มตัวอย่างชุดเดียว (คือสุ่มตัวอย่างจากประชากรกลุ่มเดียว) แต่ทำการทดลอง 2 ครั้ง ก่อนและหลังจากการใช้ทรีทเมนต์หนึ่งเพื่อดูว่ามีการเปลี่ยนแปลงระหว่างก่อนและหลังการทดลองหรือไม่ เช่น

การทดสอบความคิดเห็นของคนกลุ่มหนึ่งก่อนและหลังการฟังอภิปราย

การทดสอบอาการเปลี่ยนแปลงของคนกลุ่มหนึ่งก่อนและหลังการใช้ยาชนิดหนึ่ง

การทดสอบความคิดเห็นของคนกลุ่มหนึ่งก่อนและหลังการอ่านบทความในหนังสือพิมพ์

ดังนั้น สามารถวัดประสิทธิภาพของทรีทเมนต์หนึ่งๆ (การอภิปราย, ยาชนิดหนึ่ง, บทความในหนังสือพิมพ์) ว่ามีผลทำให้เกิดการเปลี่ยนแปลงหรือไม่

ข้อกำหนดเบื้องต้น

1. ข้อมูลประกอบด้วยข้อมูลคู่ n คู่ และสามารถจัดลงตาราง 2 x 2 ได้
2. มาตรวัดของข้อมูลเป็นแบบนามบัญญัติหรือแบบเรียงลำดับที่มีเพียง 2 กลุ่มย่อยเท่านั้น เช่น ใช่ กับ ไม่ใช่ หรือ + กับ
3. ข้อมูลคู่แต่ละคู่เป็นอิสระกัน แต่ภายในคู่สัมพันธ์กัน

