

การเปรียบเทียบตัวลัทธิที่ใช้วัดความสัมพันธ์ของตัวแปรแบบแบ่งกลุ่ม



นายประวิทย์ วชิระจงกล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาตรีศึกษาศาสตร์บัณฑิต

ภาควิชาลัทธิ

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2530

ISBN 974-567-992-5

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

012894

T 10299622

A COMPARISON OF STATISTICS FOR MEASURING
CORRELATION OF CATEGORICAL VARIABLES

Mr, Prawit Vachirajongkol

A Thesis submitted in Partial Fulfilment of the Requirements
for the Degree of Master of Science
Department of Statistics
Graduate School

Chulalongkorn University

1987

หัวข้อวิทยานิพนธ์

การเปรียบเทียบตัวสถิติที่ใช้วัดความสัมพันธ์ของตัวแปรแบบแบ่งกลุ่ม

โดย

นายประวิทย์ วัชรจงกล

ภาควิชา

สถิติ

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร



บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

[Handwritten signature]
.....คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.ถาวร วัชรารักษ์)

คณะกรรมการสอบวิทยานิพนธ์

[Handwritten signature]
.....ประธานกรรมการ
(รองศาสตราจารย์ มณฑา พิววิไล)

[Handwritten signature]
.....กรรมการ
(รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร)

[Handwritten signature]
.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ธีระพร วีระถาวร)

[Handwritten signature]
.....กรรมการ
(อาจารย์ ดร.กัลยา วานิชย์บัญชา)

หัวข้อวิทยานิพนธ์ การเปรียบเทียบตัวสถิติที่ใช้วัดความสัมพันธ์ ของตัวแปรแบบแบ่ง กลุ่ม
 ชื่อผู้คิด นายประวิทย์ วชิระจงกล
 อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร. ลีระชัย พิศาลบุตร
 ภาควิชา สถิติ
 ปีการศึกษา 2529



บทคัดย่อ

ในการศึกษาความสัมพันธ์ระหว่างตัวแปรว่ามีระดับความสัมพันธ์กัน มากหรือน้อยเพียงไร นั้น มีมาตรวัดความสัมพันธ์ได้หลายแบบ ซึ่งขึ้นอยู่กับข้อมูลที่น่าสนใจศึกษาว่า เป็นตัวแปรประเภทใด ซึ่งสามารถแบ่งประเภทของตัวแปรออกได้เป็น 4 แบบคือ แบบอัตราส่วน แบบช่วง แบบอันดับ และแบบแบ่งกลุ่ม ซึ่งตัวแปรแต่ละประเภทจะมีมาตรวัดระดับความสัมพันธ์ที่แตกต่างกันออกไป และข้อมูลแต่ละประเภทก็จะมีมาตรวัดความสัมพันธ์กันหลายแบบ ซึ่งตัวแปรแบบแบ่งกลุ่มนี้มีมาตรวัดความสัมพันธ์ได้หลายแบบเช่น สัมประสิทธิ์ฟาย สัมประสิทธิ์เงื่อนโยชของเพียร์สัน สัมประสิทธิ์เงื่อนโยชของยูโพร สัมประสิทธิ์เงื่อนโยชของคราแมร์ สัมประสิทธิ์การทนายของกัทแมน และสัมประสิทธิ์แบบกุตแมนและครัลคัล วิทยานิพนธ์นี้จะทำการศึกษ เปรียบเทียบดูว่าตัวสถิติตัวใดจะใช้วัดความสัมพันธ์ของตัวแปรแบบแบ่งกลุ่มได้ดีกว่ากัน ในการศึกษาวิจัยได้ทำการจำลองข้อมูลขึ้นโดยวิธีการจำลอง โดยให้อุณหภูมิจำลองขึ้นมาโดยเครื่องคอมพิวเตอร์มีการแจกแจงแบบปกติสองตัวแปร ที่มีค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร (ρ) ที่ระดับ 0.1 0.3 0.5 0.7 และ 0.9 โดยให้อุณหภูมิตัวอย่างแต่ละชุดมีขนาด 20 30 50 100 200 และ 500 และขนาดของตารางการฉีกรังที่จะใช้ในการศึกษาวิจัยมีขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5 โดยแต่ละขนาดของตัวอย่าง และแต่ละขนาดของตารางการฉีกรังที่มีค่าสัมประสิทธิ์สหสัมพันธ์ระดับต่าง ๆ จะทำการจำลองซ้ำ ๆ กัน 500 ครั้ง โดยแต่ละครั้งจะทำการคำนวณหาตัวสถิติที่ใช้วัดความสัมพันธ์ทั้ง 6 ตัว แล้วนำค่าสถิติที่ได้มาคำนวณหาความคลาดเคลื่อนกำลังสองเฉลี่ย เพื่อเปรียบเทียบดูว่าตัวสถิติทั้ง 6 ตัวนั้น ตัวใดจะให้ผลที่ดีที่สุด ซึ่งผลการวิจัยพบว่า ที่ระดับของสัมประสิทธิ์สหสัมพันธ์ 0.1 และ 0.3 นั้น ตัวสถิติที่ใช้วัดความสัมพันธ์ได้ดีพอ ๆ กันคือ สัมประสิทธิ์ฟาย สัมประสิทธิ์เงื่อนโยชของยูโพร สัมประสิทธิ์เงื่อนโยชของคราแมร์ และสัมประสิทธิ์แบบกุตแมน

