การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท สำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง

นางสาวสุภิญญา คำมั่น

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555

THE CUT-OFF POINT ESTIMATION OF BINARY LOGISTIC REGRESSION MODEL FOR PREDICTIVE CLASSIFICATION USING PROBIT FUNCTION AS A LINK FUNCTION

Miss Supinya Khammun

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Statistics
Department of Statistics
Faculty of Commerce and Accountancy
Chulalongkorn University
Academic Year 2012

หัวข้อวิทยานิพนธ์	การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติก แบบ 2 ประเภท สำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง
โดย	นางสาวสุภิญญา คำมั่น
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

……………………………………คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร.พสุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

…………………………………………ประธานกรรมการ

(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

…………………………………………... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

…………………………………………กรรมการ

(อาจารย์ ดร.อนันตณัฐ กันต์ธัญญูรัตน์)

…………………………………………กรรมการภายนอกมหาวิทยาลัย

(อาจารย์ ดร.อรุณี กำลัง)

สุภิญญา คำมั่น: การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง. (THE CUT-OFF POINT ESTIMATION OF BINARY LOGISTIC REGRESSION MODEL FOR PREDICTIVE CLASSIFICATION USING PROBIT FUNCTION AS A LINK FUNCTION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ.ดร. สุพล ดุรงค์วัฒนา, 101 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมที่สุดสำหรับตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการจำแนกข้อมูลโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง โดยปัจจัยที่สนใจศึกษาในงานการวิจัยครั้งนี้ คือ จำนวนตัวแปรอิสระเป็น 1, 2, 3, 4 และ 5 ขนาดตัวอย่างเป็น 50, 100, 150, 200 และ 250 สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเป็น 0.1, 0.5 และ 0.9 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น 3 ระดับ คือ ความสัมพันธ์กันในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) และระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) ซึ่งข้อมูลทั้งหมดจำลองโดยเทคนิคมอนติคาร์โล ด้วยโปรแกรม R การหาค่าจุดแบ่งจะใช้ทฤษฎีของ Hadjicostas P. (2006) ผลการวิจัยสรุปได้ดังนี้

กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง แต่ปัจจัยอื่นๆ คงที่ พบว่า ที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ เท่ากับ 0.5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 แต่ค่าเฉลี่ยจะมีค่าต่ำกว่าค่า 0.5 เมื่อระดับความสัมพันธ์สูงและขนาดตัวอย่างใหญ่ และที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจอื่นๆ ส่วนใหญ่ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลง แต่ปัจจัยอื่นๆ คงที่ พบว่า ที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ เท่ากับ 0.5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 และที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจอื่นๆ ส่วนใหญ่ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลง แต่ปัจจัยอื่นๆ คงที่ พบว่า ส่วนใหญ่ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ พบว่า ส่วนใหญ่ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ภาควิชาสถิติ........สถิติ.................... ลายมือชื่อนิสิต..

สาขาวิชา..................สถิติ.................... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก....................................

ปีการศึกษา.........2555...................

5281936326: MAJOR STATISTICS
KEYWORDS: CUT-OFF POINT / BINARY LOGISTIC REGRESSION / CLASSIFICATION ERROR RATE / PROBIT FUNCTION

SUPINYA KHAMMUN: THE CUT-OFF POINT ESTIMATION OF BINARY LOGISTIC REGRESSTION MODEL FOR PREDICTIVE CLASSIFICATION USING PROBIT FUNCTION AS A LINK FUNCTION. ADVISOR: ASSOC. PROF. SUPOL DURONGWATANA, Ph.D., 101 pp.

The object of this study is to find out the optimal cut-off point estimation of binary logistic regression model for predictive classification using probit function. The interesting factors are the numbers of independent variables (p) are 1, 2, 3, 4 and 5, the sample size (n) are 50, 100, 150, 200 and 250, the failure rate (a) are 0.1, 0.5 and 0.9 and the degree of multicollinearity among independent variables with 3 levels; low level ($0 < Max\{|r_{ij}|\} \leq 0.30$), medium level ($0.30 < Max\{|r_{ij}|\} \leq 0.60$) and high level ($0.60 < Max\{|r_{ij}|\} \leq 0.90$). The data in all situations are generated using Monte Carlo technique through R-program. The cut-off point is captured using Hadjicostas P. (2006) theory. The results can be summarized as follow:

As the number of independent variables change and the other factors are kept constant, with the failure rate equal to 0.5, the mean value of the cut-off point converges to 0.5 and the mean value of the cut-off point less than 0.5 when the sample size is big and the degree of multicollinearity among independent variables is high level and the other failure rates, the mean value of the cut-off point mostly converge to value of 0.5. As sample size change and the other factors are kept constant, with the failure rate equals to 0.5, the mean value of the cut-off point converges to 0.5 and with other failure rates, the mean value of the cut-off point mostly converges to 0.5. As the failure rate change and the other factors are kept constant, the mean value of the cut-off point mostly converges to 0.5. As the degree of multicollinearity among independent variables changed and the other factors are kept constant, the mean value of the cut-off point mostly converges to 0.5.

Department: Statistics Student's Signature……………………………………...

Field of Study: Statistics Advisor's Signature…………………………………......

Academic Year: 2012

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยความช่วยเหลือเป็นอย่างดีจาก รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำ/ปรึกษา ตลอดจนช่วยตรวจสอบ แก้ไขข้อบกพร่องต่างๆ จนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความเคารพไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาตราจารย์ ดร. ธีระพร วีระถาวร ประธานกรรมการ และอาจารย์ ดร. อนันตณัฐ กันต์ธัญญรัตน์ กรรมการ ที่ให้คำแนะนำในการทำงานวิจัยครั้งนี้ ซึ่งทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์

ผู้วิจัยขอกราบขอบพระคุณ อาจารย์ ดร.อรุณี กำลัง ที่ท่านเสียสละเวลาอันมีค่ามาเป็นกรรมการภายนอกมหาวิทยาลัย ช่วยตรวจสอบวิทยานิพนธ์ ซึ่งทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากยิ่งขึ้น

ท้ายนี้ ผู้วิจัยขอขอบพระคุณ บิดา มารดา ผู้ให้การสนับสนุนทั้งทุนการศึกษา และกำลังใจเสมอมาจนสำเร็จการศึกษา ขอบคุณพี่ๆ และน้องๆ และเพื่อนๆ ทุกคนที่ให้คำปรึกษาและเป็นกำลังใจให้เสมอ

สารบัญ

หน้า

สารบัญตาราง

สารบัญภาพ

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันงานวิจัยในด้านต่างๆ ไม่ว่าจะเป็นด้านเศรษฐศาสตร์ ด้านบริหารธุรกิจ ด้านสังคมศาสตร์ รวมถึงด้านการแพทย์ ได้นำเอาเทคนิคการวิเคราะห์ทางสถิติมาเป็นเครื่องมือในการวิเคราะห์ข้อมูล เพื่อประกอบการตัดสินใจ ซึ่งข้อมูลส่วนใหญ่ที่นำมาวิเคราะห์มักจะมีข้อมูลเชิงคุณภาพเข้ามาเกี่ยวข้อง ซึ่งก็คือ การเกิดเหตุการณ์ที่สนใจ (Success) และการไม่เกิดเหตุการณ์ที่สนใจ (Failure) ดังนั้น ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท (Binary Logistic Regression Model) ซึ่งใช้ในการพยากรณ์ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า (Dichotomy or Binary Variable) คือค่า 0 และ 1 ว่าจะอยู่ในกลุ่มใดกลุ่มหนึ่งใน 2 กลุ่มโดยใช้ตัวแปรอิสระเป็นตัวพยากรณ์ จึงถูกนำมาใช้ประโยชน์อย่างแพร่หลาย เช่น การจำแนกกลุ่มการเป็นโรคหัวใจหรือการไม่เป็นโรคหัวใจ การเลือกซื้อสินค้าหรือการไม่เลือกซื้อสินค้า และลูกค้าเป็นลูกหนี้ที่มีปัญหาหรือลูกค้าเป็นลูกหนี้ที่ดี เป็นต้น

ในการวิเคราะห์การถดถอยโลจิสติกแบบ 2 ประเภทนั้นจะมีการใช้ฟังก์ชันเชื่อมโยง (Link function) ในการวิเคราะห์ ซึ่งฟังก์ชันเชื่อมโยงที่นิยมใช้ได้แก่ ฟังก์ชันโพรบิต (Probit Function) ฟังก์ชันโลจิท (Logit Function) และฟังก์ชันคอมพลีเมนทารี ล็อก-ล็อก (Complementary log-log Function) ซึ่งการพิจารณาว่าจะใช้ฟังก์ชันใดในการวิเคราะห์การถดถอยโลจิสติก เราจะพิจารณาจากการกระจายของข้อมูลในแต่ละกลุ่มเป็นหลัก โดยฟังก์ชันโพรบิตจะใช้เมื่อตัวแปรแฝงมีการแจกแจงแบบปกติ ฟังก์ชันโลจิทจะใช้เมื่อข้อมูลแต่ละกลุ่มมีการกระจายเท่าๆกัน ส่วนฟังก์ชันคอมพลีเมนทารี ล็อก-ล็อกจะใช้เมื่อข้อมูลในอันดับสูงมีการกระจายสูงกว่ากลุ่มอื่นๆ โดยในงานวิจัยครั้งนี้ผู้วิจัยเลือกใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง เนื่องจากมีข้อจำกัดที่น้อยกว่าฟังก์ชันอื่นๆ และสามารถนำไปประยุกต์ใช้กับงานวิจัยได้หลากหลายสาขา

งานวิจัยส่วนใหญ่ที่ใช้ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทในการวิเคราะห์ข้อมูลมักจะมีข้อตกลงเบื้องต้นโดยให้กลุ่มที่เกิดเหตุการณ์ที่สนใจมีโอกาสเกิดขึ้นเท่ากับกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจ โดยใช้จุดแบ่ง (cut-off point) ที่ 0.5 หรืออาจใช้จุดแบ่งค่าหนึ่งที่จะทำให้อัตราความผิดพลาดในการจำแนกกลุ่ม (Classification error rate) มีค่าต่ำสุดหรือสัดส่วนของความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด เนื่องจากข้อมูลถูกเลือกอย่างสุ่มจากลักษณะที่สนใจศึกษาแล้ว โดยผู้วิจัยส่วนใหญ่มักไม่คำนึงถึงจำนวนของตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งลักษณะของชุดข้อมูลเหล่านี้อาจมีผลกระทบ

ต่อจุดแบ่งสำหรับการประเมินการพยากรณ์การจำแนกกลุ่มและถ้ามีผลต่อการคัดเลือกจุดแบ่งแล้ว รูปแบบเหล่านั้นคืออะไร

ดังนั้น ผู้วิจัยจึงสนใจทำการศึกษาหาจุดแบ่งที่เหมาะสมที่สุดสำหรับตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยงที่ทำให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุด โดยพิจารณาจากปัจจัยต่างๆ คือ จำนวนของตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระ

1.2 วัตถุประสงค์ของการวิจัย

เพื่อหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง เมื่อชุดข้อมูลมีลักษณะดังนี้

- จำนวนตัวแปรอิสระเพิ่มขึ้น
- ขนาดตัวอย่างเพิ่มขึ้น
- สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น
- ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

และนำปัจจัยที่กล่าวข้างต้นมาพิจารณารวมกัน ซึ่งแปรเปลี่ยนไปพร้อมกัน

1.3 ขอบเขตของการวิจัย

การวิจัยครั้งนี้ มีขอบเขตของการวิจัยสำหรับการดำเนินการวิจัย ดังนี้

1. ศึกษาตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยงเพื่อหาจุดแบ่งที่เหมาะสมที่สุด

2. ตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มี 2 ค่า คือ 0 และ 1 โดยกำหนดสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) ในการวิจัยครั้งนี้คือ 0.1, 0.5 และ 0.9

3. จำนวนของตัวแปรอิสระ (p) ในการวิจัยครั้งนี้คือ 1, 2, 3, 4 และ 5

4. ขนาดตัวอย่าง (n) ในการวิจัยครั้งนี้คือ 50, 100, 150, 200 และ 250

5. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ใช้ค่าสูงสุดของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j ในการวัดระดับความสัมพันธ์ระหว่างตัวแปรอิสระ

$$Max \quad \{|r_{ij}|\} \; ; \; i=1,2,\ldots,p_1 \quad j=1,2,\ldots,p_2$$

โดยที่ r_{ij} คือ สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

ระดับความสัมพันธ์ระหว่างตัวแปรอิสระแบ่งเป็น 3 ระดับ ดังนี้

- ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ เมื่อ 0 < $Max\{|r_{ij}|\} \le$ 0.30
- ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง เมื่อ 0.30 < $Max\{|r_{ij}|\} \le$ 0.60
- ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง เมื่อ 0.60 < $Max\{|r_{ij}|\} \le$ 0.90

6. การแจกแจงของตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอร์ม

7. กำหนดค่าพารามิเตอร์เริ่มต้นของสมการการถดถอยเป็นค่าใดๆ ในการวิจัยครั้งนี้ กำหนดให้ $\beta_i = 10\,; i = 0,1,2,\ldots,p$ และ $\varepsilon_k \overset{iid}{\sim} N(0,500)$; $k = 1,2,\ldots,n$

8. กำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05

9. จำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation) โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

1.4 ข้อตกลงเบื้องต้น

การวิจัยครั้งนี้ มีข้อตกลงเบื้องต้นสำหรับการดำเนินการวิจัย ดังนี้

1. ศึกษาตัวแบบโพรบิตแบบ 2 ประเภท โดยมีรูปแบบคือ

จาก

$$
\begin{aligned}
\pi_i &= P(Y_i = 1) \\
&= P(Y_i^* > 0) \\
&= P(\varepsilon_i < \beta' X_i) \qquad ; i = 1,2,\ldots,n \\
&= \Phi(\beta' X_i) \\
&= \int_{-\infty}^{\beta X_i'} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon^2}{2}\right) d\varepsilon
\end{aligned}
$$

โดยที่ π_i คือ ค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

Y_i คือ ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1

Y_i^* คือ ตัวแปรแฝง (Latent Variable) เป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริงทราบเพียงแต่ผลที่เกิดขึ้น

$\Phi(.)$ คือ ฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน

ε คือ ตัวแปรสุ่มปกติมาตรฐาน โดย $\varepsilon_i \overset{iid}{\sim} N(0,1)$

β' คือ เวกเตอร์ของสัมประสิทธ์การถดถอยจำนวน k+1 ตัว ขนาด 1 × (k+1)

X_i คือ เวกเตอร์ของตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณจำนวน k ตัว ขนาด (k+1) ×1

2. ศึกษาจากตัวอย่างที่มีการแจกแจงแบบยูนิฟอร์ม ซึ่งมีฟังก์ชันความน่าจะเป็นของ x อยู่ในรูปของ

$$f(x;a,b) = \frac{1}{b-a} \quad ;a < x < b$$

เมื่อ a , b เป็นค่าคงที่ และ $a < b$

โดยมีค่าเฉลี่ยและค่าความแปรปรวนคือ

$$E(X) = \mu = \frac{a+b}{2}$$

$$Var(X) = \sigma^2 = \frac{(b-a)^2}{12}$$

1.5 **คำจำกัดความที่ใช้ในการวิจัย**

การวิจัยครั้งนี้ มีคำจำกัดความที่ใช้ในการวิจัย ดังนี้

1. ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท (Binary Logistic Regression Model) หมายถึง ตัวแบบที่ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม โดยที่ตัวแปรอิสระเป็นตัวแปรเชิงปริมาณหรือตัวแปรเชิงคุณภาพก็ได้ ส่วนตัวแปรตามจะเป็นตัวแปรเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า (dichotomous variable) เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรแล้ว จะนำไปใช้ในการพยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่ง

2. จุดแบ่ง (Cut-off point) หมายถึง ค่าความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนกกลุ่มของข้อมูลว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มที่เกิดเหตุการณ์ที่สนใจและกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจ

3. ความสัมพันธ์ระหว่างตัวแปรอิสระ(multicollinearity) หมายถึง สถานการณ์ที่ตัวแปรอิสระมีความสัมพันธ์กัน

4. การแจกแจงแบบเบอร์นูลลี (Bernoulli Distribution) หมายถึง ตัวแปรสุ่ม Y เรียกว่าตัวแปรสุ่มเบอร์นูลลี กล่าวคือ

ถ้า $Y = 0$ เมื่อ การทดลองไม่เกิดเหตุการณ์ที่สนใจ

และถ้า $Y = 1$ เมื่อ การทดลองเกิดเหตุการณ์ที่สนใจ

เมื่อ $0 < p < 1$ โดยที่ $P(Y=1) = p$ และ $P(Y=0) = 1-p$

อาจเขียนแทนด้วย $Y \sim Ber(p)$ ซึ่งฟังก์ชันความน่าจะเป็นอยู่ในรูปของ

$$P(Y = y) = p^{y}(1-p)^{1-y}$$

5. ช่วงความเชื่อมั่น (Confidence Interval) คือ ช่วงของการประมาณค่าเป็นช่วงที่อยู่รอบจุดของค่าประมาณ ซึ่งเป็นค่าเฉพาะที่บ่งบอกความเชื่อมั่นว่ามีค่าพารามิเตอร์อยู่ในช่วงนี้

6. ฟังก์ชันเชื่อมโยง (link function) คือ ส่วนที่ใช้อธิบายความสัมพันธ์ระหว่างองค์ประกอบเกี่ยวกับตัวแปรสุ่มและองค์ประกอบเชิงระบบเป็นการเชื่อมโยงระหว่างส่วนตัวแปรสุ่มและส่วนเชิงระบบ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท โดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง เมื่อชุดข้อมูลมีลักษณะดังนี้

- จำนวนตัวแปรอิสระเพิ่มขึ้น
- ขนาดตัวอย่างเพิ่มขึ้น
- สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น
- ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

2. สามารถใช้เพื่อเป็นแนวทางในการวิจัยต่อไป

บทที่ 2

แนวคิด ทฤษฎีและสถิติที่เกี่ยวข้อง

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง ซึ่งในบทนี้จะมีรายละเอียดเกี่ยวกับตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท ตัวแบบโพรบิตแบบ 2 ประเภท การประมาณค่าสัมประสิทธิ์การถดถอยในตัวแบบโพรบิตแบบ 2 ประเภทด้วยวิธีภาวะความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) และการหาจุดแบ่งโดยใช้ทฤษฎี Hadjicostas P.

2.1 ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท

ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์ในการจัดประเภท ดังนี้

$$Y_i = \tilde{X}_i^T \tilde{\beta} + \varepsilon_i ; i = 1,2,\ldots,n$$

$$\tilde{X}_i^T = (1 \quad X_{1i} \quad X_{2i} \quad \ldots \quad X_{pi}), \tilde{\beta}^T = (\beta_0 \quad \beta_1 \quad \beta_2 \quad \ldots \quad \beta_p)$$

เมื่อ

$$\Pr\left(Y_i = 1 \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \pi_i \, , \Pr\left(Y_i = 0 \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = 1 - \pi_i ; 0 < \pi_i < 1$$

เมื่อ Y_i มีการแจกแจงแบบเบอร์นูลลีที่เป็นอิสระกัน มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$$E\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \tilde{x}_i^T \tilde{\beta} \, , Var\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \pi_i (1 - \pi_i) ; i = 1,2,\ldots,n$$

ถ้าตัวแบบการถดถอยเป็น

$$Y_i = \tilde{X}_i^T \tilde{\beta} + \varepsilon_i \, , E\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \tilde{x}_i^T \tilde{\beta} ; i = 1,2,\ldots,n$$

เมื่อ $E\left(\varepsilon_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = 0 ; i = 1,2,\ldots,n$

2.2 ตัวแบบโพรบิตและการประมาณค่า

ตัวแบบโพรบิตเป็นตัวแบบที่ศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ เมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่มี 2 ลักษณะ และตัวแปรอิสระเป็นตัวแปรเชิงปริมาณหรือตัวแปรหุ่น นำไปใช้หาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ และใช้ในการพยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่งได้

จากตัวแปรตาม (Y_i) ซึ่งเป็นตัวแปรเชิงคุณภาพที่เป็นได้ 2 ค่า คือ 0 กับ 1 และ $X_{i1}, X_{i2},..., X_{ik}$ เป็นตัวแปรอิสระของค่าสังเกตที่ได้จากหน่วยตัวอย่างที่ i โดยที่ $i = 1,2,...,n$ โดยมีตัวแปรแฝง คือ Y_i^* ซึ่งเป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริง ทราบเพียงแต่ผลที่เกิดขึ้น โดย Y_i^* ของหน่วยตัวอย่างที่ i เป็นฟังก์ชันเชิงเส้นของตัวแปรอิสระ $X_{i1}, X_{i2},..., X_{ik}$ และนั่นคือ

$$Y_i^* = \beta' X_i + \varepsilon_i$$

เมื่อ ε_i คือค่าความคลาดเคลื่อนของหน่วยตัวอย่าง และ $\varepsilon_i \overset{iid}{\sim} N(0,1)$

จะได้

$$\mathrm{Y_i} = \begin{cases} 1 \quad \text{if} \quad \mathrm{Y_i^*} > 0 \;\; \text{or} \;\; \varepsilon_i < \beta' \mathrm{X_i} \\ 0 \quad \text{if} \quad \mathrm{Y_i^*} \le 0 \;\; \text{or} \;\; \varepsilon_i \ge \beta' \mathrm{X_i} \end{cases}$$

จาก
$$\begin{aligned} \pi_i &= P(Y_i = 1) \\ &= P(Y_i^* > 0) \\ &= P(\varepsilon_i < \beta' X_i) \qquad ; i = 1,2,...,n \\ &= \Phi(\beta' X_i) \\ &= \int_{-\infty}^{\beta X_i'} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\varepsilon^2}{2}\right) d\varepsilon \end{aligned}$$

โดยที่ π_i คือ ค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

Y_i คือ ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1

Y_i^* คือ ตัวแปรแฝง (Latent Variable) เป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริงทราบเพียงแต่ผลที่เกิดขึ้น

$\Phi(.)$ คือ ฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน

ε คือ ตัวแปรสุ่มปกติมาตรฐาน โดย $\varepsilon_i \overset{iid}{\sim} N(0,1)$

β' คือ เวกเตอร์ของสัมประสิทธ์การถดถอยจำนวน k+1 ตัว ขนาด 1x (k+1)

X_i คือ เวกเตอร์ของตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณจำนวน k ตัว ขนาด (k+1) ×1

ดังนั้น ตัวแบบโพรบิต สามารถเขียนแปลงให้อยู่ในรูป

$$\begin{aligned} \Phi^{-1}(\pi_i) &= \beta' X_i \\ &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_k x_{ki} \end{aligned} \qquad \text{.......... (1)}$$

จากสมการที่ (1) พบว่า การแปลงดังกล่าว จะทำให้เราได้ความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปร

ฟังก์ชันโพรบิต คือ

$$Probit\left(E\left[Y_i|\tilde{X}_i^T = \tilde{x}_i^T\right]\right) = \Phi^{-1}\left(E\left[Y_i|\tilde{X}_i^T = \tilde{x}_i^T\right]\right) = \Phi^{-1}(\pi_i)$$

ซึ่งฟังก์ชันโพรบิตจะทำการแปลงค่า π_i จากช่วง (0, 1) เป็นค่าที่อยู่ในช่วง $(-\infty,\infty)$

2.3 ฟังก์ชันภาวะน่าจะเป็นของข้อมูลการถดถอยโพรบิต

เนื่องจากตัวแปรตาม (Y_i) ที่ทำการศึกษามีเพียง 2 ค่า คือ 0 กับ 1 จึงใช้ฟังก์ชันการแจกแจงแบบเบอร์นูลลี่

$$P(Y_i = y_i) = p^{y_i}(1-p)^{1-y_i} \quad ; y_i = 0,1$$

สร้างฟังก์ชันของการแจกแจงความน่าจะเป็นร่วม (Joint Probability Density Function) ของหน่วยตัวอย่างอิสระ n ค่า โดยการคูณฟังก์ชันการแจกแจงความน่าจะเป็นของทุกหน่วยตัวอย่าง $(g(Y_i))$

$$\begin{aligned} g(Y_1, Y_2, \ldots, Y_n) &= \prod_{i=1}^{n} g(Y_i) \\ &= \prod_{i=1}^{n} P_i^{y_i}(1-P)^{1-y_i} \\ &= \prod_{i=1}^{n} [\Phi(\beta X_i^{'})]^{y_i}[1-\Phi(\beta X_i^{'})]^{1-y_i} \end{aligned}$$

ดังนั้น

$$L(\beta) = \prod_{i=1}^{n} [\Phi(x_i^{'}\beta)]^{y_i}[1-\Phi(x_i^{'}\beta)]^{1-y_i} \qquad \ldots\ldots\ldots. (2)$$

2.4 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation)

การหาค่าตัวประมาณค่าสัมประสิทธิ์การถดถอย $(\hat{\beta})$ ด้วยวิธีภาวะความน่าจะเป็นสูงสุดสามารถเขียนให้อยู่ในรูปของลอการิทึม (Logarithm) หรือความควรจะเป็นลอการึทึม (Log-Likelihood) ได้ดังนี้

$$\ln L(\beta) = \sum_{i=1}^{n} \{y_i \ln[\Phi(x_i^{'}\beta)] + (1-y_i)\ln[1-\Phi(x_i^{'}\beta)]\}$$
$$= \sum_{y_i=0} \ln[1-\Phi(x_i^{'}\beta)] + \sum_{y_i=1} \ln[\Phi(x_i^{'}\beta)] \qquad \text{..........}(3)$$

และเงื่อนไขอันดับแรก (First Order) สำหรับการให้สมการที่ (3) มีค่าสูงสุด (Maximization) และสามารถแก้สมการหาค่าเงื่อนไขอันดับแรกของฟังก์ชันโพรบิต ได้ดังนี้

$$\frac{\partial \ln L(\beta)}{\partial \beta} = \sum_{i=1}^{n} \left\{ \frac{y_i \phi(.)}{\Phi(.)} + (1-y_i)\left[\frac{-\phi(.)}{1-\Phi(.)}\right]\right\} X_i = 0 \qquad \text{..........}(4)$$

$$= \sum_{y_i=0} \left[\frac{-\phi(X_i^{'}\beta)}{1-\Phi(X_i^{'}\beta)}\right] X_i + \sum_{y_i=1} \left[\frac{\phi(X_i^{'}\beta)}{\Phi(X_i^{'}\beta)}\right] X_i = 0 \qquad \text{..........}(5)$$

เมื่อ $\phi(.)$ คือ ฟังก์ชันความหนาแน่นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

$\Phi(.)$ คือ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

ค่าประมาณของ π_i ; i = 1,2,...,n เป็นดังนี้

$$\hat{\pi}_i = \Phi(\beta^{'} X_i) \qquad ; i = 1,2,...,n$$

เมื่อประมาณค่าพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุดในตัวแบบโพรบิตแบบ 2 ประเภท แล้วจะสามารถนำไปใช้ในการพยากรณ์การจำแนกกลุ่มของตัวแบบ ดังนี้