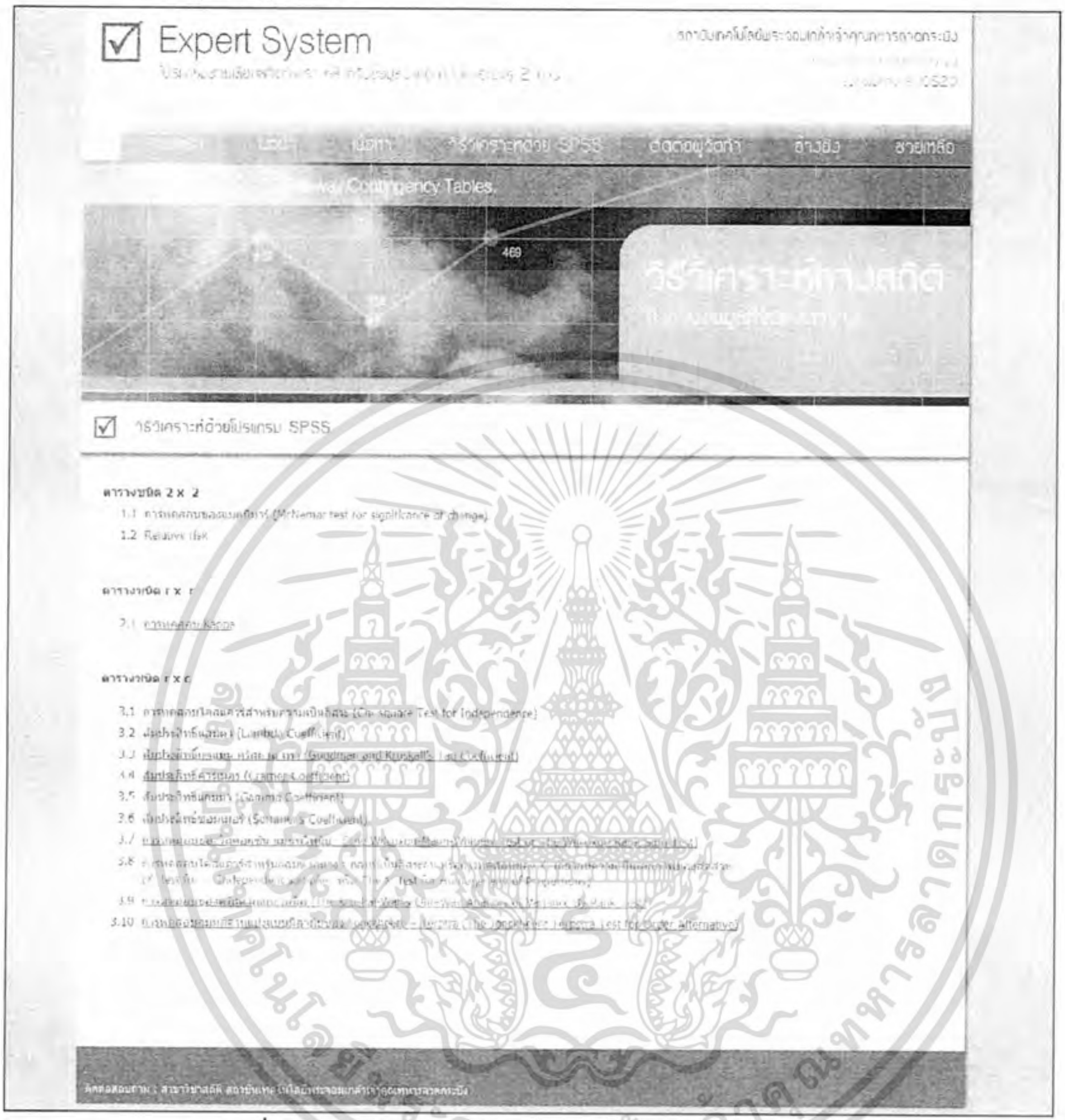
ข้อมูล

< เลือกเนื้อหาใหม่

รูปที่ 4-29 หน้าต่างของเนื้อหาในหัวข้อเรื่อง การทดสอบของแมคนีมาร์

รูปที่ 4-29 แสดงวิธีการใช้การทดสอบของแมคนีมาร์ โดยจะอธิบายรายละเอียดต่าง ๆ สำหรับการทดสอบนี้

4.4 วิธีวิเคราะห์ด้วยโปรแกรม SPSS เมื่อเลือกหัวข้อนี้ จะแสดงขั้นตอนการใช้ SPSS สำหรับการทดสอบต่าง ๆ เพื่อวิเคราะห์ข้อมูล ดังรูป 4-30



ตารางชนิด 2 x 2

1.1 การทดสอบของแมคนีมาร์ (McNemar test for significance of change)

1.2 Relative risk

ตารางชนิด r x r

2.1 การทดสอบ Kappa

ตารางชนิด r x c

3.1 การทดสอบไคสแควร์สำหรับความเป็นอิสระ (Chi-square Test for Independence)

3.2 สัมประสิทธิ์แลมด้า (Lambda Coefficient)

3.3 สัมประสิทธิ์กูดแมน และ ครัสคอล เทา (Goodman and Kruskal's Tau Coefficient)

3.4 สัมประสิทธิ์คาร์เมอร์ (Cramer's Coefficient)

3.5 สัมประสิทธิ์แกมมา (Gamma Coefficient)

3.6 สัมประสิทธิ์ซอมเมอร์ (Somer's Coefficient)

3.7 [illegible]

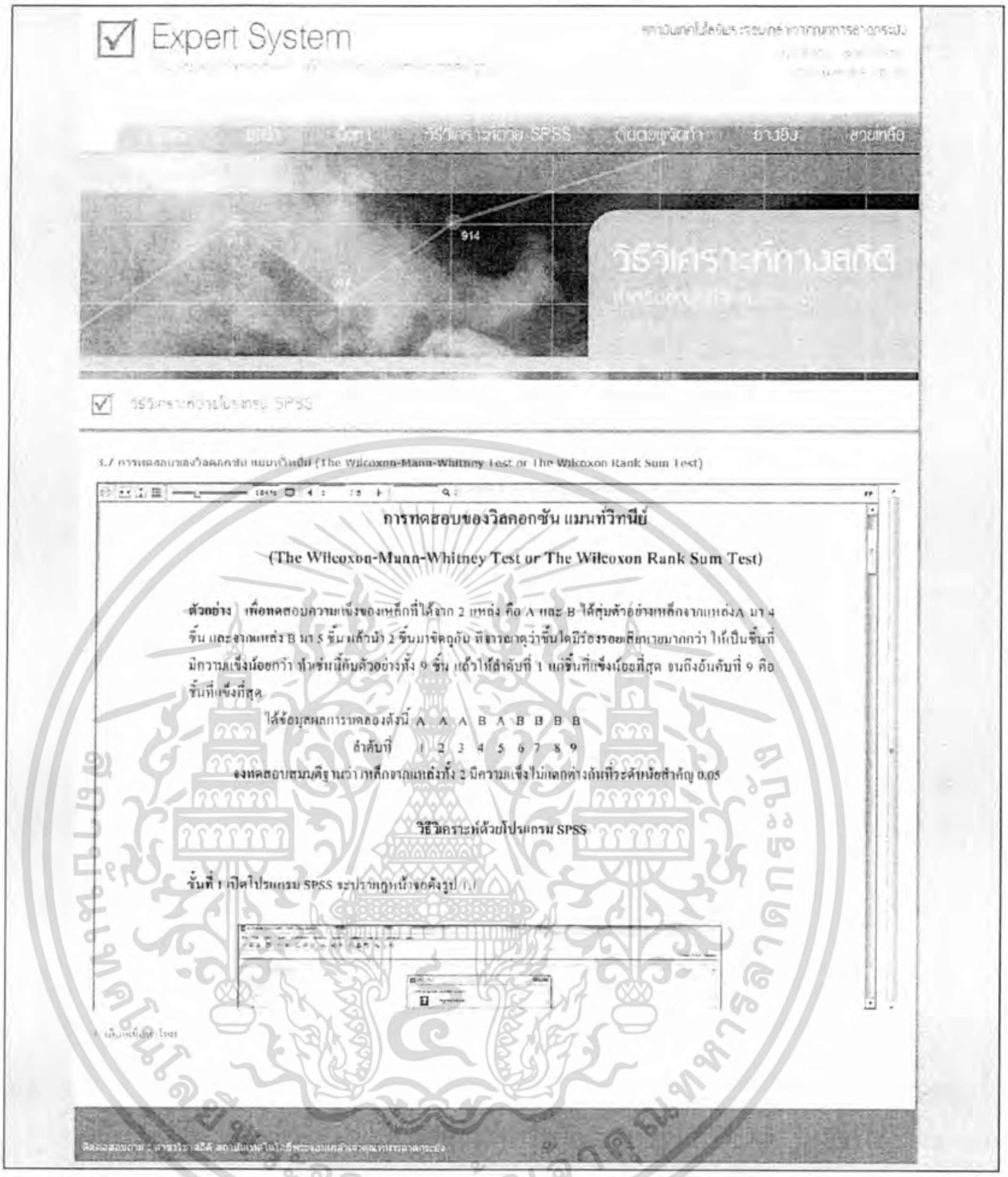
3.8 [illegible]

3.9 [illegible]

3.10 [illegible]

รูปที่ 4-30 หน้าต่างของวิธีวิเคราะห์ด้วยโปรแกรม SPSS

เมื่อเลือกหัวข้อ **การทดสอบของวิลคอกซัน แมนท์วิทนีย์** สามารถแสดงได้ดังรูปที่ 4-31



3.7 การทดสอบของวิลคอกซัน แมนท์วิทนีย์ (The Wilcoxon-Mann-Whitney Test or The Wilcoxon Rank Sum Test)

การทดสอบของวิลคอกซัน แมนท์วิทนีย์

(The Wilcoxon-Mann-Whitney Test or The Wilcoxon Rank Sum Test)

ตัวอย่าง **เพื่อทดสอบ**ความแข็งของเหล็กที่ได้จาก 2 แหล่ง คือ A และ B ได้สุ่มตัวอย่างเหล็กจากแหล่งA มา 4 ชิ้น และจากแหล่ง B มา 5 ชิ้น แล้วนำ 2 ชิ้นมาขีดถูกัน พิจารณาดูว่าชิ้นใดมีร่องรอยเสียหายมากกว่า ให้เป็นชิ้นที่มีความแข็งน้อยกว่า ทำเช่นนี้กับตัวอย่างทั้ง 9 ชิ้น แล้วให้ลำดับที่ 1 แก่ชิ้นที่แข็งน้อยที่สุด จนถึงอันดับที่ 9 คือ ชิ้นที่แข็งที่สุด

ได้ข้อมูลผลการทดลองดังนี้ A A A B A B B B B
ลำดับที่ 1 2 3 4 5 6 7 8 9

จงทดสอบสมมติฐานว่า เหล็กจากแหล่งทั้ง 2 มีความแข็งไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05

