

170719

ประกาศิต สิงคะตริยะ : การทดสอบสมมติฐานเกี่ยวกับความเป็นอิสระของตัวแบบล็อกการิทึมเชิงเส้นที่มีการแจกแจงพหุนาม. (HYPOTHESES TESTING OF INDEPENDENCE FOR MULTINOMIAL LOG-LINEAR MODELS) อ.ที่ปรึกษา : รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา, 254 หน้า. ISBN 974-17-7098-7

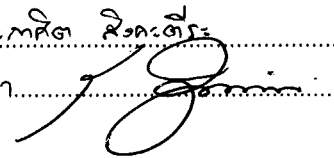
การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการทดสอบสมมติฐานเกี่ยวกับความเป็นอิสระของตัวแปรเชิงกลุ่ม 3 ตัวแปร ในตารางการถ้อย 3 ทาง ขนาด $2 \times 2 \times 2$ และศึกษาความเป็นไปได้ที่จะนำวิธีการมอนติคาร์โลมาใช้แทนวิธีการทดสอบด้วยตัวสถิติอัตราส่วนภาวะน่าจะเป็น วิธีการทดสอบที่นำมาศึกษาเปรียบเทียบกับ 3 วิธี ได้แก่ วิธีการทดสอบด้วยตัวสถิติอัตราส่วนภาวะน่าจะเป็น วิธีการทดสอบด้วยตัวสถิติของเซลเทอร์แมน และวิธีการมอนติคาร์โล โดยวิธีการมอนติคาร์โลเป็นวิธีการทดสอบสมมติฐานที่ไม่อาศัยค่าวิกฤติของตัวสถิติทดสอบ แต่จะทำการเปรียบเทียบค่าสถิติที่คำนวณได้จากข้อมูลตัวอย่างที่ต้องการศึกษากับค่าสถิติที่คำนวณได้จากข้อมูลเทียมแต่ละชุดที่จำลองขึ้นโดยการสุ่มตัวอย่างซ้ำภายใต้ข้อกำหนดตามสมมติฐานว่าง เพื่อนำไปสู่การหาค่า P-value ผู้วิจัยได้ทำการศึกษาในแต่ละตัวแบบที่เป็นไปได้และเป็นตัวแบบที่มีรูปแบบปิดของค่าความถี่คาดหวัง เพื่อให้สามารถจำลองข้อมูลได้ถูกต้อง ซึ่งตัวแบบต่าง ๆ ประกอบด้วยตัวแบบที่เป็นอิสระอย่างสมบูรณ์ ตัวแบบความเป็นอิสระร่วม และตัวแบบความเป็นอิสระอย่างมีเงื่อนไข และได้กำหนดระดับความสัมพันธ์ให้กับตัวแปรที่มีความสัมพันธ์ร่วมกัน จากนั้นไปหามากคือ 0.25, 0.50 และ 0.75 ขนาดตัวอย่างที่ใช้คือ 40, 60, 80, 120, 160, 200 และ 240 และระดับนัยสำคัญที่ใช้คือ 0.05 และ 0.01 เกณฑ์ที่ใช้ในการเปรียบเทียบคือการพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และ ค่าอำนาจการทดสอบ การวิจัยครั้งนี้ได้จำลองข้อมูลโดยกระทำซ้ำแบบเทคนิคมอนติคาร์โล (Monte Carlo simulation technique) จำนวน 500 ครั้ง ซึ่งในแต่ละครั้งจะกระทำซ้ำแบบมอนติคาร์โลอีก 500 ครั้ง เพื่อสร้างชุดข้อมูล 500 ชุด เพื่อคำนวณค่า P-value ของวิธีการทดสอบแบบมอนติคาร์โล ผลการวิจัยสามารถสรุปได้ดังนี้

วิธีการทดสอบทั้ง 3 วิธีสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ในทุกสถานการณ์

วิธีการมอนติคาร์โลมีแนวโน้มที่จะให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาคือวิธีการทดสอบด้วยตัวสถิติอัตราส่วนภาวะน่าจะเป็น และวิธีการทดสอบด้วยตัวสถิติของเซลเทอร์แมน และโดยเฉพาะอย่างยิ่งในกรณีตัวอย่างขนาดเล็ก วิธีการมอนติคาร์โลจะมีค่าอำนาจการทดสอบสูงกว่าวิธีการอื่นอย่างชัดเจน ทั้งนี้เนื่องจากแนวคิดของวิธีการมอนติคาร์โลที่มีการจำลองข้อมูลเทียมหลายชุดภายใต้สมมติฐานว่างขึ้นมาพิจารณานั้น เสมือนเป็นการสร้างขอบเขตที่เป็นไปได้ของข้อมูลตามข้อกำหนดในสมมติฐานว่าง เพื่อเปรียบเทียบกับข้อมูลตัวอย่างที่ศึกษา จึงมีความถูกต้องมากกว่าการทดสอบสมมติฐานโดยทั่วไปซึ่งเป็นการเปรียบเทียบระหว่างค่าสถิติที่คำนวณได้กับค่าวิกฤติเพียงครั้งเดียว นอกจากนี้ยังพบว่าค่าอำนาจการทดสอบของวิธีการทดสอบทั้ง 3 วิธีแปรผันตามขนาดตัวอย่าง ระดับความสัมพันธ์ของตัวแปรร่วม และระดับนัยสำคัญ

โดยสรุป ผู้วิจัยพบว่ามีความเป็นไปได้ที่จะนำเอาการทดสอบสมมติฐานด้วยวิธีการมอนติคาร์โลมาใช้ทดสอบสมมติฐานเกี่ยวกับความเป็นอิสระของตัวแบบล็อกการิทึมเชิงเส้น แทนการทดสอบทั้งสองวิธีดังกล่าวใน 3 ตัวแบบข้างต้นได้ โดยเฉพาะอย่างยิ่งในกรณีตัวอย่างขนาดเล็ก

ภาควิชา.....สถิติ.....
สาขาวิชา.....สถิติ.....
ปีการศึกษา.....2547.....

ลายมือชื่อนิสิต.....ประกาศิต สิงคะตริยะ.....
ลายมือชื่ออาจารย์ที่ปรึกษา.....

#4582277626 : MAJOR STATISTICS

KEY WORDS : Log-linear models / Three-way contingency table / Multinomial sampling scheme / Hypotheses testing of independence / Monte Carlo hypothesis testing / Likelihood ratio statistic / Zelterman's statistic.

PRAKASIT SINGKATEERA : HYPOTHESES TESTING OF INDEPENDENCE FOR MULTINOMIAL LOG-LINEAR MODELS. THESIS ADVISOR : ASSOC.PROF. SUPOL DURONGWATANA , Ph.D. , 254 pp. ISBN 974-17-7098-7

The objective of this research is to study the hypotheses testing of independence of 3 categorical variables in three-way contingency table , of which size is $2 \times 2 \times 2$, using Log-linear models. The study also investigates whether the Monte Carlo hypothesis testing is appropriate to use instead of the Likelihood ratio statistic. There are three testing methods to be compared, the Likelihood ratio statistic, the Zelterman's statistic, and the Monte Carlo hypothesis testing. The Monte Carlo method does not use any critical value of statistic; the comparison of the calculated Likelihood ratio statistic from the sampled data of interest and the Likelihood ratio statistics from pseudo-samples, resampled from a model under the null hypothesis is used instead. This leads to obtain the P – value. We consider only the possible models that could be shown in closed forms of expected frequency in order to avoid the simulation problem. The models in this study are the Model of mutual independence, the Models of joint independence, and the Models of conditional independence. The association coefficient of any two variables is set to three levels , 0.25 , 0.50 , and 0.75. The sample sizes used in the study are 40, 60, 80, 120, 160, 200, and 240. The significance levels of the study are 0.05 and 0.01. The comparison criterion is their ability to control the probability of type I error and their power of the test. The data in this study are simulated 500 times using Monte Carlo simulation technique concept, in each time, the data is simulated 500 times in order to compute Monte Carlo P – value. The results could be concluded as follows:

In all situations , all the three methods can completely control the probability of type I error.

The Monte Carlo method tends to give the highest power, while the Likelihood ratio statistic and the Zelterman's statistic give the lower power respectively especially with the small data sampling. Since the Monte Carlo resampling concept, as a finding technique of the possible set of data under the null hypothesis can enhance a possible parameter when comparing with the sample data. Hence, the Monte Carlo method will give more precision than the any other methods that perform a test by comparing the calculated statistic with the critical value only once. Moreover, it is found that the power of the test of all three methods vary in accordance with the sample sizes , the association coefficients of two variables, and the significance levels.

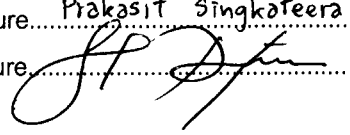
In conclusion, it is possible to use the Monte Carlo hypothesis testing for testing of independence for the three models of the Log-linear models instead of the two tests , in particular the cases of small sample sizes.

Department.....Statistics.....

Field of Study.....Statistics.....

Academic Year.....2004.....

Student's signature.....Prakasit Singkateera.....

Advisor's signature..........