- หน่วยที่ i จะถูกจัดให้อยู่ในกลุ่มที่เกิดเหตุการณ์ที่สนใจ (Y=1) ถ้า

$$\hat{\pi}_i = \Phi(\beta^{'} X_i) > c \qquad ;0 \le c \le 1$$

- หน่วยที่ i จะถูกจัดให้อยู่ในกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจ (Y=0) ถ้า

$$\hat{\pi}_i = \Phi(\beta^{'} X_i) \le c \qquad ;0 \le c \le 1$$

เมื่อ c คือ จุดแบ่งหรือระดับของความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนกกลุ่มว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มที่เกิดเหตุการณ์ที่สนใจ และกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจ

2.5 เปรียบเทียบตัวแบบโพรบิต โลจิท และคอมพรีเมนทารีล็อก-ล็อก

การหาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ เมื่อตัวแปรตามเป็นตัวแปรเชิงกลุ่มที่มีได้เพียง 2 ค่า และตัวแปรอิสระเป็นตัวแปรเชิงปริมาณ สามารถใช้การวิเคราะห์ที่มีลักษณะไม่เป็นเชิงเส้นของตัวแบบโพรบิต และตัวแบบคอมพรีเมนทารีล็อก-ล็อก ที่มีการแจกแจงสะสมปกติมาตรฐาน หรือตัวแบบโลจิทที่มีการแจกแจงโลจิสติกได้ ซึ่งตัวแบบทั้ง 3 ตัวแบบดังกล่าวมีคุณสมบัติคล้ายกัน การเลือกใช้ตัวแบบใดขึ้นอยู่กับพื้นฐานทางทฤษฎีซึ่งมีข้อแตกต่างดังนี้

1. ลักษณะฟังก์ชันการแจกแจงสะสม กราฟจะมีลักษณะรูปตัวเอส ทั้งตัวแบบโพรบิตและตัวแบบโลจิสติกมีค่าความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจและความน่าจะเป็นของการไม่เกิดเหตุการณ์ที่สนใจใกล้กันมากจนเกือบจะเท่ากัน ดังนั้นจะใช้ตัวแบบใดก็ไม่ก่อให้เกิดความแตกต่าง ยกเว้นแต่กรณีที่มีข้อมูลหนาแน่นในช่วงหาง ในขณะที่ฟังก์ชันคอมพรีเมนทารีล็อก-ล็อก มีความชันสูงกว่าฟังก์ชันโลจิสติกและฟังก์ชันโพรบิตที่ปรับแล้ว

2. ด้านการคำนวณตัวแบบโพรบิตและตัวแบบโลจิสติกจะให้ผลใกล้เคียงกัน โดยตัวแบบโพรบิตอยู่บนรากฐานของการแจกแจงสะสมแบบปกติ คำนวณได้ง่าย แต่ตัวแบบโพรบิตมีการคำนวณที่ยุ่งยากกว่า กล่าวคือ ในการกำหนดค่าความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจจะต้องติดอยู่ในรูปแบบของการอินทิเกรตเสมอ

2.6 ช่วงความเชื่อมั่น

ช่วงความเชื่อมั่นเป็นการประมาณค่าพารามิเตอร์ในรูปแบบช่วงโดยใช้ข้อมูลตัวอย่าง การประมาณแบบช่วงนั้นจะบอกถึงค่าต่ำสุดและค่าสูงสุดของพารามิเตอร์ที่เป็นไปได้

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ จะได้ว่า

$$P(L < \theta < U) = 1-\alpha$$

เรียก L ว่าขีดจำกัดความเชื่อมั่นล่าง (lower confidence limit)

U ว่าขีดจำกัดความเชื่อมั่นบน (upper confidence limit)

ในงานวิจัยนี้จะคำนวณช่วงความเชื่อมั่นของจุดแบ่งดังนี้

$$P(L < c < U) = 1-\alpha$$

L คือ ค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100\ (\alpha/2)$

U คือ ค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100\ (1-\alpha/2)$

2.7 เปอร์เซ็นต์ไทล์

เปอร์เซ็นต์ไทล์เป็นค่าที่แบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน เมื่อข้อมูลถูกเรียงจากน้อยไปหามาก เนื่องจากค่าที่แบ่งจำนวนข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน มีอยู่ 99 ค่า ดังนั้นเราจึงเรียกแต่ละค่าว่า เปอร์เซ็นต์ไทล์ที่หนึ่ง (P_1) เปอร์เซ็นต์ไทล์ที่สอง (P_2) .... และเปอร์เซ็นต์ไทล์ที่เก้าสิบเก้า (P_{99}) ตามลำดับ

การหาเปอร์เซ็นต์ไทล์ คือต้องหาตำแหน่งของเปอร์เซ็นต์ไทล์ก่อน ให้ N เป็นจำนวนข้อมูลทั้งหมด ตำแหน่งต่างๆ ของเปอร์เซ็นต์ไทล์หาได้ดังนี้

P_1 อยู่ในตำแหน่งที่ คือ $\frac{1(N+1)}{100}$

P_2 อยู่ในตำแหน่งที่ คือ $\frac{2(N+1)}{100}$

$\vdots$

P_{99} อยู่ในตำแหน่งที่ คือ $\frac{99(N+1)}{100}$

โดยทั่วไป ตำแหน่งของเปอร์เซ็นต์ไทล์ที่ r คือ

P_r อยู่ในตำแหน่งที่ คือ $\frac{r(N+1)}{100}$

2.8 เทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นการจำลองระบบที่ไม่เปลี่ยนแปลงตามเวลา ซึ่งตัวแบบของการจำลองจะมีลักษณะเป็นตัวแบบทางคณิตศาสตร์ โดยการนำตัวเลขสุ่ม มาประยุกต์ใช้ในการแก้ปัญหาให้กับระบบที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น ซึ่งมีขั้นตอนที่สำคัญดังนี้

1. การสร้างเลขสุ่ม (Generate Random Number) โดยกำหนดให้มีการแจกแจงแบบยูนิฟอร์มในช่วง [0, 1] และเป็นอิสระกัน จากนั้นจึงนำเลขสุ่มนี้ไปสร้างตัวแปรตามลักษณะการแจกแจงที่ต้องการศึกษา เพื่อเป็นข้อมูลของปัญหานั้น ๆ

2. การประยุกต์ใช้เลขสุ่มในการแก้ปัญหา ซึ่งเป็นการนำตัวแปรที่ได้จากขั้นตอนแรกมาใช้ในการหาค่าต่างๆ ตามปัญหาที่ต้องการศึกษา

3. การทดลอง เป็นการทำวิธีนั้นซ้ำๆ กัน (Replication) จำนวนหลายๆ ครั้ง โดยถือว่าการทำซ้ำๆ กันนั้น เป็นวิธีการเก็บรวบรวมข้อมูลให้มีจำนวนมาก เพื่อลดความไม่แน่นอนของคำตอบในการวิเคราะห์หาค่าต่างๆ ได้

จากหลักการของเทคนิคมอนติคาร์โลข้างต้น จะเห็นว่าการใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีที่จะนำไปสู่แนวคิดในทางทฤษฎีที่เกี่ยวข้องกับการคำนวณ โดยเฉพาะทฤษฎีความน่าจะเป็นที่จะนำไปอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริง เพราะไม่มีผลกระทบจากเรื่องอื่นๆ เข้ามาเกี่ยวข้องในการทดลอง เมื่อทำซ้ำๆ กันเป็นจำนวนมากแล้ว ความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่างๆ ในแต่ละครั้งจะมีน้อยจนหมดไป

2.9 **ทฤษฎี** Hadjicostas P (2006)

การหาจุดแบ่งที่เหมาะสมที่สุดโดยทฤษฎีของ Hadjicostas P. (2006) นั้น เป็นวิธีการนี้ใช้ผลลัพธ์ทางคณิตศาสตร์ที่ง่ายแต่มีความถูกต้องแม่นยำในการหาจุดแบ่งที่ทำให้สัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด โดยมีวิธีการหาจุดแบ่งดังนี้

บทตั้ง $N(c) = \sum_{j=1}^{M(i)} (1 - y_j) + \sum_{j=M(i)+1}^{n} y_j$ สำหรับ $i \in \{0,1,2,...,n\}$ ใดๆ และ $c \in A_i$

ทฤษฎีบท ให้ $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ สำหรับ $i = 0,1,2,...,n$ ให้ I_0 เป็นเซตของ j ทั้งหมด $j \in \{0,1,2,...,n\}$ ซึ่ง $a_j = \max_{0 \le i \le n} a_i$ และให้ C_0 เป็นเซตของ c_0 ทั้งหมด $c_0 \in [0,1]$ ซึ่ง $p(c_0) = \max_{c \in [0,1]} p(c)$ แล้ว $C_0 = \bigcup_{i \in I_0} A_i$

จากบทตั้งและทฤษฎีข้างต้น จะได้ขั้นตอนการหาจุดแบ่งดังนี้

1. เรียงอันดับค่า $\hat{\pi}_i$ จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < ... < \hat{\pi}_n$
2. สำหรับแต่ละ $i \in \{1,2,...,n\}$ ใดๆ หาค่า $M(i)$ ซึ่ง $M(i)$ คือ
   $\max j \in \{1,2,...,n\}$ ถ้า $\hat{\pi}_i = \hat{\pi}_j$ โดย $M(0) = 0\ ; i \le M(i) \le n$
3. สำหรับ $i = 0,1,2,...,n$ หา a_i โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ ซึ่งแบ่งเป็น 2 กรณี คือ

   3.1 $a_{i+1} = a_i + \sum_{k=M(i)+1}^{M(i+1)} (-1)^{y_k}$ ถ้า $M(i) < i+1$

   3.2 $a_{i+1} = a_i$ ถ้า $i+1 \le M(i)$
4. หา I_0 ซึ่งเป็นเซตของ j ทั้งหมด โดย $j \in \{0,1,2,...,n\}$ ซึ่ง $a_j = \max_{0 \le i \le n} a_i$

5. หา C_0 ซึ่งเป็นเซตของ c_0 ทั้งหมด โดย $c_0 \in [0,1]$ ซึ่ง $p(c_0) = \max_{c \in [0,1]} p(c)$ แล้ว $C_0 = \bigcup_{i \in I_0} A_i$ สำหรับ $i \in \{0,1,2,...,n\}$ จะได้ว่า

$$A_i = [0, \hat{\pi}_1) \quad \text{ถ้า } i = 0$$

$$A_i = [\hat{\pi}_i, \hat{\pi}_{i+1}) \quad \text{ถ้า } \hat{\pi}_i < \hat{\pi}_{i+1} \text{ และ } 1 \le i < n$$

$$A_i = \{\hat{\pi}_i\} = \{\hat{\pi}_{M(i)}\} \quad \text{ถ้า } \hat{\pi}_i = \hat{\pi}_{i+1} \text{ และ } 1 \le i < n$$

$$A_i = [\hat{\pi}_n, 1] \quad \text{ถ้า } i = n$$

ข้อสังเกต $\bigcup_{i=0}^{n} A_i = [0,1]$

6. เลือกค่าของจุดแบ่ง c โดย $c \in C_0$ และ $c \in [0,1]$ ซึ่งเป็นค่า c ที่ทำให้สัดส่วนของความถูกต้องในการจำแนกกลุ่มที่จุด c มีค่ามากที่สุด

$$p(c) = \frac{N(c)}{n}$$

เมื่อ $p(c)$ คือ สัดส่วนความถูกต้องในการจำแนกกลุ่มที่จุด c

$N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

2.10 งานวิจัยที่เกี่ยวข้อง

งานวิจัยครั้งนี้มีงานวิจัยที่เกี่ยวข้องดังนี้

อรุณรัตน์ โพธิ์คำ (2554) ได้ทำการศึกษาหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันโลจิทเป็นฟังก์ชันเชื่อมโยง โดยมีปัจจัยที่สนใจศึกษาคือ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ขนาดตัวอย่าง และจำนวนตัวแปรอิสระ ผลการวิจัยพบว่า กรณีสัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา เปลี่ยนแปลง แต่ปัจจัยอื่นๆคงที่ พบว่า ที่สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษามีค่าเท่ากับ 0.5 ค่าจุดแบ่งมีค่าลู่เข้าสู่ 0.5 แต่ที่ค่าอื่นๆ ค่าจุดแบ่งมีค่ามากกว่า 0.5 กรณีระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่ปัจจัยอื่นๆ คงที่ พบว่า ที่สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษามีค่าเท่ากับ 0.1 ค่าจุดแบ่งมีแนวโน้มเพิ่มขึ้นจาก 0.5 กรณีขนาดตัวอย่างเพิ่มขึ้น แต่ปัจจัยอื่นๆ คงที่ พบว่า ที่ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษามี ค่าต่ำ เมื่อขนาดตัวอย่างเพิ่มขึ้นค่าจุดแบ่งมีแนวโน้มเพิ่มขึ้น กรณีจำนวนตัวแปรอิสระเพิ่มขึ้นแต่ปัจจัยอื่นๆ คงที่ พบว่า ที่สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษามีค่าต่ำ จำนวนตัวแปรอิสระเพิ่มขึ้น ค่าจุดแบ่งมีค่าลดลงจาก 0.5 แต่ที่สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษามีค่าสูง จำนวนตัวแปรอิสระเพิ่มขึ้น ค่าจุดแบ่งมีแนวโน้มเพิ่มขึ้นจาก จาก 0.5

นิภาพรรณ ไพจินดา (2554) ได้ทำการศึกษาหาจุดแบ่งที่เหมาะสมที่สุดของตัวแบบถดถอยโลจิสติก 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง โดยปัจจัยที่สนใจศึกษาคือ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา จำนวนตัวแปรอิสระ ขนาดของตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ 3 ระดับ คือ ต่ำ กลาง สูง ผลการวิจัยพบว่า กรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งไม่เกินค่า 0.5 สลับสูงต่ำไปมา และที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มขึ้น แต่ไม่มีค่าใดเกินค่า 0.5 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่ พบว่าที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 2 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งไม่เกินค่า 0.5 สลับสูงต่ำไปมา และที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มขึ้น แต่ไม่เกินค่า 0.5 กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่ พบว่า ที่ทุกขนาดตัวอย่างจะให้ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง และสุดท้ายกรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งคงที่แต่ไม่เกิน 0.5 และ ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มสูงขึ้น

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง สำหรับแต่ละสถานการณ์ที่ต้องการศึกษา จะทำการเปรียบเทียบสถานการณ์ที่จำนวนตัวแปรอิสระเป็น 1,2,3,4 และ 5 ขนาดตัวอย่างมีค่าเป็น 50,100,150,200 และ 250 สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเป็น 0.1, 0.5 และ 0.9 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น 3 ระดับ คือ ความสัมพันธ์กันในระดับต่ำ ปานกลาง และสูง ซึ่งทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรม R ซึ่งรายละเอียดของแผนการดำเนินการวิจัย ขั้นตอนการดำเนินการวิจัยมีดังนี้

3.1 แผนดำเนินการวิจัย

ในการวิจัยครั้งนี้ได้จำลองข้อมูลขึ้นภายใต้สถานการณ์ต่างๆ ดังนี้

1. จำนวนของตัวแปรอิสระ (p) ในการวิจัยครั้งนี้คือ 1, 2, 3, 4 และ 5

2. ขนาดตัวอย่าง (n) ในการวิจัยครั้งนี้คือ 50, 100, 150, 200 และ 250

3. กำหนดสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) ในการวิจัยครั้งนี้คือ 0.1, 0.5 และ 0.9

4. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ แบ่งเป็น 3 ระดับ คือ
   - ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ เมื่อ $0 < Max\{|r_{ij}|\} \leq 0.30$
   - ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง เมื่อ $0.30 < Max\{|r_{ij}|\} \leq 0.60$
   - ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง เมื่อ $0.60 < Max\{|r_{ij}|\} \leq 0.90$

5. กำหนดค่าพารามิเตอร์เริ่มต้นของสมการการถดถอยเป็นค่าใดๆ ในการวิจัยครั้งนี้กำหนดให้ $\beta_i = 10\,;\, i = 0,1,2,\ldots,p$ และ $\varepsilon_k \overset{iid}{\sim} N(0,500)$; $k = 1,2,\ldots,n$

6. กำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05

7. จำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation) โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

3.2 ขั้นตอนการดำเนินการวิจัย

สำหรับขั้นตอนการดำเนินการวิจัยมีขั้นตอนดังนี้

3.2.1 จำลองข้อมูลที่ใช้ในการวิจัย

1. ในการจำลองข้อมูลในการวิจัยครั้งนี้ ต้องสร้างข้อมูลตัวแปรอิสระ เริ่มต้นให้มีการแจกแจงยูนิฟอร์ม ซึ่งกำหนดให้มีค่าของช่วงบวกและช่วงลบที่เท่ากัน คือ X~U(−1,1)

2. สร้างจำนวนตัวแปรอิสระตามที่กำหนดไว้ โดยให้ตัวแปรอิสระมีความสัมพันธ์กันตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่กำหนดไว้ 3 ระดับ คือ ต่ำ ปานกลางและสูง และให้มีขนาดตัวอย่างตามที่กำหนดไว้

3. สร้างค่าตัวแปรตาม $\left(Y^*\right)$ โดยสร้างให้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระและความคลาดเคลื่อนใดๆ ตามตัวแบบดังนี้

$$Y^* = X\tilde{\beta} + \varepsilon_k$$

โดยที่ Y^* คือ เมทริกซ์ของตัวแปรตามที่ทำการพยากรณ์เพื่อกำหนดค่าเบื้องต้น

X คือ เมทริกซ์ของตัวแปรอิสระ

$\tilde{\beta}$ คือ เวกเตอร์ของพารามิเตอร์ของตัวแบบ กำหนดให้มีค่าเริ่มต้นที่ 10

ε_k คือ ค่าความคลาดเคลื่อน $\varepsilon_k \overset{iid}{\sim} N(0,500)$; k = 1, 2, …, n

4. สร้างค่าตัวแปรตาม (Y) ที่มีค่าเป็น 0 หรือ 1 โดยแปลงจากค่าตัวแปร Y^* ตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและขนาดตัวอย่างที่กำหนดไว้

5. ประมาณค่าพารามิเตอร์โดยใช้ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท ด้วยวิธีภาวะน่าจะเป็นสูงสุด

6. หาค่าประมาณของ $\hat{\pi}_i$ โดยการนำค่าพารามิเตอร์ที่ได้จากการประมาณค่าข้างต้นและค่าของตัวแปรอิสระ แทนค่ากลับลงไปในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท

3.2.2 คำนวณค่าจุดแบ่งโดยทฤษฎี Hadjicostas P. (2006)

1. เรียงอันดับค่า $\hat{\pi}_i$ จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < \ldots < \hat{\pi}_n$

2. หาค่า $M(i)$; $i \in \{1,2,\ldots,n\}$ ซึ่ง $M(i)$ คือ อันดับของ $\hat{\pi}_i$ ถ้า $\hat{\pi}_i = \hat{\pi}_j$ จะเลือก $M(i)$ ที่มีค่ามากที่สุดเป็นอันดับของ $\hat{\pi}_i$ และ $\hat{\pi}_j$

3. หาค่า a_i ; $i = 0,1,2,...,n$ โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ แบ่งเป็น 2 กรณีดังนี้

กรณี 1: $a_{i+1} = a_i + \sum_{k=M(i)+1}^{M(i+1)} (-1)^{y_k}$ ถ้า $M(i) < i+1$

กรณี 2: $a_{i+1} = a_i$ ถ้า $i+1 \le M(i)$

4. หาค่า I_0 ซึ่งเป็นเซตของ j ทั้งหมด โดย $j \in \{0,1,2,...,n\}$ ที่ซึ่ง $a_j = \max_{0 \le i \le n} a_i$

5. หาค่า C_0 ซึ่งเป็นเซตของ c_0 ทั้งหมด โดย $c_0 \in [0,1]$ ที่ซึ่ง $C_0 = \bigcup_{i \in I_0} A_i$; $i \in \{0,1,2,...,n\}$ โดยพิจารณาตามเงื่อนไขต่อไปนี้

$A_i = [0, \hat{\pi}_1)$ ถ้า $i = 0$

$A_i = [\hat{\pi}_i, \hat{\pi}_{i+1})$ ถ้า $\hat{\pi}_i < \hat{\pi}_{i+1}$ และ $1 \le i < n$

$A_i = \{\hat{\pi}_i\} = \{\hat{\pi}_{M(i)}\}$ ถ้า $\hat{\pi}_i = \hat{\pi}_{i+1}$ และ $1 \le i < n$

$A_i = [\hat{\pi}_n, 1]$ ถ้า $i = n$

6. เลือกค่าของจุดแบ่ง (c) ที่ทำให้สัดส่วนของความถูกต้องในการจำแนกกลุ่มมีค่ามากที่สุดโดย โดยที่ $c \in C_0$ และ $c \in [0,1]$

$$p(c) = \frac{N(c)}{n}$$

เมื่อ $p(c)$ คือ สัดส่วนความถูกต้องในการจำแนกกลุ่มที่จุด c

$N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

3.2.3 คำนวณค่าเฉลี่ยและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสมที่สุด

1. การหาค่าเฉลี่ยของจุดแบ่ง ($\bar{\hat{c}}$) หาค่าเฉลี่ยของจุดแบ่งในแต่ละสถานการณ์ จากการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์

กำหนดให้ $\bar{\hat{c}}$ เป็นตัวประมาณค่าของพารามิเตอร์ c

$$\bar{\hat{c}} = \frac{\sum_{k=1}^{N} \hat{c}_{(k)}}{N} \qquad ; k = 1, 2, ..., N$$

โดย N คือ จำนวนรอบที่กระทำซ้ำ (N=500)

เนื่องจากการหาค่าจุดแบ่งในแต่ละรอบในทุกสถานการณ์จะได้ค่าออกมาเป็นช่วง ซึ่งค่าที่อยู่ในช่วงดังกล่าวนั้นเป็นค่าที่ทำให้สัดส่วนของความถูกต้องในการจำแนกกลุ่มมีค่ามากที่สุดทั้งหมด ในงานวิจัยครั้งนี้ผู้วิจัยจึงใช้ค่ามัธยฐานของค่าจุดแบ่งในแต่ละรอบ (ทั้งหมด 500 รอบ) จากนั้นจึงนำค่าจุดแบ่งทั้ง 500 รอบมาหาค่าเฉลี่ยของจุดแบ่งตามที่กล่าวข้างต้น

2. ช่วงความเชื่อมั่นของจุดแบ่ง โดยช่วงความเชื่อมั่นนี้จะให้ค่าของจุดแบ่งที่ต่ำสุดและสูงสุดที่เป็นไปได้ในแต่ละสถานการณ์ กำหนดค่า $\alpha = 0.05$

$$P(L < \bar{\hat{c}} < U) = 1 - \alpha$$

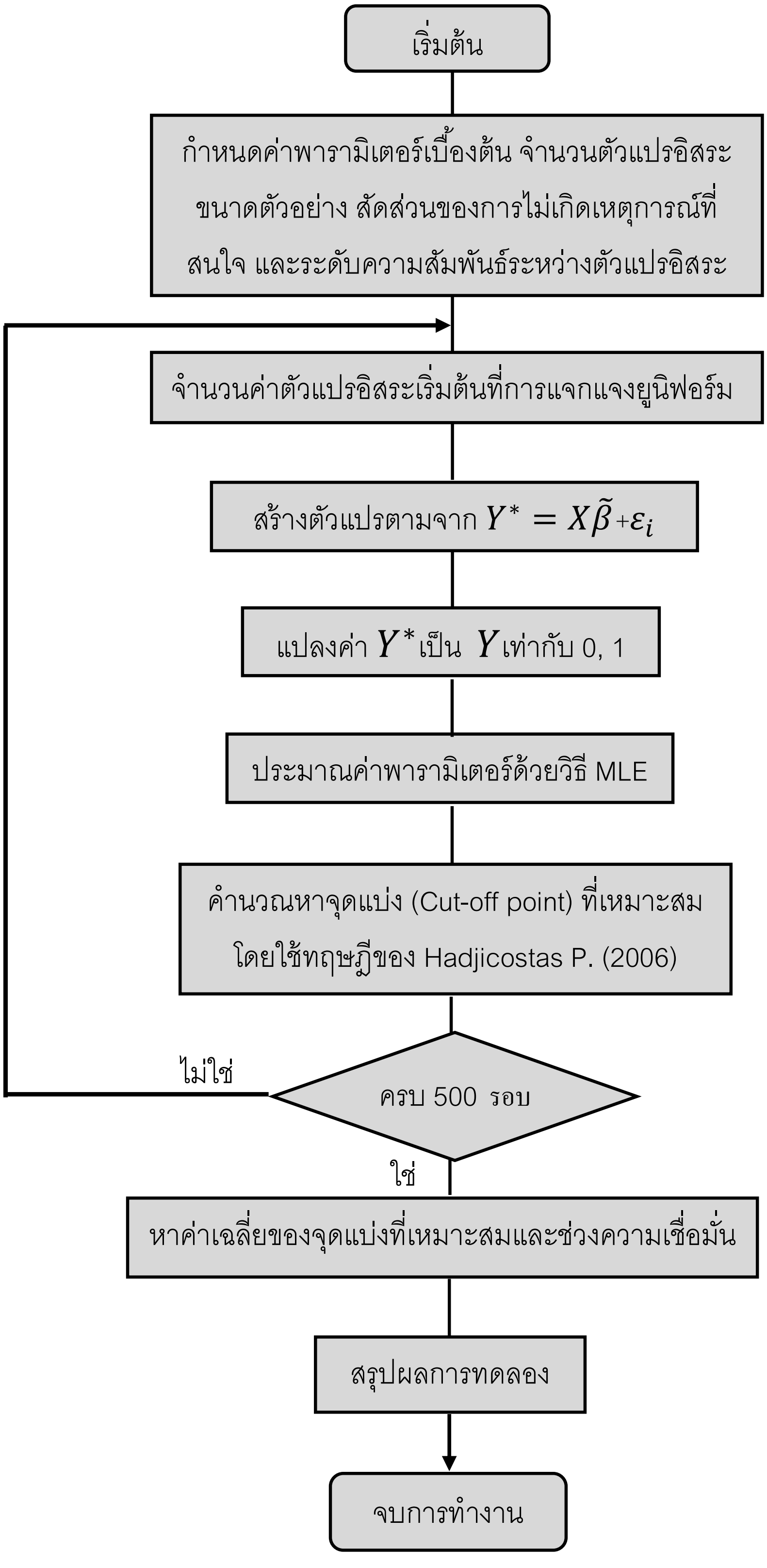
โดยที่ L คือ ขีดจำกัดความเชื่อมั่นล่าง ที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 2.5

U คือ ขีดจำกัดความเชื่อมั่นบน ที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 97.5

3.2.4 **สรุปการวิจัยในแต่ละสถานการณ์**

เมื่อหาค่าจุดแบ่งที่เหมาะสมที่สุดครบทุกสถานการณ์ที่ต้องการศึกษาแล้ว นำผลการทดลองมาสรุปผลในรูปตาราง เพื่อดูว่าปัจจัยที่ต้องการศึกษามีผลต่อค่าจุดแบ่งอย่างไร

3.3 ขั้นตอนการทำงานของโปรแกรม



ภาพที่ 3.1 แสดงขั้นตอนของการทำงานของโปรแกรม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง เมื่อตัวแปรอิสระมีการแจกแจงเริ่มต้นแบบยูนิฟอร์ม จากนั้นพิจารณาปัจจัยอื่นว่ามีผลต่อการหาจุดแบ่งหรือไม่ ซึ่งปัจจัยเหล่านั้นได้แก่ จำนวนตัวแปรอิสระที่เพิ่มขึ้น ขนาดตัวอย่างที่เพิ่มขึ้น สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจที่เพิ่มขึ้น และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่เพิ่มขึ้น ซึ่งศึกษาภายใต้สถานการณ์ดังต่อไปนี้

1. กำหนดให้จำนวนของตัวแปรอิสระเพิ่มขึ้น เป็น 1,2,3,4 และ 5 ตามลำดับ เมื่อขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ มีค่าคงที่
2. กำหนดให้ขนาดตัวอย่างเพิ่มขึ้น เป็น 50,100,150, 200 และ 250 ตามลำดับ เมื่อจำนวนตัวแปรอิสระ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ มีค่าคงที่
3. กำหนดให้สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น เป็น 0.1, 0.5 และ 0.9 ตามลำดับ เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ มีค่าคงที่
4. กำหนดให้ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น เป็นระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ตามลำดับ เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง และสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ มีค่าคงที่

การวิจัยนี้จะนำเสนอผลของการวิเคราะห์ข้อมูลในแต่ละสถานการณ์ที่ต้องการศึกษาในรูปแบบตาราง ซึ่งผู้วิจัยกำหนดสัญลักษณ์แทนความหมายต่างๆ ดังนี้

p หมายถึง จำนวนตัวแปรอิสระ

n หมายถึง ขนาดตัวอย่าง

ρ หมายถึง ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ

a หมายถึง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ

$\bar{\hat{c}}$ หมายถึง ค่าเฉลี่ยของจุดแบ่ง

Cl.Lower of $\bar{\hat{c}}$ หมายถึง ค่าต่ำสุดของช่วงความเชื่อมั่นของจุดแบ่ง

Cl.Upper of $\bar{\hat{c}}$ หมายถึง ค่าสูงสุดของช่วงความเชื่อมั่นของจุดแบ่ง

การหาค่าของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม โดยใช้ทฤษฎี Hadjicostas P. (2006) เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.1.– 4.43 ดังนี้

4.1 กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.1-4.9

ตารางที่ 4.1 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\}$ ≤ 0.30) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	**ค่าเฉลี่ยของ** $Max\{|r_{ij}|\}$	n	p	$\bar{\hat{c}}$	Cl. Lower of $\bar{\hat{c}}$	Cl. Upper of $\bar{\hat{c}}$
0.1	0	50	1	0.498256	0.348717	0.659052
	0.1886804		2	0.500867	0.22381	0.774946
	0.2350503		3	0.501172	0.250077	0.753873
	0.254126		4	0.499337	0.260935	0.746921
	0.2695499		5	0.496327	0.285974	0.715822

ตารางที่ 4.1 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0	100	1	0.493823	0.378832	0.612861
	0.2078653		2	0.4994	0.236924	0.765269
	0.2529071		3	0.502182	0.240616	0.770389
	0.2655062		4	0.490926	0.216475	0.772606
	0.2734939		5	0.498461	0.244965	0.758303
	0	150	1	0.495794	0.386127	0.596221
	0.2177278		2	0.495476	0.247742	0.743235
	0.2549317		3	0.502931	0.242619	0.768415
	0.2674501		4	0.499868	0.216776	0.771611
	0.2787541		5	0.511795	0.240792	0.764196
	0	200	1	0.495721	0.406011	0.588072
	0.2239275		2	0.502327	0.272957	0.72627
	0.2584532		3	0.499529	0.252766	0.749504
	0.2726991		4	0.505795	0.257207	0.746059
	0.2781472		5	0.508619	0.223032	0.766436
	0	250	1	0.494895	0.406844	0.579196
	0.2287889		2	0.510853	0.291907	0.733491
	0.2609532		3	0.507128	0.27192	0.726269
	0.2731542		4	0.490329	0.247489	0.745463
	0.2807881		5	0.490082	0.227694	0.761813

จากตารางที่ 4.1 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.2 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\}\leq$ 0.60) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	**ค่าเฉลี่ยของ** $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.4994576	50	2	0.500675	0.235503	0.772616
	0.5385911		3	0.501844	0.242707	0.74924
	0.5565136		4	0.498926	0.278862	0.707956
	0.5645582		5	0.499102	0.377179	0.514213
	0.5219221	100	2	0.49528	0.227029	0.758286
	0.5554746		3	0.495576	0.228455	0.779218
	0.5640089		4	0.503611	0.235779	0.757799
	0.570552		5	0.50368	0.254696	0.73254
	0.5299841	150	2	0.507824	0.252215	0.756141
	0.556986		3	0.497141	0.221334	0.773431
	0.566902		4	0.509403	0.240063	0.782342
	0.5725225		5	0.501993	0.238995	0.773778
	0.5359944	200	2	0.491696	0.256523	0.731223
	0.5587158		3	0.486614	0.228105	0.75554
	0.5693879		4	0.508991	0.236898	0.778514
	0.5739922		5	0.501506	0.233406	0.774567
	0.5372966	250	2	0.502143	0.266988	0.743305
	0.5607767		3	0.490277	0.266541	0.748254
	0.5701132		4	0.49896	0.22798	0.757359
	0.5768426		5	0.497107	0.227003	0.78334

จากตารางที่ 4.2 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\}\leq$ 0.60)

ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.3 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.8393408	50	2	0.497366	0.22066	0.771432
	0.8620921		3	0.495291	0.268113	0.719601
	0.8699663		4	0.501512	0.402216	0.647958
	0.8747025		5	0.502021	0.50000	0.575944*
	0.847068	100	2	0.4899	0.208386	0.777063
	0.863123		3	0.497898	0.250699	0.766899
	0.8722286		4	0.503028	0.252933	0.76234
	0.8758678		5	0.495558	0.267639	0.702735
	0.8487206	150	2	0.509393	0.237081	0.791051
	0.8607247		3	0.497446	0.230128	0.775159
	0.8680124		4	0.503083	0.241903	0.778957
	0.8730344		5	0.505336	0.280116	0.724207
	0.849504	200	2	0.508489	0.242717	0.770344
	0.8603123		3	0.499144	0.220767	0.760301
	0.8675026		4	0.505198	0.225411	0.783045
	0.871125		5	0.498684	0.24358	0.758786
	0.8491737	250	2	0.496598	0.264818	0.731034
	0.8600407		3	0.497591	0.23843	0.747535
	0.8667469		4	0.511508	0.227845	0.789503
	0.869478		5	0.50073	0.230196	0.769607

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.3 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.4 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0	50	1	0.503665	0.365606	0.656167
	0.1864547		2	0.500494	0.237315	0.767111
	0.2348507		3	0.500676	0.2304	0.752722
	0.2601783		4	0.507427	0.273588	0.740752
	0.2685089		5	0.497784	0.289586	0.662065
	0	100	1	0.497691	0.387556	0.610248
	0.2047484		2	0.50044	0.237995	0.770704
	0.251319		3	0.48944	0.231358	0.769243
	0.2672831		4	0.489803	0.224208	0.762047
	0.2754783		5	0.499431	0.250611	0.758003
	0	150	1	0.495785	0.396597	0.591528
	0.2190896		2	0.502754	0.270441	0.729218
	0.2568403		3	0.496919	0.233292	0.745431
	0.2680613		4	0.495113	0.248984	0.777624
	0.2775778		5	0.502148	0.222499	0.788029
	0	200	1	0.500222	0.411185	0.589516
	0.2222356		2	0.495216	0.258498	0.717659
	0.2585383		3	0.498575	0.258514	0.754046
	0.271567		4	0.487352	0.234282	0.770728

ตารางที่ 4.4 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\hat{\bar{c}}$	CI. Lower of $\hat{\bar{c}}$	CI. Upper of $\hat{\bar{c}}$
0.5	0.2798083	200	5	0.500622	0.226709	0.772803
	0	250	1	0.496572	0.412955	0.578279
	0.2338323		2	0.50075	0.271328	0.71322
	0.2616638		3	0.508402	0.265561	0.740355
	0.2744325		4	0.49877	0.230719	0.754654
	0.2792333		5	0.501861	0.254699	0.751812

จากตารางที่ 4.4 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.5 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\hat{\bar{c}}$	CI. Lower of $\hat{\bar{c}}$	CI. Upper of $\hat{\bar{c}}$
0.5	0.5007751	50	2	0.501119	0.228249	0.771341
	0.535131		3	0.500478	0.257131	0.735763
	0.557933		4	0.50057	0.287409	0.731583
	0.5668828		5	0.501685	0.493776*	0.510152**
	0.5164418	100	2	0.504465	0.226986	0.775911
	0.5511936		3	0.505789	0.228973	0.771086
	0.5632398		4	0.494576	0.238845	0.76584
	0.5708508		5	0.505952	0.274687	0.759896

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

ตารางที่ 4.5 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\bar{c}}$	CI. Lower of $\bar{\bar{c}}$	CI. Upper of $\bar{\bar{c}}$
0.5	0.5309888	150	2	0.501573	0.239628	0.766404
	0.5559691		3	0.499629	0.232451	0.774833
	0.5662719		4	0.487577	0.231274	0.764482
	0.5745328		5	0.500495	0.249938	0.7688
	0.5339025	200	2	0.493578	0.238064	0.735588
	0.5603839		3	0.499552	0.236095	0.772088
	0.5686589		4	0.493009	0.220505	0.755097
	0.5754855		5	0.497713	0.236977	0.758642
	0.5391257	250	2	0.497347	0.250164	0.720739
	0.5596843		3	0.501579	0.252983	0.735728
	0.5716462		4	0.499859	0.237691	0.77363
	0.5750038		5	0.5033	0.227734	0.774936

จากตารางที่ 4.5 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.6 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\bar{c}}$	CI. Lower of $\bar{\bar{c}}$	CI. Upper of $\bar{\bar{c}}$
0.5	0.8419899	50	2	0.502854	0.22429	0.75836
	0.8605197		3	0.508168	0.270338	0.738057
	0.8696018		4	0.500205	0.369926	0.638865
	0.8758058		5	0.499978	0.377643*	0.620358**

ตารางที่ 4.6 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.8471553	100	2	0.499313	0.228393	0.769043
	0.8632002		3	0.494965	0.232188	0.774918
	0.8714558		4	0.505334	0.267485	0.73557
	0.8749598		5	0.497911	0.266185	0.680201
	0.8487382	150	2	0.493233	0.224216	0.759169
	0.8622969		3	0.505509	0.239588	0.766192
	0.8695362		4	0.490099	0.219119	0.778358
	0.8732464		5	0.506921	0.266837	0.751978
	0.8489849	200	2	0.499217	0.248745	0.74121
	0.8630131		3	0.501496	0.230034	0.757002
	0.8671073		4	0.498765	0.219758	0.765132
	0.8718601		5	0.493159	0.257281	0.742426
	0.8497455	250	2	0.500984	0.269106	0.73085
	0.8614076		3	0.501652	0.239267	0.764877
	0.8661106		4	0.502993	0.242588	0.760802
	0.8693057		5	0.511941	0.23163	0.78244

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.6 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง (0.60 $< Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่างเท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งเพิ่มขึ้นจากค่า 0.5 เพียงเล็กน้อย เมื่อตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.7 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	**ค่าเฉลี่ยของ** $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0	50	1	0.509422	0.37607	0.648653
	0.1821284		2	0.507494	0.217843	0.775718
	0.2333454		3	0.498069	0.232993	0.754885
	0.2570051		4	0.497186	0.260612	0.757711
	0.2689239		5	0.500073	0.292309	0.679598
	0	100	1	0.503668	0.397134	0.613246
	0.2042958		2	0.489978	0.231292	0.755361
	0.2481321		3	0.502313	0.227633	0.768928
	0.264764		4	0.510789	0.233674	0.774722
	0.2750965		5	0.493672	0.238426	0.761804
	0	150	1	0.501341	0.400452	0.59376
	0.2160963		2	0.492633	0.249681	0.749831
	0.253179		3	0.4983	0.233279	0.772768
	0.2711085		4	0.507487	0.225612	0.770483
	0.2760721		5	0.493163	0.237269	0.772208
	0	200	1	0.503505	0.410572	0.590867
	0.2226641		2	0.49615	0.2638	0.730161
	0.2561202		3	0.509764	0.233451	0.756943
	0.2731122		4	0.500595	0.240592	0.770429
	0.2793923		5	0.495289	0.238201	0.769221

ตารางที่ 4.7 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0	250	1	0.49921	0.415682	0.582303
	0.2318075		2	0.499183	0.289666	0.704873
	0.2617862		3	0.503039	0.257576	0.724219
	0.2754604		4	0.499596	0.256919	0.761046
	0.2803318		5	0.483096	0.241996	0.740038

จากตารางที่ 4.7 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.8 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.4953218	50	2	0.507027	0.237697	0.767406
	0.5354911		3	0.494613	0.234522	0.73211
	0.5532036		4	0.498156	0.30295	0.679195
	0.5625658		5	0.500792	0.469167	0.60962
	0.5206238	100	2	0.499105	0.230647	0.74143
	0.5506931		3	0.494173	0.222366	0.766775
	0.5657834		4	0.505651	0.234338	0.75323
	0.5710766		5	0.500203	0.244628	0.736526
	0.5275973	150	2	0.486646	0.238289	0.754875
	0.5588926		3	0.501856	0.216915	0.787535

ตารางที่ 4.8 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.5673897	150	4	0.494009	0.21649	0.75806
	0.5727453		5	0.493036	0.236913	0.763234
	0.5393825	200	2	0.503666	0.265697	0.735765
	0.5577018		3	0.502397	0.245572	0.751992
	0.5717861		4	0.496242	0.215561	0.783827
	0.5753501		5	0.492654	0.236066	0.771419
	0.5416306	250	2	0.498816	0.250122	0.726159
	0.562273		3	0.499273	0.247986	0.747294
	0.5706408		4	0.490808	0.222459	0.758633
	0.5757055		5	0.489227	0.230261	0.768856

จากตารางที่ 4.8 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.9 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ (p)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.8425875	50	2	0.508642	0.254318	0.764763
	0.8605327		3	0.495267	0.269261	0.698625
	0.8700897		4	0.50136	0.379242	0.641594
	0.8745183		5	0.501531	0.495975*	0.640162**

ตารางที่ 4.9 (ต่อ)

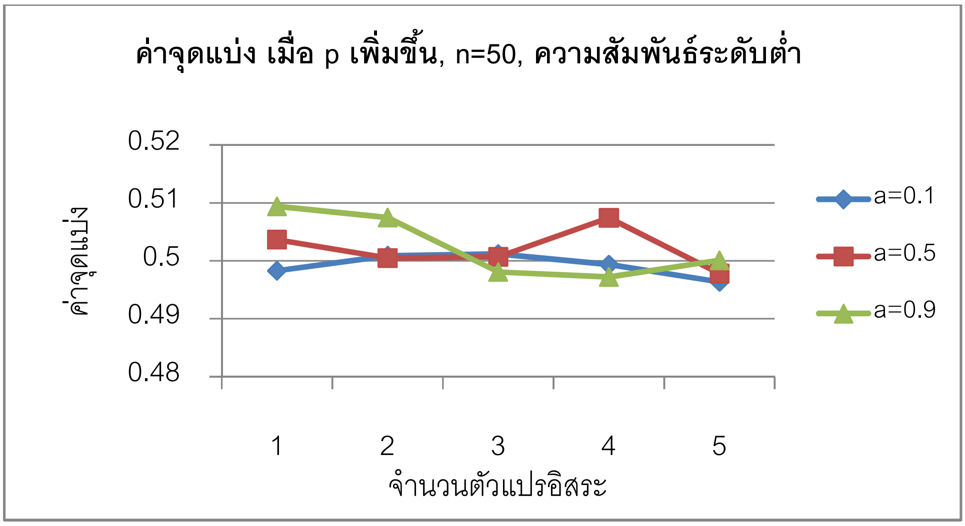
a	**ค่าเฉลี่ยของ** $Max\{\|r_{ij}\|\}$	n	p	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.8438331	100	2	0.488845	0.214951	0.76145
	0.8638203		3	0.496422	0.227551	0.762967
	0.8706262		4	0.503881	0.253366	0.751583
	0.87495		5	0.498957	0.298155	0.714686
	0.8492548	150	2	0.497619	0.22453	0.764602
	0.8616446		3	0.497891	0.236194	0.759695
	0.8698966		4	0.4973	0.231858	0.777655
	0.8730918		5	0.507266	0.256175	0.753777
	0.8494049	200	2	0.50235	0.237538	0.762992
	0.8613328		3	0.492758	0.221162	0.764929
	0.8672836		4	0.491078	0.227777	0.763907
	0.8710947		5	0.494261	0.231991	0.755092
	0.8490023	250	2	0.500899	0.262279	0.751556
	0.8603164		3	0.493694	0.227981	0.775343
	0.8656913		4	0.499081	0.219553	0.780762
	0.8700021		5	0.50278	0.256239	0.767216

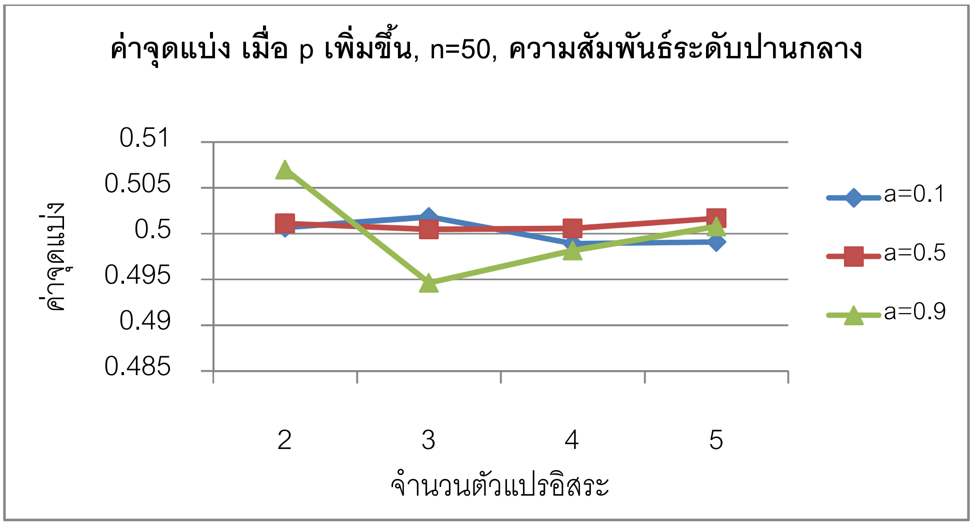
* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

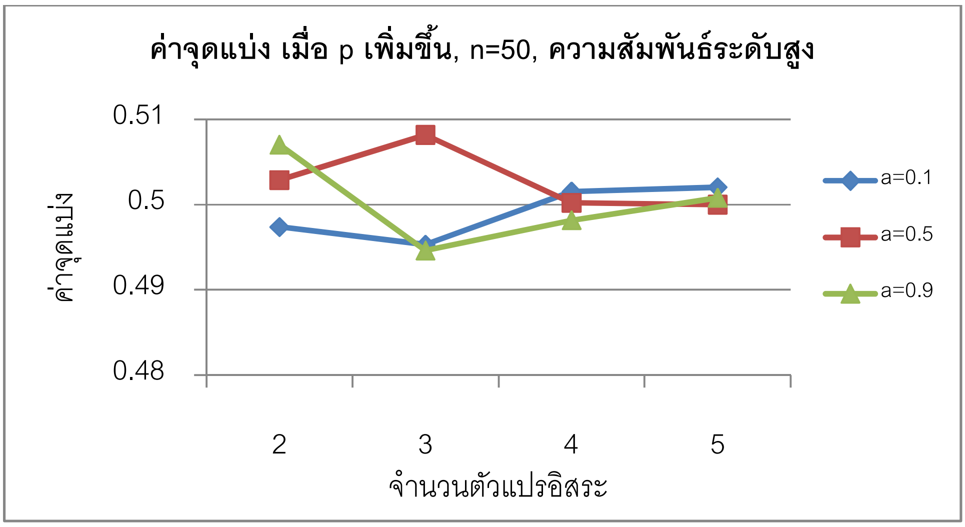
** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.9 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง (0.60 $< Max\{|r_{ij}|\} \leq$ 0.90) ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่จำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

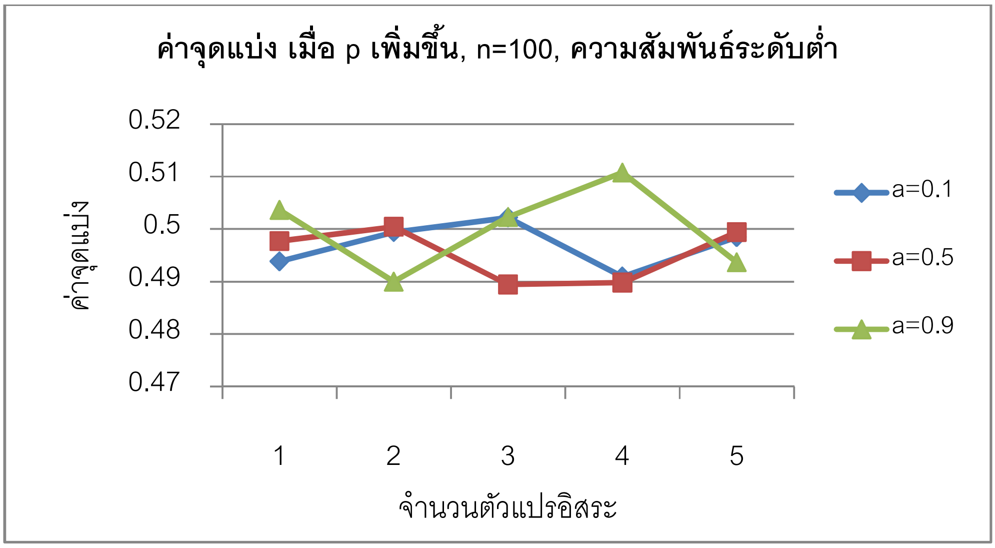
ภาพที่ 4.1 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระขนาดตัวอย่างเปลี่ยนแปลง เมื่อขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

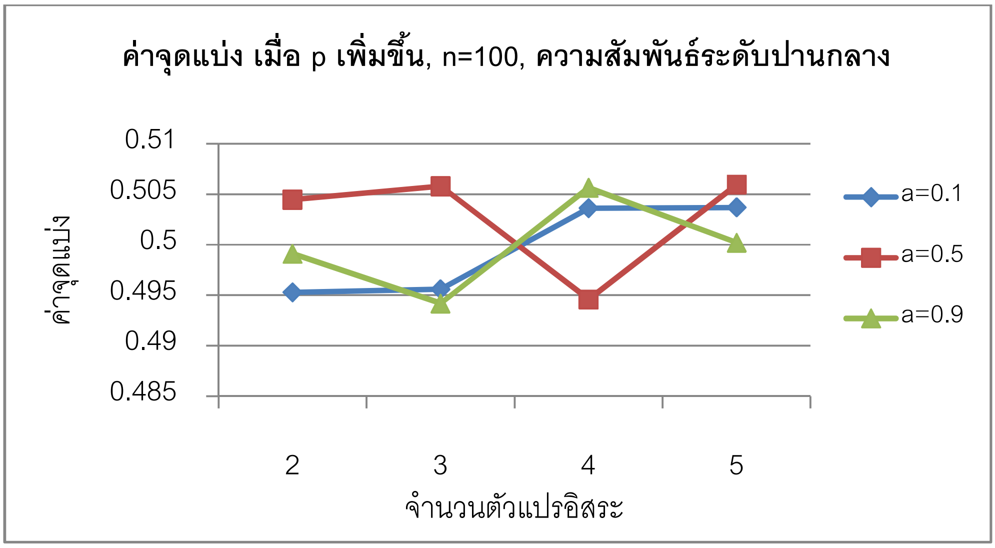


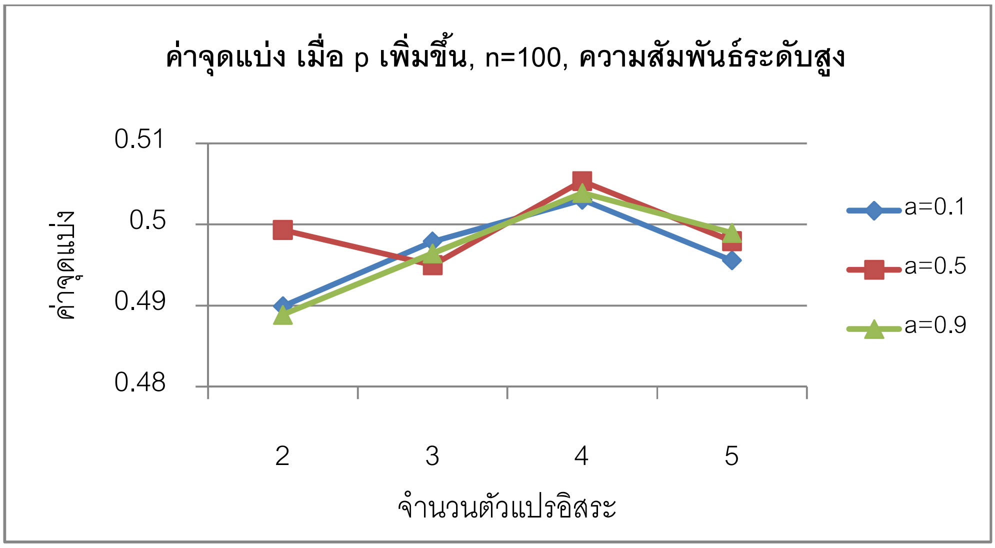




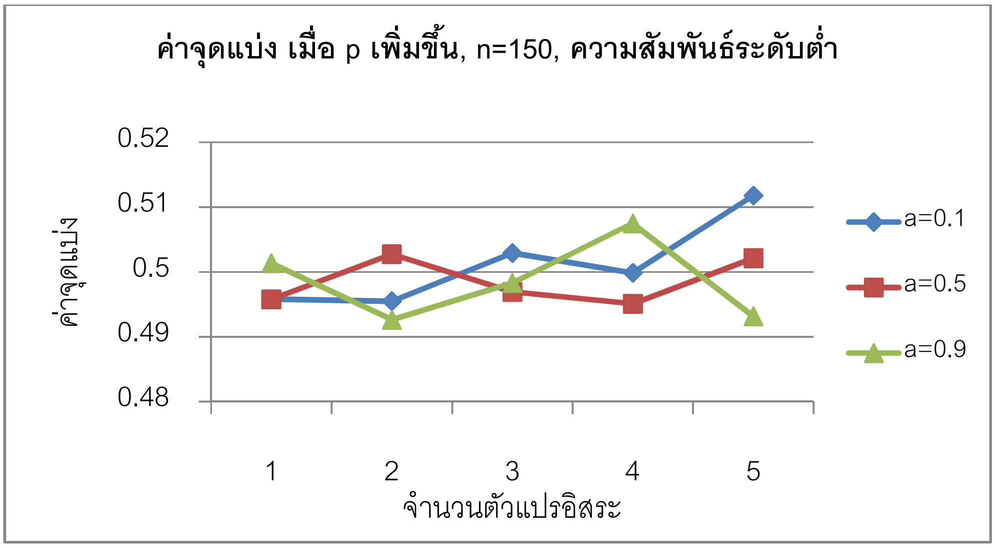
ภาพที่ 4.1 (ต่อ)

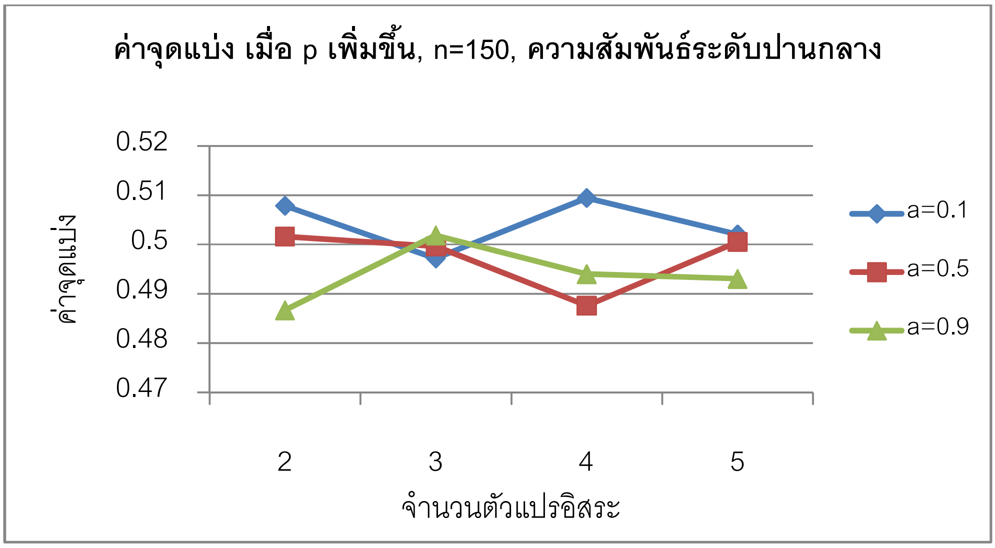

ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=100, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
1
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

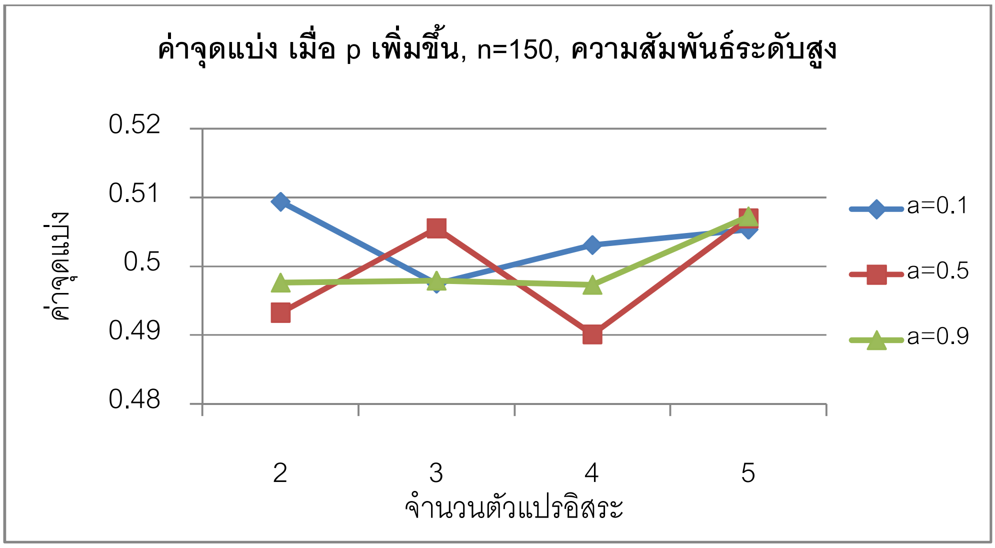

ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=100, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.51
0.505
0.5
0.495
0.49
0.485
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=100, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

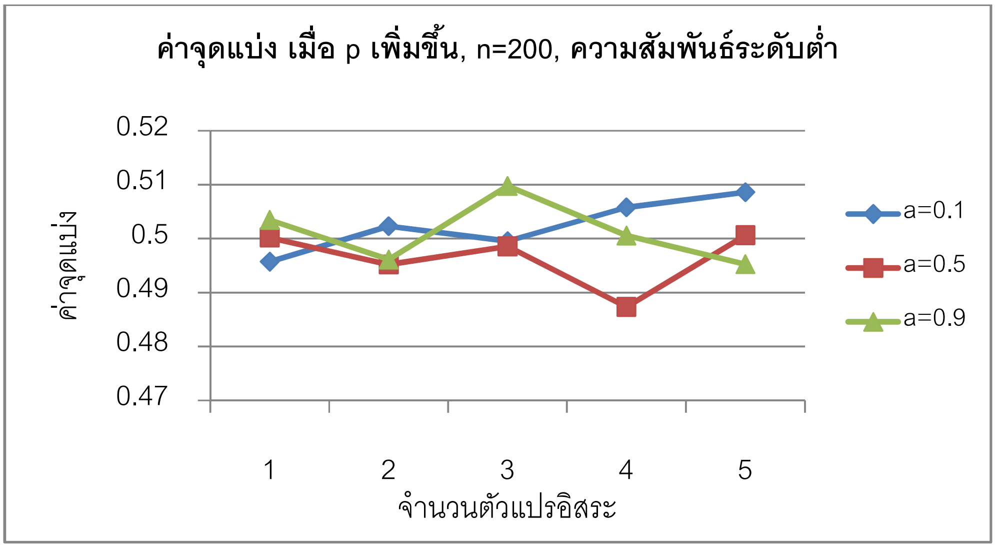
ภาพที่ 4.1 (ต่อ)


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=150, ความสัมพันธ์ระดับต่ำ
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
1
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

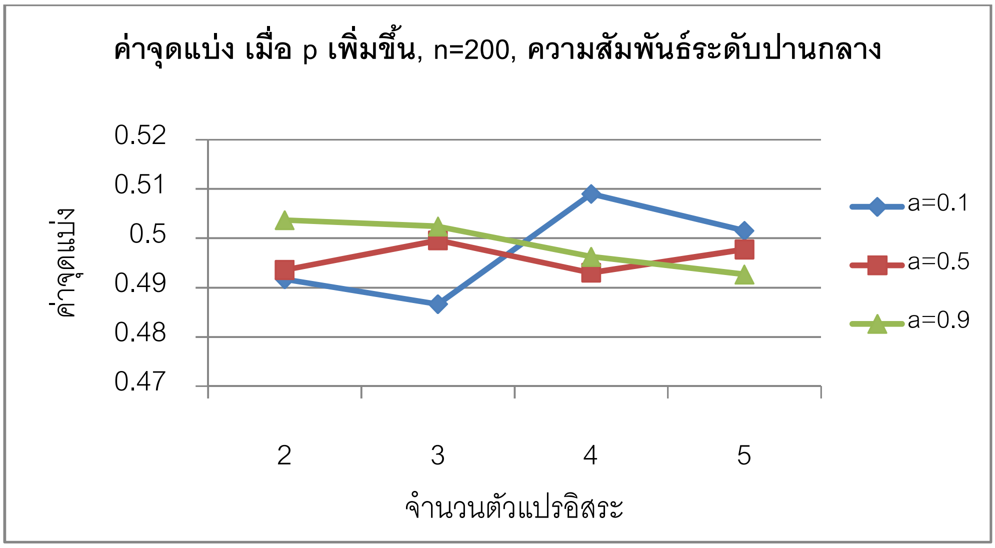

ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=150, ความสัมพันธ์ระดับปานกลาง
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=150, ความสัมพันธ์ระดับสูง
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

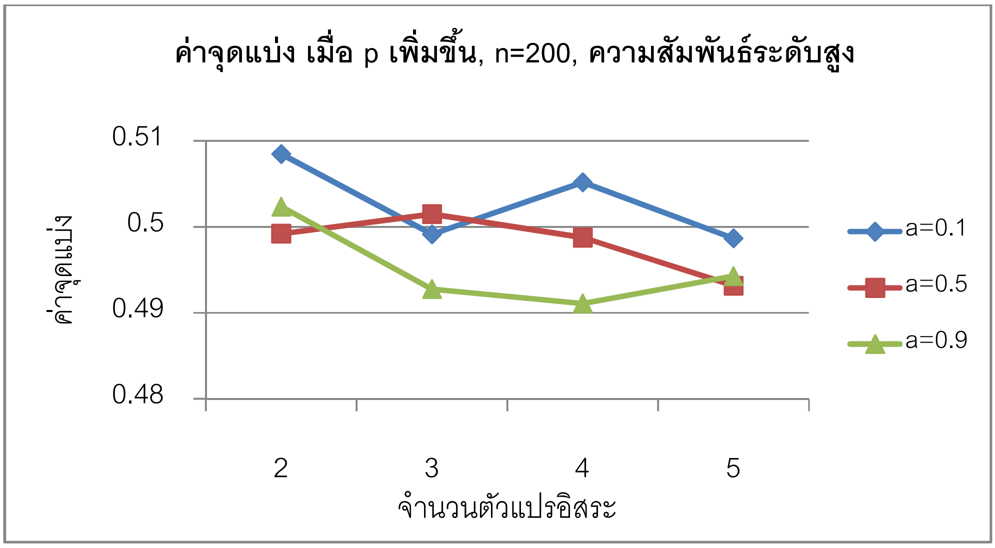
ภาพที่ 4.1 (ต่อ)



ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=200, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
1
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

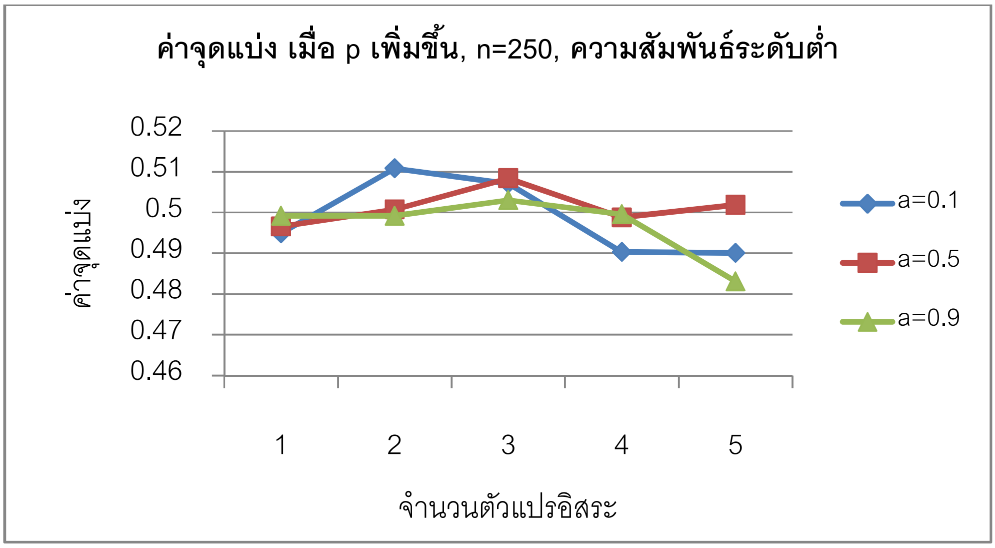


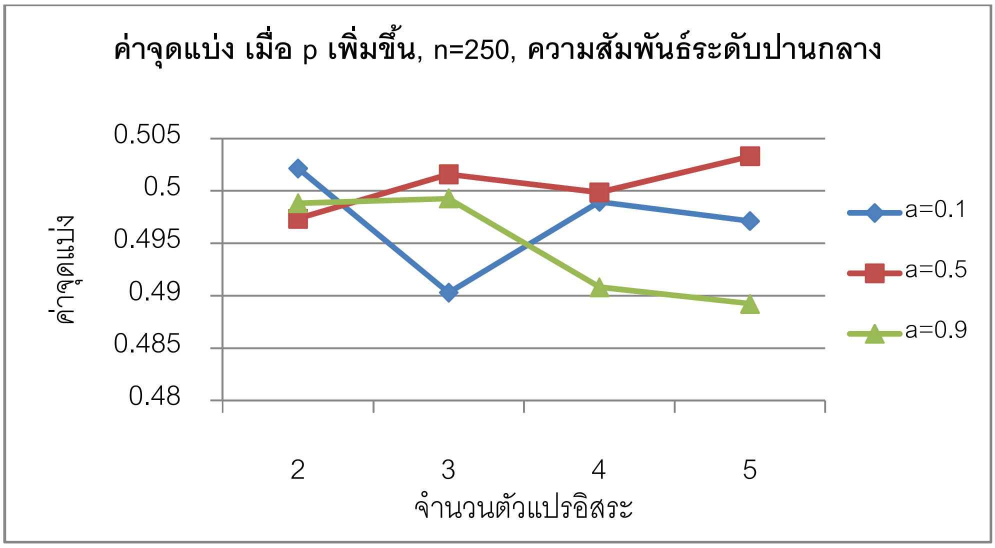
ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=200, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

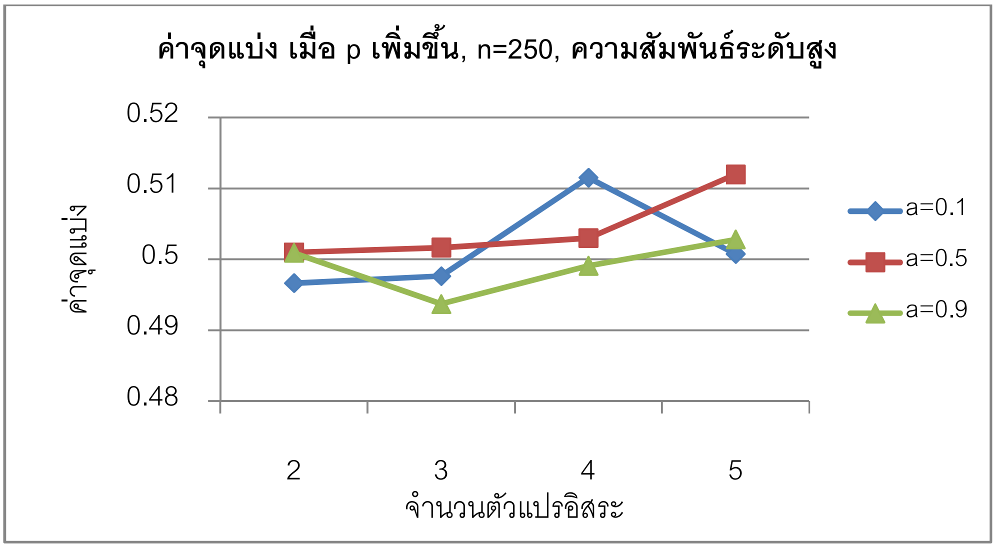


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=200, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

ภาพที่ 4.1 (ต่อ)


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=250, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.46
1
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=250, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.505
0.5
0.495
0.49
0.485
0.48
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ p เพิ่มขึ้น, n=250, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
2
3
4
5
จำนวนตัวแปรอิสระ
a=0.1
a=0.5
a=0.9

4.2 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.10-4.18

ตารางที่ 4.10 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\}\leq$ 0.30) จำนวนตัวแปรอิสระ (p) เท่ากับ 1, 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0	1	50	0.498256	0.348717	0.659052
	0		100	0.493823	0.378832	0.612861
	0		150	0.495794	0.386127	0.596221
	0		200	0.495721	0.406011	0.588072
	0		250	0.494895	0.406844	0.579196
	0.1886804	2	50	0.500867	0.22381	0.774946
	0.2078653		100	0.4994	0.236924	0.765269
	0.2177278		150	0.495476	0.247742	0.743235
	0.2239275		200	0.502327	0.272957	0.72627
	0.2287889		250	0.510853	0.291907	0.733491
	0.2350503	3	50	0.501172	0.250077	0.753873
	0.2529071		100	0.502182	0.240616	0.770389
	0.2549317		150	0.502931	0.242619	0.768415
	0.2584532		200	0.499529	0.252766	0.749504
	0.2609532		250	0.507128	0.27192	0.726269
	0.254126	4	50	0.499337	0.260935	0.746921
	0.2655062		100	0.490926	0.216475	0.772606
	0.2674501		150	0.499868	0.216776	0.771611
	0.2726991		200	0.505795	0.257207	0.746059
	0.2731542		250	0.490329	0.247489	0.745463

ตารางที่ 4.10 (ต่อ)

a	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.2695499	5	50	0.496327	0.285974	0.715822
	0.2734939		100	0.498461	0.244965	0.758303
	0.2787541		150	0.511795	0.240792	0.764196
	0.2781472		200	0.508619	0.223032	0.766436
	0.2807881		250	0.490082	0.227694	0.761813

จากตารางที่ 4.10 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.30) จำนวนตัวแปรอิสระ เท่ากับ 1, 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่จำนวนตัวแปร เท่ากับ 1 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.11 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.60) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.4994576	2	50	0.500675	0.235503	0.772616
	0.5219221		100	0.49528	0.227029	0.758286
	0.5299841		150	0.507824	0.252215	0.756141
	0.5359944		200	0.491696	0.256523	0.731223
	0.5372966		250	0.502143	0.266988	0.743305
	0.5385911	3	50	0.501844	0.242707	0.74924
	0.5554746		100	0.495576	0.228455	0.779218
	0.556986		150	0.497141	0.221334	0.773431
	0.5587158		200	0.486614	0.228105	0.75554
	0.5607767		250	0.490277	0.266541	0.748254

ตารางที่ 4.11 (ต่อ)

a	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.5565136	4	50	0.498926	0.278862	0.707956
	0.5640089		100	0.503611	0.235779	0.757799
	0.566902		150	0.509403	0.240063	0.782342
	0.5693879		200	0.508991	0.236898	0.778514
	0.5701132		250	0.49896	0.22798	0.757359
	0.5645582	5	50	0.499102	0.377179	0.514213
	0.570552		100	0.50368	0.254696	0.73254
	0.5725225		150	0.501993	0.238995	0.773778
	0.5739922		200	0.501506	0.233406	0.774567
	0.5768426		250	0.497107	0.227003	0.78334

จากตารางที่ 4.11 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.60) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.12 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.90) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.8393408	2	50	0.497366	0.22066	0.771432
	0.847068		100	0.4899	0.208386	0.777063
	0.8487206		150	0.509393	0.237081	0.791051
	0.8491737		200	0.496598	0.264818	0.731034
	0.849504		250	0.508489	0.242717	0.770344

ตารางที่ 4.12 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	0.8620921	3	50	0.495291	0.268113	0.719601
	0.863123		100	0.497898	0.250699	0.766899
	0.8607247		150	0.497446	0.230128	0.775159
	0.8603123		200	0.499144	0.220767	0.760301
	0.8600407		250	0.497591	0.23843	0.747535
	0.8699663	4	50	0.501512	0.402216	0.647958
	0.8722286		100	0.503028	0.252933	0.76234
	0.8680124		150	0.503083	0.241903	0.778957
	0.8675026		200	0.505198	0.225411	0.783045
	0.8667469		250	0.511508	0.227845	0.789503
	0.8747025	5	50	0.502021	0.50000	0.575944*
	0.8758678		100	0.495558	0.267639	0.702735
	0.8730344		150	0.505336	0.280116	0.724207
	0.871125		200	0.498684	0.24358	0.758786
	0.869478		250	0.50073	0.230196	0.769607

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.12 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง (0.60 $<$ $Max\{|r_{ij}|\} \leq$ 0.90) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่จำนวนตัวแปร เท่ากับ 3 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย และที่จำนวนตัวแปรอิสระ เท่ากับ 4ค่าเฉลี่ยของจุดแบ่งเพิ่มขึ้นจากค่า 0.5 เพียงเล็กน้อย เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.13 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) จำนวนตัวแปรอิสระ (p) เท่ากับ 1, 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0	1	50	0.503665	0.365606	0.656167
	0		100	0.497691	0.387556	0.610248
	0		150	0.495785	0.396597	0.591528
	0		200	0.500222	0.411185	0.589516
	0		250	0.496572	0.412955	0.578279
	0.1864547	2	50	0.500494	0.237315	0.767111
	0.2047484		100	0.50044	0.237995	0.770704
	0.2190896		150	0.502754	0.270441	0.729218
	0.2222356		200	0.495216	0.258498	0.717659
	0.2338323		250	0.50075	0.271328	0.71322
	0.2348507	3	50	0.500676	0.2304	0.752722
	0.251319		100	0.48944	0.231358	0.769243
	0.2568403		150	0.496919	0.233292	0.745431
	0.2585383		200	0.498575	0.258514	0.754046
	0.2616638		250	0.508402	0.265561	0.740355
	0.2601783	4	50	0.507427	0.273588	0.740752
	0.2672831		100	0.489803	0.224208	0.762047
	0.2680613		150	0.495113	0.248984	0.777624
	0.271567		200	0.487352	0.234282	0.770728
	0.2744325		250	0.49877	0.230719	0.754654

ตารางที่ 4.13 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.2685089	5	50	0.497784	0.289586	0.662065
	0.2754783		100	0.499431	0.250611	0.758003
	0.2775778		150	0.502148	0.222499	0.788029
	0.2798083		200	0.500622	0.226709	0.772803
	0.2792333		250	0.501861	0.254699	0.751812

จากตารางที่ 4.13 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) จำนวนตัวแปรอิสระ เท่ากับ 1, 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 1, 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.14 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.5007751	2	50	0.501119	0.228249	0.771341
	0.5164418		100	0.504465	0.226986	0.775911
	0.5309888		150	0.501573	0.239628	0.766404
	0.5339025		200	0.493578	0.238064	0.735588
	0.5391257		250	0.497347	0.250164	0.720739
	0.535131	3	50	0.500478	0.257131	0.735763
	0.5511936		100	0.505789	0.228973	0.771086
	0.5559691		150	0.499629	0.232451	0.774833
	0.5603839		200	0.499552	0.236095	0.772088
	0.5596843		250	0.501579	0.252983	0.735728

ตารางที่ 4.14 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.557933	4	50	0.50057	0.287409	0.731583
	0.5632398		100	0.494576	0.238845	0.76584
	0.5662719		150	0.487577	0.231274	0.764482
	0.5686589		200	0.493009	0.220505	0.755097
	0.5716462		250	0.499859	0.237691	0.77363
	0.5668828	5	50	0.501685	0.493776*	0.510152**
	0.5708508		100	0.505952	0.274687	0.759896
	0.5745328		150	0.500495	0.249938	0.7688
	0.5754855		200	0.497713	0.236977	0.758642
	0.5750038		250	0.5033	0.227734	0.774936

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.14 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 $< Max\{\lvert r_{ij}\rvert\} \leq$ 0.60) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.15 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 $< Max\{\lvert r_{ij}\rvert\} \leq$ 0.90) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.8419899	2	50	0.502854	0.22429	0.75836
	0.8471553		100	0.499313	0.228393	0.769043
	0.8487382		150	0.493233	0.224216	0.759169
	0.8489849		200	0.499217	0.248745	0.74121
	0.8497455		250	0.500984	0.269106	0.73085

ตารางที่ 4.15 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	0.8605197	3	50	0.508168	0.270338	0.738057
	0.8632002		100	0.494965	0.232188	0.774918
	0.8622969		150	0.505509	0.239588	0.766192
	0.8630131		200	0.501496	0.230034	0.757002
	0.8614076		250	0.501652	0.239267	0.764877
	0.8696018	4	50	0.500205	0.369926	0.638865
	0.8714558		100	0.505334	0.267485	0.73557
	0.8695362		150	0.490099	0.219119	0.778358
	0.8671073		200	0.498765	0.219758	0.765132
	0.8661106		250	0.502993	0.242588	0.760802
	0.8758058	5	50	0.499978	0.377643*	0.620358**
	0.8749598		100	0.497911	0.266185	0.680201
	0.8732464		150	0.506921	0.266837	0.751978
	0.8718601		200	0.493159	0.257281	0.742426
	0.8693057		250	0.511941	0.23163	0.78244

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.15 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.16 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\left\{\left|r_{ij}\right|\right\} \leq$ 0.30) จำนวนตัวแปรอิสระ (p) เท่ากับ 1, 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\left\{\left|r_{ij}\right|\right\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0	1	50	0.509422	0.37607	0.648653
	0		100	0.503668	0.397134	0.613246
	0		150	0.501341	0.400452	0.59376
	0		200	0.503505	0.410572	0.590867
	0		250	0.49921	0.415682	0.582303
	0.1821284	2	50	0.507494	0.217843	0.775718
	0.2042958		100	0.489978	0.231292	0.755361
	0.2160963		150	0.492633	0.249681	0.749831
	0.2226641		200	0.49615	0.2638	0.730161
	0.2318075		250	0.499183	0.289666	0.704873
	0.2333454	3	50	0.498069	0.232993	0.754885
	0.2481321		100	0.502313	0.227633	0.768928
	0.253179		150	0.4983	0.233279	0.772768
	0.2561202		200	0.509764	0.233451	0.756943
	0.2617862		250	0.503039	0.257576	0.724219
	0.2570051	4	50	0.497186	0.260612	0.757711
	0.264764		100	0.510789	0.233674	0.774722
	0.2711085		150	0.507487	0.225612	0.770483
	0.2731122		200	0.500595	0.240592	0.770429
	0.2754604		250	0.499596	0.256919	0.761046

ตารางที่ 4.16 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.2689239	5	50	0.500073	0.292309	0.679598
	0.2750965		100	0.493672	0.238426	0.761804
	0.2760721		150	0.493163	0.237269	0.772208
	0.2793923		200	0.495289	0.238201	0.769221
	0.2803318		250	0.483096	0.241996	0.740038

จากตารางที่ 4.16 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ (0 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.30) จำนวนตัวแปรอิสระ เท่ากับ 1, 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 1, 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.17 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.60) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.4953218	2	50	0.507027	0.237697	0.767406
	0.5206238		100	0.499105	0.230647	0.74143
	0.5275973		150	0.486646	0.238289	0.754875
	0.5393825		200	0.503666	0.265697	0.735765
	0.5416306		250	0.498816	0.250122	0.726159
	0.5354911	3	50	0.494613	0.234522	0.73211
	0.5506931		100	0.494173	0.222366	0.766775
	0.5588926		150	0.501856	0.216915	0.787535
	0.5577018		200	0.502397	0.245572	0.751992
	0.562273		250	0.499273	0.247986	0.747294