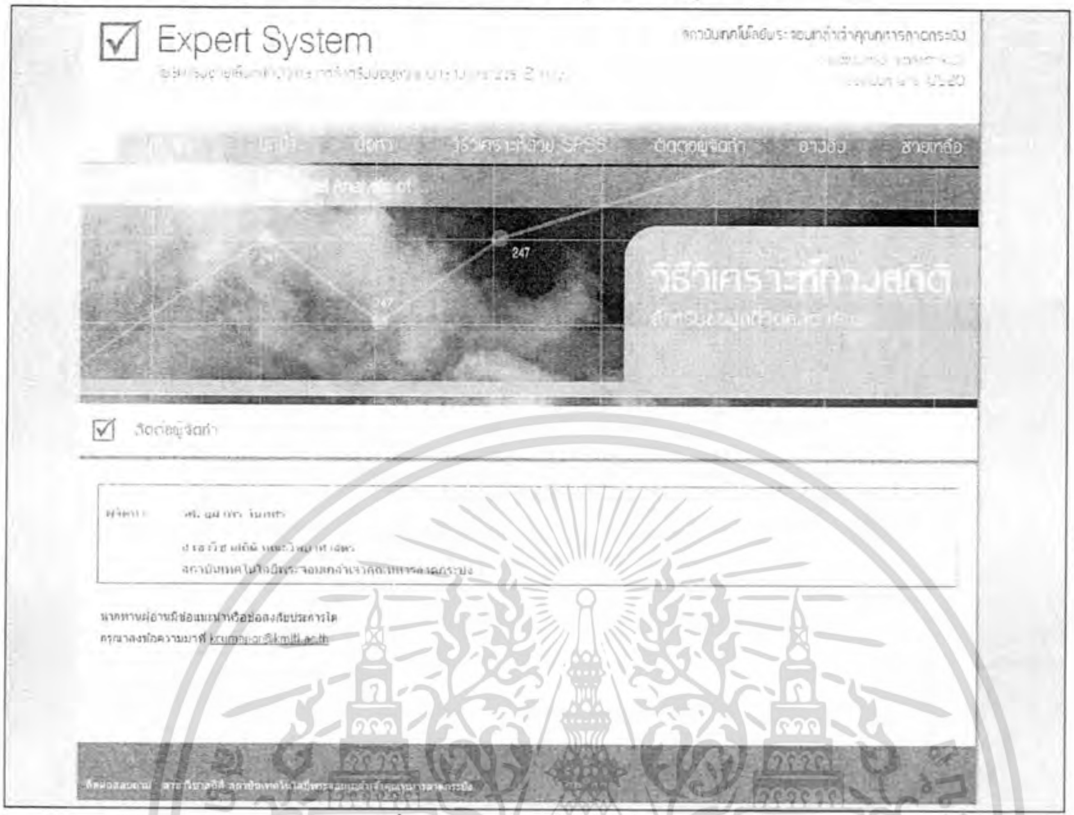
วิธีวิเคราะห์ด้วยโปรแกรม SPSS

ขั้นที่ 1 เปิดโปรแกรม SPSS จะปรากฏหน้าจอดังรูป 1.1

รูปที่ 4-31 หน้าต่างของวิธีวิเคราะห์ด้วยโปรแกรม SPSSในหัวข้อการทดสอบของวิลคอกซัน แมนท์วิทนีย์

รูปที่ 4-31 แสดงวิธีวิเคราะห์ด้วยโปรแกรม SPSS โดยใช้การทดสอบของวิลคอกซัน แมนท์วิทนีย์ โดยจะอธิบายขั้นตอนต่าง ๆ สำหรับการทดสอบนี้

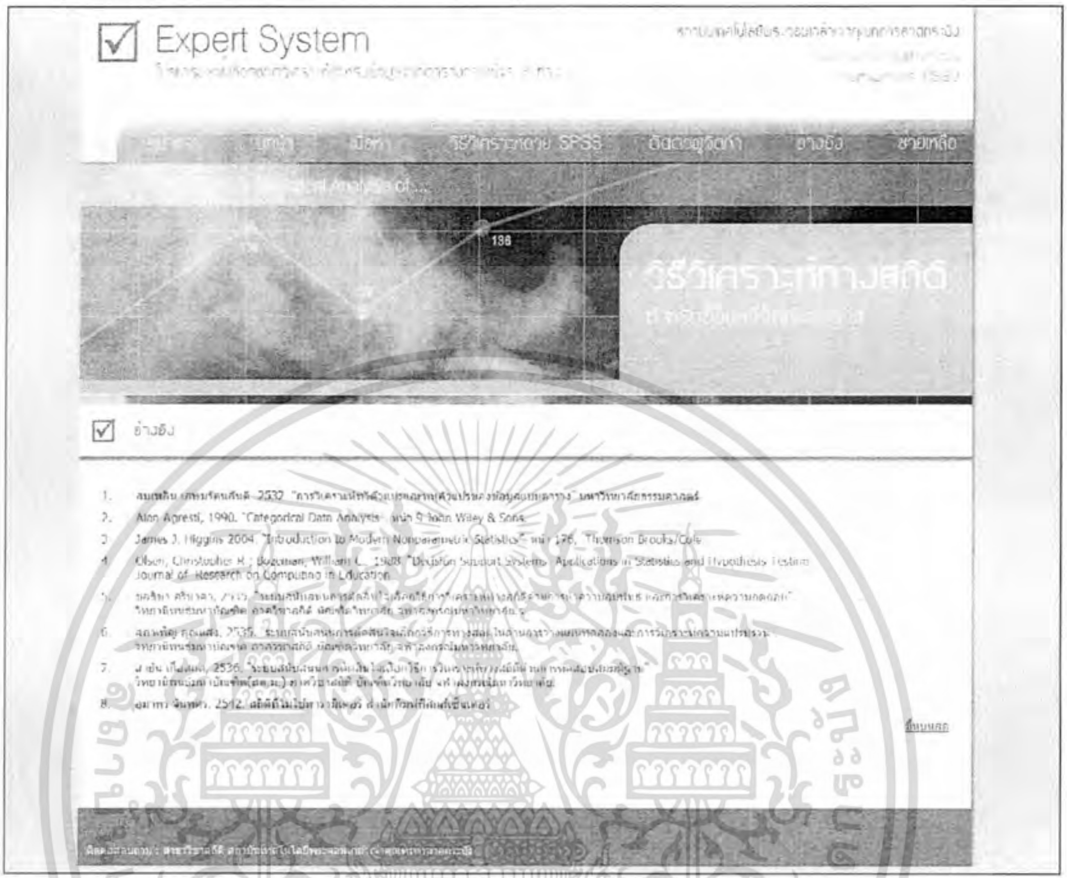
4.5 ติดต่อผู้จัดทำ เมื่อเลือกหัวข้อนี้จะแสดงการติดต่อกับผู้จัดทำดังรูป 4-32



รูปที่ 4-32 หน้าต่างติดต่อผู้จัดทำ

รูปที่ 4-32 จะแสดง e-mail address ของผู้จัดทำหากผู้ใช้มีข้อแนะนำหรือข้อสงสัยประการอย่างไร สามารถส่งข้อความเข้ามาได้ตามที่อยู่นี้

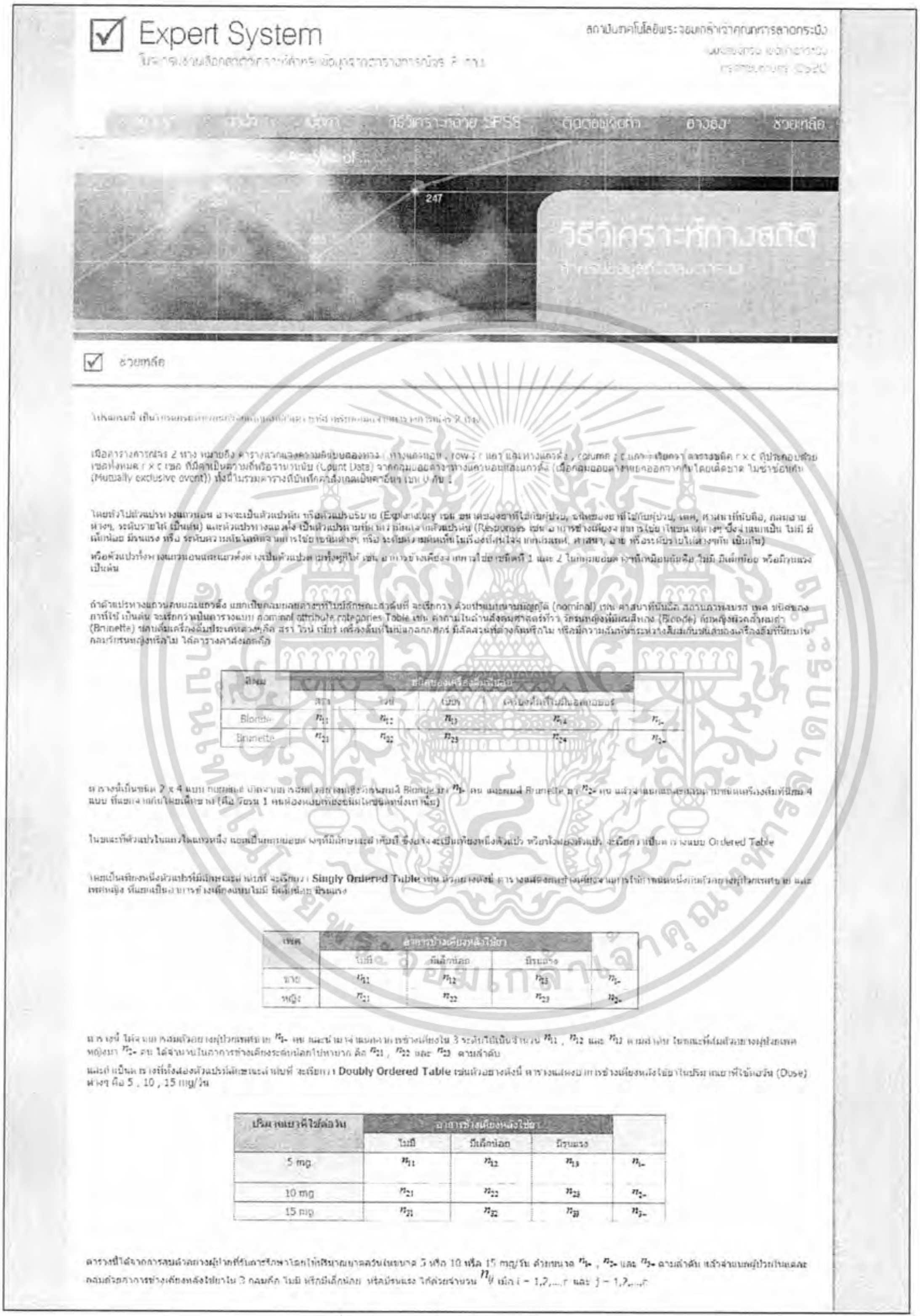
4.6 อ้างอิง เมื่อเลือกหัวข้อนี้จะแสดงเอกสารอ้างอิงที่ใช้ในการจัดทำโปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ในครั้งนี้ ดังรูป 4-33



1. [illegible] 2532. "การวิเคราะห์[illegible]" มหาวิทยาลัยธรรมศาสตร์
2. Alan Agresti, 1990. "Categorical Data Analysis" หน้า 9 John Wiley & Sons.
3. James J. Higgins 2004. "Introduction to Modern Nonparametric Statistics" หน้า 176, Thomson Brooks/Cole
4. Olsen, Christopher R.; Bozeman, William C. 1988. "Decision Support Systems: Applications in Statistics and Hypothesis Testing" Journal of Research on Computing in Education.
5. [illegible]
6. [illegible] 2535. [illegible]
7. [illegible] 2536. [illegible]
8. [illegible] 2512. [illegible]

รูปที่ 4-33 หน้าต่างของเอกสารอ้างอิง

4.7 ช่วยเหลือ เมื่อเลือกหัวข้อนี้จะแสดงรายละเอียดเพิ่มเติมจากส่วนของบทนำ ซึ่งเกี่ยวข้องกับการใช้โปรแกรมนี้ ดังรูป 4-34



รูปที่ 4-34 หน้าต่างของช่วยเหลือ

สรุปผลการศึกษาและข้อเสนอแนะ

สรุปผล

การศึกษาครั้งนี้คือการสร้างโปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทาง ซึ่งส่วนใหญ่เป็นสถิติวิเคราะห์แบบไม่ใช้พารามิเตอร์ โดยคำนึงถึงข้อกำหนดเบื้องต้น มาตราวัดข้อมูล ขนาดตัวอย่าง และอื่น ๆ ที่เหมาะสมกับการทดสอบนั้น ๆ เพื่อให้ผลสรุปที่ได้ สามารถเชื่อมั่นในคำตอบได้อย่างแท้จริง เนื่องจากการวิเคราะห์นั้น ๆ จะถูกนำไปใช้ได้ถูกต้องตามทฤษฎีทางสถิติ โปรแกรมนี้จะช่วยทำให้ผู้ใช้ได้รับความสะดวกในการเลือกวิธีทางสถิติให้ถูกต้อง รวมทั้งได้รับความรู้เกี่ยวกับเนื้อหาของวิธีวิเคราะห์นั้น ๆ โดยผู้จัดทำได้เสนอเนื้อหา วิธีการ ตลอดจนตัวอย่าง ให้ผู้ใช้สามารถทำความเข้าใจด้วยตนเองก่อน และทำให้อยู่ในรูปที่น่าสนใจ

หลังจากที่เลือกวิธีวิเคราะห์ได้ตามข้อมูลที่มีอยู่แล้ว(หรือตามวัตถุประสงค์ของงานวิจัยแล้ว) โปรแกรมนี้ยังจะแนะนำวิธีการวิเคราะห์ด้วยโปรแกรมสำเร็จรูป SPSS ไว้เป็นลำดับขั้นตอน เพื่อให้ผู้ใช้สามารถวิเคราะห์จนได้ผลสรุปตามวัตถุประสงค์ของงานวิจัยด้วย

ข้อเสนอแนะ

โปรแกรมช่วยเลือกสถิติวิเคราะห์สำหรับข้อมูลจากตารางการณ์จร 2 ทางที่จัดทำขึ้นนี้ สามารถนำไปพัฒนาต่อ โดยเน้นการนำเสนอที่น่าสนใจมากยิ่งขึ้น เช่น ทำให้อยู่ในรูปเคลื่อนไหวที่สามารถสื่อให้ผู้ใช้เข้าใจได้ง่ายขึ้น นอกจากนี้โปรแกรมยังเน้นการวิเคราะห์ด้วยโปรแกรม SPSS เท่านั้น ซึ่งจะทำให้ไม่ครอบคลุมทุกการทดสอบ อาจพัฒนาเพิ่มขึ้นโดยแนะนำให้ผู้ใช้วิเคราะห์ด้วยโปรแกรมอื่นๆ เช่น MINITAB เป็นต้น

บรรณานุกรม

1. สมเพลิน เกษมรัตนสันติ 2532 “การวิเคราะห์ทวิตัวแปรและพหุตัวแปรของข้อมูลแบบตาราง” มหาวิทยาลัยธรรมศาสตร์
2. Alan Agresti, 1990. “Categorical Data Analysis”. หน้า 9 John Wiley & Sons.
3. James J. Higgins 2004. “Introduction to Modern Nonparametric Statistics”. หน้า 176, Thomson Brooks/Cole.
4. Olsen, Christopher R.; Bozeman, William C., 1988. “Decision Support Systems: Applications in Statistics and Hypothesis Testing. Journal of Research on Computing in Education.
5. ชลธิชา ศรีนาคา, 2535. “ระบบสนับสนุนการตัดสินใจเลือกวิธีการวิเคราะห์ทางสถิติด้านการหาความสัมพันธ์ และการวิเคราะห์ความถดถอย”. วิทยานิพนธ์มหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
6. สุภาเพ็ญ คูณแสง, 2535. “ระบบสนับสนุนการตัดสินใจเลือกวิธีการทางสถิติในด้านการวางแผนทดลองและการวิเคราะห์ความแปรปรวน”, วิทยานิพนธ์มหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
7. สายัน เกื้อสกุล, 2536. “ระบบสนับสนุนการตัดสินใจเลือกวิธีการวิเคราะห์ทางสถิติด้านการทดสอบสมมติฐาน” วิทยานิพนธ์มหาบัณฑิต(สต.ม.) ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
8. อุมาพร จันทศร. 2542. สถิติที่ไม่ใช้พารามิเตอร์ สำนักพิมพ์ฟิสิกส์เซ็นเตอร์