และครัสคัล แต่ถ้ำระดับของสัมประสิทธิ์สหสัมพันธ์เป็น 0.5 และขนาดตัวอย่างต่ำกว่า 50 ลงไปแล้ว
ในบางกรณี สัมประสิทธิ์ฟาย สัมประสิทธิ์เจอนโยของชู่โพร และสัมประสิทธิ์เจอนโยของคราแมร์จะ
ใช้วัดความสัมพันธ์ได้ดี แต่ส่วนมากแล้วที่ระดับของสัมประสิทธิ์สหสัมพันธ์ที่ต่ำกว่า 0.5
สัมประสิทธิ์เจอนโยของเพียร์สันจะให้ผลดีกว่าตัวสถิติอื่น ๆ และที่ระดับของสัมประสิทธิ์สหสัมพันธ์
ที่สูงกว่า 0.5 สัมประสิทธิ์เจอนโยของเพียร์สันจะให้ผลดีกว่าตัวสถิติอื่น ๆ ในทุกกรณี



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Thesis Title A Comparison of statistics for measuring
correlation of categorical variables

Name Prawit Vachirajongkol

Thesis Advisor Associate Professor Sorachai Phisalbutra ; Ph.D.

Department Statistics

Academic Year 1986



ABSTRACT

In measuring the strength of the relationship between variables, there are various statistics available. Some statistics are more appropriate than the others of which depends on the level of measurement of the study variables. In general, there are four levels of measurement: ratio scale, interval scale, ordinal scale and nominal scale. For any level of measurement, there are also various statistics available.

This thesis is focus on comparing the efficient of six statistics; Phi coefficient, Pearson's contingency coefficient Tschuprow's contingency coefficient, Cramer's contingency coefficient, Guttman's coefficient of optimal predictability and Goodman and Kruskal's Tau, which are used for measuring the strength of the relationship between variables in nominal scale. This study makes use of computer in generating data from bivariate normal distributions which have correlation coefficient at levels of 0.1, 0.3, 0.5, 0.7 and 0.9 At each level of correlation coefficient, the samples of size 20, 30, 50, 100, 200 and 500 are set in the forms of 2x2, 2x3, 2x4, 3x3, 3x4, 3x5, 4x4, 4x5 and 5x5 contingency tables. Each case repeat by 500 times and each time calculates all the six statistics. The researcher use mean square error to compare the six statistics which is the best.

The result of research find that: For all cases of correlation coefficient at levels of 0.1 and 0.3; Phi coefficient, Tschuprow's contingency coefficient, Cramer's contingency coefficient and Guttman coefficient of optimal predictability are equally appropriate. For some cases of correlation coefficient at level of 0.5 and sample size less than 50, Phi coefficient, Tschuprow's contingency coefficient and Cramer's contingency coefficient are appropriate. But in most case at this level of correlation coefficient, Pearson's contingency coefficient is most appropriate. And for all cases of correlation coefficient at levels higher than 0.5, Pearson's contingency coefficient is more appropriate than any other.



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จเรียบร้อยลงไปได้ด้วยความกรุณาอย่างสูงจาก รองศาสตราจารย์ ดร.สรชัย พิศาลบุตร ที่ได้กรุณารับเป็นที่ปรึกษาพร้อมทั้งให้คำแนะนำ และแก้ไขข้อบกพร่องในการทำวิทยานิพนธ์ด้วยดีตลอดมา ผู้เขียนขอกราบขอบพระคุณมา ณ ที่นี้

ขอกราบขอบพระคุณ คุณพ่อ คุณแม่ ครู และอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาให้ ขอขอบคุณ คุณบุญรับ วชิระจงกล และ ต.ญ.อชิรา วชิระจงกล พร้อมทั้งพี่น้องและเพื่อนร่วมงานทุกท่านที่ได้ให้กำลังใจ แก่ผู้เขียน จนทำให้วิทยานิพนธ์นี้สำเร็จลุล่วงไปด้วยดี