ตารางที่ 4.17 (ต่อ)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.5532036	4	50	0.498156	0.30295	0.679195
	0.5657834		100	0.505651	0.234338	0.75323
	0.5673897		150	0.494009	0.21649	0.75806
	0.5706408		200	0.490808	0.222459	0.758633
	0.5717861		250	0.496242	0.215561	0.783827
	0.5625658	5	50	0.500792	0.469167	0.60962
	0.5710766		100	0.500203	0.244628	0.736526
	0.5727453		150	0.493036	0.236913	0.763234
	0.5753501		200	0.492654	0.236066	0.771419
	0.5757055		250	0.489227	0.230261	0.768856

จากตารางที่ 4.17 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง (0.30 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.60) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2, 3, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.18 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{\lvert r_{ij}\rvert\} \leq$ 0.90) จำนวนตัวแปรอิสระ (p) เท่ากับ 2, 3, 4 และ 5 โดยจำแนกตามขนาดตัวอย่าง (n)

a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.8425875	2	50	0.508642	0.254318	0.764763
	0.8438331		100	0.488845	0.214951	0.76145
	0.8490023		150	0.500899	0.262279	0.751556
	0.8492548		200	0.497619	0.22453	0.764602
	0.8494049		250	0.50235	0.237538	0.762992

ตารางที่ 4.18 (ต่อ)

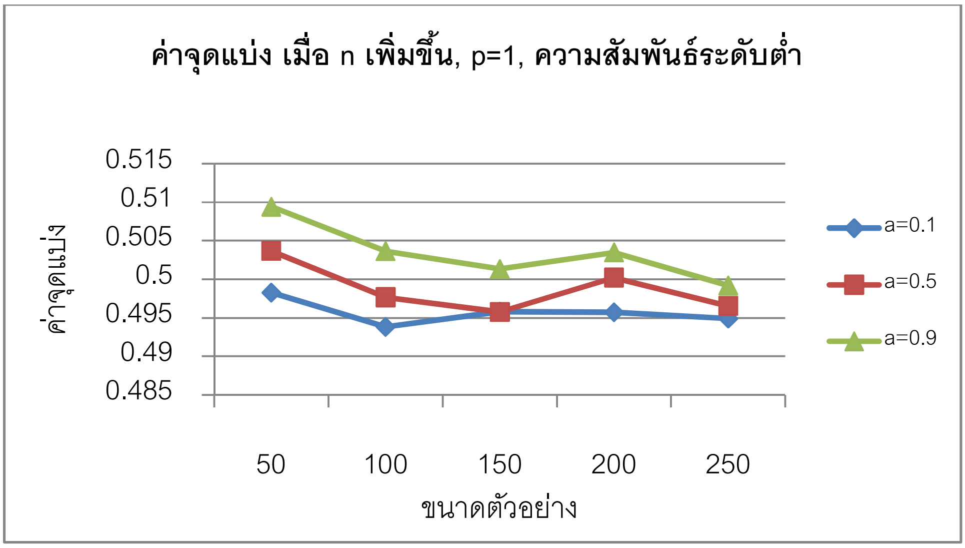
a	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	p	n	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	0.8605327	3	50	0.495267	0.269261	0.698625
	0.8638203		100	0.496422	0.227551	0.762967
	0.8616446		150	0.497891	0.236194	0.759695
	0.8613328		200	0.492758	0.221162	0.764929
	0.8603164		250	0.493694	0.227981	0.775343
	0.8700897	4	50	0.50136	0.379242	0.641594
	0.8706262		100	0.503881	0.253366	0.751583
	0.8698966		150	0.4973	0.231858	0.777655
	0.8672836		200	0.491078	0.227777	0.763907
	0.8656913		250	0.499081	0.219553	0.780762
	0.8745183	5	50	0.501531	0.495975*	0.640162**
	0.87495		100	0.498957	0.298155	0.714686
	0.8730918		150	0.507266	0.256175	0.753777
	0.8710947		200	0.494261	0.231991	0.755092
	0.8700021		250	0.50278	0.256239	0.767216

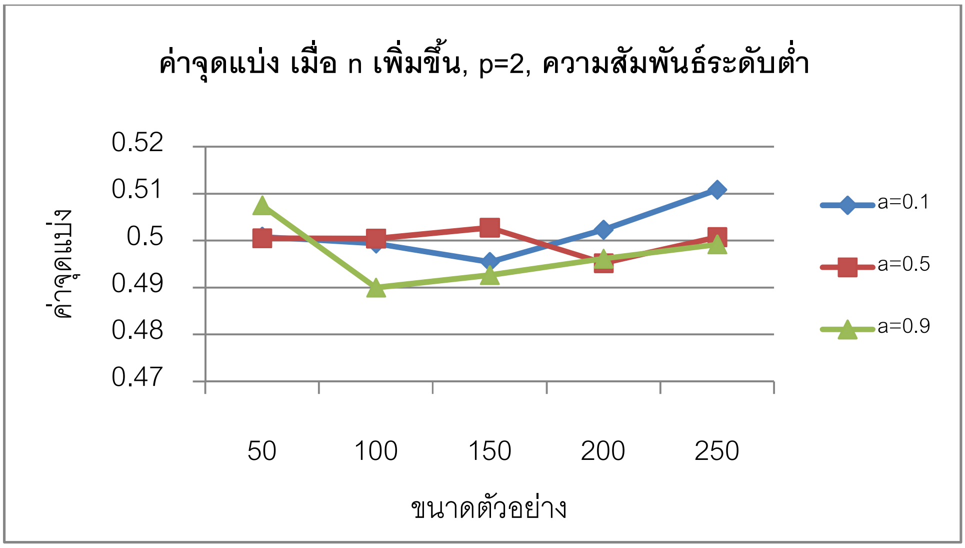
* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

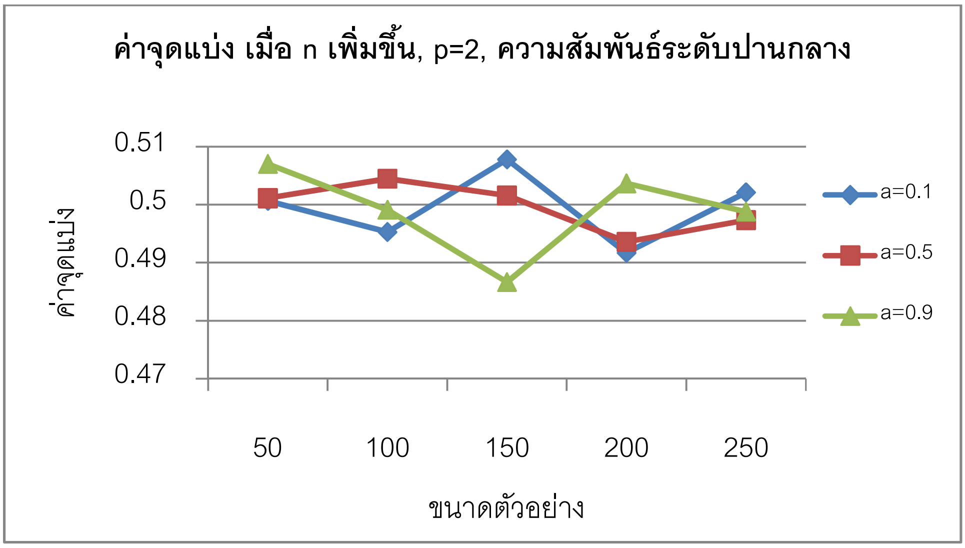
** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.18 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง (0.60 $< Max\{\lvert r_{ij}\rvert\} \leq$ 0.90) จำนวนตัวแปรอิสระ เท่ากับ 2, 3, 4 และ 5 แต่ขนาดตัวอย่างเพิ่มขึ้น พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2, 4, 5 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่จำนวนตัวแปร เท่ากับ 3 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อขนาดตัวอย่างเพิ่มขึ้น

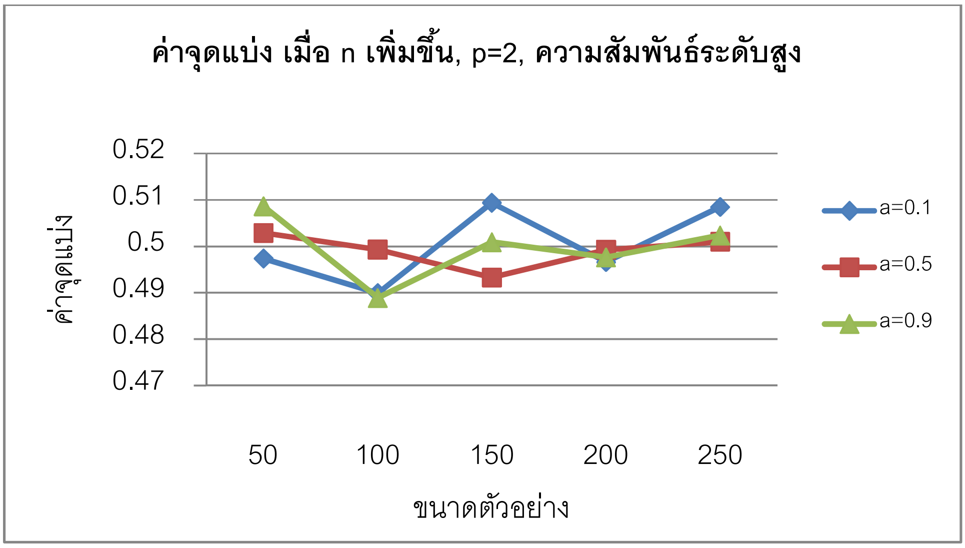
ภาพที่ 4.2 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่ขนาดตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

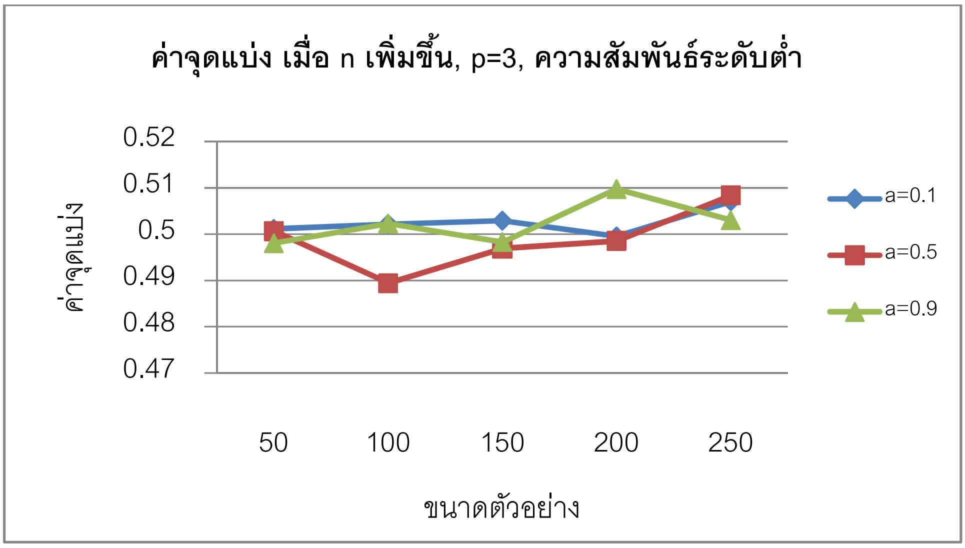


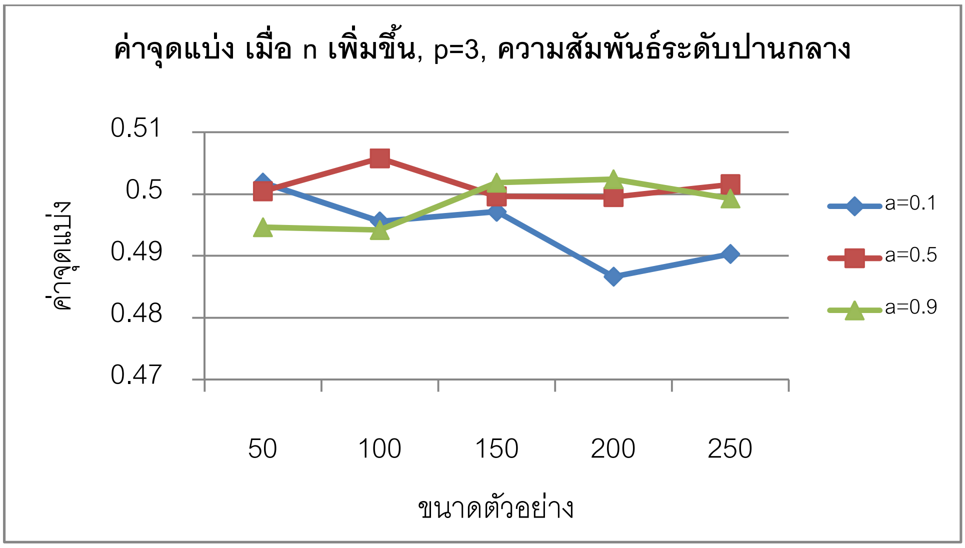




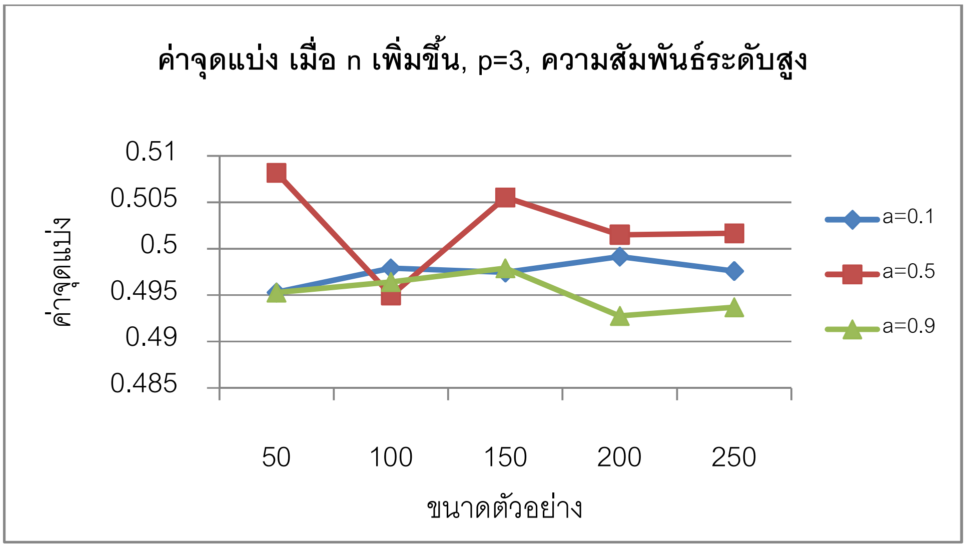
ภาพที่ 4.2 (ต่อ)

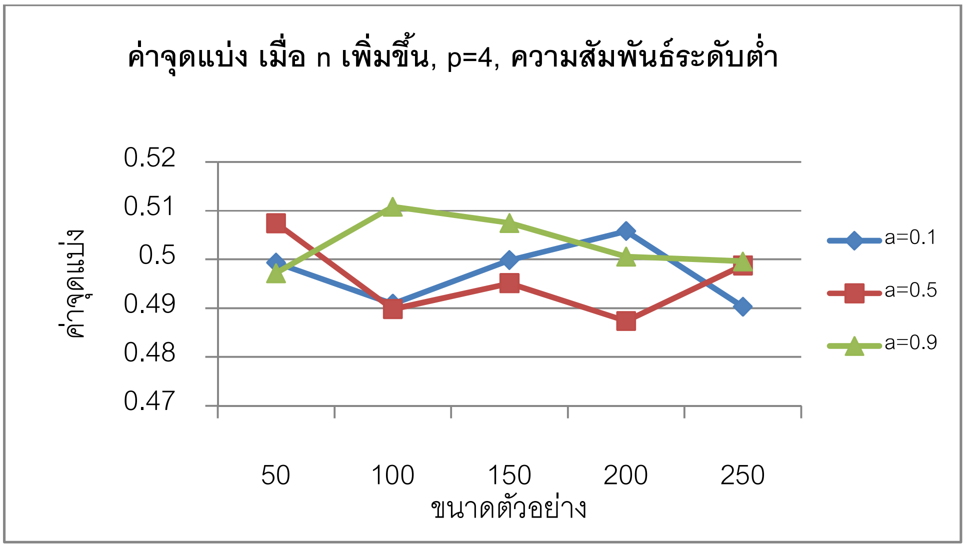

ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=2, ความสัมพันธ์ระดับสูง
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

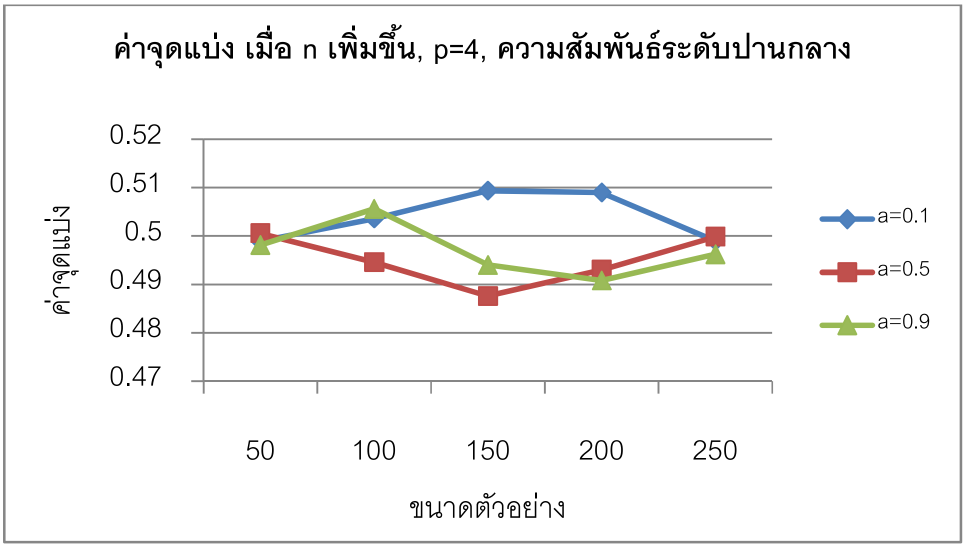

ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับต่ำ
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับปานกลาง
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

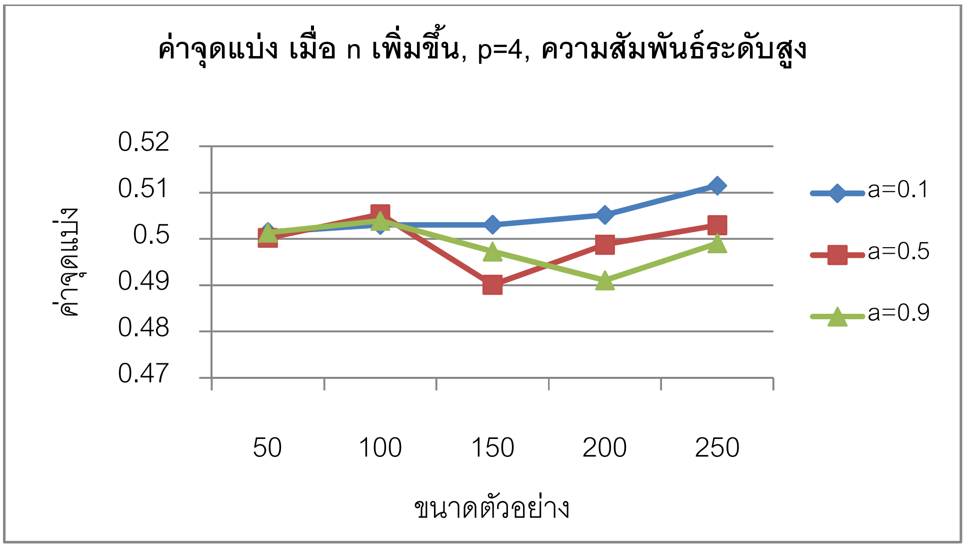
ภาพที่ 4.2 (ต่อ)

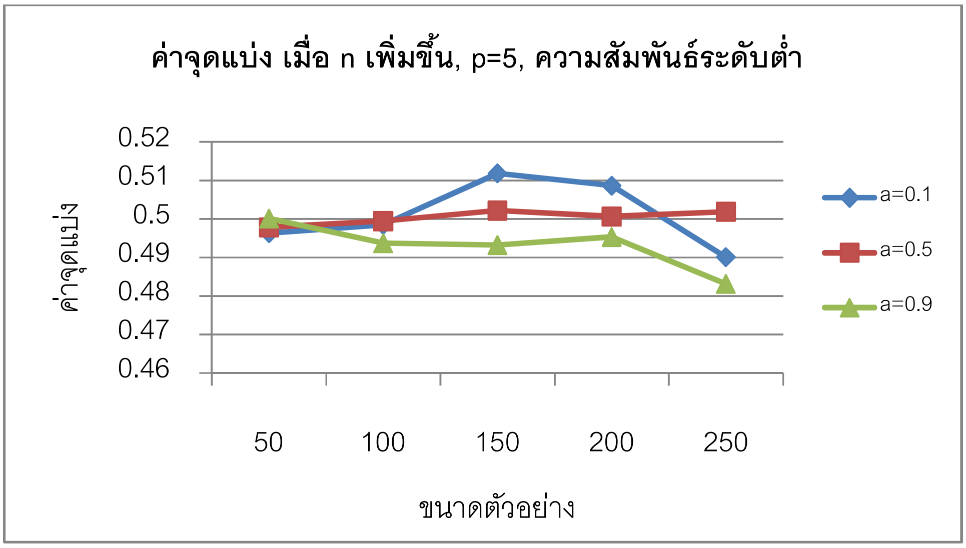

ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.51
0.505
0.5
0.495
0.49
0.485
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

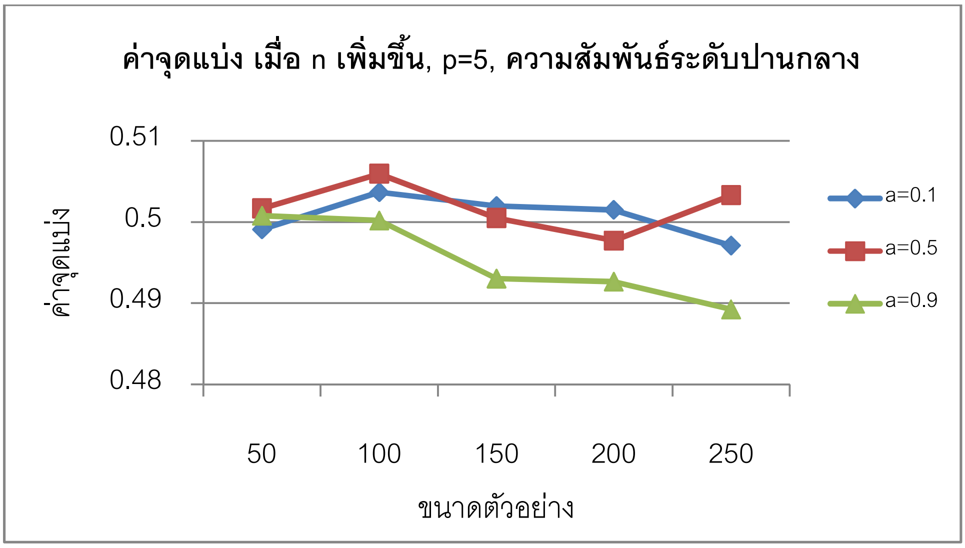

ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

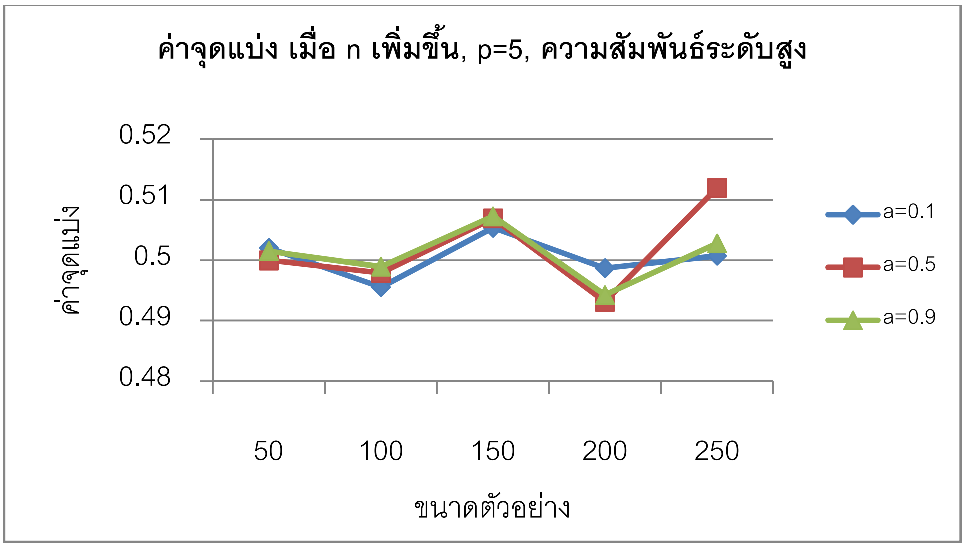
ภาพที่ 4.2 (ต่อ)


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.46
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

ภาพที่ 4.2 (ต่อ)


ค่าจุดแบ่ง เมื่อ n เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับสูง
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
50
100
150
200
250
ขนาดตัวอย่าง
a=0.1
a=0.5
a=0.9

4.3 กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่างและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.19-4.31

ตารางที่ 4.19 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 1 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
1	0	50	0.1	0.4982556	0.348717	0.659052
			0.5	0.503665	0.365606	0.656167
			0.9	0.509422	0.37607	0.648653
		100	0.1	0.493823	0.378832	0.612861
			0.5	0.497691	0.387556	0.610248
			0.9	0.503668	0.397134	0.613246
		150	0.1	0.495794	0.386127	0.596221
			0.5	0.495785	0.396597	0.591528
			0.9	0.501341	0.400452	0.59376
		200	0.1	0.495721	0.406011	0.588072
			0.5	0.500222	0.411185	0.589516
			0.9	0.503505	0.410572	0.590867
		250	0.1	0.494895	0.406844	0.579196
			0.5	0.496572	0.412955	0.578279
			0.9	0.49921	0.415682	0.582303

จากตารางที่ 4.19 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 1 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง

เท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.20 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
2	0.1886804	50	0.1	0.500867	0.22381	0.774946
	0.1864547		0.5	0.500494	0.237315	0.767111
	0.1821284		0.9	0.507494	0.217843	0.775718
	0.2078653	100	0.1	0.4994	0.236924	0.765269
	0.2047484		0.5	0.50044	0.237995	0.770704
	0.2042958		0.9	0.489978	0.231292	0.755361
	0.2177278	150	0.1	0.495476	0.247742	0.743235
	0.2190896		0.5	0.502754	0.270441	0.729218
	0.2160963		0.9	0.492633	0.249681	0.749831
	0.2239275	200	0.1	0.502327	0.272957	0.72627
	0.2222356		0.5	0.495216	0.258498	0.717659
	0.2226641		0.9	0.49615	0.2638	0.730161
	0.2287889	250	0.1	0.510853	0.291907	0.733491
	0.2338323		0.5	0.50075	0.271328	0.71322
	0.2318075		0.9	0.499183	0.289666	0.704873

จากตารางที่ 4.20 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.21 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
2	0.4994576	50	0.1	0.500675	0.235503	0.772616
	0.5007751		0.5	0.501119	0.228249	0.771341
	0.4953218		0.9	0.507027	0.237697	0.767406
	0.5219221	100	0.1	0.49528	0.227029	0.758286
	0.5164418		0.5	0.504465	0.226986	0.775911
	0.5206238		0.9	0.499105	0.230647	0.74143
	0.5299841	150	0.1	0.507824	0.252215	0.756141
	0.5309888		0.5	0.501573	0.239628	0.766404
	0.5275973		0.9	0.486646	0.238289	0.754875
	0.5359944	200	0.1	0.491696	0.256523	0.731223
	0.5339025		0.5	0.493578	0.238064	0.735588
	0.5393825		0.9	0.503666	0.265697	0.735765
	0.5372966	250	0.1	0.502143	0.266988	0.743305
	0.5391257		0.5	0.497347	0.250164	0.720739
	0.5416306		0.9	0.498816	0.250122	0.726159

จากตารางที่ 4.21 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่าที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.22 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
2	0.8393408	50	0.1	0.497366	0.22066	0.771432
	0.8419899		0.5	0.502854	0.22429	0.75836
	0.8425875		0.9	0.508642	0.254318	0.764763
	0.847068	100	0.1	0.4899	0.208386	0.777063
	0.8471553		0.5	0.499313	0.228393	0.769043
	0.8438331		0.9	0.488845	0.214951	0.76145
	0.8487206	150	0.1	0.509393	0.237081	0.791051
	0.8487382		0.5	0.493233	0.224216	0.759169
	0.8492548		0.9	0.497619	0.22453	0.764602
	0.849504	200	0.1	0.508489	0.242717	0.770344
	0.8489849		0.5	0.499217	0.248745	0.74121
	0.8494049		0.9	0.50235	0.237538	0.762992
	0.8491737	250	0.1	0.496598	0.264818	0.731034
	0.8497455		0.5	0.500984	0.269106	0.73085
	0.8490023		0.9	0.500899	0.262279	0.751556

จากตารางที่ 4.22 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 100 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.23 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{|r_{ij}|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
3	0.2350503	50	0.1	0.501172	0.250077	0.753873
	0.2348507		0.5	0.500676	0.2304	0.752722
	0.2333454		0.9	0.498069	0.232993	0.754885
	0.2529071	100	0.1	0.502182	0.240616	0.770389
	0.251319		0.5	0.48944	0.231358	0.769243
	0.2481321		0.9	0.502313	0.227633	0.768928
	0.2549317	150	0.1	0.502931	0.242619	0.768415
	0.2568403		0.5	0.496919	0.233292	0.745431
	0.253179		0.9	0.4983	0.233279	0.772768
	0.2584532	200	0.1	0.499529	0.252766	0.749504
	0.2585383		0.5	0.498575	0.258514	0.754046
	0.2561202		0.9	0.509764	0.233451	0.756943
	0.2609532	250	0.1	0.507128	0.27192	0.726269
	0.2616638		0.5	0.508402	0.265561	0.740355
	0.2617862		0.9	0.503039	0.257576	0.724219

จากตารางที่ 4.23 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.24 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 $< Max\{|r_{ij}|\} \leq$ 0.60) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
3	0.5385911	50	0.1	0.501844	0.242707	0.74924
	0.535131		0.5	0.500478	0.257131	0.735763
	0.5354911		0.9	0.494613	0.234522	0.73211
	0.5554746	100	0.1	0.495576	0.228455	0.779218
	0.5511936		0.5	0.505789	0.228973	0.771086
	0.5506931		0.9	0.494173	0.222366	0.766775
	0.556986	150	0.1	0.497141	0.221334	0.773431
	0.5559691		0.5	0.499629	0.232451	0.774833
	0.5588926		0.9	0.501856	0.216915	0.787535
	0.5587158	200	0.1	0.486614	0.228105	0.75554
	0.5603839		0.5	0.499552	0.236095	0.772088
	0.5577018		0.9	0.502397	0.245572	0.751992
	0.5607767	250	0.1	0.490277	0.266541	0.748254
	0.5596843		0.5	0.501579	0.252983	0.735728
	0.562273		0.9	0.499273	0.247986	0.747294

จากตารางที่ 4.24 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 $< Max\{|r_{ij}|\} \leq$ 0.60) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่าที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.25 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	**ค่าเฉลี่ยของ** $Max\{|r_{ij}|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
3	0.8620921	50	0.1	0.495291	0.268113	0.719601
	0.8605197		0.5	0.508168	0.270338	0.738057
	0.8605327		0.9	0.495267	0.269261	0.698625
	0.863123	100	0.1	0.497898	0.250699	0.766899
	0.8632002		0.5	0.494965	0.232188	0.774918
	0.8638203		0.9	0.496422	0.227551	0.762967
	0.8607247	150	0.1	0.497446	0.230128	0.775159
	0.8622969		0.5	0.505509	0.239588	0.766192
	0.8616446		0.9	0.497891	0.236194	0.759695
	0.8603123	200	0.1	0.499144	0.220767	0.760301
	0.8630131		0.5	0.501496	0.230034	0.757002
	0.8613328		0.9	0.492758	0.221162	0.764929
	0.8600407	250	0.1	0.497591	0.23843	0.747535
	0.8614076		0.5	0.501652	0.239267	0.764877
	0.8603164		0.9	0.493694	0.227981	0.775343

จากตารางที่ 4.25 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 100 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.26 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
4	0.254126	50	0.1	0.499337	0.260935	0.746921
	0.2601783		0.5	0.507427	0.273588	0.740752
	0.2570051		0.9	0.497186	0.260612	0.757711
	0.2655062	100	0.1	0.490926	0.216475	0.772606
	0.2672831		0.5	0.489803	0.224208	0.762047
	0.264764		0.9	0.510789	0.233674	0.774722
	0.2674501	150	0.1	0.499868	0.216776	0.771611
	0.2680613		0.5	0.495113	0.248984	0.777624
	0.2711085		0.9	0.507487	0.225612	0.770483
	0.2726991	200	0.1	0.505795	0.257207	0.746059
	0.271567		0.5	0.487352	0.234282	0.770728
	0.2731122		0.9	0.500595	0.240592	0.770429
	0.2731542	250	0.1	0.490329	0.247489	0.745463
	0.2744325		0.5	0.49877	0.230719	0.754654
	0.2754604		0.9	0.499596	0.256919	0.761046

จากตารางที่ 4.26 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่างเท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.27 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
4	0.5565136	50	0.1	0.498926	0.278862	0.707956
	0.557933		0.5	0.50057	0.287409	0.731583
	0.5532036		0.9	0.498156	0.30295	0.679195
	0.5640089	100	0.1	0.503611	0.235779	0.757799
	0.5632398		0.5	0.494576	0.238845	0.76584
	0.5657834		0.9	0.505651	0.234338	0.75323
	0.566902	150	0.1	0.509403	0.240063	0.782342
	0.5662719		0.5	0.487577	0.231274	0.764482
	0.5673897		0.9	0.494009	0.21649	0.75806
	0.5693879	200	0.1	0.508991	0.236898	0.778514
	0.5686589		0.5	0.493009	0.220505	0.755097
	0.5717861		0.9	0.496242	0.215561	0.783827
	0.5701132	250	0.1	0.49896	0.22798	0.757359
	0.5716462		0.5	0.499859	0.237691	0.77363
	0.5706408		0.9	0.490808	0.222459	0.758633

จากตารางที่ 4.27 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.28 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	**ค่าเฉลี่ยของ** $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
4	0.8699663	50	0.1	0.501512	0.402216	0.647958
	0.8696018		0.5	0.500205	0.369926	0.638865
	0.8700897		0.9	0.50136	0.379242	0.641594
	0.8722286	100	0.1	0.503028	0.252933	0.76234
	0.8714558		0.5	0.505334	0.267485	0.73557
	0.8706262		0.9	0.503881	0.253366	0.751583
	0.8680124	150	0.1	0.503083	0.241903	0.778957
	0.8695362		0.5	0.490099	0.219119	0.778358
	0.8698966		0.9	0.4973	0.231858	0.777655
	0.8675026	200	0.1	0.505198	0.225411	0.783045
	0.8671073		0.5	0.498765	0.219758	0.765132
	0.8672836		0.9	0.491078	0.227777	0.763907
	0.8667469	250	0.1	0.511508	0.227845	0.789503
	0.8661106		0.5	0.502993	0.242588	0.760802
	0.8656913		0.9	0.499081	0.219553	0.780762

จากตารางที่ 4.28 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.29 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
5	0.2695499	50	0.1	0.496327	0.285974	0.715822
	0.2685089		0.5	0.497784	0.289586	0.662065
	0.2689239		0.9	0.500073	0.292309	0.679598
	0.2734939	100	0.1	0.498461	0.244965	0.758303
	0.2754783		0.5	0.499431	0.250611	0.758003
	0.2750965		0.9	0.493672	0.238426	0.761804
	0.2787541	150	0.1	0.511795	0.240792	0.764196
	0.2775778		0.5	0.502148	0.222499	0.788029
	0.2760721		0.9	0.493163	0.237269	0.772208
	0.2781472	200	0.1	0.508619	0.223032	0.766436
	0.2798083		0.5	0.500622	0.226709	0.772803
	0.2793923		0.9	0.495289	0.238201	0.769221
	0.2807881	250	0.1	0.490082	0.227694	0.761813
	0.2792333		0.5	0.501861	0.254699	0.751812
	0.2803318		0.9	0.483096	0.241996	0.740038

จากตารางที่ 4.29 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับต่ำ (0 < $Max\{|r_{ij}|\} \leq$ 0.30) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่างเท่ากับ 50, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่างเท่ากับ 100 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.30 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

p	ค่าเฉลี่ยของ $Max\{|r_{ij}|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
5	0.5645582	50	0.1	0.499102	0.377179	0.514213
	0.5668828		0.5	0.501685	0.493776*	0.510152**
	0.5625658		0.9	0.500792	0.469167	0.60962
	0.570552	100	0.1	0.50368	0.254696	0.73254
	0.5708508		0.5	0.505952	0.274687	0.759896
	0.5710766		0.9	0.500203	0.244628	0.736526
	0.5725225	150	0.1	0.501993	0.238995	0.773778
	0.5745328		0.5	0.500495	0.249938	0.7688
	0.5727453		0.9	0.493036	0.236913	0.763234
	0.5739922	200	0.1	0.501506	0.233406	0.774567
	0.5754855		0.5	0.497713	0.236977	0.758642
	0.5753501		0.9	0.492654	0.236066	0.771419
	0.5768426	250	0.1	0.497107	0.227003	0.78334
	0.5750038		0.5	0.5033	0.227734	0.774936
	0.5757055		0.9	0.489227	0.230261	0.768856

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.30 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับปานกลาง (0.30 < $Max\{|r_{ij}|\} \leq$ 0.60) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่าที่ขนาดตัวอย่าง เท่ากับ 50,100,150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.31 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) โดยจำแนกตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a)

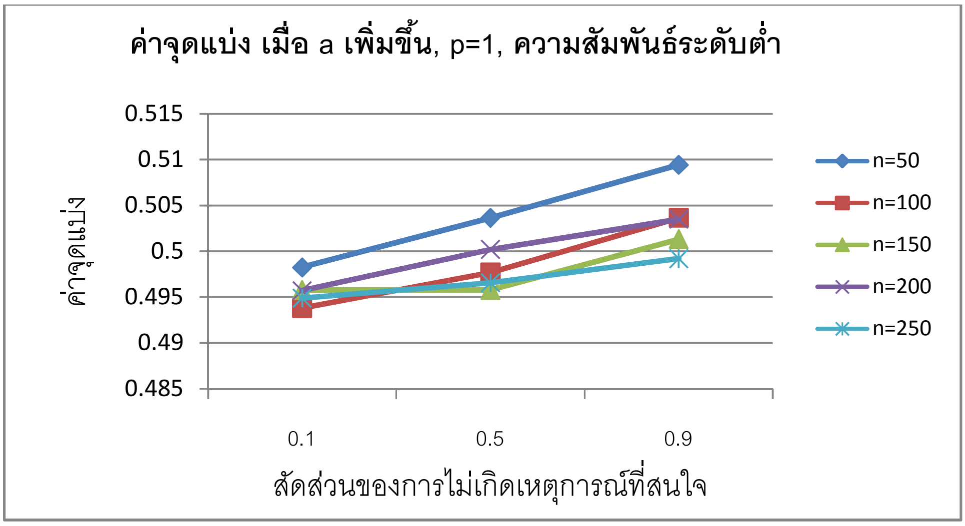
p	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	n	a	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
5	0.8747025	50	0.1	0.502021	0.500000	0.575944**
	0.8758058		0.5	0.499978	0.377643*	0.620358**
	0.8745183		0.9	0.501531	0.495975*	0.640162**
	0.8758678	100	0.1	0.495558	0.267639	0.702735
	0.8749598		0.5	0.497911	0.266185	0.680201
	0.87495		0.9	0.498957	0.298155	0.714686
	0.8730344	150	0.1	0.505336	0.280116	0.724207
	0.8732464		0.5	0.506921	0.266837	0.751978
	0.8730918		0.9	0.507266	0.256175	0.753777
	0.871125	200	0.1	0.498684	0.24358	0.758786
	0.8718601		0.5	0.493159	0.257281	0.742426
	0.8710947		0.9	0.494261	0.231991	0.755092
	0.869478	250	0.1	0.50073	0.230196	0.769607
	0.8693057		0.5	0.511941	0.23163	0.78244
	0.8700021		0.9	0.50278	0.256239	0.767216

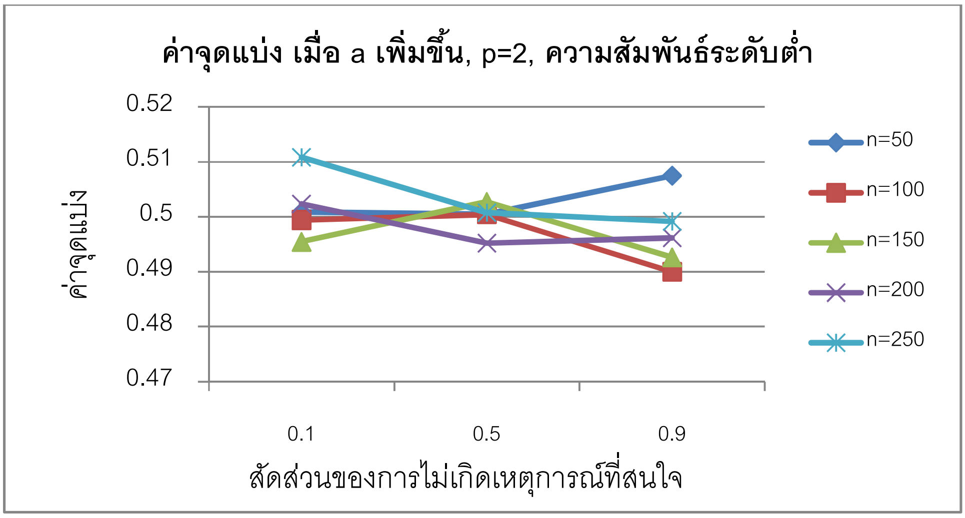
* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

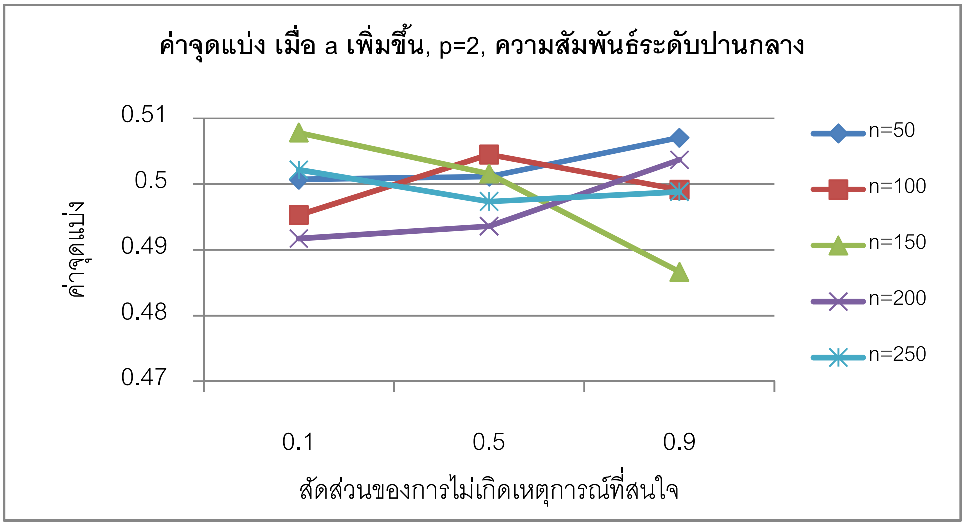
** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.31 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 ความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) อยู่ในระดับสูง (0.60 < $Max\{|r_{ij}|\} \leq$ 0.90) แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 100 และ 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

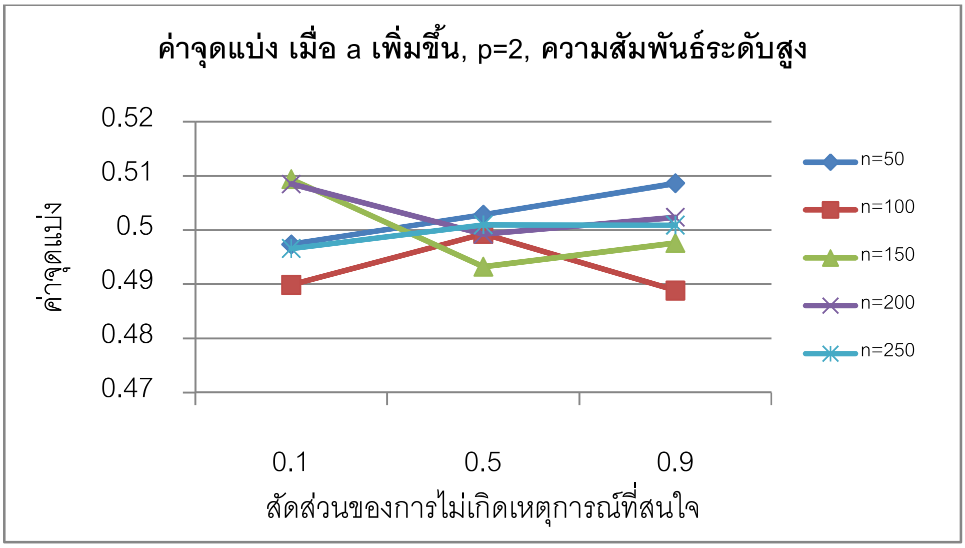
ภาพที่ 4.3 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่างและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

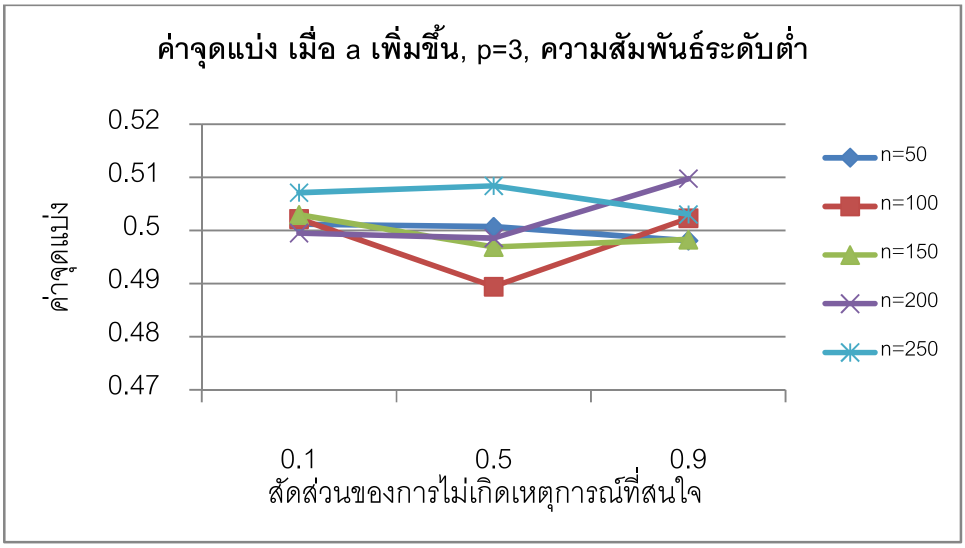


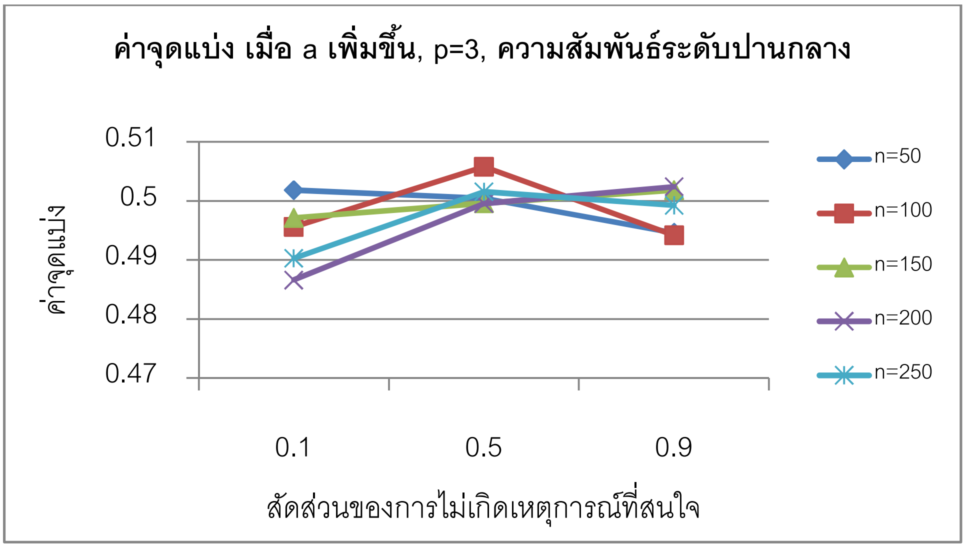




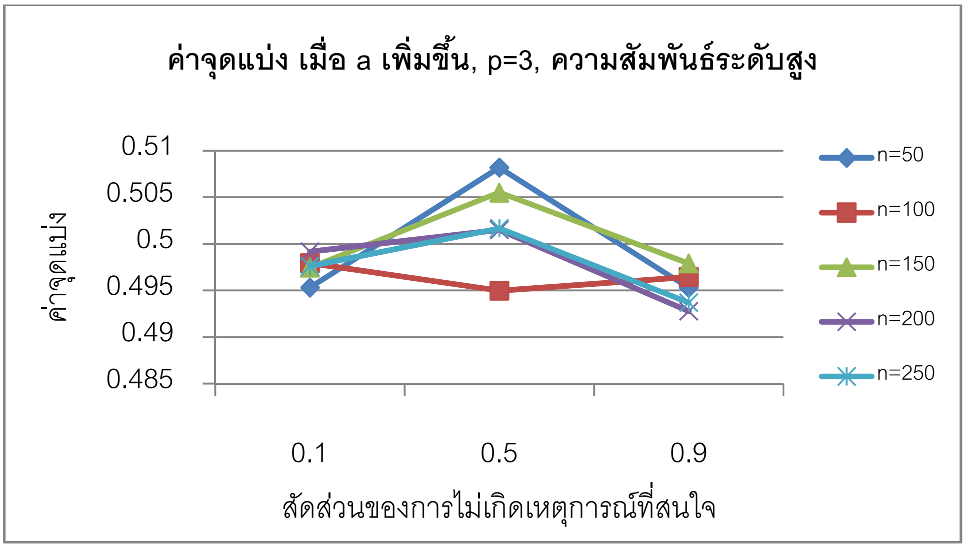
ภาพที่ 4.3 (ต่อ)

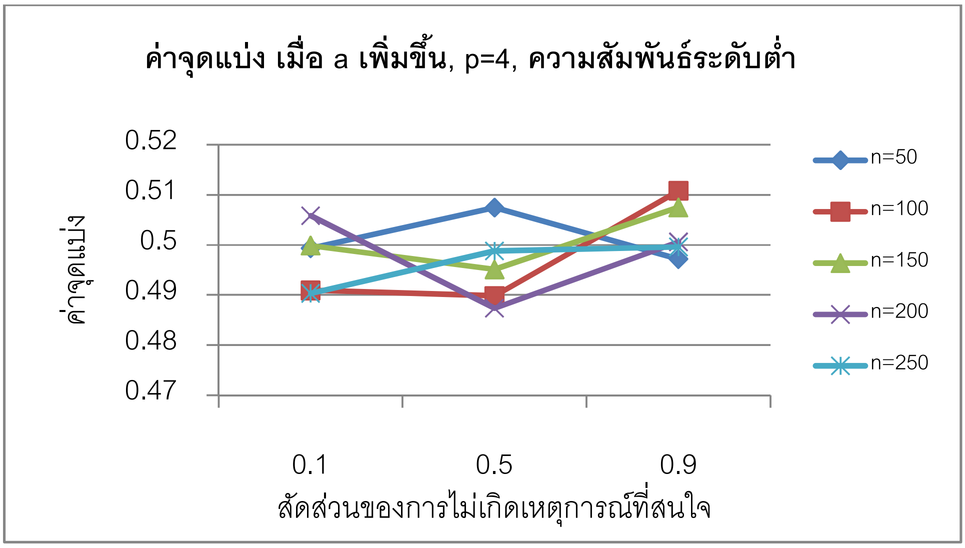

ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=2, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

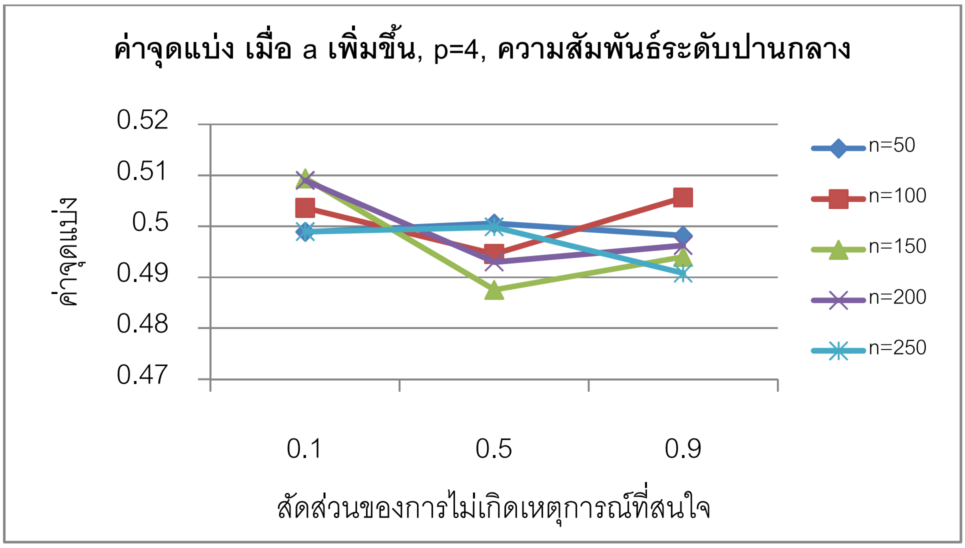

ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
0.47
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

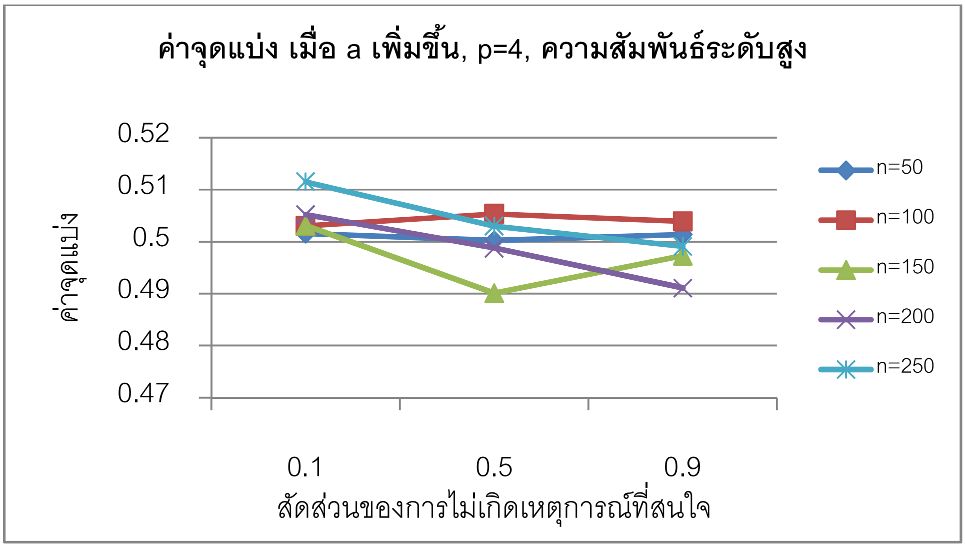
ภาพที่ 4.3 (ต่อ)

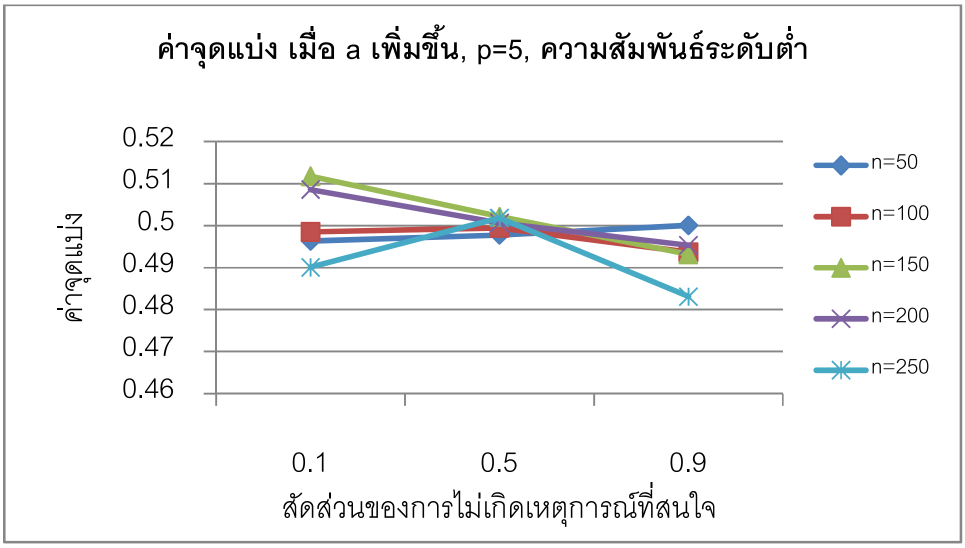

ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=3, ความสัมพันธ์ระดับสูง
0.51
0.505
0.5
0.495
0.49
0.485
ค่าจุดแบ่ง
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

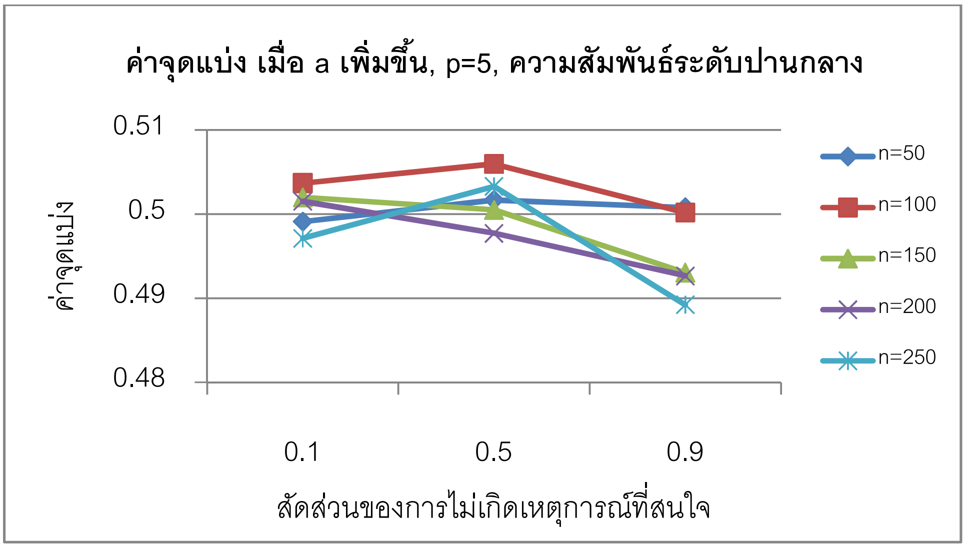

ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับต่ำ
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับปานกลาง
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

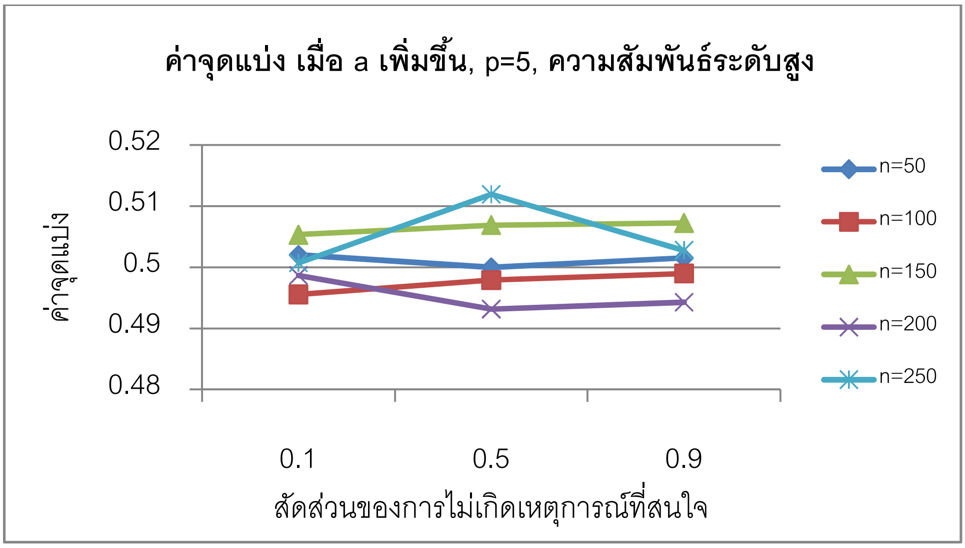
ภาพที่ 4.3 (ต่อ)


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=4, ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
0.46
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับปานกลาง
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

ภาพที่ 4.3 (ต่อ)


ค่าจุดแบ่ง เมื่อ a เพิ่มขึ้น, p=5, ความสัมพันธ์ระดับสูง
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
0.1
0.5
0.9
สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ
n=50
n=100
n=150
n=200
n=250

4.4 กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่างและสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.32-4.43

ตารางที่ 4.32 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 จำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	2	50	0.1886804	0.500867	0.22381	0.774946
			0.4994576	0.500675	0.235503	0.772616
			0.8393408	0.497366	0.22066	0.771432
		100	0.2078653	0.4994	0.236924	0.765269
			0.5219221	0.49528	0.227029	0.758286
			0.847068	0.4899	0.208386	0.777063
		150	0.2177278	0.495476	0.247742	0.743235
			0.5299841	0.507824	0.252215	0.756141
			0.8487206	0.509393	0.237081	0.791051
		200	0.2239275	0.502327	0.272957	0.72627
			0.5359944	0.491696	0.256523	0.731223
			0.849504	0.508489	0.242717	0.770344
		250	0.2287889	0.510853	0.291907	0.733491
			0.5372966	0.502143	0.266988	0.743305
			0.8491737	0.496598	0.264818	0.731034

จากตารางที่ 4.32 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 100 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.33 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1 จำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	3	50	0.2350503	0.501172	0.250077	0.753873
			0.5385911	0.501844	0.242707	0.74924
			0.8620921	0.495291	0.268113	0.719601
		100	0.2529071	0.502182	0.240616	0.770389
			0.5554746	0.495576	0.228455	0.779218
			0.863123	0.497898	0.250699	0.766899
		150	0.2549317	0.502931	0.242619	0.768415
			0.556986	0.497141	0.221334	0.773431
			0.8607247	0.497446	0.230128	0.775159
		200	0.2584532	0.499529	0.252766	0.749504
			0.5587158	0.486614	0.228105	0.75554
			0.8603123	0.499144	0.220767	0.760301
		250	0.2609532	0.507128	0.27192	0.726269
			0.5607767	0.490277	0.266541	0.748254
			0.8600407	0.497591	0.23843	0.747535

จากตารางที่ 4.33 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.34 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1จำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	4	50	0.254126	0.499337	0.260935	0.746921
			0.5565136	0.498926	0.278862	0.707956
			0.8699663	0.501512	0.402216	0.647958
		100	0.2655062	0.490926	0.216475	0.772606
			0.5640089	0.503611	0.235779	0.757799
			0.8722286	0.503028	0.252933	0.76234
		150	0.2674501	0.499868	0.216776	0.771611
			0.566902	0.509403	0.240063	0.782342
			0.8680124	0.503083	0.241903	0.778957
		200	0.2726991	0.505795	0.257207	0.746059
			0.5693879	0.508991	0.236898	0.778514
			0.8675026	0.505198	0.225411	0.783045
		250	0.2731542	0.490329	0.247489	0.745463
			0.5701132	0.49896	0.22798	0.757359
			0.8667469	0.511508	0.227845	0.789503

จากตารางที่ 4.34 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.35 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1จำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.1	5	50	0.2695499	0.496327	0.285974	0.715822
			0.5645582	0.499102	0.377179	0.514213
			0.8747025	0.502021	0.500000	0.575944*
		100	0.2734939	0.498461	0.244965	0.758303
			0.570552	0.50368	0.254696	0.73254
			0.8758678	0.495558	0.267639	0.702735
		150	0.2787541	0.511795	0.240792	0.764196
			0.5725225	0.501993	0.238995	0.773778
			0.8730344	0.505336	0.280116	0.724207
		200	0.2781472	0.508619	0.223032	0.766436
			0.5739922	0.501506	0.233406	0.774567
			0.871125	0.498684	0.24358	0.758786
		250	0.2807881	0.490082	0.227694	0.761813
			0.5768426	0.497107	0.227003	0.78334
			0.869478	0.50073	0.230196	0.769607

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.35 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.36 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 จำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	2	50	0.1864547	0.500494	0.237315	0.767111
			0.5007751	0.501119	0.228249	0.771341
			0.8419899	0.502854	0.22429	0.75836
		100	0.2047484	0.50044	0.237995	0.770704
			0.5164418	0.504465	0.226986	0.775911
			0.8471553	0.499313	0.228393	0.769043
		150	0.2190896	0.502754	0.270441	0.729218
			0.5309888	0.501573	0.239628	0.766404
			0.8487382	0.493233	0.224216	0.759169
		200	0.2222356	0.495216	0.258498	0.717659
			0.5339025	0.493578	0.238064	0.735588
			0.8489849	0.499217	0.248745	0.74121
		250	0.2338323	0.50075	0.271328	0.71322
			0.5391257	0.497347	0.250164	0.720739
			0.8497455	0.500984	0.269106	0.73085

จากตารางที่ 4.36 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100,150, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.37 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 จำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	3	50	0.2348507	0.500676	0.2304	0.752722
			0.535131	0.500478	0.257131	0.735763
			0.8605197	0.508168	0.270338	0.738057
		100	0.251319	0.48944	0.231358	0.769243
			0.5511936	0.505789	0.228973	0.771086
			0.8632002	0.494965	0.232188	0.774918
		150	0.2568403	0.496919	0.233292	0.745431
			0.5559691	0.499629	0.232451	0.774833
			0.8622969	0.505509	0.239588	0.766192
		200	0.2585383	0.498575	0.258514	0.754046
			0.5603839	0.499552	0.236095	0.772088
			0.8630131	0.501496	0.230034	0.757002
		250	0.2616638	0.508402	0.265561	0.740355
			0.5596843	0.501579	0.252983	0.735728
			0.8614076	0.501652	0.239267	0.764877

จากตารางที่ 4.37 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100,150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.38 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 จำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	4	50	0.2601783	0.507427	0.273588	0.740752
			0.557933	0.50057	0.287409	0.731583
			0.8696018	0.500205	0.369926	0.638865
		100	0.2672831	0.489803	0.224208	0.762047
			0.5632398	0.494576	0.238845	0.76584
			0.8714558	0.505334	0.267485	0.73557
		150	0.2680613	0.495113	0.248984	0.777624
			0.5662719	0.487577	0.231274	0.764482
			0.8695362	0.490099	0.219119	0.778358
		200	0.271567	0.487352	0.234282	0.770728
			0.5686589	0.493009	0.220505	0.755097
			0.8671073	0.498765	0.219758	0.765132
		250	0.2744325	0.49877	0.230719	0.754654
			0.5716462	0.499859	0.237691	0.77363
			0.8661106	0.502993	0.242588	0.760802

จากตารางที่ 4.38 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 150, 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.39 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.5 จำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.5	5	50	0.2685089	0.497784	0.289586	0.662065
			0.5668828	0.501685	0.493776*	0.510152**
			0.8758058	0.499978	0.377643*	0.620358**
		100	0.2754783	0.499431	0.250611	0.758003
			0.5708508	0.505952	0.274687	0.759896
			0.8749598	0.497911	0.266185	0.680201
		150	0.2775778	0.502148	0.222499	0.788029
			0.5745328	0.500495	0.249938	0.7688
			0.8732464	0.506921	0.266837	0.751978
		200	0.2798083	0.500622	0.226709	0.772803
			0.5754855	0.497713	0.236977	0.758642
			0.8718601	0.493159	0.257281	0.742426
		250	0.2792333	0.501861	0.254699	0.751812
			0.5750038	0.5033	0.227734	0.774936
			0.8693057	0.511941	0.23163	0.78244

* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.39 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ เพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5

ตารางที่ 4.40 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 จำนวนตัวแปรอิสระ (p) เท่ากับ 2 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\lvert r_{ij} \rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	2	50	0.1821284	0.507494	0.217843	0.775718
			0.4953218	0.507027	0.237697	0.767406
			0.8425875	0.508642	0.254318	0.764763
		100	0.2042958	0.489978	0.231292	0.755361
			0.5206238	0.499105	0.230647	0.74143
			0.8438331	0.488845	0.214951	0.76145
		150	0.2160963	0.492633	0.249681	0.749831
			0.5275973	0.486646	0.238289	0.754875
			0.8492548	0.497619	0.22453	0.764602
		200	0.2226641	0.49615	0.2638	0.730161
			0.5393825	0.503666	0.265697	0.735765
			0.8494049	0.50235	0.237538	0.762992
		250	0.2318075	0.499183	0.289666	0.704873
			0.5416306	0.498816	0.250122	0.726159
			0.8490023	0.500899	0.262279	0.751556

จากตารางที่ 4.40 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 100, 150 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.41 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 จำนวนตัวแปรอิสระ (p) เท่ากับ 3 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	3	50	0.2333454	0.498069	0.232993	0.754885
			0.5354911	0.494613	0.234522	0.73211
			0.8605327	0.495267	0.269261	0.698625
		100	0.2481321	0.502313	0.227633	0.768928
			0.5506931	0.494173	0.222366	0.766775
			0.8638203	0.496422	0.227551	0.762967
		150	0.253179	0.4983	0.233279	0.772768
			0.5588926	0.501856	0.216915	0.787535
			0.8616446	0.497891	0.236194	0.759695
		200	0.2561202	0.509764	0.233451	0.756943
			0.5577018	0.502397	0.245572	0.751992
			0.8613328	0.492758	0.221162	0.764929
		250	0.2617862	0.503039	0.257576	0.724219
			0.562273	0.499273	0.247986	0.747294
			0.8603164	0.493694	0.227981	0.775343

จากตารางที่ 4.41 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 100, 150, 200, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 50 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.42 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 จำนวนตัวแปรอิสระ (p) เท่ากับ 4 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

a	p	n	ค่าเฉลี่ยของ $Max\{\|r_{ij}\|\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	4	50	0.2570051	0.497186	0.260612	0.757711
			0.5532036	0.498156	0.30295	0.679195
			0.8700897	0.50136	0.379242	0.641594
		100	0.264764	0.510789	0.233674	0.774722
			0.5657834	0.505651	0.234338	0.75323
			0.8706262	0.503881	0.253366	0.751583
		150	0.2711085	0.507487	0.225612	0.770483
			0.5673897	0.494009	0.21649	0.75806
			0.8698966	0.4973	0.231858	0.777655
		200	0.2731122	0.500595	0.240592	0.770429
			0.5717861	0.496242	0.215561	0.783827
			0.8672836	0.491078	0.227777	0.763907
		250	0.2754604	0.499596	0.256919	0.761046
			0.5706408	0.490808	0.222459	0.758633
			0.8656913	0.499081	0.219553	0.780762

จากตารางที่ 4.42 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ เพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 250 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.43 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.9 จำนวนตัวแปรอิสระ (p) เท่ากับ 5 ขนาดตัวอย่าง (n) เท่ากับ 50, 100, 150, 200 และ 250 โดยจำแนกตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ)

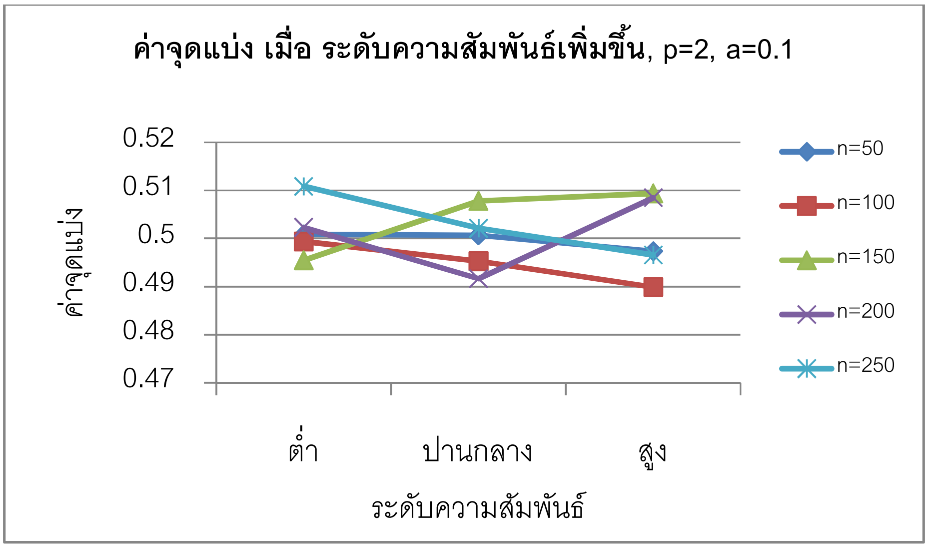
a	p	n	**ค่าเฉลี่ยของ** $Max\{\lvert r_{ij}\rvert\}$	$\bar{\hat{c}}$	CI. Lower of $\bar{\hat{c}}$	CI. Upper of $\bar{\hat{c}}$
0.9	5	50	0.2689239	0.500073	0.292309	0.679598
			0.5625658	0.500792	0.469167	0.60962
			0.8745183	0.501531	0.495975*	0.640162**
		100	0.2750965	0.493672	0.238426	0.761804
			0.5710766	0.500203	0.244628	0.736526
			0.87495	0.498957	0.298155	0.714686
		150	0.2760721	0.493163	0.237269	0.772208
			0.5727453	0.493036	0.236913	0.763234
			0.8730918	0.507266	0.256175	0.753777
		200	0.2793923	0.495289	0.238201	0.769221
			0.5753501	0.492654	0.236066	0.771419
			0.8710947	0.494261	0.231991	0.755092
		250	0.2803318	0.483096	0.241996	0.740038
			0.5757055	0.489227	0.230261	0.768856
			0.8700021	0.50278	0.256239	0.767216

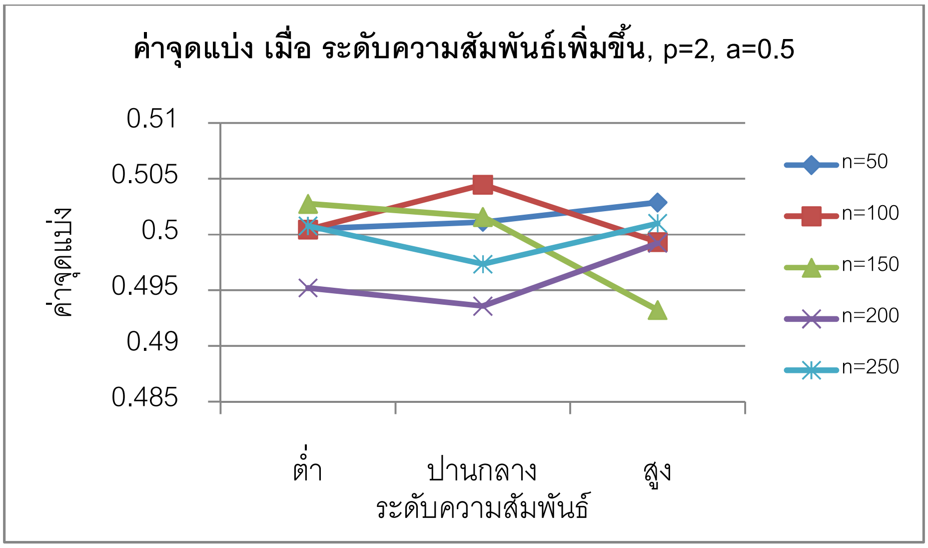
* ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 2.5

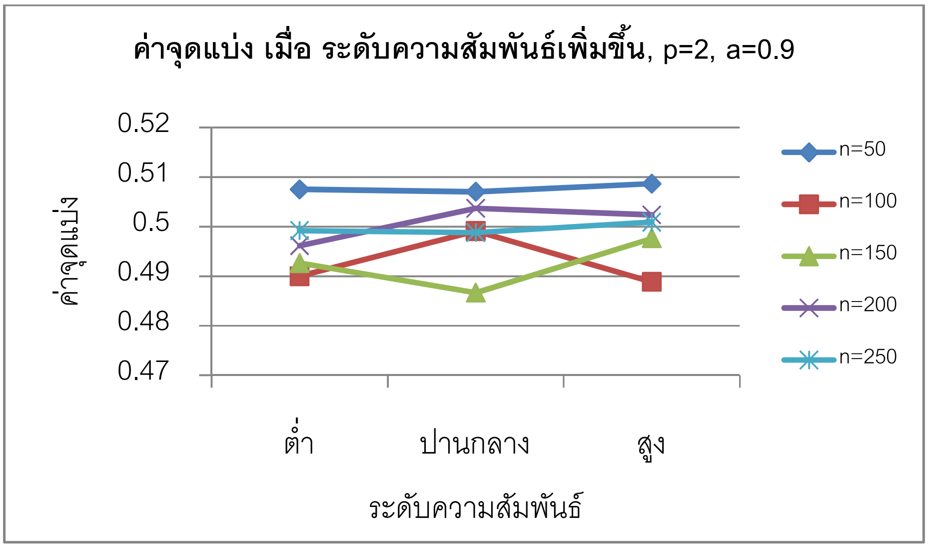
** ใช้หลักเกณฑ์ในการหาเปอร์เซ็นต์ไทล์ที่ใกล้เปอร์เซ็นต์ไทล์ที่ 97.5

จากตารางที่ 4.43 เมื่อพิจารณาค่าของจุดแบ่ง กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.9 จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200 และ 250 แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 250 ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 ส่วนที่ขนาดตัวอย่าง เท่ากับ 200 ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

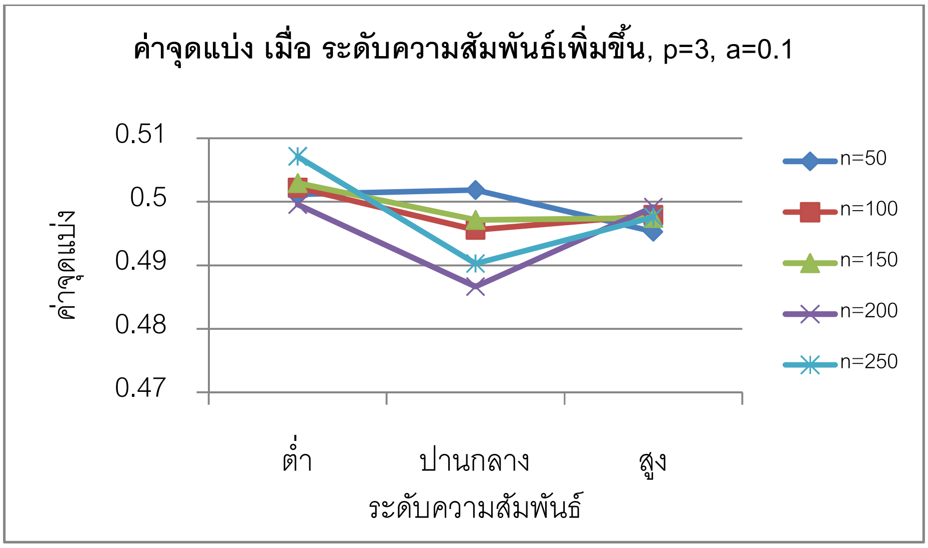
ภาพที่ 4.4 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่างและสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจคงที่

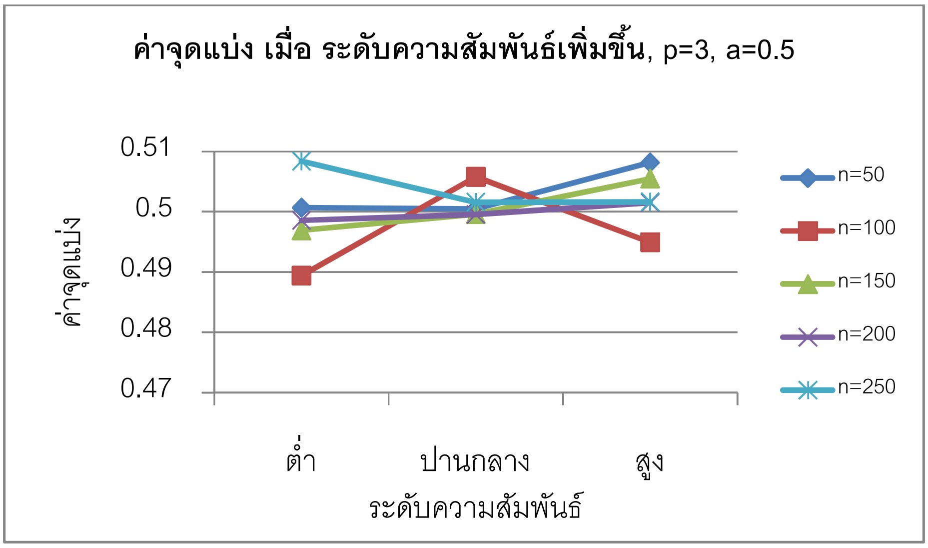


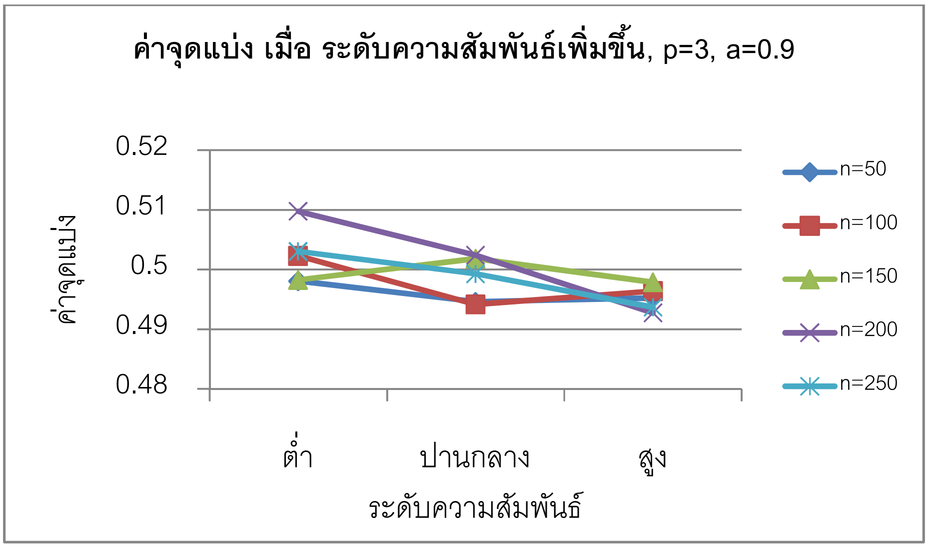




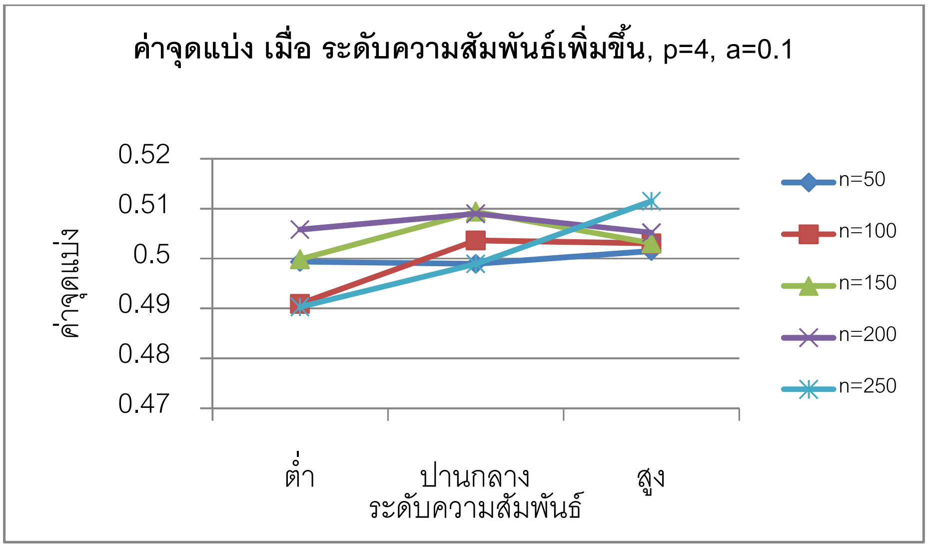
ภาพที่ 4.4 (ต่อ)


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=3, a=0.1
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

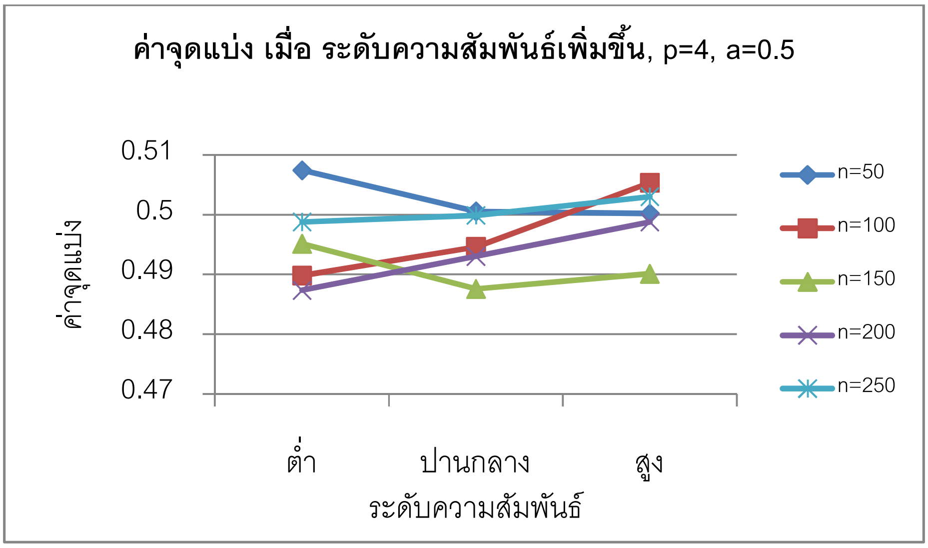

ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=3, a=0.5
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=3, a=0.9
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

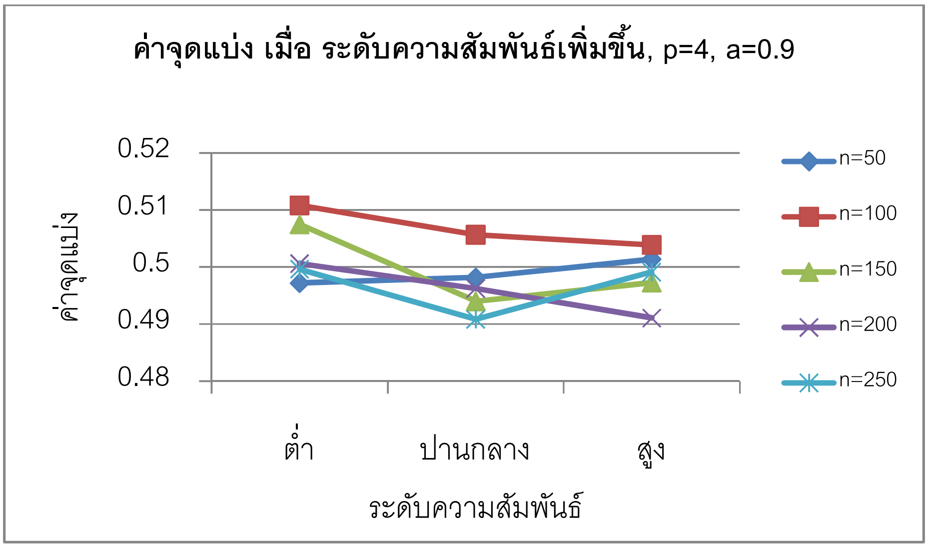
ภาพที่ 4.4 (ต่อ)



ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=4, a=0.1
0.52
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

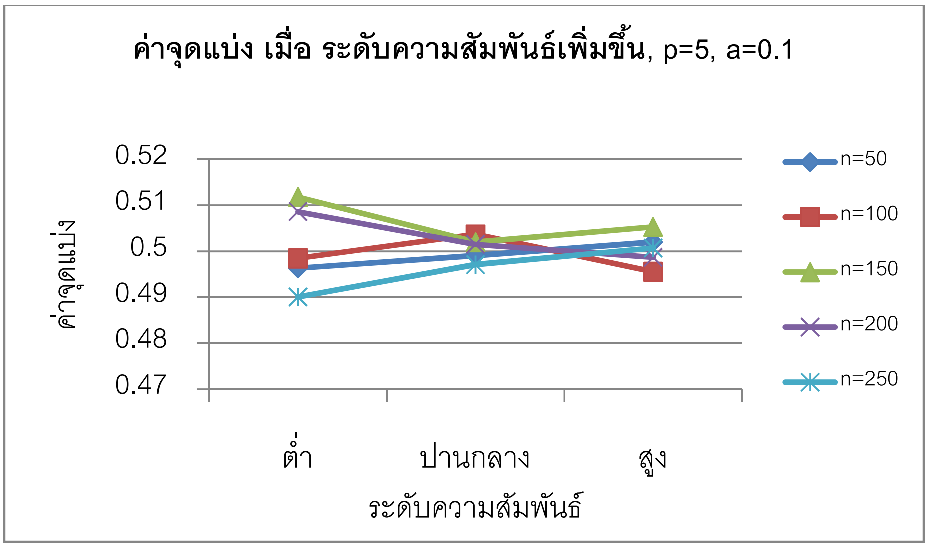


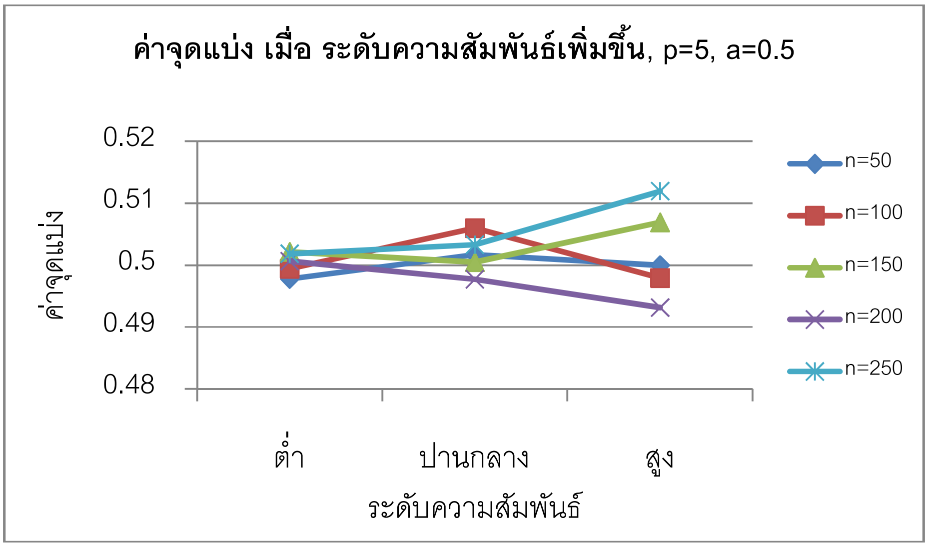
ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=4, a=0.5
0.51
0.5
0.49
0.48
0.47
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

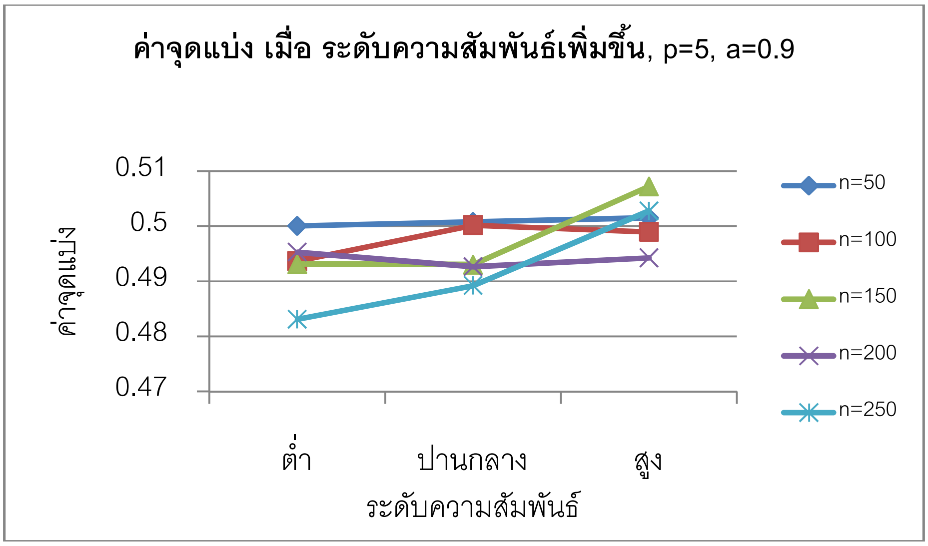


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=4, a=0.9
0.52
0.51
0.5
0.49
0.48
ค่าจุดแบ่ง
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

ภาพที่ 4.4 (ต่อ)


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=5, a=0.1
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
0.47
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=5, a=0.5
ค่าจุดแบ่ง
0.52
0.51
0.5
0.49
0.48
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่ง เมื่อ ระดับความสัมพันธ์เพิ่มขึ้น, p=5, a=0.9
ค่าจุดแบ่ง
0.51
0.5
0.49
0.48
0.47
ต่ำ
ปานกลาง
สูง
ระดับความสัมพันธ์
n=50
n=100
n=150
n=200
n=250

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยงในแต่ละสถานการณ์ที่ต้องการศึกษา โดยปัจจัยที่สนใจศึกษา ได้แก่ จำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ การจำลองข้อมูลในแต่ละสถานการณ์จำลองขึ้นโดยใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรม R

5.1 สรุปผลการวิจัย

ผลการวิจัยเรื่องการหาจุดแบ่งสำหรับการพยากรณ์การจำแนกข้อมูลสำหรับตัวแบบถดถอยโลจิสติก 2 ประเภท โดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง ให้ผลของค่าเฉลี่ยของจุดแบ่งในแต่ละสถานการณ์ ดังนี้

1) กรณีที่จำนวนตัวแปรอิสระเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

2) กรณีที่ขนาดของตัวอย่างเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

3) กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

4) กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

1) กรณีที่จำนวนตัวแปรอิสระเพิ่มขึ้น แต่ปัจจัยอื่น ๆคงที่

ในงานวิจัยครั้งนี้ กำหนดให้จำนวนของตัวแปรอิสระเพิ่มขึ้น เป็น 1,2,3,4 และ 5 ตามลำดับ เมื่อขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ มีค่าคงที่ ผลการวิจัยพบว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ปานกลาง และสูง ขนาดตัวอย่าง เท่ากับ 50, 100, 150, 200, 250 พบว่าทุกสถานการณ์ให้ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง ($0.30 < Max\{|r_{ij}|\} \leq 0.60$) ขนาดตัวอย่าง เท่ากับ 200 พบว่าค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) ขนาดตัวอย่าง เท่ากับ 250 พบว่า ค่าเฉลี่ยของจุดแบ่งเพิ่มขึ้นจากค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง ($0.30 < Max\{|r_{ij}|\} \leq 0.60$) ขนาดตัวอย่าง เท่ากับ 250 พบว่าค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

2) กรณีที่ขนาดของตัวอย่างเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

ในงานวิจัยครั้งนี้ กำหนดให้ขนาดตัวอย่างเพิ่มขึ้น เป็น 50,100,150, 200 และ 250 ตามลำดับ เมื่อจำนวนตัวแปรอิสระ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ มีค่าคงที่ ผลการวิจัยพบว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ($0 < Max\{|r_{ij}|\} \leq 0.30$) จำนวนตัวแปรอิสระ เท่ากับ 1 พบว่า ค่าเฉลี่ยของจุดแบ่ง ต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) จำนวนตัวแปรอิสระ เท่ากับ 3 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย และที่จำนวนตัวแปรอิสระ เท่ากับ 4 พบว่า ค่าเฉลี่ยของจุดแบ่งเพิ่มขึ้นจากค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อขนาดตัวอย่างเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ปานกลาง และสูง ตัวแปรอิสระ เท่ากับ 1, 2, 3, 4, 5 พบว่า ทุกสถานการณ์ให้ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อขนาดตัวอย่างเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) จำนวนตัวแปรอิสระ เท่ากับ 3 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อขนาดตัวอย่างเพิ่มขึ้น

3) กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

ในงานวิจัยครั้งนี้ กำหนดให้สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น เป็น 0.1, 0.5 และ 0.9 ตามลำดับ เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระมีค่าคงที่ ผลการวิจัยพบว่า

เมื่อจำนวนตัวแปรอิสระ เท่ากับ 1 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ($0 < Max\{|r_{ij}|\} \leq 0.30$) ขนาดตัวอย่าง เท่ากับ 250 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระ เท่ากับ 2 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) ขนาดตัวอย่าง เท่ากับ 100 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระ เท่ากับ 3 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) ขนาดตัวอย่าง เท่ากับ 100 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระ เท่ากับ 4 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ($0 < Max\{|r_{ij}|\} \leq 0.30$) ขนาดตัวอย่าง เท่ากับ 250 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง ($0.30 < Max\{|r_{ij}|\} \leq 0.60$) ขนาดตัวอย่าง เท่ากับ 250 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระ เท่ากับ 5 ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ ($0 < Max\{|r_{ij}|\} \leq 0.30$) ขนาดตัวอย่าง เท่ากับ 100 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ($0.60 < Max\{|r_{ij}|\} \leq 0.90$) ขนาดตัวอย่าง เท่ากับ 100, 200 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

4) กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่ปัจจัยอื่นๆคงที่

ในงานวิจัยครั้งนี้ กำหนดให้ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น เป็นระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับปานกลาง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง ตามลำดับ เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง และสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ มีค่าคงที่ ผลการวิจัยพบว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 ที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 100 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 200 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 200 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ส่วนที่จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 150, 200 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่จำนวนตัวแปรอิสระ เท่ากับ 2 ขนาดตัวอย่าง เท่ากับ 100, 150 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ที่จำนวนตัวแปรอิสระ เท่ากับ 3 ขนาดตัวอย่าง เท่ากับ 50 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย ที่จำนวนตัวแปรอิสระ เท่ากับ 4 ขนาดตัวอย่าง เท่ากับ 250 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย และที่จำนวนตัวแปรอิสระ เท่ากับ 5 ขนาดตัวอย่าง เท่ากับ 200 พบว่า ค่าเฉลี่ยของจุดแบ่งต่ำกว่าค่า 0.5 เพียงเล็กน้อย สำหรับสถานการณ์อื่นๆ พบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

5.2 อภิปรายผล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง สำหรับแต่ละสถานการณ์ที่ต้องการศึกษา จะทำการเปรียบเทียบสถานการณ์ที่จำนวนตัวแปรอิสระเป็น 1,2,3,4 และ 5 ขนาดตัวอย่างมีค่าเป็น 50,100,150,200 และ 250 สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเป็น 0.1, 0.5 และ 0.9 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น 3 ระดับ คือ ความสัมพันธ์กันในระดับต่ำ ปานกลาง และสูง

ผลการวิจัยสามารถสรุปได้ว่า จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระมีผลต่อค่าของจุดแบ่งเพียงเล็กน้อย เนื่องจากค่าจุดแบ่งในเกือบทุกสถานการณ์มีค่าขึ้นๆ ลงๆ และลู่เข้าสู่ค่า 0.5 แสดงว่าค่าจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง คือ ค่า 0.5

ผลการวิจัยครั้งนี้สอดคล้องกับงานวิจัยของอรุณรัตน์ โพธิ์คำ ที่หาค่าจุดแบ่งโดยใช้ฟังก์ชันโลจิทเป็นฟังก์ชันเชื่อมโยง และนิภาพรรณ ไพจินดา ที่หาค่าจุดแบ่งโดยใช้ฟังก์ชันคอมพลีเมนทารีล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง ซึ่งได้ค่าจุดแบ่งที่เหมาะสมคือ ค่า 0.5 ดังนั้น แม้ว่าจะใช้ฟังก์ชันเชื่อมโยงที่ต่างกันแต่จุดแบ่งที่เหมาะสมที่สุดคือ ค่า 0.5

5.3 ข้อเสนอแนะ

1. ในงานวิจัยครั้งนี้ได้ศึกษาปัจจัยที่มีผลต่อค่าจุดแบ่งเพียง 4 ปัจจัย คือ จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระในการวิจัยครั้งต่อไปอาจหาปัจจัยอื่นๆ เพิ่มเติม

2. ในการวิจัยครั้งนี้ได้กำหนดให้แต่ละปัจจัยคงที่ เมื่อมีปัจจัยหนึ่งเปลี่ยนแปลง ในงานวิจัยครั้งต่อไปอาจใช้กรณีที่ปัจจัยเปลี่ยนไปพร้อมกันหลายๆ ปัจจัย

3. เมื่อต้องการหาค่าจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันโพรบิตเป็นฟังก์ชันเชื่อมโยง สามารถนำค่าจุดแบ่งนี้ไปใช้ได้ตามแต่ละสถานการณ์ที่ต้องการศึกษา ภายใต้สถานการณ์ที่เหมือนกัน

รายการอ้างอิง

ภาษาไทย

กัลยา วานิชย์บัญชา . การวิเคราะห์ข้อมูลหลายตัวแปร . พิมพ์ครั้งที่ 3. กรุงเทพฯ : สำนักพิมพ์ ธรรมสาร, 2551.

กัลยา วานิชย์บัญชา . การวิเคราะห์สถิติขั้นสูงด้วย SPSS for Windows . พิมพ์ครั้งที่ 7. กรุงเทพฯ : สำนักพิมพ์ ธรรมสาร, 2552.

อรุณรัตน์ โพธิ์คำ. การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันโลจิทเป็นฟังก์ชันเชื่อมโยง. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2554.

นิภาพรรณ ไพจินดา. การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2554.

ภาษาอังกฤษ

Hosmer, D.W. & Lemeshow, S. Applied Logistic Regression. New York: Wiley, 2000.Walter

Hadjicostas, P. Maximizing proportions of correct classifications in binary logistic regression. Journal of Applied Statistics 33 (2006) : 629-640.

Hadijicostas, P. & Hadjinicola, G.C. The asymptotic distribution of the proportion of Correct classification for a holdout sample in logistic regression. Journal of Statistical Planning and Inference 92(2001) : 193-211

.

ภาคผนวก

ตัวอย่างการใช้โปรแกรม R ในการวิจัย

```
library(MASS)

x.fun<-function(n,p=1,r=0,u=10,d=0.05){
	if(p==1|r==0)x<-matrix(runif(n*p,-u,u),n);
	if(p>1&r>0){
		r0<-ifelse(r==0.3,0,ifelse(r==0.6,0.3,ifelse(r==0.9,0.6,NA)));
		if(is.na(r0))stop("กำหนดค่า r ไม่ถูกต้อง")
		norm.r<-2*sin(pi/6*(r-d));
		norm.cor<-matrix(norm.r,p,p);
		diag(norm.cor)<-1;
		R<-Inf;
		while(max(R)>r|max(R)<=r0){
			x<-mvrnorm(n,rep(0,p),norm.cor);
			x<-pnorm(x);
			rho.mat<-cor(x);
			R<-rho.mat[row(rho.mat)>col(rho.mat)];
		}
		x<-x*u*2-u;
	}
	return(list(x,max(R)))
}

ai.fun<-function(midata){
	MI<-c(0,midata$MI);
	Yn<-c(midata$Y,NA);
	n<-nrow(midata);
	n1<-n+1;
	Co<-rep(-1,n1);
```

```
        Case2nd<-c(MI[-n1]<1:n,NA);
        Pos1st<-rep(NA,n1)->Pos2nd;
        Pos1st[which(!is.na(Case2nd)&Case2nd==T)]<-MI[which(!is.na(Case2nd)&
                Case2nd==T)]+1;
        Pos2nd[which(!is.na(Case2nd)&Case2nd==T)+1]<-MI[which(!is.na(Case2nd)&
                Case2nd==T)+1];
        AI<-c(0,rep(NA,n));
        for(j in 2:n1){
                j1<-j-1;
                if(Case2nd[j1]==T){
                        AI[j]<-
AI[j1]+sum(Co[Pos1st[j1]:Pos2nd[j]]^Yn[Pos1st[j1]:Pos2nd[j]]);
                }
                else{
                        AI[j]<-AI[j1];
                }
        }
        return(AI[-1])
}

data.fun<-function(xsim,a.){
        if(!is.numeric(xsim))stop("ข้อมูลตัวแปรอิสระไม่ถูกต้อง")
        n<-nrow(xsim);
        errorTerm<-rnorm(n,0,sqrt(500));
        u<-runif(n);
        beta0<-qnorm(a.);
        Ystar<-beta0+beta*rowSums(xsim)+errorTerm;
        Prob<-pnorm(Ystar);
        Y<-ifelse(u>=Prob,0,1);
        data1<-data.frame(xsim,Y);
```

```
	fitmodel<-glm(Y~.,data=data1,family=binomial(link="probit"));
	pihat<-fitmodel$fitted;
	MI<-rank(pihat,ties.method="max");
	data1<-cbind(Ystar,data1,pihat,MI);
	data1<-data1[order(MI),];
	AI<-ai.fun(data1);
	result<-cbind(subset(data1,select=c("Ystar","Y","pihat","MI")),AI);
	return(result)
}

midcut.fun<-function(aidata){
	ai<-aidata$pihat;
	I0<-which(aidata$AI==max(aidata$AI));
	yi<-ai[I0];
	yi1<-ai[I0+1];
	yi1[I0==length(ai)]<-1;
	d<-yi1-yi;
	p<-prop.table(d);
	P<-cumsum(p);
	i<-min(which(P>0.5));
	midcut<-yi[i]+(0.5-ifelse(i>1,P[i-1],0))/p[i]*d[i];
	return(midcut)
}

a<-c(0.1,0.5,0.9);
beta<-10;

n<-50;
p<-2;
r<-0.3;
```

```
set.seed(555)
cutoff<-list();
length(cutoff)<-na<-length(a);
maxcor<-matrix(0,500,na);
names(cutoff)<-colnames(maxcor)<-a;
for(j in 1:na){
	a.<-a[j];
	for(i in 1:500){
		x<-x.fun(n,p,r);
		maxcor[i,j]<-x[[2]];
		data1<-data.fun(x[[1]],a.);
		cutoff[[j]][i]<-midcut.fun(data1);
		pie(c(i,500-i),main=a.,labels=c(i,500-i),radius=1,clockwise=T)
	}
}

result<-matrix(unlist(lapply(cutoff,function(x)c(mean(x),quantile(x,c(0.025,0.975))))),
	3,byrow=T);
dimnames(result)<-list(a,c("C.hatbar","Lower","Upper"));
result

mean(maxcor[,1])
mean(maxcor[,2])
mean(maxcor[,3])
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวสุภิญญา คำมั่น เกิดวันที่ 18 พฤศจิกายน พ.ศ. 2528 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ในปีการศึกษา 2551 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตรมหาบัณฑิต (สต.ม.) สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2552