ประวิทย์ วชิระจงกล



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	ฉ
กิตติกรรมประกาศ.....	ช
รายการตารางประกอบ.....	ฅ
รายการแผนผังประกอบ.....	ฉ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	6
1.3 สัมมติฐานของการวิจัย.....	6
1.4 ขอบเขตของการวิจัย.....	6
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	6
บทที่ 2. ทฤษฎีที่เกี่ยวข้องในการวิจัย.....	7
2.1 ตัวสถิติที่ใช้สำหรับทดสอบความเป็นอิสระกันของข้อมูล ในตารางการถักร.....	7
2.2 สัมประสิทธิ์ฟาย.....	7
2.3 สัมประสิทธิ์เจอนโยของเพียร์สัน.....	8
2.4 สัมประสิทธิ์เจอนโยของชูโพร.....	9
2.5 สัมประสิทธิ์เจอนโยของคราแมร์.....	10
2.6 มาตราวัดเชิงลดสัดส่วนความคลาดเคลื่อน.....	11
2.7 สัมประสิทธิ์การทำนายของกัทแมน.....	12
2.8 สัมประสิทธิ์แบบกุดแมนและครัลล์.....	14
2.9 ความสัมพันธ์ระหว่างสัมประสิทธิ์ฟาย สัมประสิทธิ์เจอนโย ของชูโพรและสัมประสิทธิ์เจอนโยของคราแมร์.....	16
2.10 ความคลาดเคลื่อนกำลังสองเฉลี่ย.....	21

สารบัญ (ต่อ)

	หน้า
2.11 การสร้างข้อมูลที่มีการแจกแจงแบบปกติสองตัวแปรโดยวิธีการ ฮิมูเลชัน.....	21
2.12 การแจกแจงแบบปกติสองตัวแปร.....	23
บทที่ 3 วิธีดำเนินการวิจัย และผลการวิจัย.....	25
3.1 วิธีดำเนินการวิจัย.....	25
3.2 ผลการวิจัย.....	26
บทที่ 4 สรุปผลการวิจัยและข้อเสนอแนะ.....	54
4.1 สรุปผลการวิจัย.....	54
4.2 ข้อเสนอแนะ.....	60
บรรณานุกรม.....	61
ภาคผนวก.....	63
ภาคผนวก ก.....	64
ภาคผนวก ข.....	71
ประวัติผู้เขียน.....	101

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการตารางประกอบ (ต่อ)

ตารางที่		หน้า
3.40	แสดงค่าคลาดเคลื่อนกำลังส่องเจสีย่าแนกตาม สัมประสิทธิ์สี่เหลี่ยม และตัวสถิติที่ใช้ เมื่อใช้ตัวอย่างขนาด 500 และตารางการแจกแจงขนาด 4x4	51
3.41	แสดงค่าคลาดเคลื่อนกำลังส่องเจสีย่าแนกตาม สัมประสิทธิ์สี่เหลี่ยม และตัวสถิติที่ใช้ เมื่อใช้ตัวอย่างขนาด 500 และตารางการแจกแจงขนาด 4x5	52
3.42	แสดงค่าคลาดเคลื่อนกำลังส่องเจสีย่าแนกตาม สัมประสิทธิ์สี่เหลี่ยม และตัวสถิติที่ใช้ เมื่อใช้ตัวอย่างขนาด 500 และตารางการแจกแจงขนาด 5x5	52
4.1	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 20 และตารางการแจกแจงขนาด 2x2	55
4.2	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 30 และตารางการแจกแจงขนาด 2x2 2x3 2x4 3x3 และ 3x4	55
4.3	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 50 และตารางการแจกแจงขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5	56
4.4	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 100 และตารางการแจกแจงขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5	57
4.5	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 200 และตารางการแจกแจงขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5	58
4.6	แสดงตัวสถิติที่ให้ค่าคลาดเคลื่อนกำลังส่องเจสีย่าที่ลู่ต่ำ แนกตาม สัมประสิทธิ์สี่เหลี่ยม ตัวอย่างขนาด 500 และตารางการแจกแจงขนาด 2x2 2x3 2x4 3x3 3x4 3x5 4x4 4x5 และ 5x5	59

รายการแผนผังประกอบ

แผนผังที่

หน้า

1	การสร้างตัวแปรคู่ที่มีการแจกแจงแบบปกติสองตัวแปร.....	22
---	--	----



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย