

ตัวแบบโพธิ์แบบ 2 ประเภท สำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่ม



นางสาวชลลดา กล้วยไม้

ศูนย์วิทยพัทยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต


สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2553

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

BINARY PROBIT MODEL FOR UNGROUPED-DATA PREDICTIVE CLASSIFICATION



Miss Cholada Kluaymai

ศูนย์วิทยพัทยากร

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 2010

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

ตัวแบบโพรบิตแบบ 2 ประเภท สำหรับการพยากรณ์การ

จำแนกข้อมูลไม่จัดกลุ่ม

โดย

นางสาวชลลดา กล้วยไม้


สาขาวิชา

สถิติ


อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา

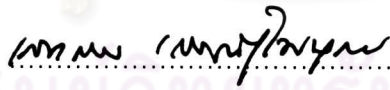
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย
ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบริหารธุรกิจ

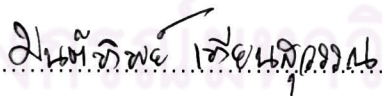
.......... คณบดีคณะพาณิชยศาสตร์และการบัญชี
(รองศาสตราจารย์ ดร. อรรถนพ ตันละมัย)

คณะกรรมการสอบวิทยานิพนธ์

.......... ประธานกรรมการ
(รองศาสตราจารย์ ดร. วีระพร วีระถาวร)

.......... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา)

.......... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. เสกสรร เกียรติสุไพบูลย์)


.......... กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ)

ชลลดา กล้วยไม้ : ตัวแบบโพรบิตแบบ 2 ประเภท สำหรับการพยากรณ์การจำแนก
 ข้อมูลไม่จัดกลุ่ม (BINARY PROBIT MODEL FOR UNGROUPED-DATA
 PREDICTIVE CLASSIFICATION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รองศาสตราจารย์
 ดร.สุพล ดุรงค์วัฒนา, 138 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาค่าของจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การ
 จำแนกข้อมูลไม่จัดกลุ่มโดยใช้ตัวแบบโพรบิตแบบ 2 ประเภท ปัจจัยที่นำมาพิจารณาคือ จำนวนตัว
 แปรอิสระ (p) แบ่งเป็น 3 ระดับ คือ น้อย(p=1, 2) ปานกลาง(p= 3, 4) และ มาก(p= 5,6) ขนาด
 ตัวอย่าง (n) แบ่งเป็น 3 ระดับ คือ เล็ก(n=20,40) ปานกลาง (n=60,80) และใหญ่(p=100 ,120)
 สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) เท่ากับ 0.1,0.5 และ 0.9 ระดับความสัมพันธ์ระหว่าง
 ของตัวแปรอิสระ (M) เท่ากับ 0, 0.33,0.67 และ 0.99 ข้อมูลทั้งหมดนี้ใช้การจำลองโดยเทคนิคมอน
 ทีคาร์โล ด้วยโปรแกรม R การหาค่าของจุดแบ่งใช้ทฤษฎีของ Hadjicostas P. (2006) จากการวิจัย
 สรุปผลได้ดังนี้

กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ จะพบว่าค่าจุดของแบ่งที่
 สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 และ 0.9 จะลู่ออกหาค่าจุดของแบ่งที่สัดส่วนของ
 การไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ซึ่งมีค่าประมาณ 0.5 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลงไป
 แต่ปัจจัยอื่น ๆ คงที่ จะพบว่าค่าจุดของแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จะ
 ลู่ออกหาค่า 0.5 แต่ที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจอื่นๆ ค่าของจุดแบ่งมีค่าต่ำกว่าค่า 0.5
 กรณีที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ พบว่าที่สัดส่วนของ
 การไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ค่าของจุดแบ่งมีค่าสูงที่สุดและลู่ออกหาค่า 0.5 กรณีที่ระดับ
 ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ พบว่าที่สัดส่วนของการไม่
 เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ค่าจุดแบ่งจะมีค่าลดลงจากค่า 0.5 แต่ที่ค่าอื่น ๆ ค่าจุดแบ่งจะมี
 ค่าลดลงจนถึงระดับความสัมพันธ์ระหว่างตัวแปรอิสระเท่ากับ 0.67 และจะเพิ่มขึ้นอีกเล็กน้อย จาก
 การประมาณค่าของจุดแบ่งจากตัวแบบโพรบิตแบบ 2 ประเภท ที่มีผลอันตรกิริยา (Interaction)
 พบว่าค่าของสัมประสิทธิ์การตัดสินใจ (R^2) มีค่าสูง แสดงว่าสมการการถดถอยมีความเหมาะสม
 สามารถนำไปใช้ประมาณค่าของจุดแบ่งที่เหมาะสมในสถานการณ์อื่นๆ ได้

ภาควิชา..... สถิติ.....
 สาขาวิชา..... สถิติ.....
 ปีการศึกษา..... 2553.....

ลายมือชื่อนิสิต ชลลดา กล้วยไม้
 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก 

5181785826 : MAJOR STATISTICS

KEYWORDS : BINARY PROBIT MODEL / CUT-OFF POINTS / CLASSIFICATION ERROR RATE / UNGROUPED-DATA

CHOLADA KLUAYMAI: BINARY PROBIT MODEL FOR UNGROUPED-DATA PREDICTIVE CLASSIFICATION. THESIS ADVISOR: ASSOC.PROF.SUPOL DURONGKAWATANA, Ph.D., 138 pp.

The objective of this study is to find out the optimal cut-off point for predictive classification of ungrouped data using binary probit model. The interesting factors are the number of independent variables (p) with 3 levels; low level (p=1, 2), medium level (p= 3, 4) and high level (p=5,6), the sample size (n) with 3 levels ; low level (n=20,40), medium level (n=60,80) and high level (n=100,120), the failure rate (a) of the values 0.1,0.5 and 0.9 and the degree of multicollinearity among independent variables (M) of the values 0, 0.33, 0.67 and 0.99. 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 changed but the other factors are kept constant, with the failure rate equal to 0.1 and 0.9, the optimal cut-off point converges approximately to 0.5 as the failure rate equals to 0.5. As sample size changed but the other factors are kept constant, with the failure rate equals to 0.5, the optimal cut-off point converges approximately to 0.5. For the other situations, the optimal cut-off point are lower than 0.5. As the failure rate in data set changed but the other factors are kept constant, with the failure rate equals to 0.5, the optimal cut-off point is the highest and converges approximately to 0.5. As the degree of multicollinearity among independent variables changed but the other factors are kept constant, with the failure rate equals to 0.5, the optimal cut-off point decreases from 0.5. For the other situations, the optimal cut-off point decreases to M =0.67 then slightly increases. Finally the estimated binary probit model with all interaction terms is needed to find the estimated cut-off point for all situations. The R² is needed to measure the goodness-of-fit of the estimated model. From the estimated regression equation, the optimal cut-off point for any situation can be predicted.

Department : Statistics
Field of Study : Statistics
Academic Year : 2010

Student's Signature *ชอลดา คล้ายไม*
Advisor's Signature *สุพล ดุรงคawatana*

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลือเป็นอย่างดีจาก รองศาสตราจารย์ ดร. สุปล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำปรึกษา ตลอดจนช่วยเหลือตรวจสอบ แก้ไขข้อบกพร่องต่างๆ เป็นอย่างดีจนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความรู้สึกซาบซึ้ง เคารพและสำนึกในพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ธีระพร วีระถาวร ประธานกรรมการสอบวิทยานิพนธ์ และผู้ช่วยศาสตราจารย์ ดร. เสกสรร เกียรติสุไพบูรณ์ กรรมการสอบวิทยานิพนธ์ ที่ให้คำแนะนำในการทำวิจัยครั้งนี้ และกรุณาตรวจสอบแก้ไขวิทยานิพนธ์ฉบับนี้ให้มีความสมบูรณ์ยิ่งขึ้น

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ ที่ท่านได้เสียสละเวลาอันมีค่ามาเป็นกรรมการภายนอกมหาวิทยาลัย ซึ่งทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากยิ่งขึ้น

สุดท้ายนี้ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา มารดา ที่ให้การส่งเสริม สนับสนุนด้านทุนการศึกษา และให้ความรัก ให้กำลังใจแก่ผู้วิจัยมาโดยตลอดจนสำเร็จการศึกษา ตลอดจนเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่ให้คำปรึกษาและคอยช่วยเหลือด้วยดีมาตลอด

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญภาพ.....	ณ
บทที่	1
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 ขอบเขตการวิจัย.....	3
1.4 ข้อตกลงเบื้องต้น.....	4
1.5 คำจำกัดความที่ใช้ในการวิจัย.....	5
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	6
1.7 วิธีการดำเนินการวิจัย.....	6
2 แนวคิด ทฤษฎีและสถิติที่เกี่ยวข้อง.....	8
2.1 ตัวแบบโพรบิตแบบ 2 ประเภท.....	9
2.2 ฟังก์ชันความน่าจะเป็นของข้อมูลการถดถอยโพรบิต.....	9
2.3 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด.....	10
2.4 สถิติ Kaiser-Meyer-Olkin (KMO).....	11
2.5 การตรวจสอบความเหมาะสมของตัวแบบ.....	11
2.6 ช่วงความเชื่อมั่น.....	13
2.7 เปอร์เซ็นไทล์.....	13
2.8 การหาจุดแบ่งโดยทฤษฎีHadjicostas P.....	13
3 วิธีดำเนินการวิจัย.....	15
3.1 เทคนิคมอนติคาร์โล.....	15
3.2 แผนการดำเนินการวิจัย.....	16

บทที่	ช หน้า
3.3	ขั้นตอนในการดำเนินการวิจัย..... 17
3.4	การจำลองข้อมูลที่ใช้ในการวิจัย..... 17
3.5	การคำนวณค่าของจุดแบ่งโดยทฤษฎีHadjicostas P..... 19
3.6	คำนวณค่าเฉลี่ย ค่าร้อยละ และช่วงความเชื่อมั่นของจุดแบ่ง..... 20
3.7	การวิเคราะห์โดยใช้ตัวแบบการถดถอยเชิงพหุ..... 20
3.8	สรุปผลการวิจัยในแต่ละสถานการณ์..... 21
3.9	ขั้นตอนการทำงานของโปรแกรม..... 22
4	ผลการวิเคราะห์ข้อมูล..... 24
4.1	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม..... 25
4.2	แสดงสมการถดถอยเชิงพหุ..... 100
5	สรุปผลการวิจัย และข้อเสนอแนะ..... 102
5.1	สรุปของค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบพหุคูณแบบ 2 ประเภท ภายใต้สถานการณ์ต่าง ๆ 102
5.2	สรุปผลการวิเคราะห์การถดถอยเชิงพหุ..... 110
5.3	ข้อเสนอแนะ..... 111
	รายการอ้างอิง..... 112
	ภาคผนวก..... 113
	ประวัติผู้เขียนวิทยานิพนธ์..... 138

ตารางที่	หน้า
4.19 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ.....	49
4.20 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ.....	50
4.21 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 ตามลำดับ.....	51
4.22 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ.....	52
4.23 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ.....	53
4.24 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ.....	54
4.25 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=1$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	59
4.26 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=2$, $M=0$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	60
4.27 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=2$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	61

ตารางที่		หน้า
4.37	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=4$, $M=0.99$, $n= 80, 100, 120$ และ a เปลี่ยนแปลงไป เป็น 0.1, 0.5, 0.9 ตามลำดับ.....	71
4.38	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=5$, $M=0$, $n= 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	72
4.39	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=5$, $M=0.33$, $n= 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	72
4.40	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=5$, $M=0.67$, $n= 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	73
4.41	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=5$, $M=0.99$, $n= 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	73
4.42	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $p=6$, $M=0, 0.33, 0.67, 0.99$, $n= 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ.....	74
4.43	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.1$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	81
4.44	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.1$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไป เป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	82
4.45	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.1$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	83

ตารางที่		หน้า
4.46	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.1, p=5, n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	84
4.47	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.1, p=6, n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	85
4.48	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.5, p=2, n= 40, 60, 80,100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	86
4.49	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.5, p=3, n= 60, 80,100, 120$ และ M เปลี่ยนแปลงไป เป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	87
4.50	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.5, p=4, n= 80,100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	88
4.51	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.5, p=5, n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	89
4.52	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.5, p=6, n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	90
4.53	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.9, p=2, n= 40, 60, 80,100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	91
4.54	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง ที่เหมาะสม เมื่อ $a=0.9, p=3, n=60, 80,100, 120$ และ M เปลี่ยนแปลงไป เป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	92

ตารางที่		หน้า
4.55	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=4$, $n=80,100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	93
4.56	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	94
4.57	แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=6$ $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ.....	95
4.58	แสดงการประมาณค่าพารามิเตอร์จากตัวแบบการถดถอยเชิงพหุ.....	100
4.59	แสดงค่าสัมประสิทธิ์การตัดสินใจเชิงซ้อน.....	101
5.1	แสดงผลสรุปค่าของจุดแบ่งเมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป (p) แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M), สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) และขนาดตัวอย่าง (n) คงที่.....	103
5.2	แสดงผลสรุปค่าของจุดแบ่งเมื่อขนาดตัวอย่างเปลี่ยนแปลงไป (n) แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M), สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a) และจำนวนตัวแปรอิสระ (p) คงที่.....	105
5.3	แสดงผลสรุปค่าของจุดแบ่งเมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป (a) แต่ขนาดตัวอย่าง (n), จำนวนตัวแปรอิสระ (p) และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M) คงที่.....	107
5.4	แสดงผลสรุปค่าของจุดแบ่งเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป (M) เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ (a), ขนาดตัวอย่าง (n) และจำนวนตัวแปรอิสระ (p) คงที่.....	109

สารบัญภาพ

ภาพที่		หน้า
3.1	แสดงขั้นตอนการทำงานของโปรแกรม.....	22
4.1	แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0$, $a = 0.1, 0.5, 0.9$ โดยแยกตามขนาด $n = 40, 60, 80, 100, 120$	39
4.2	แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0.33$, $a = 0.1, 0.5, 0.9$ โดยแยกตามขนาด $n = 60, 80, 100, 120$	40
4.3	แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0.67$, $a = 0.1, 0.5, 0.9$ โดยแยกตามขนาด $n = 60, 80, 100, 120$	41
4.4	แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0.99$, $a = 0.1, 0.5, 0.9$ โดยแยกตามขนาด $n = 60, 80, 100, 120$	42
4.5	แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 1, 2, 3, 4, 5$ และ 6	55
4.6	แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.33$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4, 5$ และ 6	56
4.7	แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.67$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4, 5$ และ 6	57
4.8	แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.99$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4$ และ 5	58
4.9	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 1$, $n = 20, 40, 60, 80, 100, 120$ และ $M = 0$	75
4.10	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 2$, $n = 40, 60, 80, 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99	76
4.11	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 3$, $n = 60, 80, 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99	77
4.12	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 4$, $n = 80, 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99	78

ภาพที่		หน้า
4.13	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลง ไป ที่ $p = 5, n = 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99	79
4.14	แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลง ไป ที่ $p = 6, n = 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99	80
4.15	แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง ไป ที่ $p = 2, a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 40, 60, 80, 100$ และ 120	96
4.16	แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง ไป ที่ $p = 3, a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 60, 80, 100$ และ 120	97
4.17	แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง ไป ที่ $p = 4, a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 80, 100$ และ 120	98
4.18	แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง ไป ที่ $p = 5, a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 100$ และ 120	99
4.19	แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง ไป ที่ $p = 6, a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 120$	99

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันนี้ไม่ว่าจะเป็น ด้านเศรษฐศาสตร์ ด้านธุรกิจ ด้านสังคมศาสตร์ หรือแม้แต่ทางการแพทย์ล้วนแล้วแต่จะต้องอาศัยวิธีการทางสถิติมาเป็นเครื่องมือในการวิเคราะห์ข้อมูลเพื่อประกอบการตัดสินใจทั้งสิ้น โดยการเลือกใช้วิธีการทางสถิตินั้นจะต้องพิจารณาจากชนิดของข้อมูลวัตถุประสงค์ที่ต้องการและข้อสมมติเบื้องต้น (Basic Assumption) อีกด้วย เพื่อผลการวิเคราะห์ทางสถิติที่ถูกต้องและแม่นยำ

แต่ในการวิจัยจำนวนไม่น้อยจะพบว่าตัวแปรตาม (Dependent Variable) ที่สนใจศึกษามักมีลักษณะข้อมูลเป็นเชิงคุณภาพ กล่าวคือข้อมูลไม่ต่อเนื่อง เช่น การเป็นโรคหัวใจหรือการไม่เป็นโรคหัวใจ การสอบได้หรือสอบตก และการเลือกซื้อสินค้าหรือไม่เลือกซื้อสินค้า เป็นต้น ดังนั้นตัวแบบโพรบิตแบบ 2 ประเภท (Binary Probit Model) จึงเป็นอีกตัวแบบที่นิยมกันแพร่หลายสำหรับตัวแบบที่มีข้อจำกัดเกี่ยวกับตัวแปรตามที่เป็นตัวแปรเชิงคุณภาพมีค่าได้เพียง 2 ค่า (Dichotomy or Binary Variable) เท่านั้น คือค่า 0 และ 1 โดยกำหนดให้ตัวแปรตามมีค่าเท่ากับ 1 เมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i และ ตัวแปรตามมีค่าเท่ากับ 0 เมื่อไม่เกิดเหตุการณ์ที่สนใจในหน่วยที่ i และยิ่งเนื่องมาจากทฤษฎีการถดถอยที่มีข้อจำกัดเกี่ยวกับตัวแปรตามลักษณะดังกล่าวข้างต้นนี้จะไม่สามารถนำมาทำการวิเคราะห์การถดถอยโดยใช้ตัวแบบความน่าจะเป็นเชิงเส้น (Linear Probability Model) เพื่อประมาณค่าความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจได้ ด้วยวิธีกำลังสองน้อยที่สุด (Method of Least Squares) ได้ เนื่องจากข้อสมมติเบื้องต้นของตัวแบบไม่เป็นไปตามข้อกำหนด คือ การแจกแจงของค่าความคลาดเคลื่อนไม่เป็นแบบปกติ (Nonnormality) ค่าความแปรปรวนของพจน์ความคลาดเคลื่อน (Error Terms) มีลักษณะแตกต่างกัน (Heteroscedastic) ค่าประมาณของการเกิดเหตุการณ์ที่สนใจอาจอยู่นอกช่วง 0 ถึง 1 และการประมาณค่าความชัน (Slope) ที่สูงเกินจริง (Overestimated Slope) หรือต่ำเกินจริง (Underestimated Slope) (Guijarati ,1995,Pindyck and Rubinfeld,1998,Johnston and Dinardo,1997) อีกด้วย

ตัวแบบโพรบิตแบบ 2 ประเภท เป็นตัวแบบที่นำมาใช้วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ แล้วยังสามารถนำมาใช้ในการพยากรณ์ตัวแปรตามเชิงคุณภาพว่าจะอยู่ในกลุ่มใดใน 2 กลุ่ม คือ กลุ่มที่เกิดเหตุการณ์ที่สนใจ (Success) และกลุ่มไม่เกิดเหตุการณ์ที่สนใจ (Failure) ของสิ่งที่สนใจศึกษา โดยส่วนใหญ่งานวิจัยเกือบทั้งหมดจะมีการจัดให้แต่ละหน่วยของสิ่งที่สนใจ อยู่ในกลุ่มใดกลุ่มหนึ่ง มักจะใช้จุดแบ่ง (Cut-off Point) หรือระดับความน่าจะเป็นที่

ระดับ 0.5 โดยให้เหตุผลว่ากลุ่มที่เกิดเหตุการณ์ที่สนใจมีโอกาสเกิดขึ้นเท่ากับ กลุ่มที่ไม่เกิด เหตุการณ์ที่สนใจของสิ่งที่สนใจศึกษา หรือบางงานวิจัยอาจใช้จุดแบ่งค่าหนึ่งที่จะทำให้สัดส่วน ของความถูกต้องของการจัดกลุ่มทั้งหมดมีค่าสูงสุดหรือทำให้อัตราความผิดพลาดในการจำแนก กลุ่มมีค่าต่ำ (Classification Error Rate) มีค่าต่ำสุดโดยให้เหตุผลว่า ข้อมูลถูกเลือกอย่างสุ่มจาก สิ่งที่สนใจศึกษาแล้ว ผู้วิจัยส่วนใหญ่มักไม่คำนึงถึงจำนวนของตัวแปรอิสระ, ขนาดตัวอย่าง, สัดส่วนของการเกิดเหตุการณ์ที่สนใจและไม่เกิดเหตุการณ์ที่สนใจในชุดข้อมูล และระดับ ความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) ซึ่งผู้วิจัยได้เห็นถึงความสำคัญที่จะทำการ วิจัยการคัดเลือกจุดแบ่งของตัวแบบโพรบิตแบบ 2 ประเภทที่จะทำให้อัตราส่วนความผิดพลาดใน การจำแนกกลุ่มมีค่าต่ำสุดหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มมีความถูกต้องสูงสุด สำหรับชุดข้อมูลในสถานการณ์ต่างๆ โดยลักษณะเหล่านั้นของชุดข้อมูล คือ จำนวนของตัวแปรอิสระ, ขนาดตัวอย่าง, สัดส่วนของการเกิดเหตุการณ์ที่สนใจและการไม่เกิด เหตุการณ์ที่สนใจในชุดข้อมูลและระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งลักษณะเหล่านี้จะมี ผลกระทบต่อจุดแบ่งสำหรับการประเมินการพยากรณ์การจำแนกกลุ่มของข้อมูลหรือไม่ ถ้ามีผลต่อ การคัดเลือกจุดแบ่งแล้วรูปแบบเหล่านั้นคืออะไร ที่จะทำให้สามารถหาจุดแบ่งที่เหมาะสมสำหรับ ตัวแบบโพรบิตแบบ 2 ประเภทมากที่สุด

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อหาของจุดแบ่ง ที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท เมื่อ จำนวนตัวแปรอิสระเปลี่ยนแปลงไป, ขนาดตัวอย่างเปลี่ยนแปลงไป, สัดส่วนของ การไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูลเปลี่ยนแปลงไป, ระดับความสัมพันธ์ของตัวแปร อิสระเปลี่ยนแปลงไป และนำปัจจัยต่างๆเหล่านี้มาพิจารณารวมกัน ซึ่งแปรเปลี่ยนไปพร้อมกัน

2. เพื่อหาสมการถดถอยพหุคูณ (Multiple Regression Model) สำหรับใช้ในการ ประเมินค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่นๆ ต่อไป

1.3 ขอบเขตของการวิจัย

1. ทำการศึกษาตัวแบบโพรบิตแบบ 2 ประเภทเพื่อหาจุดแบ่งที่เหมาะสมที่สุด
2. ตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มี 2 ค่า คือ 0 และ 1 โดยกำหนดสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ(a) ของลักษณะที่สนใจศึกษา เท่ากับ 0.1 ,0.5 และ 0.9
3. จำนวนของตัวแปรอิสระ (p) ในการวิจัยครั้งนี้แบ่งเป็น 3 ระดับ คือ
 - จำนวนตัวแปรอิสระน้อย คือ 1 และ 2 ตัว
 - จำนวนตัวแปรอิสระปานกลาง คือ 3 และ 4 ตัว
 - จำนวนตัวแปรอิสระมาก คือ 5 และ 6 ตัว
4. ขนาดตัวอย่าง (n) ในการวิจัยครั้งนี้แบ่งเป็น 3 ระดับ คือ
 - ขนาดตัวอย่างเล็ก คือ 20 และ 40
 - ขนาดตัวอย่างปานกลาง คือ 60 และ 80
 - ขนาดตัวอย่างใหญ่ คือ 100 และ 120

โดยขนาดตัวอย่างและจำนวนตัวแปรอิสระที่ใช้ในการสร้างชุดข้อมูลจะขึ้นอยู่กับ

$$\text{ความสัมพันธ์ ดังนี้ } \frac{n}{p} \geq 20$$

5. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M)
 - ตัวแปรอิสระไม่มีความสัมพันธ์กัน ($\rho = 0$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่ำ ($\rho = 0.33$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับปานกลาง ($\rho = 0.67$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับสูง ($\rho = 0.99$)

โดยกำหนดให้รูปแบบเมทริกซ์สหสัมพันธ์ (Correlation Matrix) คือ

$$\rho_{p \times p} = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \rho & \cdots & \rho^{p-1} \\ \rho & 1 & \cdots & \rho^{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \cdots & 1 \end{pmatrix}$$

โดยที่ ρ_{ij} ; $i, j = 1, 2, \dots, p$ คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j ซึ่งจะมีทั้งหมด $\frac{p(p-1)}{2}$ คู่

6. กำหนดค่าพารามิเตอร์เริ่มต้นของสมการการถดถอยเป็นค่าใดๆ ในการวิจัยครั้งนี้ กำหนดให้ $\beta_i = 0.1; i = 0, 1, 2, \dots, p$ และ $\varepsilon_i \sim N(0, 25); i = 1, 2, \dots, n$
7. กำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05
8. ตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอรม์
9. ในการวิจัยครั้งนี้ทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation) โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

1.4 ข้อตกลงเบื้องต้น

สำหรับการวิจัยครั้งนี้ศึกษามาจากตัวอย่างที่จำลองขึ้นมาจากตัวแบบ และการแจกแจง ดังนี้

1. ศึกษาตัวแบบโพรบิตแบบ 2 โดยมีรูปแบบ คือ

$$\pi_i = P(Y_i = 1)$$

$$= P(Y_i^* > 0)$$

$$\text{จาก} \quad = P(\varepsilon_i < \beta' X_i) \quad ; i = 1, 2, \dots, n$$

$$= \int_{-\alpha}^{\beta X_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2}\right) d\varepsilon$$

โดยที่ π_i คือ ค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

Y_i คือ ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1

Y_i^* คือ ตัวแปรแฝง (Latent Variable) เป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริง

ทราบเพียงแต่ผลที่เกิดขึ้น

$\Phi(\cdot)$ คือ ฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน

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

β' คือ เวกเตอร์ของสัมประสิทธิ์การถดถอย จำนวน $k+1$ ตัว ขนาด $1 \times (k+1)$

X_i คือ เวกเตอร์ของตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณจำนวน k ตัว ขนาด

$$(k+1) \times 1$$

2. ตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอรม์ ที่มีฟังก์ชันความน่าจะเป็นของ x อยู่ในรูปของ

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

โดยมี

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

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

1.5 คำจำกัดความที่ใช้ในการวิจัย

1. พารามิเตอร์ (Parameter) คือค่าคงที่ที่แสดงคุณลักษณะบางประการของประชากร
2. ค่าเฉลี่ย (Mean) คือ ค่าที่แสดงว่าข้อมูลส่วนใหญ่อยู่ที่ไหน
3. ความแปรปรวน (Variance) คือ ค่าที่บอกว่าค่าของตัวแปรสุ่ม จะมีความผิดพลาดกา
ลังสอง โดยเฉลี่ยจากค่าของข้อมูลส่วนใหญ่ เท่ากับเท่าไร กล่าวคือ เป็นค่าที่วัด
แนวโน้มที่ตัวแปรสุ่ม จะห่างจากค่าเฉลี่ยของมัน
4. ตัวแบบโพรบิตแบบ 2 ประเภท (Binary Probit Model) หมายถึง ตัวแบบที่ศึกษา
ความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ เมื่อตัวแปรตามเป็นตัวแปรเชิง
คุณภาพที่มี 2 ลักษณะ คือมีค่าเท่ากับ 1 หรือ 0 และตัวแปรอิสระเป็นตัวแปรเชิง
ปริมาณหรือตัวแปรหุ่น เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรแล้ว จะนำไปใช้ใน
การประมาณค่าตัวแปรตาม
5. จุดแบ่ง (Cut-off Point) หมายถึง ค่าความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนก
กลุ่มของข้อมูลว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มของเหตุการณ์ที่สนใจ
(Success) และกลุ่มไม่เกิดเหตุการณ์ที่สนใจ(Failure)
6. ข้อมูลดิบ (Ungrouped Data) คือ ข้อมูลที่เพิ่งได้จากการเก็บรวบรวมข้อมูล ยังไม่ได้
จัดแบ่งกลุ่ม จำแนกกลุ่ม แยกประเภทและไม่อยู่ในรูปตารางความถี่
7. ตัวแปรแฝง (Latent Variable) เป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริง ทราบ
เพียงแต่ผลที่เกิดขึ้น
8. ความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) คือสถานการณ์ที่ตัวแปรอิสระ
มีความสัมพันธ์กัน โดยแบ่งเป็น 5 ระดับ ดังนี้
 - ตัวแปรอิสระไม่มีความสัมพันธ์กัน
 - ตัวแปรอิสระมีความสัมพันธ์ในระดับต่ำ
 - ตัวแปรอิสระมีความสัมพันธ์ในระดับปานกลาง
 - ตัวแปรอิสระมีความสัมพันธ์ในระดับสูง
 - ตัวแปรอิสระมีความสัมพันธ์กันอย่างสมบูรณ์

9. การแจกแจงแบบเบอร์นูลลี (Bernoulli Distribution) หมายถึงตัวแปรสุ่ม Y เรียกว่าตัวแปรสุ่มเบอร์นูลลี กล่าวคือ

ถ้า $Y = 0$ เมื่อ การทดลองไม่เกิดเหตุการณ์ที่สนใจ
 และถ้า $Y = 1$ เมื่อ การทดลองเกิดเหตุการณ์ที่สนใจ
 โดยที่ $P(Y = 1) = p$, $0 < p < 1$ และ $P(Y = 0) = 1 - p$

เราอาจเขียนแทนด้วย $Y \sim \text{Ber}(p)$ ซึ่งฟังก์ชันความน่าจะเป็นอยู่ในรูปของ

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

10. การประมาณค่าแบบช่วงหรือ ช่วงความเชื่อมั่น (Confidence Intervals) เป็นการประมาณค่าพารามิเตอร์ของประชากรในรูปแบบช่วงโดยใช้ข้อมูลตัวอย่าง การประมาณแบบช่วงนั้นจะบอกถึงค่าต่ำสุดและค่าสูงสุดของพารามิเตอร์ที่เป็นไปได้

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท เมื่อ จำนวนตัวแปรอิสระเปลี่ยนแปลงไป, ขนาดตัวอย่างเปลี่ยนแปลงไป, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูลเปลี่ยนแปลงไป, ระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนแปลงไป และนำปัจจัยต่างๆเหล่านี้มาพิจารณารวมกัน ซึ่งแปรเปลี่ยนไปพร้อมกัน
2. สามารถหาสมการถดถอยพหุคูณ สำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่นๆ ต่อไป
3. สามารถใช้เพื่อเป็นแนวทางในการศึกษาวิจัยต่อไป

1.7 วิธีดำเนินการวิจัย

1. ศึกษาค้นคว้าเอกสารและข้อมูลที่เกี่ยวข้องกับการวิจัย
2. จำลองข้อมูลให้มีการแจกแจงตามลักษณะที่ต้องการศึกษาโดยใช้เทคนิคมอนติคาร์โล
3. คำนวณหาจุดแบ่งที่เหมาะสมที่สุดในแต่ละสถานการณ์สำหรับตัวแบบโพรบิตแบบ 2 ประเภท ซึ่งทำให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุดหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด
4. ทำการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์
5. คำนวณหาค่าเฉลี่ยของจุดแบ่งที่เหมาะสมที่สุดเพื่อนำมาคำนวณหาค่าเปอร์เซ็นต์ (Percent) และช่วงความเชื่อมั่น

6. ใช้ตัวแบบการถดถอยพหุคูณ เพื่อประมาณค่าพารามิเตอร์ สำหรับใช้ในการหาจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป
7. สรุปผลที่ได้จากการวิจัย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

แนวคิด ทฤษฎีและสถิติที่เกี่ยวข้อง

การหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่ม สำหรับการวิจัยครั้งนี้จะใช้หาจุดแบ่งที่เหมาะสมที่สุดในตัวแบบโพรบิตแบบ 2 ประเภท ซึ่งจุดแบ่งที่เหมาะสมที่สุดจะให้ค่าอัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุด และใช้ตัวแบบการถดถอยพหุคูณเพื่อประมาณค่าพารามิเตอร์สำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป ซึ่งในบทนี้จะกล่าวถึงรายละเอียดเกี่ยวกับตัวแบบโพรบิตแบบ 2 ประเภท การประมาณค่าสัมประสิทธิ์การถดถอยในตัวแบบโพรบิตแบบ 2 ประเภท ด้วยวิธีภาวะความน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) การหาจุดแบ่งโดยใช้ทฤษฎี Hadjicostas P. (2006) การหาอัตราความผิดพลาดในการจำแนกกลุ่มหรือสัดส่วนความถูกต้องในการจำแนกกลุ่ม

2.1 ตัวแบบโพรบิตแบบ 2 ประเภท (Binary Probit Model)

ตัวแบบที่ศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ เมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่มี 2 ลักษณะ และตัวแปรอิสระเป็นตัวแปรเชิงปริมาณหรือตัวแปรหุ่น เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรแล้ว จะนำไปใช้ในการประมาณค่าตัวแปรตาม และพยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่งได้

จากตัวแปรตาม (Y_i) ซึ่งเป็นตัวแปรเชิงคุณภาพที่เป็นได้ 2 ค่า คือ 0 กับ 1 และ $X_{i1}, X_{i2}, \dots, X_{ik}$ เป็นตัวแปรอิสระของค่าสังเกตที่ได้จากหน่วยตัวอย่างที่ i โดยที่ $i = 1, 2, 3, \dots, n$

โดยมี ตัวแปรแฝง คือ Y_i^* ซึ่งเป็นค่าที่วัดไม่ได้ จึงไม่ทราบค่าที่แท้จริง ทราบเพียงแต่ผลที่เกิดขึ้น โดย Y_i^* ของหน่วยตัวอย่างที่ i เป็นฟังก์ชันเชิงเส้นของตัวแปรอิสระ $X_{i1}, X_{i2}, \dots, X_{ik}$ และนั่นคือ

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

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

$$\text{จะได้ } Y_i = \begin{cases} 1 & \text{if } Y_i^* > 0 \text{ or } \varepsilon_i < \beta'X_i \\ 0 & \text{if } Y_i^* \leq 0 \text{ or } \varepsilon_i \geq \beta'X_i \end{cases}$$

จาก

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

- โดยที่ π_i คือ ค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i
- Y_i คือ ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1
- Y_i^* คือ ตัวแปรแฝง เป็นค่าที่วัดไม่ได้ ไม่ทราบค่าที่แท้จริง ทราบเพียงแต่ผล
- $\Phi(\cdot)$ คือ ฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน
- ε คือ ตัวแปรสุ่มปกติมาตรฐาน โดย $\varepsilon_i \sim N(0,1)$ ^{iid}
- β' คือ เวกเตอร์ของสัมประสิทธิ์การถดถอย จำนวน $k+1$ ตัว ขนาด $1 \times (k+1)$
- X_i คือ เวกเตอร์ของตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณจำนวน k ตัว ขนาด $(k+1) \times 1$

ดังนั้น Probit Model สามารถเขียนแปลงให้อยู่ในรูป

$$\begin{aligned}
 \Phi^{-1}(\pi_i) &= \beta'X_i \\
 &= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} \dots \dots \dots (1)
 \end{aligned}$$

จากสมการที่(1) พบว่า การแปลงดังกล่าว จะทำให้เราได้ความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปร

2.2 ฟังก์ชันความน่าจะเป็น (Likelihood function) ของข้อมูลการถดถอยโพรบิต

จากประชากรที่ทำการศึกษาคือตัวแปรตาม (Y_i) ที่มีเพียง 2 ค่า คือ 0 กับ 1 จึงใช้ฟังก์ชันการแจกแจงแบบเบอร์นูลลี

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

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

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

$$\text{ดังนั้น} \quad L(\beta) = \prod_i \Phi(x_i'\beta)^{y_i} [1 - \Phi(x_i'\beta)]^{1-y_i} \quad \dots\dots\dots (2)$$

2.3 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation)

การหาค่าตัวประมาณค่าสัมประสิทธิ์การถดถอย ($\hat{\beta}$) ด้วยวิธีภาวะความน่าจะเป็นสูงสุด คือต้องทำให้ L มีค่ามากที่สุดโดยทำการหาอนุพันธ์เทียบกับ β ต่างๆ โดยเราสามารถเขียนให้อยู่ในรูปของลอการิทึม (Logarithm) หรือความควรจะเป็นลอการิทึม (Log-Likelihood) ได้ดังนี้

$$\ln L(\beta) = \sum_{i=1}^n \{y_i \cdot \ln[\Phi(x_i'\beta)] + (1 - y_i) \cdot \ln[1 - \Phi(x_i'\beta)]\} \quad \dots\dots\dots (3)$$

$$= \sum_{y_i=0} \ln[1 - \Phi(x_i'\beta)] + \sum_{y_i=1} \ln \Phi(x_i'\beta) \quad \dots\dots\dots (4)$$

และเงื่อนไขอันดับแรก (First Order) สำหรับการให้สมการที่ (4) มีค่าสูงสุด (Maximization) และสามารถแก้สมการหาค่า First Order Condition ของแบบจำลองโพรบิต ได้ดังนี้

$$\begin{aligned} \frac{\partial \ln L(\beta)}{\partial \beta} &= \sum_{i=1}^n \left\{ \frac{y_i \phi(\cdot)}{\Phi(\cdot)} + (1 - y_i) \left[\frac{-\phi(\cdot)}{1 - \Phi(\cdot)} \right] \right\} x_i = 0 \\ &= \sum_{y_i=0} \left[\frac{-\phi(x_i'\beta)}{1 - \Phi(x_i'\beta)} \right] x_i + \sum_{y_i=1} \left[\frac{\phi(x_i'\beta)}{\Phi(x_i'\beta)} \right] x_i = 0 \quad \dots\dots\dots (5) \end{aligned}$$

เมื่อ ϕ_i คือ ฟังก์ชันความหนาแน่นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

Φ_i คือ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ

เมื่อประมาณค่าพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุดในตัวแบบโพรบิตแบบ 2 ประเภท แล้ว จะสามารถนำไปใช้ในการพยากรณ์การจำแนกกลุ่มของตัวแบบ ดังนี้

- หน่วยที่ i จะถูกจัดให้อยู่ในกลุ่มที่เกิดที่สนใจ ($Y=1$) ถ้า

$$\hat{\pi}_i = \Phi(\beta'X_i) > c \quad ; 0 \leq c \leq 1$$

- หน่วยที่ i จะถูกจัดให้อยู่ในกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจ ($Y=0$) ถ้า

$$\hat{\pi}_i = \Phi(\beta'X_i) \leq c \quad ; 0 \leq c \leq 1$$

เมื่อ c คือ จุดแบ่ง หรือระดับของความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนก

กลุ่มว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มการเกิดเหตุการณ์ที่สนใจ และกลุ่มไม่เกิดเหตุการณ์ที่สนใจ

2.4 สถิติ Kaiser-Meyer-Olkin (KMO)

Kaiser (1970) ได้เสนอสถิติ KMO ในการศึกษานี้ สถิติ KMO เป็นสถิติที่ใช้วัดระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity)

$$0 \leq \text{KMO} = \frac{\sum_{i=1}^p \sum_{j=1}^p \rho_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p \rho_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p \rho_{ij|1,2,\dots,(i-1),(i+1),\dots,(j-1),(j+1),\dots,p}^2} \leq 1$$

โดยที่ ρ_{ij} คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

$\rho_{ij|1,2,\dots,(i-1),(i+1),\dots,(j-1),(j+1),\dots,p}$ คือ สัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j สำหรับ $i < j = 1, 2, \dots, p$

เกณฑ์สำหรับการแบ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระ แสดงดังนี้

- ไม่มีระดับความสัมพันธ์ระหว่างตัวแปรอิสระ: $\text{KMO} = 0$
- ตัวแปรอิสระมีความสัมพันธ์กันในระดับอย่างต่ำ: $0 < \text{KMO} \leq 0.50$
- ตัวแปรอิสระมีความสัมพันธ์กันในระดับปานกลาง: $0.50 < \text{KMO} \leq 0.75$
- ตัวแปรอิสระมีความสัมพันธ์กันในระดับสูง: $0.75 < \text{KMO} < 1.00$
- ตัวแปรอิสระมีความสัมพันธ์กันอย่างสมบูรณ์: $\text{KMO} = 1$

2.5 การตรวจสอบความเหมาะสมของตัวแบบ

ตัวแบบที่เป็นไปได้มีหลายตัวแบบ ซึ่งมีวิธีการวัดทางสถิติที่หลากหลาย สำหรับใช้ในการวัดว่าตัวแบบพหุคูณแบบ 2 ประเภทมีความสามารถในการจำแนกกลุ่มให้แต่ละหน่วยอยู่ในกลุ่มใดกลุ่มหนึ่งระหว่างกลุ่มที่เกิดเหตุการณ์การที่สนใจ และกลุ่มที่ไม่เกิดเหตุการณ์ที่สนใจว่าเหมาะสมมากน้อยเพียงใด ซึ่งมีวิธีต่างๆ ดังนี้

- Chi-square Goodness of Fit Test และสถิติ Deviance
- HosMer-LoMeshow Test
- Classification Table
- Receiver Operating Characteristic Curve (ROC)
- สัมประสิทธิ์การตัดสินใจสำหรับการถดถอยพหุคูณ (R^2)

- ตรวจสอบความถูกต้องของตัวแบบ (Model validation) ด้วยวิธีการใช้ชุดข้อมูลภายนอกหรือโดยแบ่งชุดข้อมูลออกเป็น 2 ส่วน

Classification table จะสอดคล้องกับแต่ละจุดแบ่ง ที่ถูกคัดเลือกซึ่งจะถูกใช้เป็นเกณฑ์สำหรับการคัดเลือกจุดที่ทำให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำมีค่าต่ำมีค่าต่ำที่สุด โดย Classification table สำหรับแต่ละจุดแบ่ง แสดงดังนี้

		ค่าสังเกต		
		$y_i = 1$	$y_i = 0$	
ค่าพยากรณ์ (เกณฑ์ที่ใช้ในการตัดสินใจ คือ $\hat{\pi}_i \geq c$)	$\hat{y}_i = 1$	A	C	A+C
	$\hat{y}_i = 0$	B	D	B+D
		A+B	C+D	A+B+C+D

โดย

“A” คือ จำนวนของค่าสังเกตที่ถูกจัดให้อยู่ในกลุ่มของการเกิดเหตุการณ์ที่สนใจและค่าสังเกตที่แท้จริงอยู่ในกลุ่มการเกิดเหตุการณ์ที่สนใจด้วย

“B” คือ จำนวนของค่าสังเกตที่ถูกจัดให้อยู่ในกลุ่มของการไม่เกิดเหตุการณ์ที่สนใจในขณะที่ค่าสังเกตที่แท้จริงอยู่ในกลุ่มการเกิดเหตุการณ์ที่สนใจด้วย

“C” คือ จำนวนของค่าสังเกตที่ถูกจัดให้อยู่ในกลุ่มของการเกิดเหตุการณ์ที่สนใจในขณะที่ค่าสังเกตที่แท้จริงอยู่ในกลุ่มของการไม่เกิดเหตุการณ์ที่สนใจด้วย

“D” คือ จำนวนของค่าสังเกตที่ถูกจัดให้อยู่ในกลุ่มของการเกิดเหตุการณ์ที่สนใจและค่าสังเกตที่แท้จริงอยู่ในกลุ่มของการไม่เกิดเหตุการณ์ที่สนใจด้วย

การคัดเลือกจุดแบ่ง ต้องทำให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงที่สุด โดยแสดงดังนี้

อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำมีค่าต่ำ (Classification Error Rate: CER)	สัดส่วนความถูกต้องในการจำแนกกลุ่ม
$CER = \frac{B + C}{A + B + C + D}$	$1 - CER$
CER มีค่าต่ำกว่าแสดงว่ามีความเหมาะสมมากกว่า	อัตราความถูกต้องในการจำแนกกลุ่มมีค่ามากกว่าแสดงว่ามีความเหมาะสมมากกว่า

2.6 ช่วงความเชื่อมั่น (Confidence Interval)

การประมาณค่าแบบช่วงหรือช่วงความเชื่อมั่นเป็นการประมาณค่าพารามิเตอร์ของประชากรในรูปแบบช่วงโดยใช้ข้อมูลตัวอย่าง การประมาณแบบช่วงนั้นจะบอกถึงค่าต่ำสุดและค่าสูงสุดของพารามิเตอร์ที่เป็นไปได้

ระดับของค่าความเชื่อมั่นที่ใช้ในการสร้างช่วงความเชื่อมั่นนั้นจะกำหนดเป็นค่าควบคู่กับระดับนัยสำคัญ นั่นคือ $1 - \alpha$ ที่เรียกว่าช่วงความเชื่อมั่น $(1 - \alpha) 100\%$ จะได้ว่า

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

เรียก L ว่าขีดจำกัดความเชื่อมั่นล่าง (Lower Confidence Limit)

U ว่าขีดจำกัดความเชื่อมั่นบน (Upper Confidence Limit)

2.7 เปอร์เซ็นไทล์ (Percentile)

เป็นค่าที่แบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน เมื่อข้อมูลถูกเรียงจากน้อยไปหามาก เนื่องจากค่าที่แบ่งจำนวนข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน มีอยู่ 99 ค่า ดังนั้นเราจึงตั้งชื่อแต่ละค่าว่า เปอร์เซ็นไทล์ที่หนึ่ง (P_1) เปอร์เซ็นไทล์ที่สอง (P_2).... และเปอร์เซ็นไทล์ที่เก้าสิบเก้า (P_{99}) ตามลำดับ

การหาเปอร์เซ็นไทล์ คือต้องหาตำแหน่งของเปอร์เซ็นไทล์ก่อน ให้ N เป็นจำนวนข้อมูลทั้งหมด ตำแหน่งต่างๆ ของเปอร์เซ็นไทล์หาได้ดังนี้

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

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

⋮

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

โดยทั่วไป ตำแหน่งของเปอร์เซ็นไทล์ที่ r คือ

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

2.8 การหาจุดแบ่ง โดยทฤษฎีHadjicostas P. (2006)

ทฤษฎี Hadjicostas P. (2006) เป็นวิธีการนี้ใช้ผลลัพธ์ทางคณิตศาสตร์ที่ง่ายแต่มีความถูกต้องแม่นยำในการหาจุดแบ่ง ที่ทำให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุดหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุดอีกด้วย โดยมีวิธีดังนี้

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

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

จากทฤษฎีและบทตั้งดังกล่าว จะได้ว่า

1. เรียงอันดับค่า $\hat{\pi}_i$ จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < \dots < \hat{\pi}_n$
2. สำหรับแต่ละ $i \in \{1, 2, \dots, n\}$ ใดๆ ให้ $M(i)$ เป็น $\max j \in \{1, 2, \dots, n\}$ ที่ซึ่ง $\hat{\pi}_i = \hat{\pi}_j$ โดย $M(0) = 0$; $i \leq M(i) \leq n$
3. หาค่า a_i โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ สำหรับ $i = 0, 1, 2, \dots, n$ แบ่งเป็น 2 กรณี คือ
 กรณีที่ 1 $a_{i+1} = a_i + \sum_{k=M(i)+1}^{M(i+1)} (-1)^{y_k}$ ถ้า $M(i) < i+1$
 กรณีที่ 2 $a_{i+1} = a_i$ ถ้า $i+1 \leq M(i)$
4. หา I_0 ซึ่งเป็นเซตของ j ทั้งหมด $j \in \{0, 1, 2, \dots, n\}$ ที่ซึ่ง $a_j = \max_{0 \leq i \leq n} a_i$
5. หา C_0 เป็นเซตของ c_0 ทั้งหมด $c_0 \in [0, 1]$ ที่ซึ่ง $p(c_0) = \max_{c \in [0, 1]} p(c)$ แล้ว $C_0 = \bigcup_{i \in I_0} A_i$ โดย $i \in \{0, 1, 2, \dots, n\}$ จะได้ว่า

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

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

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

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

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

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

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

โดย $p(c)$ คือ สัดส่วนของความถูกต้องในการจำแนกกลุ่มที่จุด c

$N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบโพรบิตแบบ 2 ประเภทสำหรับแต่ละสถานการณ์ที่ต้องการศึกษา โดยจุดแบ่งที่เหมาะสมที่สุดจะให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุดหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด จากนั้นจะนำผลลัพธ์ของทุกสถานการณ์มาวิเคราะห์ด้วยตัวแบบการถดถอยพหุคูณที่มีผลอันตรกิริยา (Interaction) เพื่อประมาณค่าพารามิเตอร์สำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป

การจำลองข้อมูลในแต่ละสถานการณ์จะจำลองขึ้นด้วยการทำงานของเครื่องคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรม R เนื่องจากวิธีมอนติคาร์โลเป็นเทคนิคที่ใช้ในการวิจัยครั้งนี้ ดังนั้นในตอนแรกของบทนี้จะกล่าวถึงวิธีมอนติคาร์โลก่อน แล้วจึงแสดงรายละเอียดของแผนการดำเนินการวิจัย ขั้นตอนในแผนการดำเนินการวิจัย ตลอดจนโปรแกรมที่ใช้ในการวิจัย ซึ่งรายละเอียดต่างๆเป็นดังนี้

3.1 เทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นการจำลองระบบที่ไม่เปลี่ยนแปลงตามเวลา ซึ่งตัวแบบของการจำลองจะมีลักษณะเป็นตัวแทนทางคณิตศาสตร์ โดยการนำตัวเลขสุ่ม มาประยุกต์ใช้ในการแก้ปัญหาหรือหาคำตอบให้กับระบบที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น ซึ่งมีขั้นตอนที่สำคัญ 3 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การสร้างเลขสุ่ม (Generate Random Number) การสร้างเลขสุ่มจะกำหนดให้มีการแจกแจงแบบยูนิฟอร์มในช่วง $[0, 1]$ และเป็นอิสระซึ่งกันและกัน จากนั้นนำเลขสุ่มนี้ไปสร้างตัวแปรตามลักษณะการแจกแจงที่ต้องการศึกษา เพื่อเป็นข้อมูลของปัญหานั้นๆ

ขั้นตอนที่ 2 การประยุกต์ใช้เลขสุ่มในการแก้ปัญหา ขั้นตอนนี้เป็นการนำตัวแปรที่ได้จากขั้นตอนแรกมาใช้ในการหาค่าต่างๆ ตามปัญหาที่ต้องการศึกษา

ขั้นตอนที่ 3 การทดลอง ขั้นตอนนี้เป็นการทำวิธีนั้นซ้ำๆกัน (Replication) จำนวนหลายครั้ง โดยถือว่าการทำซ้ำๆ กันนั้น เป็นวิธีการเก็บรวบรวมข้อมูลให้มีจำนวนมาก เพื่อลดความไม่แน่นอนของคำตอบ ในการวิเคราะห์หาค่าต่างๆ ได้

จากหลักการของเทคนิคมอนติคาร์โล จะเห็นว่าการใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีที่จะนำไปสู่แนวคิดในทางทฤษฎีที่เกี่ยวข้องกับการคำนวณโดยเฉพาะทฤษฎีความน่าจะเป็นที่จะนำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริงเพราะไม่มี

ผลกระทบจากเรื่องอื่นๆ เข้ามาเกี่ยวข้องในการทดลอง เมื่อทำซ้ำๆกันเป็นจำนวนมากแล้ว ความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่างๆ ในแต่ละครั้งให้หมดไป

3.2 แผนการดำเนินการวิจัย

ในการวิจัยครั้งนี้ได้จำลองข้อมูลขึ้น โดยกำหนดสถานการณ์จำลองต่างๆ ดังนี้

1. กำหนดให้ข้อมูลของตัวแปรอิสระ (X) เริ่มต้นมีการแจกแจงแบบยูนิฟอร์ม
2. เมื่อข้อมูลตัวแปรตาม (Y^*) ให้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระและความผิดพลาด กำหนดให้ค่าพารามิเตอร์เริ่มต้นของสมการการถดถอยเป็นค่าใดๆ ในการวิจัยครั้งนี้ $\beta_i = 0.1; i = 0, 1, 2, \dots, p$ และ $\varepsilon_i \sim N(0, 25); i = 1, 2, \dots, n$
3. ตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มี 2 ค่า คือ 0 และ 1 โดยกำหนดสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a) เท่ากับ 0.1, 0.5 และ 0.9
4. กำหนดให้จำนวนของตัวแปรอิสระ (p) ในการวิจัยครั้งนี้แบ่งเป็น 3 ระดับ คือ
 - จำนวนตัวแปรอิสระน้อย คือ 1 และ 2 ตัว
 - จำนวนตัวแปรอิสระปานกลาง คือ 3 และ 4 ตัว
 - จำนวนตัวแปรอิสระมาก คือ 5 และ 6 ตัว
5. กำหนดให้ขนาดตัวอย่าง (n) ในการวิจัยครั้งนี้แบ่งเป็น 3 ระดับ คือ
 - ขนาดตัวอย่างเล็ก คือ 20 และ 40
 - ขนาดตัวอย่างปานกลาง คือ 60 และ 80
 - ขนาดตัวอย่างใหญ่ คือ 100 และ 120
6. กำหนดให้ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M) ในการวิจัยครั้งนี้มี 4 กรณี คือ
 - ไม่มีความสัมพันธ์กันระหว่างตัวแปรอิสระ ($\rho = 0$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่ำ ($\rho = 0.33$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับปานกลาง ($\rho = 0.67$)
 - ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับสูง ($\rho = 0.99$)
7. กำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05
8. ในการวิจัยครั้งนี้ทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

3.3 ขั้นตอนในการดำเนินการวิจัย

สำหรับการดำเนินการวิจัยมีขั้นตอนดังนี้

1. ศึกษาตัวแบบโพบริตและวิธีการหาค่าของจุดแบ่งจากทฤษฎี Hadjicostas P.(2006)
2. สร้างข้อมูลตามลักษณะที่กำหนดเพื่อใช้ในการวิจัย
3. หาค่า Y^* จากการสร้างความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระ
4. ทำการแปลงค่าตัวแปรตาม Y^* ที่ได้เป็น Y ที่มีค่าเป็น 0 หรือ 1 ตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและขนาดตัวอย่างตามที่กำหนดไว้
5. ประมาณค่าพารามิเตอร์ของตัวแบบโพบริตแบบ 2 ประเภทโดยใช้วิธีภาวะน่าจะเป็นสูงสุด ในกรณีที่ข้อมูลมีลักษณะที่ต้องการศึกษา
6. คำนวณหาจุดแบ่งที่เหมาะสมที่สุดในแต่ละสถานการณ์สำหรับตัวแบบโพบริตแบบ 2 ประเภท ซึ่งทำให้สัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด
7. ทำการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์
8. คำนวณหาค่าเฉลี่ยของจุดแบ่งที่เหมาะสมที่สุดของแต่ละสถานการณ์จาก 500 รอบ พร้อมทั้งหาค่าร้อยละและช่วงความเชื่อมั่น
9. ใช้ตัวแบบการถดถอยพหุคูณ เพื่อประมาณค่าพารามิเตอร์สำหรับการประมาณค่าของจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป
10. สรุปผลการวิจัยในแต่ละสถานการณ์

3.4 การจำลองข้อมูลที่ใช้ในการวิจัย

การจำลองข้อมูลที่ใช้ในการวิจัย มีขั้นตอนต่างๆดังต่อไปนี้

1. สร้างข้อมูลตัวแปรอิสระเริ่มต้นให้มีการแจกแจงแบบยูนิฟอร์มซึ่งมี equal space คือ กำหนดให้มีค่าเป็นช่วงลบและช่วงบวกเท่าๆกัน ตามขนาดตัวอย่างที่กำหนดไว้ ดังนี้

$$x \sim U\left(-\frac{n}{2}, +\frac{n}{2}\right)$$

2. สร้างจำนวนตัวแปรอิสระตามที่กำหนดไว้ และให้ตัวแปรอิสระดังกล่าวมีความสัมพันธ์กันตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่กำหนดไว้ โดยกำหนดให้รูปแบบเมทริกซ์สหสัมพันธ์ (Correlation Matrix) ดังนี้

$$\rho_{p \times p} = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \rho & \cdots & \rho^{p-1} \\ \rho & 1 & \cdots & \rho^{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \cdots & 1 \end{pmatrix}$$

โดยที่ ρ_{ij} ; $i, j = 1, 2, \dots, p$ คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j ซึ่งจะมีทั้งหมด $\frac{p(p-1)}{2}$ คู่

- ตัวแปรอิสระไม่มีความสัมพันธ์กัน ($\rho = 0$)

โดยสร้างตามเงื่อนไข $\rho_{12} : \rho_{13} : \dots : \rho_{p,p-1} = 0$

- ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่ำ ($\rho = 0.33$)

โดยสร้างตามเงื่อนไข $\rho_{12} : \rho_{13} : \dots : \rho_{p,p-1}$ คือ $(0.33)^1 : (0.33)^2 : \dots : (0.33)^{\frac{p(p-1)}{2}}$

- ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับปานกลาง ($\rho = 0.67$)

โดยสร้างตามเงื่อนไข $\rho_{12} : \rho_{13} : \dots : \rho_{p,p-1}$ คือ $(0.67)^1 : (0.67)^2 : \dots : (0.67)^{\frac{p(p-1)}{2}}$

- ความสัมพันธ์ระหว่างตัวแปรอิสระในระดับสูง ($\rho = 0.99$) โดย

สร้างตามเงื่อนไข $\rho_{12} : \rho_{13} : \dots : \rho_{p,p-1}$ คือ $(0.99)^1 : (0.99)^2 : \dots : (0.99)^{\frac{p(p-1)}{2}}$

3. สร้างค่าตัวแปรตาม (Y^*) จากการนำข้อมูลตัวแปรอิสระที่สร้างได้จากข้างต้น มาหาค่าตัวแปรตาม (Y^*) โดยสร้างให้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระและความคลาดเคลื่อน ซึ่งมีรูปแบบ ดังนี้

$$Y^* = \beta X' + \varepsilon$$

โดยที่ Y^* เป็นเมตริกซ์ของตัวแปรตามที่ทำการพยากรณ์เพื่อกำหนดค่าเบื้องต้น

เมื่อ X' เป็นเมตริกซ์ของตัวแปรอิสระ

β เป็นเวกเตอร์ของพารามิเตอร์ที่กำหนด กำหนดให้ β เริ่มต้น เท่ากับ 0.1

ε เป็นความคลาดเคลื่อนซึ่ง $\varepsilon \sim N(0, 25)$

4. สร้างค่าตัวแปรตาม (Y) ที่มีค่าเป็น 0 หรือ 1 จากค่า Y^* ที่สร้างได้จากข้างต้นโดยทำการแปลงค่าตัวแปรตาม Y^* ที่ได้เป็น Y ที่มีค่าเป็น 0 หรือ 1 ตามสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจของลักษณะที่สนใจศึกษา และค่า ขนาดตัวอย่างที่กำหนดไว้ข้างต้นดังนี้

4.1 หาค่าจำนวน Y ที่มีค่าเป็น 0 และ Y ที่มีค่าเป็น 1 โดย

- Y ที่มีค่าเป็น 0 มีจำนวน = ขนาดตัวอย่าง \times สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ

- Y ที่มีค่าเป็น 1 มีจำนวน = ขนาดตัวอย่าง - [จำนวน Y ที่มีค่าเป็น 0]

4.2 การกำหนดค่า Y^* ให้เป็นค่า Y ที่มีค่าเป็น 0 และ Y ที่มีค่าเป็น 1 โดย

- เรียงลำดับค่าของ Y^* ทั้งหมดที่ได้ จากน้อยไปมาก

- ให้ Y^* ที่มีค่าน้อยที่สุด ตามจำนวน Y ที่มีค่าเป็น 0 เป็น Y ที่มีค่าเป็น 0 และ Y^* นอกนั้นคือ Y ที่มีค่าเป็น 1 ตามจำนวน Y ที่มีค่าเป็น 1

5. ประมาณค่าพารามิเตอร์โดยใช้ตัวแบบโพรบิตแบบ 2 ประเภทด้วยวิธีภาวะความน่าจะเป็นสูงที่สุด

6. หาค่าประมาณของ $\hat{\pi}_i$ โดยนำค่าพารามิเตอร์ที่ได้จากข้อ 5 และ ค่าของตัวแปรอิสระที่สร้างขึ้น มาแทนค่ากลับลงไปในตัวแบบโพรบิตแบบ 2 ประเภท

3.5 คำนวณค่าของจุดแบ่งโดยทฤษฎี Hadjicostas P. (2006)

เมื่อได้ข้อมูลที่มีลักษณะตามต้องการแล้ว จากทฤษฎีของ Hadjicostas P. (2006) จะสามารถหาค่าของจุดแบ่งตามขั้นตอนต่อไปนี้

ขั้นที่ 1. เรียงลำดับค่า $\hat{\pi}_i$ จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < \dots < \hat{\pi}_n$

ขั้นที่ 2. หาค่า $M(i)$ สำหรับแต่ละ $i \in \{1, 2, \dots, n\}$ โดย $M(i)$ คือ อันดับของ $\hat{\pi}_i$

แต่ ถ้า $\hat{\pi}_i = \hat{\pi}_j$ จะเลือก $M(i)$ ที่มีค่ามากที่สุดเป็นอันดับของค่า $\hat{\pi}_i$ และ $\hat{\pi}_j$

ขั้นที่ 3. หาค่า a_i สำหรับ $i = 0, 1, 2, \dots, n$ โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ ซึ่งแบ่งเป็น 2 กรณี คือ

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

$$\text{กรณีที่ 2 } a_{i+1} = a_i \quad \text{ถ้า } i+1 \leq M(i)$$

ขั้นที่ 4. หาค่า I_0 คือเซตของ j ทั้งหมด $j \in \{0, 1, 2, \dots, n\}$ ที่ซึ่ง $a_j = \max_{0 \leq i \leq n} a_i$

ขั้นที่ 5. หาค่า C_0 ซึ่งเป็นเซตของ c_0 ทั้งหมด จาก $C_0 = \bigcup_{i \in I_0} A_i$; $i \in \{0, 1, 2, \dots, n\}$

โดย พิจารณาตามเงื่อนไขดังนี้

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

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

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

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

ขั้นที่ 6. เลือกค่าของจุดแบ่ง (c) ที่เหมาะสมที่สุด ซึ่งคือค่า c ที่ทำให้สัดส่วนของความ

ถูกต้องในการจำแนกกลุ่ม มีค่ามากที่สุด โดย $c \in C_0$ และ $c \in [0, 1]$

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

โดย $p(c)$ คือ สัดส่วนของความถูกต้องในการจำแนกกลุ่มที่จุด c
 $N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

3.6 คำนวณค่าเฉลี่ยของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งของแต่ละสถานการณ์

- ค่าเฉลี่ยของจุดแบ่ง (\hat{c})

การหาเฉลี่ยของจุดแบ่งของแต่ละสถานการณ์ จากการทดลองโดยการกระทำซ้ำ 500 รอบ ในแต่ละสถานการณ์ กำหนดให้ \hat{c} เป็นตัวประมาณค่าของค่าพารามิเตอร์ c จะได้ว่า

$$\text{ค่าเฉลี่ยของจุดแบ่ง } \hat{c} = \frac{\sum_{k=1}^N \hat{c}_{(k)}}{N} \quad ; k = 1, 2, \dots, N$$

โดย N คือจำนวนรอบที่กระทำซ้ำในแต่ละสถานการณ์ ($N=500$)

- ค่าร้อยละของจุดแบ่ง (Percent)

ค่าร้อยละของจุดแบ่งของแต่ละสถานการณ์ จากการทดลองโดยการกระทำซ้ำ 500 รอบ ในแต่ละสถานการณ์

$$\text{ค่าร้อยละของจุดแบ่ง } \hat{c} = \frac{\sum_{k=1}^N \hat{c}_{(k)}}{N} \times 100 \quad ; k = 1, 2, \dots, N$$

โดย N คือจำนวนรอบที่กระทำซ้ำในแต่ละสถานการณ์ ($N=500$)

- ช่วงความเชื่อมั่นของจุดแบ่ง (Confidence Interval)

ช่วงความเชื่อมั่นของจุดแบ่งจะบอกถึงค่าต่ำสุดและค่าสูงสุดของค่าของจุดแบ่งที่เป็นไปได้ในแต่ละสถานการณ์

โดยในการวิจัยนี้ กำหนดให้ L คือค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100 (\alpha/2)$ และ U คือค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100 (1-\alpha/2)$ โดยกำหนด ค่า $\alpha = 0.05$ สามารถคำนวณช่วงความเชื่อมั่นของจุดแบ่งของ ดังนี้

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

โดย L คือ คำนวณจากค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 2.5

U คือ คำนวณจากค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ 97.5

3.7 การวิเคราะห์โดยใช้ตัวแบบการถดถอยพหุคูณ

เมื่อกำหนดค่าร้อยละของจุดแบ่งครบทุกสถานการณ์ที่ต้องการศึกษาแล้ว จะใช้ตัวแบบการถดถอยพหุคูณ เพื่อประมาณค่าพารามิเตอร์ สำหรับใช้ในการหาค่าของจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป โดยตัวแบบการถดถอยพหุคูณ คือ

$$\text{Percent} = \theta_0 + \theta_1(p) + \theta_2(a) + \theta_3(n) + \theta_4(M) + \theta_5(ap) + \theta_6(an) + \theta_7(aM) + \theta_8(np) \\ + \theta_9(nM) + \theta_{10}(pM) + \theta_{11}(apn) + \theta_{12}(apM) + \theta_{13}(pnM) + \theta_{14}(apnM) + \varepsilon$$

โดย Percent คือ ค่าร้อยละของจุดแบ่ง

p คือ จำนวนตัวแปรอิสระ

n คือ ขนาดตัวอย่าง

a คือ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ

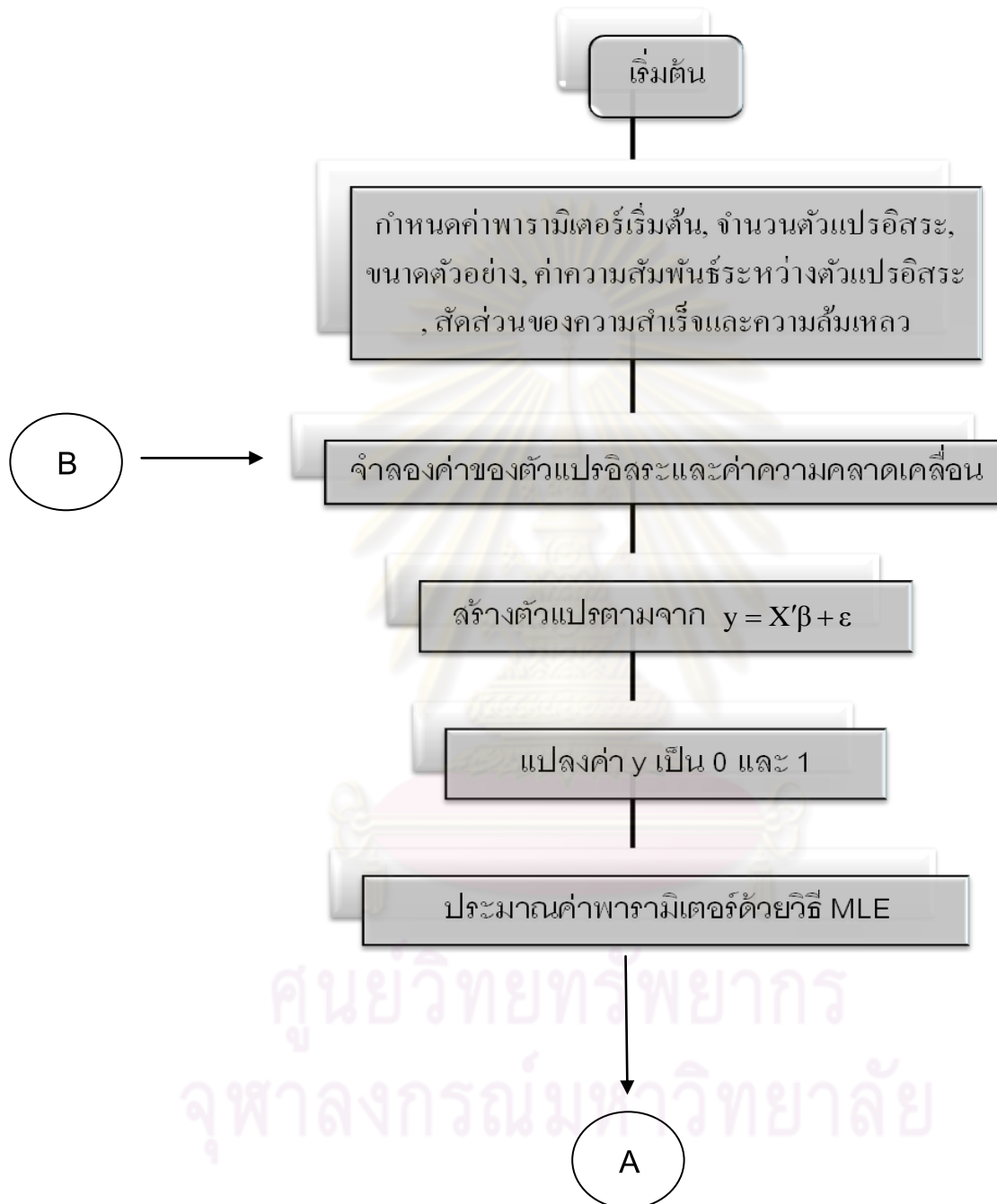
M คือ ระดับความสัมพันธ์ของตัวแปรอิสระ

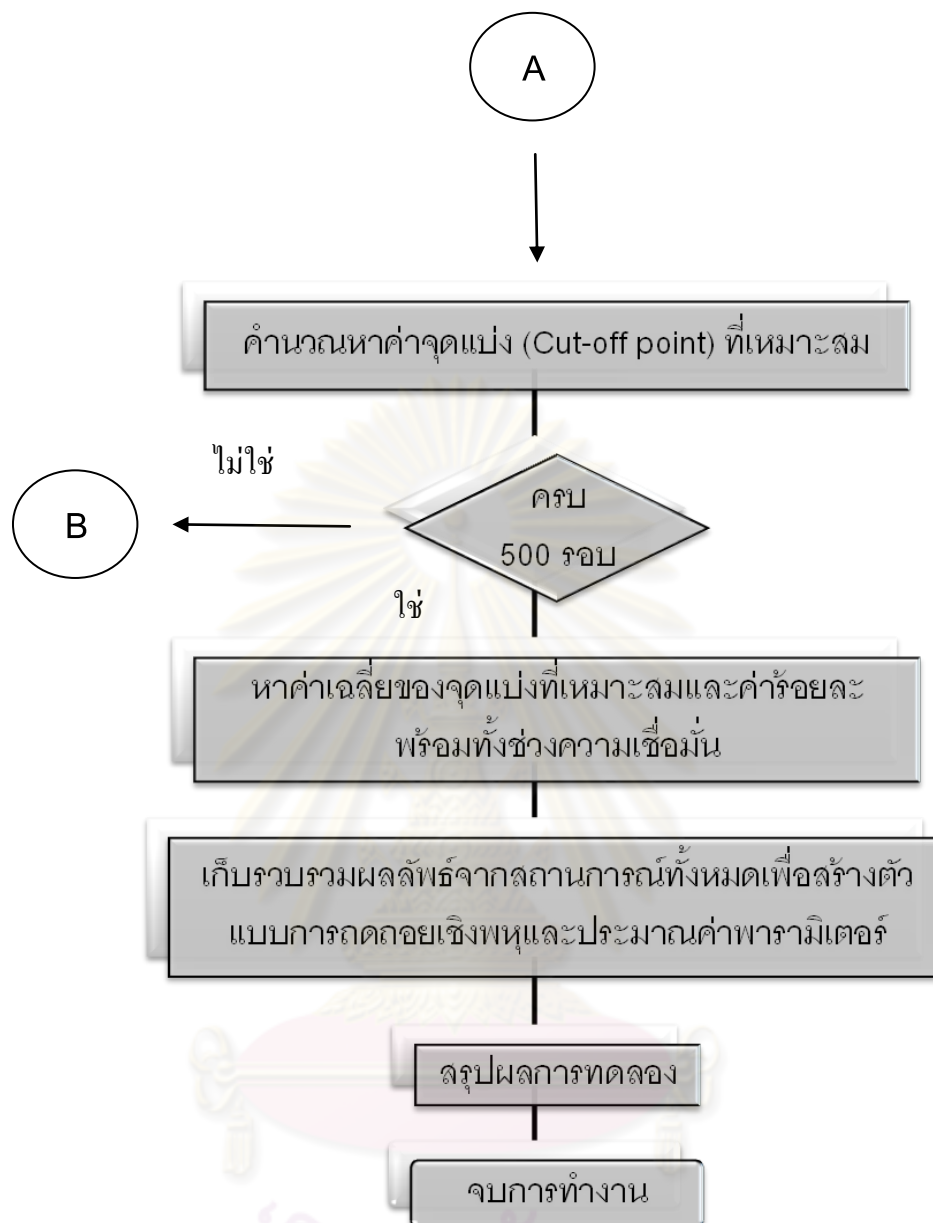
3.8 สรุปผลการวิจัยในแต่ละสถานการณ์

เมื่อทำการหาค่าของจุดแบ่งที่เหมาะสมที่สุดครบทุกสถานการณ์ที่ต้องการศึกษาแล้ว นำผลการทดลองมาสรุปในรูปตาราง เพื่อดูแนวโน้มว่าปัจจัยที่ต้องการศึกษามีผลต่อค่าของจุดแบ่งอย่างไรในแต่ละสถานการณ์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.9 ขั้นตอนการทำงานของโปรแกรม





รูปที่ 3.1 แสดงขั้นตอนการทำงานของโปรแกรม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท เมื่อ ตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอร์ม, จำนวนตัวแปรอิสระเปลี่ยนแปลงไป, ขนาดตัวอย่างเปลี่ยนแปลงไป, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูลเปลี่ยนแปลงไป, ระดับความสัมพันธ์ของตัวแปรอิสระมีค่าสูงขึ้น เมื่อนำปัจจัยต่างๆเหล่านี้มาพิจารณารวมกัน ซึ่งแปรเปลี่ยนไปพร้อมกัน และหาสมการถดถอยพหุคูณสำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่นๆ ต่อไป โดยศึกษาภายใต้สถานการณ์ดังต่อไปนี้

1. ข้อมูลที่กำหนดให้จำนวนตัวแปรอิสระเปลี่ยนแปลงไป เป็น 1, 2, 3, 4, 5 และ 6 ตามลำดับ แต่ขนาดตัวอย่าง, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่
2. ข้อมูลที่กำหนดให้ขนาดตัวอย่างเปลี่ยนแปลงไป เป็น 20, 40, 60, 80, 100 และ 120 ตามลำดับ แต่จำนวนตัวแปรอิสระ, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่
1. ข้อมูลที่กำหนดให้สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ ในแต่ละชุดข้อมูลเปลี่ยนแปลงไป เป็น 0.1, 0.5 และ 0.9 ตามลำดับ แต่จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่
2. ข้อมูลที่กำหนดให้ระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนแปลงไป เป็น ตัวแปรอิสระไม่มีความสัมพันธ์กัน, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่ำ, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับปานกลาง และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับสูง ตามลำดับ แต่จำนวนตัวแปรอิสระขนาดตัวอย่างและสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล คงที่
5. กำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05
6. ในการวิจัยครั้งนี้ทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

โดยการหาค่าของจุดแบ่งที่เหมาะสมในตัวแบบโพรบิต 2 ประเภทในแต่ละสถานการณ์ จะพิจารณาจากค่าของจุดแบ่งที่ให้ค่าอัตราความผิดพลาดในการจำแนกกลุ่มต่ำสุดหรือสัดส่วนความถูกต้องในการจำแนกกลุ่มสูงที่สุดในแต่ละรอบของแต่ละสถานการณ์ที่ต้องการศึกษา และหาค่าเฉลี่ยของจุดแบ่ง ในแต่ละสถานการณ์และนำมาเป็นจุดแบ่งที่เหมาะสมของสถานการณ์นั้นๆ

การวิจัยครั้งนี้จะนำเสนอผลการวิเคราะห์ในแต่ละสถานการณ์ที่ต้องการศึกษา ตามวัตถุประสงค์หลักของการวิจัย 2 ตอนดังนี้

4.1 ค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่งและช่วงความเชื่อมั่นที่ของจุดแบ่งที่เหมาะสมของแต่ละสถานการณ์

4.2 สมการถดถอยพหุคูณ เพื่อใช้ประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่น ๆ ต่อไป

โดยจะนำเสนอผลการวิจัยในรูปแบบตาราง เพื่อความสะดวกในการอธิบาย ผู้วิจัยขอใช้สัญลักษณ์ต่อไปนี้แทนความหมายต่างๆ ดังนี้

\hat{c} หมายถึง ค่าเฉลี่ยของจุดแบ่งที่เหมาะสม

p หมายถึง จำนวนตัวแปรอิสระ

n หมายถึง ขนาดตัวอย่าง

a หมายถึง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ

M หมายถึง ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ

Percent of \hat{c} หมายถึง ค่าร้อยละของจุดแบ่งที่เหมาะสม

CI.Lower หมายถึง ค่าต่ำสุดของช่วงความเชื่อมั่น

CI.Upper หมายถึง ค่าสูงสุดของช่วงความเชื่อมั่น

4.1 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม

การหาค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม โดยใช้ทฤษฎี Hadjicostas P. (2006) เมื่อปัจจัยที่นำมาพิจารณา คือจำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล และระดับความสัมพันธ์ของตัวแปรอิสระ ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จำลองขึ้นมาหลายสถานการณ์ด้วยกันโดยการจำลองสถานการณ์ทั้งหมดใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรม R โดยให้ตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอร์ม จำนวนตัวแปรอิสระแบ่งเป็น 3 ระดับ คือ จำนวนตัวแปรอิสระน้อย ($p=1, 2$) ปานกลาง ($p= 3, 4$) และ มาก ($p= 5,6$) ขนาดตัวอย่าง (n) แบ่งเป็น 3 ระดับ คือ ขนาดตัวอย่างเล็ก ($n=20,40$) ปานกลาง ($n=60,80$) และใหญ่ ($p=100, 120$) สัดส่วนระหว่างของการไม่เกิด

เหตุการณ์ที่สนใจเท่ากับ 0.1, 0.5 และ 0.9 และระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ 0, 0.33, 0.67 และ 0.99 จากการจำลองข้อมูลจะกระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.1. – 4.57 ดังต่อไปนี้

4.1.1 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป เป็น 1, 2, 3, 4, 5 และ 6 ตามลำดับ แต่ขนาดตัวอย่าง, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่ นำเสนอในตารางที่ 4.1-4.12

4.1.2 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป เป็น 20, 40, 60, 80, 100 และ 120 ตามลำดับ แต่จำนวนตัวแปรอิสระ, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่ นำเสนอในตารางที่ 4.13 – 4.24

4.1.3 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูลเปลี่ยนแปลงไป เป็น 0.1, 0.5 และ 0.9 ตามลำดับ แต่จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง และระดับความสัมพันธ์ของตัวแปรอิสระ คงที่ นำเสนอในตารางที่ 4.24 – 4.42

4.1.4 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนแปลงไป เป็น ตัวแปรอิสระไม่มีความสัมพันธ์กัน, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่ำ, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับปานกลาง และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระในระดับสูง ตามลำดับ แต่จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง, และสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูล คงที่ นำเสนอในตารางที่ 4.43 – 4.57

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.1.1 เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป นำเสนอในรูปแบบตารางที่ 4.1-4.12 ดังนี้

ตารางที่ 4.1 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 1, 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0	20	1	0.75237	75.2374	0.3073	0.89671
			2	0.68223	68.2231	0.38386	0.86516
		40	1	0.77917	77.9172	0.54562	0.89561
			2	0.68223	68.2231	0.38386	0.86516
		60	1	0.76083	76.0828	0.52691	0.89186
			2	0.64599	64.5994	0.33165	0.85251
			3	0.55977	55.9769	0.2122	0.7957
		80	1	0.72206	72.2064	0.4789	0.86316
			2	0.58701	58.7014	0.30018	0.7785
			3	0.51334	51.3338	0.21485	0.72966
			4	0.46002	46.0022	0.15307	0.72649
		100	1	0.68557	68.5568	0.49025	0.83014
			2	0.5278	52.7796	0.26477	0.7412
			3	0.4593	45.9296	0.17763	0.70699
			4	0.43834	43.8343	0.1413	0.71265
			5	0.43478	43.4778	0.14117	0.70837
		120	1	0.65278	65.2778	0.49943	0.78195
			2	0.49786	49.7863	0.26269	0.7013
			3	0.44807	44.8075	0.21319	0.72649
			4	0.43788	43.788	0.12485	0.71127
			5	0.42975	42.9753	0.13598	0.70241
			6	0.41997	41.9975	0.06356	0.7417

จากตารางที่ 4.1 เมื่อพิจารณาค่าของจุดแบ่งที่ $a=0.1$, $M=0$ และ $n=20, 40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไปเป็น 1, 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.2 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.33	40	2	0.67018	67.0181	0.27361	0.86004
			3	0.5252	52.5204	0.1351	0.76894
		60	2	0.60942	60.9416	0.30218	0.82894
			3	0.5252	52.5204	0.1351	0.76894
			4	0.43301	43.3013	0.09586	0.75
		80	2	0.54617	54.617	0.2756	0.73909
			3	0.47463	47.4632	0.15051	0.71359
			4	0.43301	43.3013	0.09586	0.75
		100	2	0.51218	51.2179	0.27879	0.71302
			3	0.45812	45.8125	0.18206	0.72049
			4	0.43196	43.1961	0.12907	0.72818
			5	0.41729	41.7291	1.93E-06	0.75141
		120	2	0.47717	47.7167	0.24916	0.69354
			3	0.44994	44.9937	0.16168	0.72472
			4	0.43565	43.5652	0.14489	0.72309
			5	0.41349	41.3487	0.06637	0.72074
			6	0.37392	37.3923	3.02E-08	0.73211

จากตารางที่ 4.2 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.33$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.3 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.67$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.67	40	2	0.65455	65.4549	0.28008	0.85724
			3	0.49272	49.2717	0.14982	0.74319
		60	2	0.60442	60.4422	0.28297	0.80628
			3	0.49272	49.2717	0.14982	0.74319
			4	0.40612	40.6124	2.21E-07	0.73659
		80	2	0.5405	54.0505	0.28724	0.7306
			3	0.44814	44.8144	0.15618	0.72943
			4	0.40612	40.6124	2.21E-07	0.73659
		100	2	0.49677	49.6769	0.26523	0.70925
			3	0.43638	43.6385	0.14844	0.71901
			4	0.41093	41.0934	0.08094	0.72006
			5	0.37829	37.8286	1.51E-08	0.7374
		120	2	0.47141	47.1406	0.21362	0.67901
			3	0.44083	44.0826	0.16325	0.73528
			4	0.41093	41.093	0.1014	0.73073
			5	0.37305	37.3048	2.54E-08	0.73663
			6	0.32607	32.6069	4.88E-09	0.76377

จากตารางที่ 4.3 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.67$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.4 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.99$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.99	40	2	0.64032	64.0317	0.23306	0.86059
			3	0.5028	50.2797	0.19802	0.73019
		60	2	0.60067	60.0674	0.25928	0.79092
			3	0.5028	50.2797	0.19802	0.73019
			4	0.42508	42.5079	0.10247	0.7313
		80	2	0.54926	54.9261	0.30632	0.7406
			3	0.47026	47.0258	0.18257	0.72371
			4	0.42508	42.5079	0.10247	0.7313
		100	2	0.52592	52.5915	0.29777	0.70328
			3	0.45864	45.8637	0.22063	0.69053
			4	0.43188	43.1877	0.14442	0.72998
			5	0.38561	38.5607	4.20E-08	0.72156
		120	2	0.49092	49.0915	0.2934	0.67545
			3	0.44915	44.9146	0.20672	0.68944
			4	0.4222	42.2203	0.15642	0.71711
			5	0.40095	40.0946	3.35E-07	0.70207
			6	0.35028	35.0282	9.03E-09	0.73619

จากตารางที่ 4.4 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.99$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.5 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 1, 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0	20	1	0.46822	46.8221	0.28809	0.61839
		40	1	0.47926	47.9264	0.33424	0.61628
			2	0.47233	47.2327	0.30226	0.64787
		60	1	0.48003	48.0034	0.32142	0.62747
			2	0.4771	47.7102	0.31363	0.63549
			3	0.47547	47.547	0.29924	0.65298
		80	1	0.48188	48.1878	0.32843	0.62314
			2	0.4726	47.2598	0.30776	0.63994
			3	0.46259	46.2591	0.29234	0.63993
			4	0.47273	47.2731	0.27635	0.67147
		100	1	0.47932	47.932	0.32971	0.64431
			2	0.48006	48.0061	0.31361	0.64925
			3	0.47588	47.5879	0.28671	0.65611
			4	0.46896	46.8962	0.28237	0.67624
			5	0.46161	46.1608	0.27983	0.66731
		120	1	0.47699	47.6988	0.31906	0.64881
			2	0.4741	47.4099	0.29529	0.64896
			3	0.47333	47.3333	0.28963	0.66095
			4	0.47057	47.0568	0.27234	0.65595
			5	0.46541	46.5414	0.27372	0.671
			6	0.46217	46.2169	0.24694	0.65623

จากตารางที่ 4.5 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0$ และ $n=20, 40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 1, 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.6 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.33	40	2	0.47138	47.1384	0.28741	0.65789
			3	0.46983	46.9831	0.2884	0.64878
		60	2	0.46337	46.3365	0.27359	0.66617
			3	0.46975	46.9745	0.299	0.65329
		80	2	0.4615	46.1495	0.25548	0.66171
			3	0.46814	46.8144	0.26142	0.66893
			4	0.47767	47.7666	0.3093	0.67435
		100	2	0.46452	46.4517	0.2744	0.67215
			3	0.46129	46.1286	0.25883	0.68498
			4	0.45621	45.621	0.25078	0.69079
			5	0.47333	47.3329	0.29284	0.66456
		120	2	0.46709	46.7086	0.26704	0.6597
			3	0.46232	46.2317	0.25888	0.67532
			4	0.46355	46.355	0.26699	0.70299
			5	0.46271	46.2708	0.25382	0.71527
			6				

จากตารางที่ 4.6 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.33$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.7 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.67$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.67	40	2	0.46397	46.3965	0.27176	0.65041
			3	0.46602	46.6019	0.2716	0.667
		60	2	0.46602	46.6019	0.2716	0.667
			3	0.45526	45.526	0.25517	0.67837
			4	0.47027	47.0267	0.27515	0.65225
		80	2	0.47027	47.0267	0.27515	0.65225
			3	0.46657	46.6574	0.25612	0.70047
			4	0.45859	45.8593	0.21739	0.69528
		100	2	0.4697	46.9695	0.26749	0.68788
			3	0.45583	45.5832	0.23351	0.68222
			4	0.46422	46.4224	0.24617	0.70388
			5	0.44936	44.9355	0.20872	0.69572
		120	2	0.47146	47.1455	0.28081	0.67865
			3	0.4659	46.59	0.2443	0.67324
			4	0.46009	46.0085	0.24019	0.6991
			5	0.45972	45.9721	0.22517	0.70778
			6	0.43803	43.8034	0.19376	0.69501

จากตารางที่ 4.7 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.67$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.8 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.99$, $n= 40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.99	40	2	0.46419	46.4193	0.2907	0.67189
			3	0.46436	46.4364	0.27972	0.67675
		60	2	0.45981	45.9809	0.22786	0.69569
			3	0.45426	45.4262	0.2536	0.66528
		80	2	0.45295	45.2953	0.20883	0.70894
			3	0.43344	43.3435	0.15069	0.71533
			4	0.45806	45.8055	0.25328	0.66288
		100	2	0.4544	45.4402	0.22473	0.71321
			3	0.44332	44.332	0.19475	0.72846
			4	0.40833	40.8332	0.15015	0.71905
			5	0.46609	46.6087	0.27335	0.68624
		120	2	0.46018	46.0183	0.22043	0.70256
			3	0.4417	44.1703	0.18798	0.70711
			4	0.42575	42.5751	0.09891	0.73452
			5	0.36773	36.7727	9.58E-09	0.71746
			6				

จากตารางที่ 4.8 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.99$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.9 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 1, 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0	20	1	0.20721	20.7212	0.10035	0.42655
		40	1	0.20818	20.818	0.10387	0.40344
			2	0.27727	27.7268	0.12625	0.52265
		60	1	0.22479	22.4793	0.10875	0.38986
			2	0.31754	31.754	0.15127	0.54438
			3	0.35961	35.961	0.18128	0.59866
		80	1	0.26423	26.4232	0.11717	0.45023
			2	0.36052	36.0516	0.20651	0.58619
			3	0.39379	39.3788	0.22074	0.62339
			4	0.40506	40.5057	0.18064	0.63361
		100	1	0.29848	29.8483	0.16638	0.45264
			2	0.40403	40.4034	0.23545	0.61522
			3	0.41641	41.6408	0.21114	0.65083
			4	0.41412	41.4121	0.19606	0.65972
			5	0.4148	41.4798	0.16203	0.67396
		120	1	0.33518	33.518	0.21168	0.48828
			2	0.41938	41.9375	0.23791	0.61582
			3	0.43794	43.7944	0.21962	0.69881
			4	0.43084	43.0837	0.21707	0.68736
			5	0.42674	42.6743	0.18193	0.66697
			6	0.42047	42.0469	0.14876	0.68296

จากตารางที่ 4.9 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0$ และ $n=20, 40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 1, 2, 3, 4, 5, 6 พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้นเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.10 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.33	40	2	0.28812	28.8119	0.12943	0.5327
			3	0.32304	32.3042	0.16313	0.53465
		60	2	0.36417	36.4165	0.17854	0.59963
			3	0.37216	37.2157	0.2133	0.58456
		80	2	0.41612	41.6124	0.21177	0.6522
			3	0.41193	41.1925	0.1797	0.66963
			4	0.40943	40.9425	0.24483	0.62505
		100	2	0.42704	42.7044	0.20399	0.66039
			3	0.4189	41.8897	0.15653	0.67981
			4	0.39466	39.4664	3.09E-08	0.68351
			5	0.43912	43.9124	0.25432	0.65165
		120	2	0.42929	42.9294	0.21403	0.67132
			3	0.42061	42.0607	0.15375	0.69642
			4	0.40211	40.2111	0.11899	0.70151
			5	0.38475	38.4751	1.14E-08	0.71555
			6				

จากตารางที่ 4.10 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.33$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าที่ $n=60, 80$ ค่าของจุดแบ่งมี

แนวโน้มเพิ่มขึ้น เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนที่ $n=100, 120$ ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.11 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.67$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.67	40	2	0.28925	28.925	0.12971	0.54021
			3	0.3436	34.3597	0.1905	0.56929
		60	2	0.38339	38.3391	0.16709	0.60909
			3	0.39376	39.3761	0.22817	0.59437
		80	2	0.42013	42.0127	0.18648	0.66234
			3	0.39994	39.9944	0.13297	0.66285
			4	0.4243	42.4299	0.24703	0.62776
		100	2	0.42285	42.2854	0.17225	0.65572
			3	0.4028	40.2804	0.09981	0.69822
			4	0.38456	38.4564	8.98E-09	0.70799
			5	0.44103	44.1031	0.24428	0.64052
		120	2	0.42551	42.5512	0.18479	0.67987
			3	0.40276	40.2761	0.14374	0.67524
			4	0.38981	38.981	6.29E-08	0.69617
			5	0.31568	31.5678	3.73E-09	0.69443
			6				

จากตารางที่ 4.11 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.67$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าที่ $n=60, 80$ ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนที่ $n=100, 120$ ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

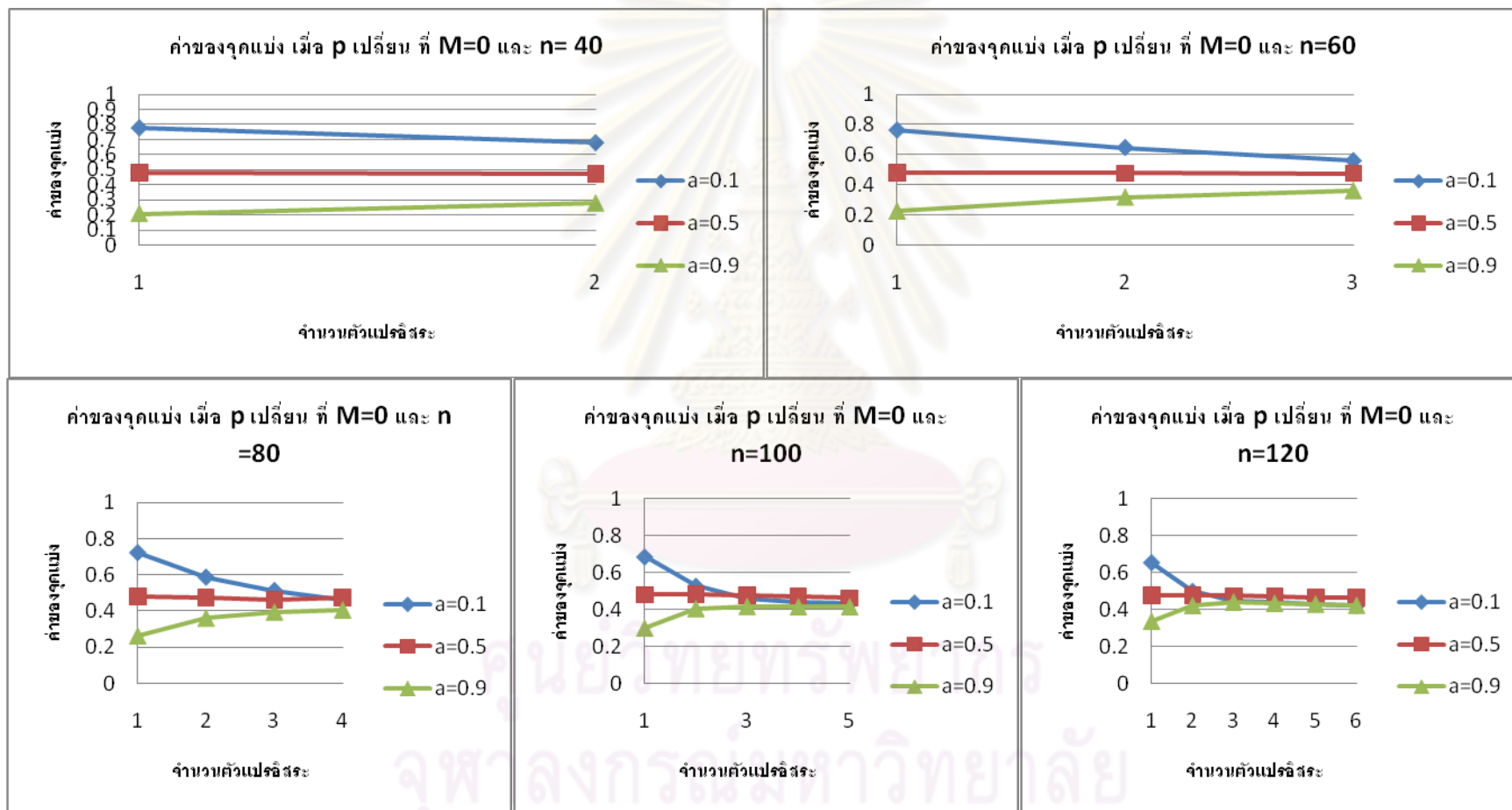
ตารางที่ 4.12 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.99$, $n=40, 60, 80, 100, 120$ และ p เปลี่ยนแปลงไปเป็น 1, 2, 3, 4, 5, 6 ตามลำดับ

a	M	n	p	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.99	40	2	0.2968	29.6804	0.13289	0.54229
			3	0.35092	35.0922	0.1819	0.56456
		60	2	0.39449	39.4486	0.17745	0.61879
			3	0.38854	38.8537	0.22029	0.5746
			4	0.41766	41.766	0.19765	0.63321
		80	2	0.41054	41.054	8.77E-07	0.68588
			3	0.41632	41.6321	0.26044	0.60219
			4	0.42307	42.307	0.18805	0.6397
		100	2	0.42009	42.0094	0.14641	0.68313
			3	0.40809	40.8089	6.83E-08	0.6834
			4	0.43298	43.2982	0.24341	0.62163
			5	0.43774	43.774	0.19636	0.67311
		120	2	0.41959	41.9588	0.15893	0.69205
			3	0.39278	39.2782	1.06E-07	0.69536
			4	0.33698	33.698	8.99E-09	0.6791
			5				
			6				

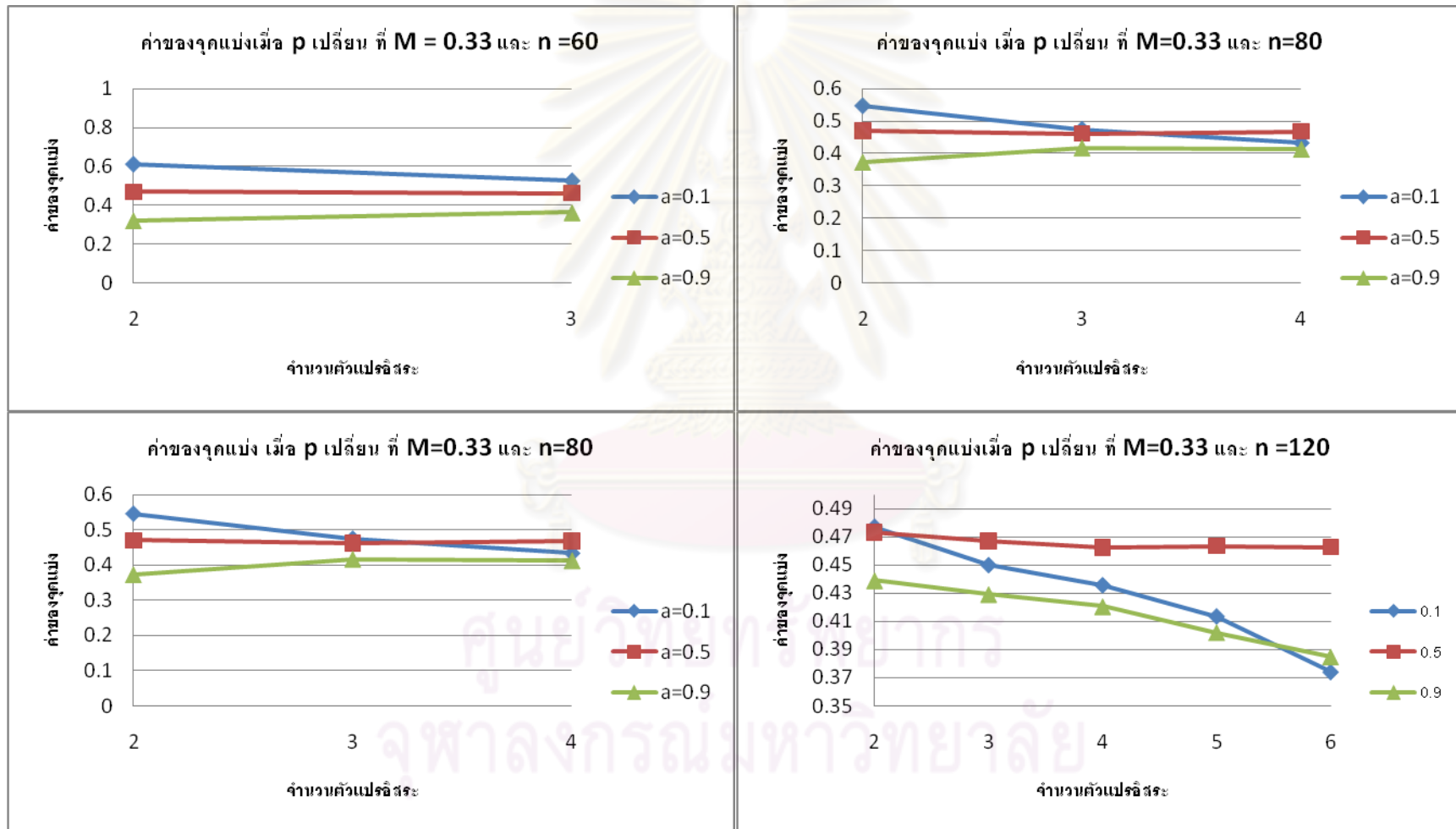
จากตารางที่ 4.12 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.99$ และ $n=40, 60, 80, 100, 120$ แต่จำนวนตัวแปรอิสระเปลี่ยนแปลงไป 2, 3, 4, 5, 6 พบว่าที่ $n=60, 80$ ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนที่ $n=100, 120$ ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

จากตารางที่ 4.1- 4.12 สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ ได้ดังรูปที่ 4.1-4.4 ต่อไปนี้

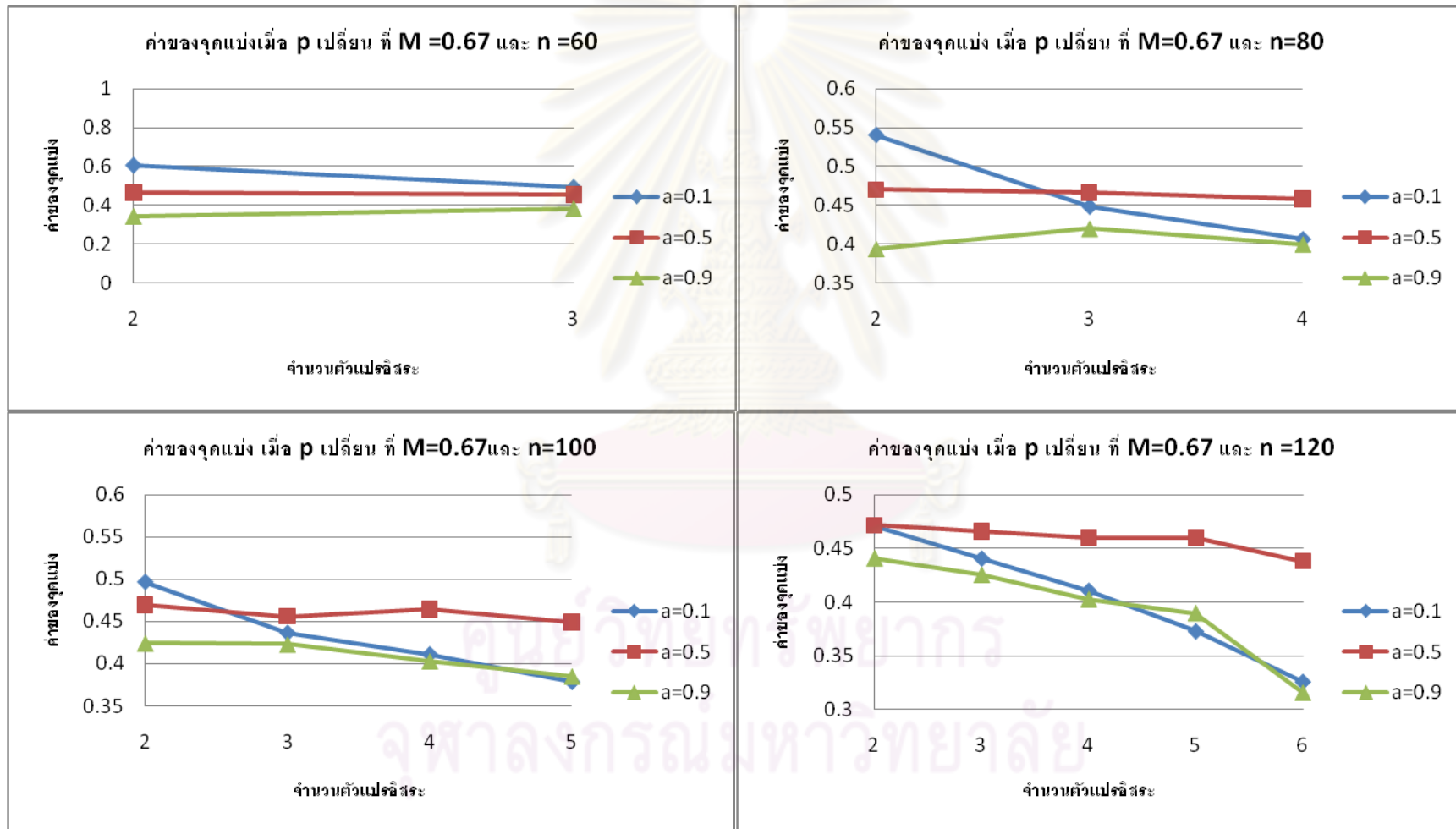
รูปที่ 4.1 แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไปที่ $M = 0$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 40, 60, 80, 100, 120$



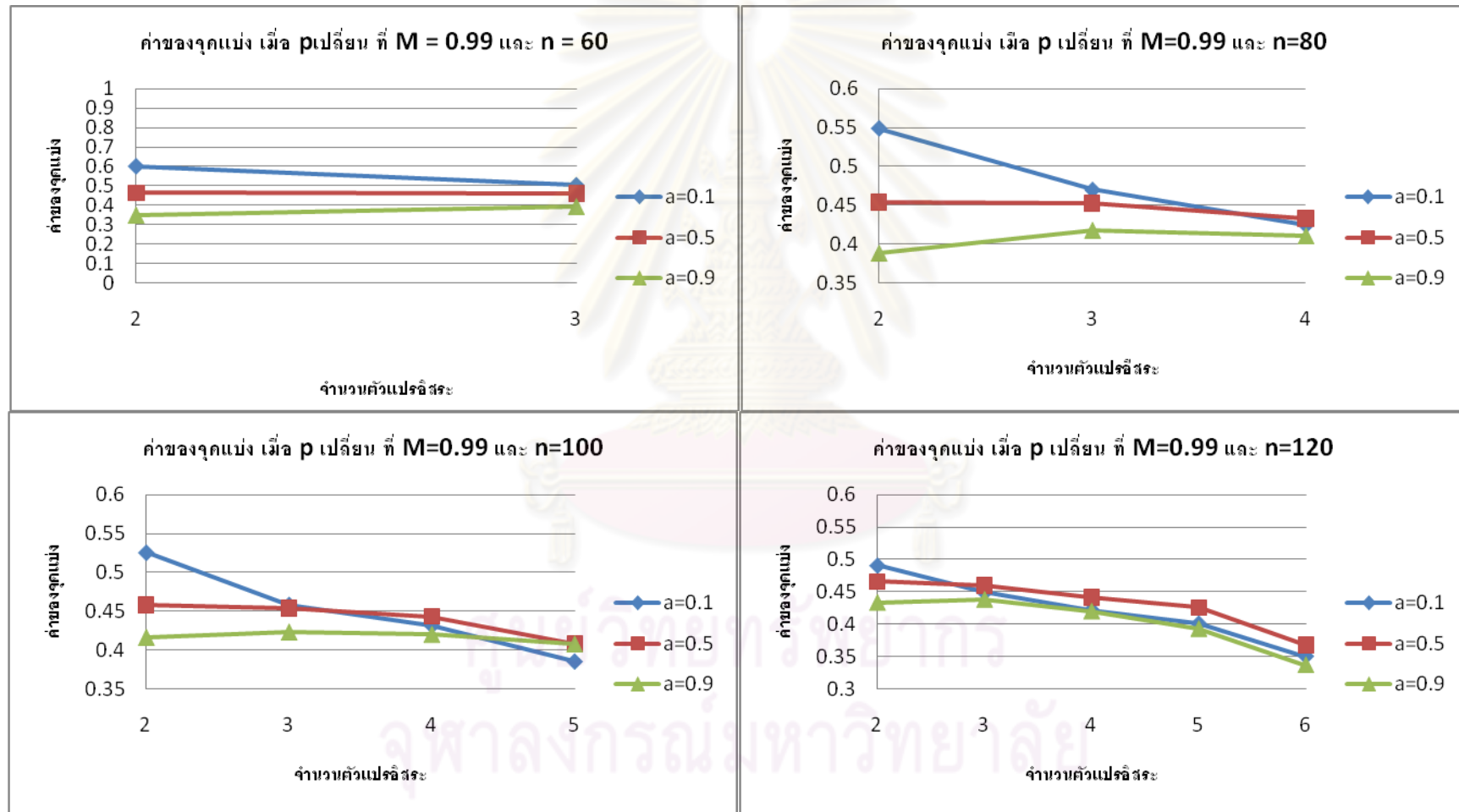
รูปที่ 4.2 แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไปที่ $M = 0.33$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 60, 80, 100, 120$



รูปที่ 4.3 แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0.67$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 60, 80, 100, 120$



รูปที่ 4.4 แสดงค่าของจุดแบ่ง เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ $M = 0.99$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 60, 80, 100, 120$



4.1.2 เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป นำเสนอในรูปแบบตารางที่ 4.13 - 4.24 ดังนี้

ตารางที่ 4.13 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0	1	20	0.75237	75.2374	0.3073	0.89671
			40	0.77917	77.9172	0.54562	0.89561
			60	0.76083	76.0828	0.52691	0.89186
			80	0.72206	72.2064	0.4789	0.86316
			100	0.68557	68.5568	0.49025	0.83014
			120	0.65278	65.2778	0.49943	0.78195
		2	40	0.68223	68.2231	0.38386	0.86516
			60	0.64599	64.5994	0.33165	0.85251
			80	0.58701	58.7014	0.30018	0.7785
			100	0.5278	52.7796	0.26477	0.7412
			120	0.49786	49.7863	0.26269	0.7013
		3	60	0.55977	55.9769	0.2122	0.7957
			80	0.51334	51.3338	0.21485	0.72966
			100	0.4593	45.9296	0.17763	0.70699
			120	0.44807	44.8075	0.21319	0.72649
		4	80	0.46002	46.0022	0.15307	0.72649
			100	0.43834	43.8343	0.1413	0.71265
			120	0.43788	43.788	0.12485	0.71127
		5	100	0.43478	43.4778	0.14117	0.70837
			120	0.42975	42.9753	0.13598	0.70241
		6	120	0.41997	41.9975	0.06356	0.7417

จากตารางที่ 4.13 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.14 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.33	2	40	0.67018	67.0181	0.27361	0.86004
			60	0.60942	60.9416	0.30218	0.82894
			80	0.54617	54.617	0.2756	0.73909
			100	0.51218	51.2179	0.27879	0.71302
			120	0.47717	47.7167	0.24916	0.69354
		3	60	0.5252	52.5204	0.1351	0.76894
			80	0.47463	47.4632	0.15051	0.71359
			100	0.45812	45.8125	0.18206	0.72049
			120	0.44994	44.9937	0.16168	0.72472
		4	80	0.43301	43.3013	0.09586	0.75
			100	0.43196	43.1961	0.12907	0.72818
			120	0.43565	43.5652	0.14489	0.72309
		5	100	0.41729	41.7291	1.93E-06	0.75141
			120	0.41349	41.3487	0.06637	0.72074
6	120	0.37392	37.3923	3.02E-08	0.73211		

จากตารางที่ 4.14 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.15 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.67	2	40	0.65455	65.4549	0.28008	0.85724
			60	0.60442	60.4422	0.28297	0.80628
			80	0.5405	54.0505	0.28724	0.7306
			100	0.49677	49.6769	0.26523	0.70925
			120	0.47141	47.1406	0.21362	0.67901
		3	60	0.49272	49.2717	0.14982	0.74319
			80	0.44814	44.8144	0.15618	0.72943
			100	0.43638	43.6385	0.14844	0.71901
			120	0.44083	44.0826	0.16325	0.73528
		4	80	0.40612	40.6124	2.21E-07	0.73659
			100	0.41093	41.0934	0.08094	0.72006
			120	0.41093	41.093	0.1014	0.73073
		5	100	0.37829	37.8286	1.51E-08	0.7374
			120	0.37305	37.3048	2.54E-08	0.73663
		6	120	0.32607	32.6069	4.88E-09	0.76377

จากตารางที่ 4.15 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.16 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	0.99	2	40	0.64032	64.0317	0.23306	0.86059
			60	0.60067	60.0674	0.25928	0.79092
			80	0.54926	54.9261	0.30632	0.7406
			100	0.52592	52.5915	0.29777	0.70328
			120	0.49092	49.0915	0.2934	0.67545
		3	60	0.5028	50.2797	0.19802	0.73019
			80	0.47026	47.0258	0.18257	0.72371
			100	0.45864	45.8637	0.22063	0.69053
			120	0.44915	44.9146	0.20672	0.68944
		4	80	0.42508	42.5079	0.10247	0.7313
			100	0.43188	43.1877	0.14442	0.72998
			120	0.4222	42.2203	0.15642	0.71711
		5	100	0.38561	38.5607	4.20E-08	0.72156
			120	0.40095	40.0946	3.35E-07	0.70207
		6	120	0.35028	35.0282	9.03E-09	0.73619

จากตารางที่ 4.16 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.17 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0	1	20	0.46822	46.8221	0.28809	0.61839
			40	0.47926	47.9264	0.33424	0.61628
			60	0.48003	48.0034	0.32142	0.62747
			80	0.48188	48.1878	0.32843	0.62314
			100	0.47932	47.932	0.32971	0.64431
			120	0.47699	47.6988	0.31906	0.64881
		2	40	0.47233	47.2327	0.30226	0.64787
			60	0.4771	47.7102	0.31363	0.63549
			80	0.4726	47.2598	0.30776	0.63994
			100	0.48006	48.0061	0.31361	0.64925
			120	0.4741	47.4099	0.29529	0.64896
		3	60	0.47547	47.547	0.29924	0.65298
			80	0.46259	46.2591	0.29234	0.63993
			100	0.47588	47.5879	0.28671	0.65611
			120	0.47333	47.3333	0.28963	0.66095
		4	80	0.47273	47.2731	0.27635	0.67147
			100	0.46896	46.8962	0.28237	0.67624
			120	0.47057	47.0568	0.27234	0.65595
		5	100	0.46161	46.1608	0.27983	0.66731
			120	0.46541	46.5414	0.27372	0.671
		6	120	0.46217	46.2169	0.24694	0.65623

จากตารางที่ 4.17 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีค่าไม่แตกต่างกันมากนัก เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.18 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.33	2	40	0.47138	47.1384	0.28741	0.65789
			60	0.46983	46.9831	0.2884	0.64878
			80	0.46975	46.9745	0.299	0.65329
			100	0.47767	47.7666	0.3093	0.67435
			120	0.47333	47.3329	0.29284	0.66456
		3	60	0.46336	46.3365	0.27359	0.66617
			80	0.46149	46.1495	0.25548	0.66171
			100	0.46452	46.4517	0.2744	0.67215
			120	0.46709	46.7086	0.26704	0.6597
		4	80	0.46814	46.8144	0.26142	0.66893
			100	0.46129	46.1286	0.25883	0.68498
			120	0.46232	46.2317	0.25888	0.67532
		5	100	0.45621	45.621	0.25078	0.69079
			120	0.46355	46.355	0.26699	0.70299
6	120	0.46271	46.2708	0.25382	0.71527		

จากตารางที่ 4.18 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีค่าไม่แตกต่างกันมากนัก เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.19 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.67	2	40	0.46397	46.3965	0.271761147	0.65041
			60	0.46602	46.6019	0.271596983	0.667
			80	0.47027	47.0267	0.275146912	0.65225
			100	0.46969	46.9695	0.267490668	0.68788
			120	0.47146	47.1455	0.28080958	0.67865
		3	60	0.45526	45.526	0.255168353	0.67837
			80	0.46657	46.6574	0.256123107	0.70047
			100	0.45583	45.5832	0.233513744	0.68222
			120	0.4659	46.59	0.244299251	0.67324
		4	80	0.45859	45.8593	0.217386924	0.69528
			100	0.46422	46.4224	0.246172556	0.70388
			120	0.46008	46.0085	0.240187818	0.6991
		5	100	0.44935	44.9355	0.20871633	0.69572
			120	0.45972	45.9721	0.225172846	0.70778
		6	120	0.43803	43.8034	0.19376	0.695010369

จากตารางที่ 4.19 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีค่าไม่แตกต่างกันมากนัก เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.20 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	0.99	2	40	0.46419	46.4193	0.2907	0.67189
			60	0.46436	46.4364	0.27972	0.67675
			80	0.45426	45.4262	0.2536	0.66528
			100	0.45806	45.8055	0.25328	0.66288
			120	0.46609	46.6087	0.27335	0.68624
		3	60	0.45981	45.9809	0.22786	0.69569
			80	0.45295	45.2953	0.20883	0.70894
			100	0.4544	45.4402	0.22473	0.71321
			120	0.46018	46.0183	0.22043	0.70256
		4	80	0.43343	43.3435	0.15069	0.71533
			100	0.44332	44.332	0.19475	0.72846
			120	0.4417	44.1703	0.18798	0.70711
		5	100	0.40833	40.8332	0.15015	0.71905
			120	0.42575	42.5751	0.09891	0.73452
		6	120	0.36773	36.7727	9.58E-09	0.71746

จากตารางที่ 4.20 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีค่าไม่แตกต่างกันมากนัก เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.21 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0	1	20	0.20721	20.7212	0.10035	0.42655
			40	0.20818	20.818	0.10387	0.40344
			60	0.22479	22.4793	0.10875	0.38986
			80	0.26423	26.4232	0.11717	0.45023
			100	0.29848	29.8483	0.16638	0.45264
			120	0.33518	33.518	0.21168	0.48828
		2	40	0.27727	27.7268	0.12625	0.52265
			60	0.31754	31.754	0.15127	0.54438
			80	0.36052	36.0516	0.20651	0.58619
			100	0.40403	40.4034	0.23545	0.61522
			120	0.41937	41.9375	0.23791	0.61582
		3	60	0.35961	35.961	0.18128	0.59866
			80	0.39379	39.3788	0.22074	0.62339
			100	0.41641	41.6408	0.21114	0.65083
			120	0.43794	43.7944	0.21962	0.69881
		4	80	0.40506	40.5057	0.18064	0.63361
			100	0.41412	41.4121	0.19606	0.65972
			120	0.43084	43.0837	0.21707	0.68736
		5	100	0.4148	41.4798	0.16203	0.67396
			120	0.42674	42.6743	0.18193	0.66697
		6	120	0.42047	42.0469	0.14876	0.68296

จากตารางที่ 4.21 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0$, $p=1, 2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 20, 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.22 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.33	2	40	0.28812	28.8119	0.12943	0.5327
			60	0.32304	32.3042	0.16313	0.53465
			80	0.37216	37.2157	0.2133	0.58456
			100	0.40942	40.9425	0.24483	0.62505
			120	0.43912	43.9124	0.25432	0.65165
		3	60	0.36416	36.4165	0.17854	0.59963
			80	0.41612	41.6124	0.21177	0.6522
			100	0.42704	42.7044	0.20399	0.66039
			120	0.42929	42.9294	0.21403	0.67132
		4	80	0.41193	41.1925	0.1797	0.66963
			100	0.4189	41.8897	0.15653	0.67981
			120	0.42061	42.0607	0.15375	0.69642
		5	100	0.39466	39.4664	3.09E-08	0.68351
			120	0.40211	40.2111	0.11899	0.70151
6	120	0.38475	38.4751	1.14E-08	0.71555		

จากตารางที่ 4.22 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.33$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 4.23 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.67	2	40	0.28925	28.925	0.12971	0.54021
			60	0.3436	34.3597	0.1905	0.56929
			80	0.39376	39.3761	0.22817	0.59437
			100	0.4243	42.4299	0.24703	0.62776
			120	0.44103	44.1031	0.24428	0.64052
		3	60	0.38339	38.3391	0.16709	0.60909
			80	0.42013	42.0127	0.18648	0.66234
			100	0.42285	42.2854	0.17225	0.65572
			120	0.42551	42.5512	0.18479	0.67987
		4	80	0.39994	39.9944	0.13297	0.66285
			100	0.4028	40.2804	0.09981	0.69822
			120	0.40276	40.2761	0.14374	0.67524
		5	100	0.38456	38.4564	8.98E-09	0.70799
			120	0.38981	38.981	6.29E-08	0.69617
		6	120	0.31568	31.5678	3.73E-09	0.69443

จากตารางที่ 4.23 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.67$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น

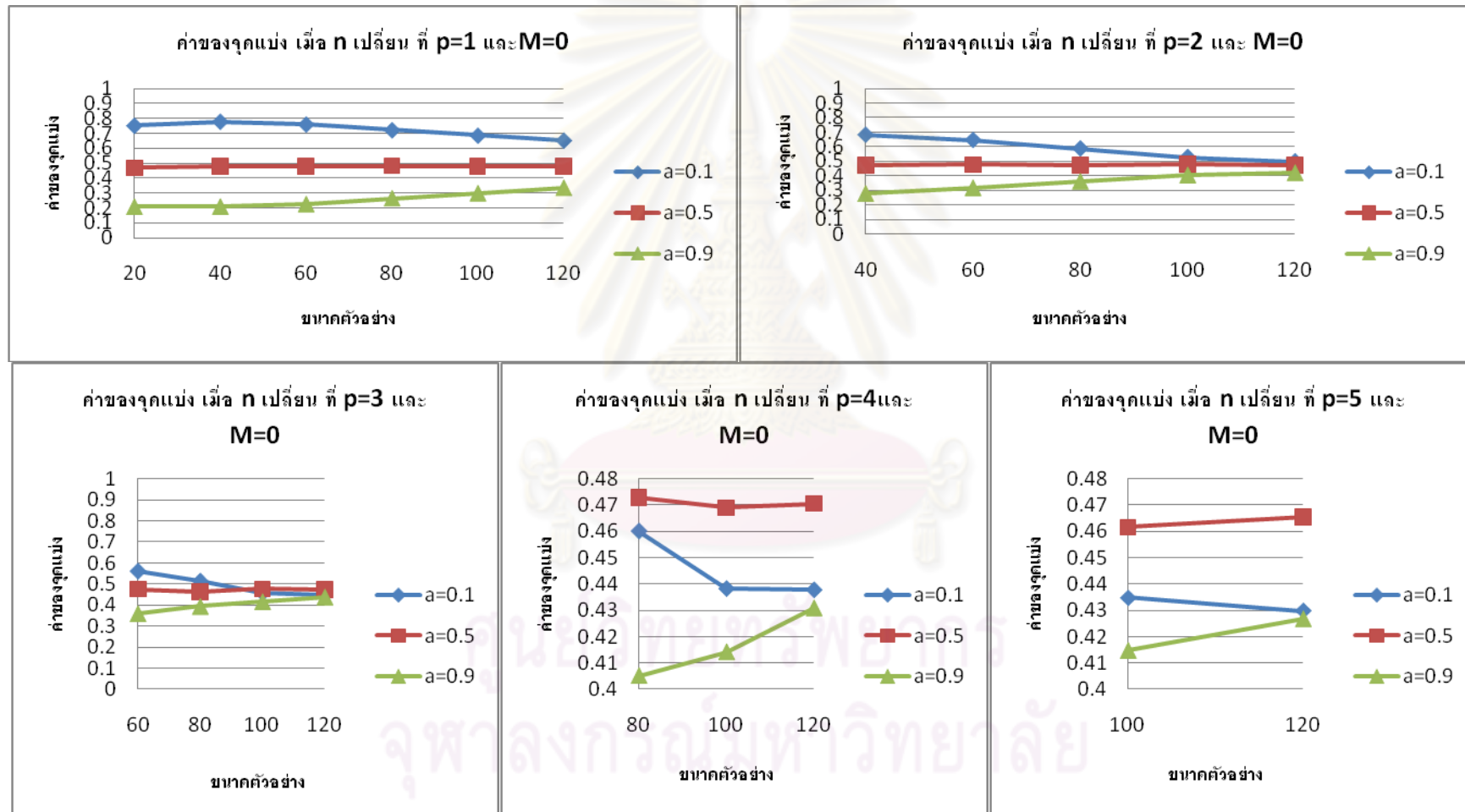
ตารางที่ 4.24 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 ตามลำดับ

a	M	p	n	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	0.99	2	40	0.2968	29.6804	0.13289	0.54229
			60	0.35092	35.0922	0.1819	0.56456
			80	0.38854	38.8537	0.22029	0.5746
			100	0.41632	41.6321	0.26044	0.60219
			120	0.43298	43.2982	0.24341	0.62163
		3	60	0.39449	39.4486	0.17745	0.61879
			80	0.41766	41.766	0.19765	0.63321
			100	0.42307	42.307	0.18805	0.6397
			120	0.43774	43.774	0.19636	0.67311
		4	80	0.41054	41.054	8.77E-07	0.68588
			100	0.42009	42.0094	0.14641	0.68313
			120	0.41959	41.9588	0.15893	0.69205
		5	100	0.40809	40.8089	6.83E-08	0.6834
			120	0.39278	39.2782	1.06E-07	0.69536
		6	120	0.33698	33.698	8.99E-09	0.6791

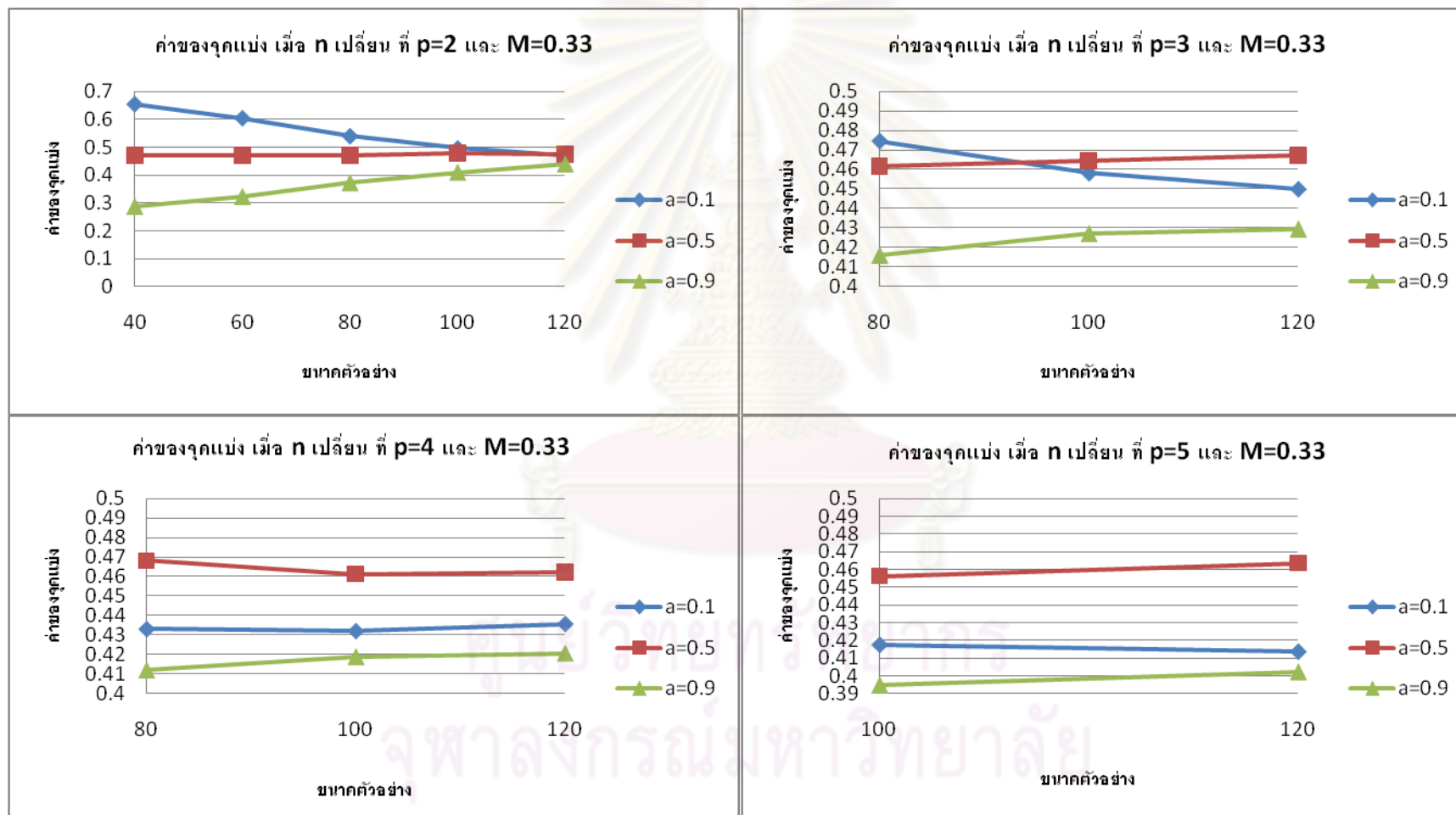
จากตารางที่ 4.24 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $M=0.99$, $p=2, 3, 4, 5, 6$ และ n เปลี่ยนแปลงไปเป็น 40, 60, 80, 100, 120 พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น

จากตารางที่ 4.13- 4.24 สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ ได้ดังรูปที่ 4.5-4.8 ต่อไปนี้

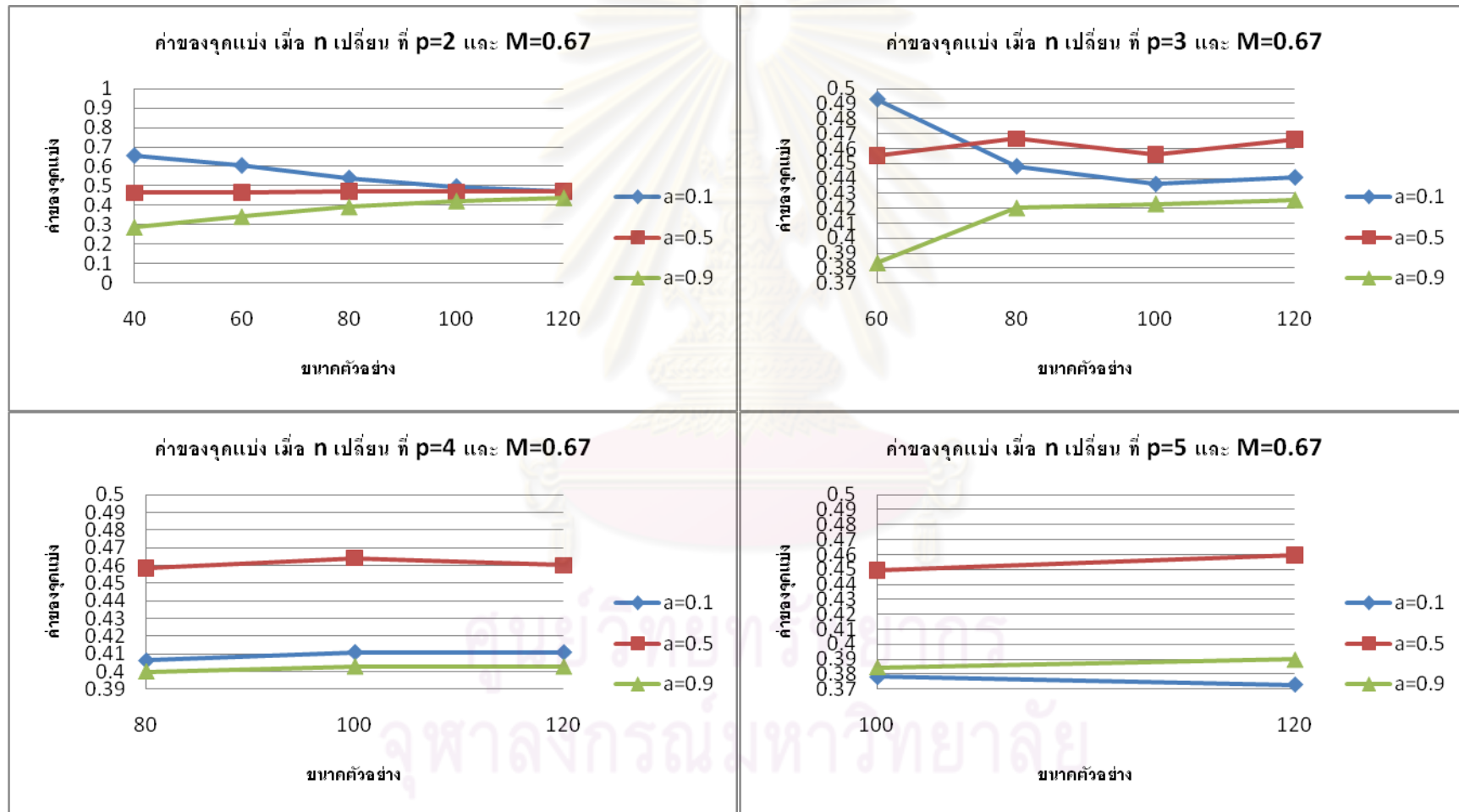
รูปที่ 4.5 แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไปที่ $M = 0$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 1, 2, 3, 4$ และ 5



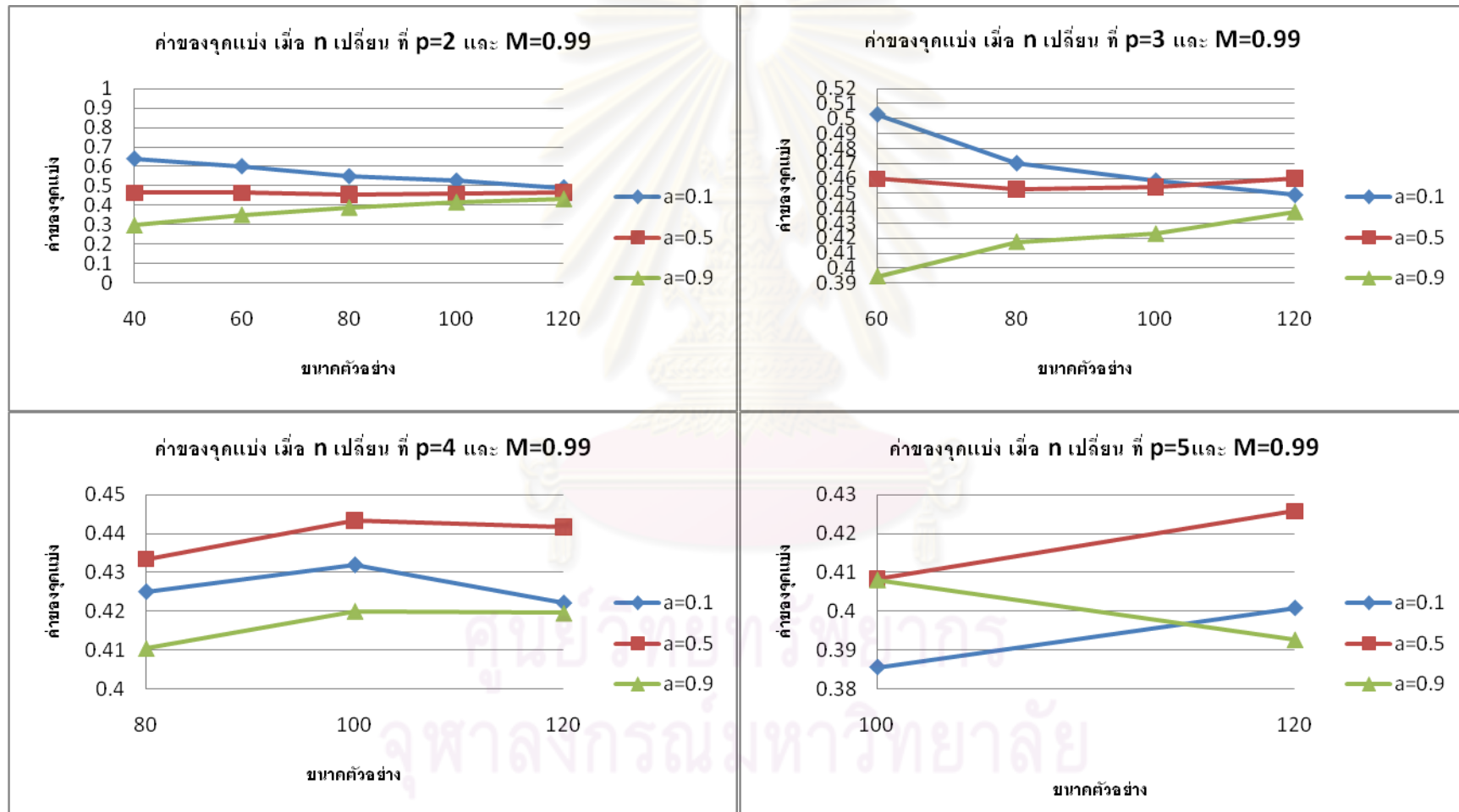
รูปที่ 4.6 แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.33$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4$ และ 5



รูปที่ 4.7 แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.67$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4$ และ 5



รูปที่ 4.8 แสดงค่าของจุดแบ่ง เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป ที่ $M = 0.99$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $p = 2, 3, 4$ และ 5



4.1.3 เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป นำเสนอในรูปแบบตารางที่ 4.25 - 4.42 ดังนี้

ตารางที่ 4.25 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=1$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
1	0	20	0.1	0.75237	75.2374	0.3073	0.89671
			0.5	0.46822	46.8221	0.28809	0.61839
			0.9	0.20721	20.7212	0.10035	0.42655
		40	0.1	0.77917	77.9172	0.54562	0.89561
			0.5	0.47926	47.9264	0.33424	0.61628
			0.9	0.20818	20.818	0.10387	0.40344
		60	0.1	0.76083	76.0828	0.52691	0.89186
			0.5	0.48003	48.0034	0.32142	0.62747
			0.9	0.22479	22.4793	0.10875	0.38986
		80	0.1	0.72206	72.2064	0.4789	0.86316
			0.5	0.48188	48.1878	0.32843	0.62314
			0.9	0.26423	26.4232	0.11717	0.45023
		100	0.1	0.68557	68.5568	0.49025	0.83014
			0.5	0.47932	47.932	0.32971	0.64431
			0.9	0.29848	29.8483	0.16638	0.45264
		120	0.1	0.65278	65.2778	0.49943	0.78195
			0.5	0.47699	47.6988	0.31906	0.64881
			0.9	0.33518	33.518	0.21168	0.48828

จากตารางที่ 4.25 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=1$, $M=0$, $n=20, 40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น $0.1, 0.5, 0.9$ พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.26 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=2$, $M=0$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น $0.1, 0.5, 0.9$ ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
2	0	40	0.1	0.68223	68.2231	0.38386	0.86516
			0.5	0.47233	47.2327	0.30226	0.64787
			0.9	0.27727	27.7268	0.12625	0.52265
		60	0.1	0.64599	64.5994	0.33165	0.85251
			0.5	0.4771	47.7102	0.31363	0.63549
			0.9	0.31754	31.754	0.15127	0.54438
		80	0.1	0.58701	58.7014	0.30018	0.7785
			0.5	0.4726	47.2598	0.30776	0.63994
			0.9	0.36052	36.0516	0.20651	0.58619
		100	0.1	0.5278	52.7796	0.26477	0.7412
			0.5	0.48006	48.0061	0.31361	0.64925
			0.9	0.40403	40.4034	0.23545	0.61522
120	0.1	0.49786	49.7863	0.26269	0.7013		
	0.5	0.4741	47.4099	0.29529	0.64896		
	0.9	0.41937	41.9375	0.23791	0.61582		

จากตารางที่ 4.26 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=2$, $M=0$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น $0.1, 0.5, 0.9$ พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.27 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=2$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
2	0.33	40	0.1	0.67018	67.0181	0.27361	0.86004
			0.5	0.47138	47.1384	0.28741	0.65789
			0.9	0.28812	28.8119	0.12943	0.5327
		60	0.1	0.60942	60.9416	0.30218	0.82894
			0.5	0.46983	46.9831	0.2884	0.64878
			0.9	0.32304	32.3042	0.16313	0.53465
		80	0.1	0.54617	54.617	0.2756	0.73909
			0.5	0.46975	46.9745	0.299	0.65329
			0.9	0.37216	37.2157	0.2133	0.58456
		100	0.1	0.51218	51.2179	0.27879	0.71302
			0.5	0.47767	47.7666	0.3093	0.67435
			0.9	0.40942	40.9425	0.24483	0.62505
		120	0.1	0.47717	47.7167	0.24916	0.69354
			0.5	0.47333	47.3329	0.29284	0.66456
			0.9	0.43912	43.9124	0.25432	0.65165

จากตารางที่ 4.27 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=2$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.28 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=1$, $M=0.67$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
2	0.67	40	0.1	0.65455	65.4549	0.28008	0.85724
			0.5	0.46397	46.3965	0.27176	0.65041
			0.9	0.28925	28.925	0.12971	0.54021
		60	0.1	0.60442	60.4422	0.28297	0.80628
			0.5	0.46602	46.6019	0.2716	0.667
			0.9	0.3436	34.3597	0.1905	0.56929
		80	0.1	0.5405	54.0505	0.28724	0.7306
			0.5	0.47027	47.0267	0.27515	0.65225
			0.9	0.39376	39.3761	0.22817	0.59437
		100	0.1	0.49677	49.6769	0.26523	0.70925
			0.5	0.46969	46.9695	0.26749	0.68788
			0.9	0.4243	42.4299	0.24703	0.62776
120	0.1	0.47141	47.1406	0.21362	0.67901		
	0.5	0.47146	47.1455	0.28081	0.67865		
	0.9	0.44103	44.1031	0.24428	0.64052		

จากตารางที่ 4.28 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=2$, $M=0.33$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.29 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=2$, $M=0.99$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
2	0.99	40	0.1	0.64032	64.0317	0.23306	0.86059
			0.5	0.46419	46.4193	0.2907	0.67189
			0.9	0.2968	29.6804	0.13289	0.54229
		60	0.1	0.60067	60.0674	0.25928	0.79092
			0.5	0.46436	46.4364	0.27972	0.67675
			0.9	0.35092	35.0922	0.1819	0.56456
		80	0.1	0.54926	54.9261	0.30632	0.7406
			0.5	0.45426	45.4262	0.2536	0.66528
			0.9	0.38854	38.8537	0.22029	0.5746
		100	0.1	0.52592	52.5915	0.29777	0.70328
			0.5	0.45806	45.8055	0.25328	0.66288
			0.9	0.41632	41.6321	0.26044	0.60219
120	0.1	0.49092	49.0915	0.2934	0.67545		
	0.5	0.46609	46.6087	0.27335	0.68624		
	0.9	0.43298	43.2982	0.24341	0.62163		

จากตารางที่ 4.29 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=2$, $M=0.99$, $n=40, 60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.30 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=3$, $M=0$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
3	0	60	0.1	0.55977	55.9769	0.2122	0.7957
			0.5	0.47547	47.547	0.29924	0.65298
			0.9	0.35961	35.961	0.18128	0.59866
		80	0.1	0.51334	51.3338	0.21485	0.72966
			0.5	0.46259	46.2591	0.29234	0.63993
			0.9	0.39379	39.3788	0.22074	0.62339
		100	0.1	0.4593	45.9296	0.17763	0.70699
			0.5	0.47588	47.5879	0.28671	0.65611
			0.9	0.41641	41.6408	0.21114	0.65083
		120	0.1	0.44807	44.8075	0.21319	0.72649
			0.5	0.47333	47.3333	0.28963	0.66095
			0.9	0.43794	43.7944	0.21962	0.69881

จากตารางที่ 4.30 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=3$, $M=0$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.31 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วง ความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=3$, $M=0.33$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
3	0.33	60	0.1	0.5252	52.5204	0.1351	0.76894
			0.5	0.46336	46.3365	0.27359	0.66617
			0.9	0.36416	36.4165	0.17854	0.59963
		80	0.1	0.47463	47.4632	0.15051	0.71359
			0.5	0.46149	46.1495	0.25548	0.66171
			0.9	0.41612	41.6124	0.21177	0.6522
		100	0.1	0.45812	45.8125	0.18206	0.72049
			0.5	0.46452	46.4517	0.2744	0.67215
			0.9	0.42704	42.7044	0.20399	0.66039
		120	0.1	0.44994	44.9937	0.16168	0.72472
			0.5	0.46709	46.7086	0.26704	0.6597
			0.9	0.42929	42.9294	0.21403	0.67132

จากตารางที่ 4.31 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=3$, $M=0.33$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.32 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=3$, $M=0.67$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
3	0.67	60	0.1	0.49272	49.2717	0.14982	0.74319
			0.5	0.45526	45.526	0.25517	0.67837
			0.9	0.38339	38.3391	0.16709	0.60909
		80	0.1	0.44814	44.8144	0.15618	0.72943
			0.5	0.46657	46.6574	0.25612	0.70047
			0.9	0.42013	42.0127	0.18648	0.66234
		100	0.1	0.43638	43.6385	0.14844	0.71901
			0.5	0.45583	45.5832	0.23351	0.68222
			0.9	0.42285	42.2854	0.17225	0.65572
		120	0.1	0.44083	44.0826	0.16325	0.73528
			0.5	0.4659	46.59	0.2443	0.67324
			0.9	0.42551	42.5512	0.18479	0.67987

จากตารางที่ 4.32 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=3$, $M=0.67$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$ ยกเว้น ที่ $n=60$ พบว่าค่าของจุดแบ่งลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

ตารางที่ 4.33 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=3$, $M=0.99$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
3	0.99	60	0.1	0.5028	50.2797	0.19802	0.73019
			0.5	0.45981	45.9809	0.22786	0.69569
			0.9	0.39449	39.4486	0.17745	0.61879
		80	0.1	0.47026	47.0258	0.18257	0.72371
			0.5	0.45295	45.2953	0.20883	0.70894
			0.9	0.41766	41.766	0.19765	0.63321
		100	0.1	0.45864	45.8637	0.22063	0.69053
			0.5	0.4544	45.4402	0.22473	0.71321
			0.9	0.42307	42.307	0.18805	0.6397
		120	0.1	0.44915	44.9146	0.20672	0.68944
			0.5	0.46018	46.0183	0.22043	0.70256
			0.9	0.43774	43.774	0.19636	0.67311

จากตารางที่ 4.33 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=3$, $M=0.99$, $n=60, 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ค่าของจุดแบ่งลดลงเมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น ยกเว้น ที่ $n=120$ พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ตารางที่ 4.34 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=4$, $M=0$, $n= 80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
4	0	80	0.1	0.46002	46.0022	0.15307	0.72649
			0.5	0.47273	47.2731	0.27635	0.67147
			0.9	0.40506	40.5057	0.18064	0.63361
		100	0.1	0.43834	43.8343	0.1413	0.71265
			0.5	0.46896	46.8962	0.28237	0.67624
			0.9	0.41412	41.4121	0.19606	0.65972
		120	0.1	0.43788	43.788	0.12485	0.71127
			0.5	0.47057	47.0568	0.27234	0.65595
			0.9	0.43084	43.0837	0.21707	0.68736

จากตารางที่ 4.34 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=4$, $M=0$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.35 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=4$, $M=0.33$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
4	0.33	80	0.1	0.43301	43.3013	0.09586	0.75
			0.5	0.46814	46.8144	0.26142	0.66893
			0.9	0.41193	41.1925	0.1797	0.66963
		100	0.1	0.43196	43.1961	0.12907	0.72818
			0.5	0.46129	46.1286	0.25883	0.68498
			0.9	0.4189	41.8897	0.15653	0.67981
		120	0.1	0.43565	43.5652	0.14489	0.72309
			0.5	0.46232	46.2317	0.25888	0.67532
			0.9	0.42061	42.0607	0.15375	0.69642

จากตารางที่ 4.35 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=4$, $M=0.33$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.36 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=4$, $M=0.67$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
4	0.67	80	0.1	0.40612	40.6124	2.21E-07	0.73659
			0.5	0.45859	45.8593	0.21739	0.69528
			0.9	0.39994	39.9944	0.13297	0.66285
		100	0.1	0.41093	41.0934	0.08094	0.72006
			0.5	0.46422	46.4224	0.24617	0.70388
			0.9	0.4028	40.2804	0.09981	0.69822
		120	0.1	0.41093	41.093	0.1014	0.73073
			0.5	0.46008	46.0085	0.24019	0.6991
			0.9	0.40276	40.2761	0.14374	0.67524

จากตารางที่ 4.36 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=4$, $M=0.67$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.37 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=4$, $M=0.99$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
4	0.99	80	0.1	0.42508	42.5079	0.10247	0.7313
			0.5	0.43343	43.3435	0.15069	0.71533
			0.9	0.41054	41.054	8.77E-07	0.68588
		100	0.1	0.43188	43.1877	0.14442	0.72998
			0.5	0.44332	44.332	0.19475	0.72846
			0.9	0.42009	42.0094	0.14641	0.68313
		120	0.1	0.4222	42.2203	0.15642	0.71711
			0.5	0.4417	44.1703	0.18798	0.70711
			0.9	0.41959	41.9588	0.15893	0.69205

จากตารางที่ 4.37 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=4$, $M=0.99$, $n=80, 100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.38 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=5$, $M=0$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
5	0	100	0.1	0.43478	43.4778	0.141165995	0.70837
			0.5	0.46161	46.1608	0.279829406	0.66731
			0.9	0.4148	41.4798	0.162029412	0.67396
		120	0.1	0.42975	42.9753	0.135982669	0.70241
			0.5	0.46541	46.5414	0.273721677	0.671
			0.9	0.42674	42.6743	0.181928079	0.66697

จากตารางที่ 4.38 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=5$, $M=0$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ตารางที่ 4.39 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=5$, $M=0.33$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
5	0.33	100	0.1	0.41729	41.7291	1.93E-06	0.75141
			0.5	0.45621	45.621	0.25078	0.69079
			0.9	0.39466	39.4664	3.09E-08	0.68351
		120	0.1	0.41349	41.3487	0.06637	0.72074
			0.5	0.46355	46.355	0.26699	0.70299
			0.9	0.40211	40.2111	0.11899	0.70151

จากตารางที่ 4.39 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=5$, $M=0.33$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ตารางที่ 4.40 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=5$, $M=0.67$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
5	0.67	100	0.1	0.37829	37.8286	1.51E-08	0.7374
			0.5	0.44935	44.9355	0.20872	0.69572
			0.9	0.38456	38.4564	8.98E-09	0.70799
		120	0.1	0.37305	37.3048	2.54E-08	0.73663
			0.5	0.45972	45.9721	0.22517	0.70778
			0.9	0.38981	38.981	6.29E-08	0.69617

จากตารางที่ 4.40 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=5$, $M=0.67$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

ตารางที่ 4.41 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=5$, $M=0.99$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
5	0.99	100	0.1	0.38561	38.5607	4.20E-08	0.72156
			0.5	0.40833	40.8332	0.15015	0.71905
			0.9	0.40809	40.8089	6.83E-08	0.6834
		120	0.1	0.40095	40.0946	3.35E-07	0.70207
			0.5	0.42575	42.5751	0.09891	0.73452
			0.9	0.39278	39.2782	1.06E-07	0.69536

จากตารางที่ 4.41 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=5$, $M=0.99$, $n=100, 120$ และ a เปลี่ยนแปลงไปเป็น 0.1, 0.5, 0.9 พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0$

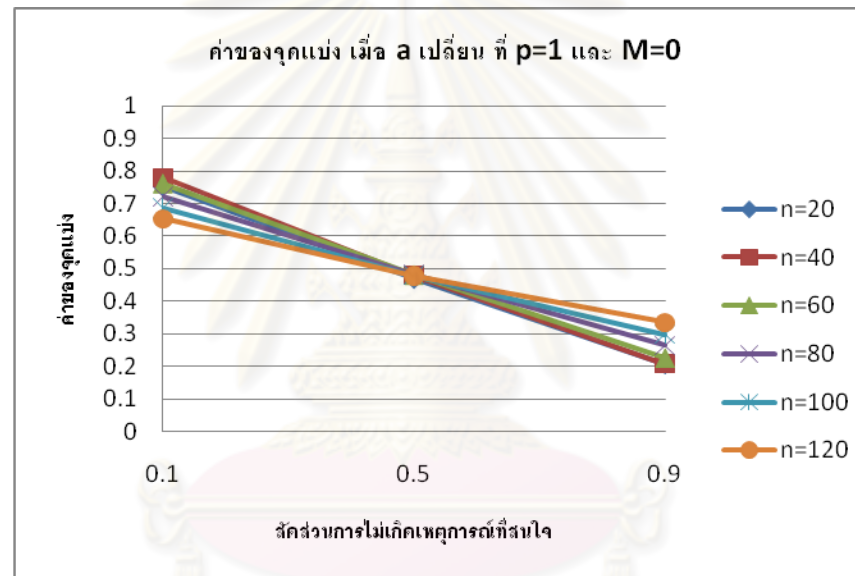
ตารางที่ 4.42 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $p=6$, $M=0, 0.33, 0.67, 0.99$, $n=120$ และ a เปลี่ยนแปลงไปเป็น $0.1, 0.5, 0.9$ ตามลำดับ

p	M	n	a	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
6	0	120	0.1	0.41997	41.9975	0.06356	0.7417
			0.5	0.46217	46.2169	0.24694	0.65623
			0.9	0.42047	42.0469	0.14876	0.68296
	0.33	120	0.1	0.37392	37.3923	3.02E-08	0.73211
			0.5	0.46271	46.2708	0.25382	0.71527
			0.9	0.38475	38.4751	1.14E-08	0.71555
	0.67	120	0.1	0.32607	32.6069	4.88E-09	0.76377
			0.5	0.43803	43.8034	0.19376	0.69501
			0.9	0.31568	31.5678	3.73E-09	0.69443
	0.99	120	0.1	0.35028	35.0282	9.03E-09	0.73619
			0.5	0.36773	36.7727	9.58E-09	0.71746
			0.9	0.33698	33.698	8.99E-09	0.6791

จากตารางที่ 4.42 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $p=6$, $M=0, 0.33, 0.67, 0.99$, $n=120$ และ a เปลี่ยนแปลงไปเป็น $0.1, 0.5, 0.9$ พบว่าค่าของจุดแบ่งจะมีค่ามากที่สุด เมื่อ $a=0.5$

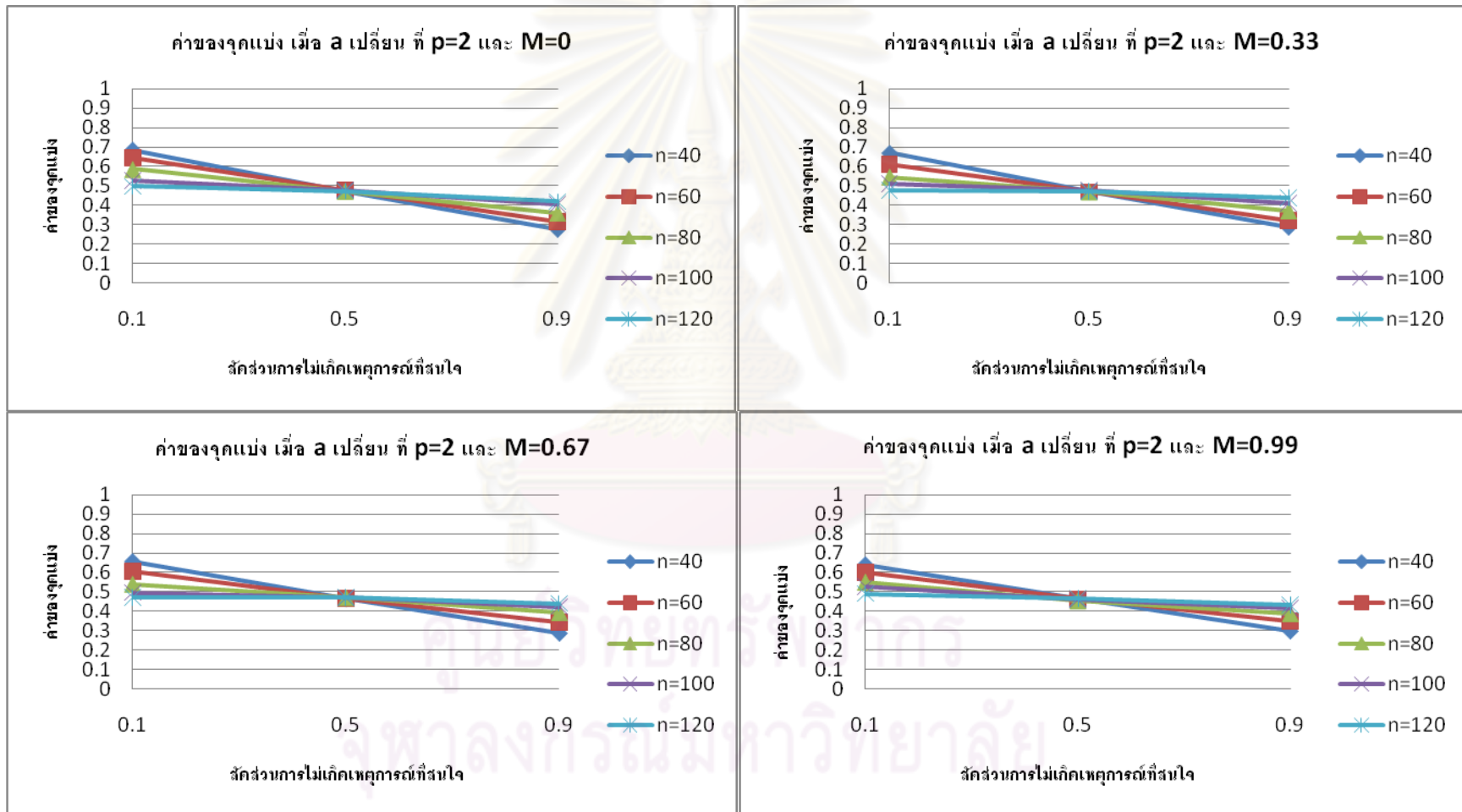
จากตารางที่ 4.25- 4.42 สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ ได้ดังรูปที่ 4.9-4.14 ต่อไปนี้

รูปที่ 4.9 แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 1$, $n = 20, 40, 60, 80, 100, 120$ และ $M = 0$

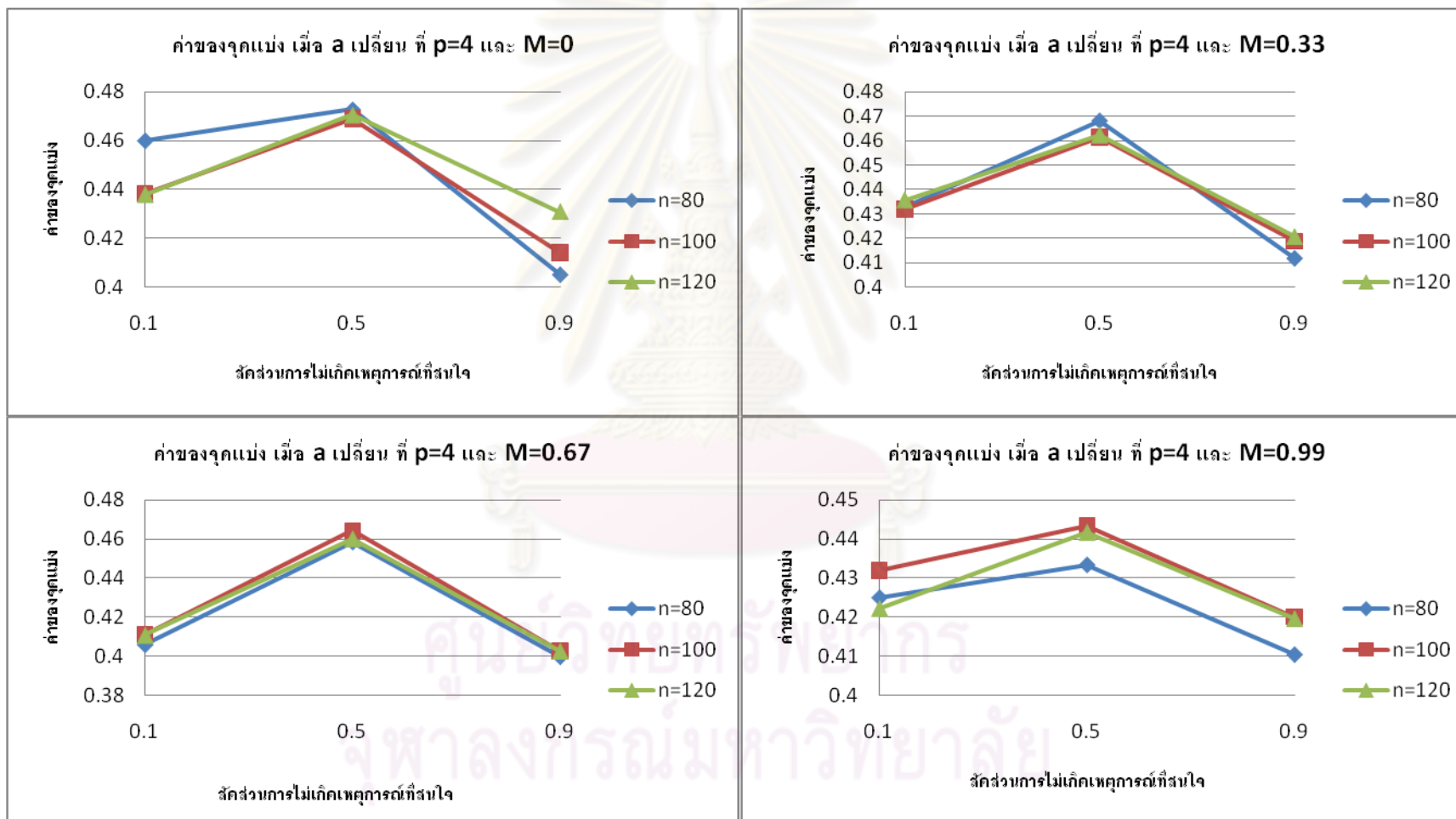


ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 4.10 แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไปที่ $p = 2$, $n = 40, 60, 80, 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99

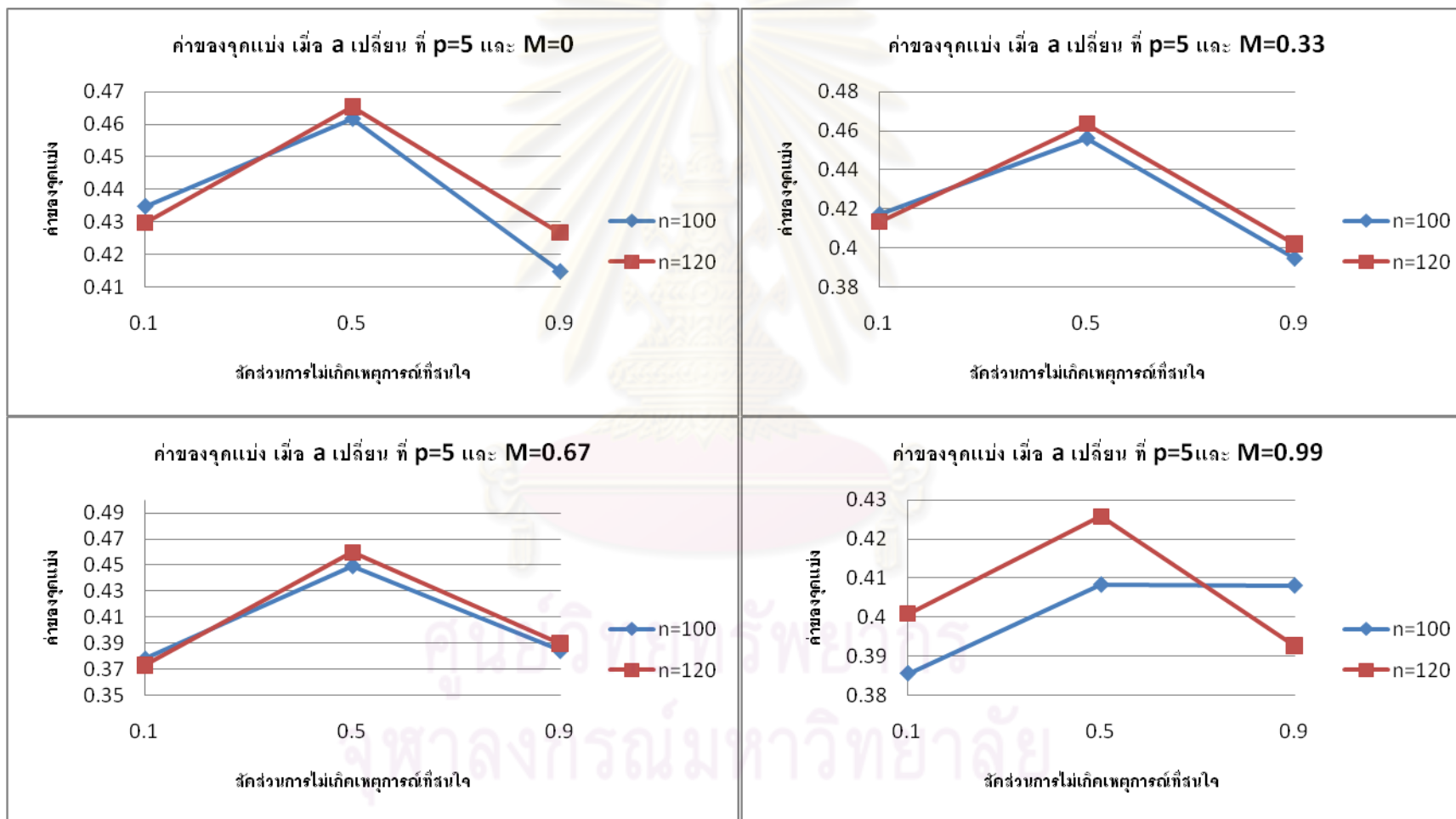


รูปที่ 4.12 แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 4$, $n = 80, 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99

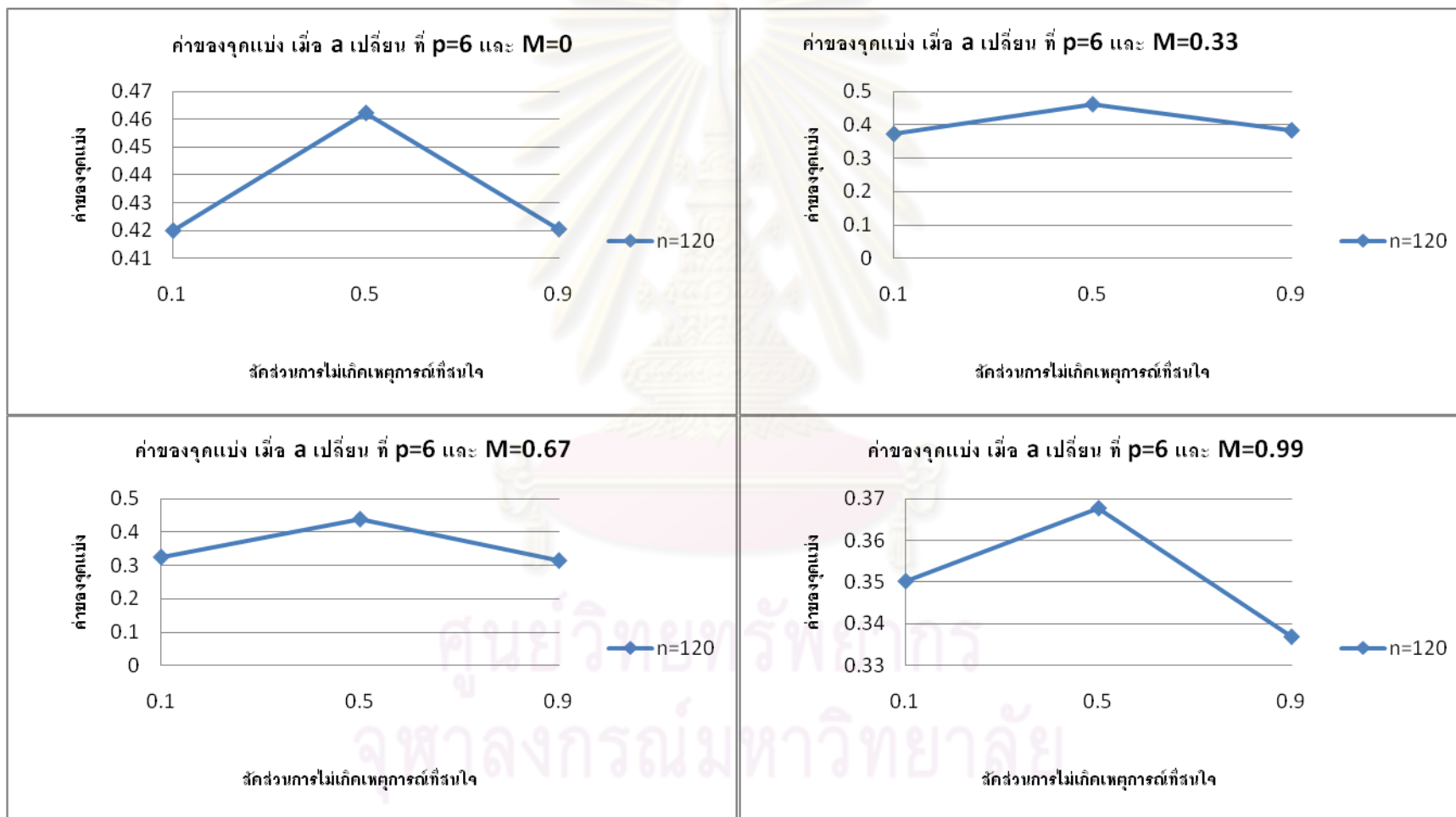


รูปที่ 4.13 แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 5$, $n = 100, 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ

0.99



รูปที่ 4.14 แสดงค่าของจุดแบ่ง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ที่ $p = 6$, $n = 120$ โดยแยกตาม $M = 0, 0.33, 0.67$ และ 0.99



4.1.4 เมื่อระดับความสัมพันธ์ระหว่างอิสระเปลี่ยนแปลงไป นำเสนอในรูปแบบตารางที่ 4.43-4.57 ดังนี้

ตารางที่ 4.43 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	2	40	0	0.68223	68.2231	0.38386	0.86516
			0.33	0.67018	67.0181	0.27361	0.86004
			0.67	0.65455	65.4549	0.28008	0.85724
			0.99	0.64032	64.0317	0.23306	0.86059
		60	0	0.64599	64.5994	0.33165	0.85251
			0.33	0.60942	60.9416	0.30218	0.82894
			0.67	0.60442	60.4422	0.28297	0.80628
			0.99	0.60067	60.0674	0.25928	0.79092
		80	0	0.58701	58.7014	0.30018	0.7785
			0.33	0.54617	54.617	0.2756	0.73909
			0.67	0.5405	54.0505	0.28724	0.7306
			0.99	0.54926	54.9261	0.30632	0.7406
		100	0	0.5278	52.7796	0.26477	0.7412
			0.33	0.51218	51.2179	0.27879	0.71302
			0.67	0.49677	49.6769	0.26523	0.70925
			0.99	0.52592	52.5915	0.29777	0.70328
		120	0	0.49786	49.7863	0.26269	0.7013
			0.33	0.47717	47.7167	0.24916	0.69354
			0.67	0.47141	47.1406	0.21362	0.67901
			0.99	0.49092	49.0915	0.2934	0.67545

จากตารางที่ 4.43 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.44 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ ตามลำดับ

a	p	n	M	c	Percent of c	CI.Lower of c	CI.Upper of c
0.1	3	60	0	0.55977	55.9769	0.2122	0.7957
			0.33	0.5252	52.5204	0.1351	0.76894
			0.67	0.49272	49.2717	0.14982	0.74319
			0.99	0.5028	50.2797	0.19802	0.73019
		80	0	0.51334	51.3338	0.21485	0.72966
			0.33	0.47463	47.4632	0.15051	0.71359
			0.67	0.44814	44.8144	0.15618	0.72943
			0.99	0.47026	47.0258	0.18257	0.72371
		100	0	0.4593	45.9296	0.17763	0.70699
			0.33	0.45812	45.8125	0.18206	0.72049
			0.67	0.43638	43.6385	0.14844	0.71901
			0.99	0.45864	45.8637	0.22063	0.69053
		120	0	0.44807	44.8075	0.21319	0.72649
			0.33	0.44994	44.9937	0.16168	0.72472
			0.67	0.44083	44.0826	0.16325	0.73528
			0.99	0.44915	44.9146	0.20672	0.68944

จากตารางที่ 4.44 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.45 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	4	80	0	0.46002223	46.0022235	0.15306672	0.72648844
			0.33	0.4330129	43.30129	0.09586183	0.75000418
			0.67	0.40612352	40.6123522	2.21E-07	0.73658948
			0.99	0.42507875	42.507875	0.10246599	0.73130133
		100	0	0.43834349	43.8343487	0.14130259	0.71265221
			0.33	0.43196149	43.1961492	0.12906761	0.72817588
			0.67	0.4109335	41.0933503	0.08094209	0.72006231
			0.99	0.43187702	43.1877018	0.14442297	0.7299817
		120	0	0.43788032	43.7880324	0.12485213	0.71127247
			0.33	0.43565217	43.565217	0.14489424	0.72309226
			0.67	0.41093005	41.0930048	0.10139677	0.73073299
			0.99	0.42220281	42.2202811	0.15642496	0.71710764

จากตารางที่ 4.45 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.46 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	5	100	0	0.43477811	43.4778114	0.141166	0.70837276
			0.33	0.41729137	41.7291366	1.93E-06	0.75140989
			0.67	0.37828563	37.8285633	1.51E-08	0.73740199
			0.99	0.38560694	38.5606942	4.20E-08	0.72156229
		120	0	0.4297535	42.9753498	0.13598267	0.70241302
			0.33	0.41348738	41.3487384	0.06636541	0.72074044
			0.67	0.37304836	37.3048355	2.54E-08	0.7366253
			0.99	0.40094575	40.0945749	3.35E-07	0.70207233

จากตารางที่ 4.46 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพัทธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.47 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.1$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.1	6	120	0	0.41997465	41.9974651	0.06355508	0.74170489
			0.33	0.37392308	37.3923082	3.02E-08	0.73210932
			0.67	0.32606937	32.6069365	4.88E-09	0.76377226
			0.99	0.35028198	35.0281981	9.03E-09	0.73619366

จากตารางที่ 4.47 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.1$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.48 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	2	40	0	0.47233	47.2327	0.30226	0.64787
			0.33	0.47138	47.1384	0.28741	0.65789
			0.67	0.46397	46.3965	0.27176	0.65041
			0.99	0.46419	46.4193	0.2907	0.67189
		60	0	0.4771	47.7102	0.31363	0.63549
			0.33	0.46983	46.9831	0.2884	0.64878
			0.67	0.46602	46.6019	0.2716	0.667
			0.99	0.46436	46.4364	0.27972	0.67675
		80	0	0.4726	47.2598	0.30776	0.63994
			0.33	0.46975	46.9745	0.299	0.65329
			0.67	0.47027	47.0267	0.27515	0.65225
			0.99	0.45426	45.4262	0.2536	0.66528
		100	0	0.48006	48.0061	0.31361	0.64925
			0.33	0.47767	47.7666	0.3093	0.67435
			0.67	0.46969	46.9695	0.26749	0.68788
			0.99	0.45806	45.8055	0.25328	0.66288
		120	0	0.4741	47.4099	0.29529	0.64896
			0.33	0.47333	47.3329	0.29284	0.66456
			0.67	0.47146	47.1455	0.28081	0.67865
			0.99	0.46609	46.6087	0.27335	0.68624

จากตารางที่ 4.48 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.49 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	3	60	0	0.47547	47.547	0.29924	0.65298
			0.33	0.46336	46.3365	0.27359	0.66617
			0.67	0.45526	45.526	0.25517	0.67837
			0.99	0.45981	45.9809	0.22786	0.69569
		80	0	0.46259	46.2591	0.29234	0.63993
			0.33	0.46149	46.1495	0.25548	0.66171
			0.67	0.46657	46.6574	0.25612	0.70047
			0.99	0.45295	45.2953	0.20883	0.70894
		100	0	0.47588	47.5879	0.28671	0.65611
			0.33	0.46452	46.4517	0.2744	0.67215
			0.67	0.45583	45.5832	0.23351	0.68222
			0.99	0.4544	45.4402	0.22473	0.71321
		120	0	0.47333	47.3333	0.28963	0.66095
			0.33	0.46709	46.7086	0.26704	0.6597
			0.67	0.4659	46.59	0.2443	0.67324
			0.99	0.46018	46.0183	0.22043	0.70256

จากตารางที่ 4.49 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.50 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	4	80	0	0.4727307	47.2730704	0.27634681	0.67146711
			0.33	0.46814426	46.8144258	0.26141809	0.66892658
			0.67	0.4585933	45.8593298	0.21738692	0.69527817
			0.99	0.43343498	43.3434983	0.15069147	0.71532545
		100	0	0.46896223	46.8962231	0.28236794	0.67623643
			0.33	0.46128633	46.1286329	0.25882998	0.68497825
			0.67	0.46422418	46.4224183	0.24617256	0.70387527
			0.99	0.44331984	44.3319838	0.19474601	0.72846172
		120	0	0.47056764	47.056764	0.27234309	0.65594627
			0.33	0.46231734	46.2317345	0.25887683	0.67532006
			0.67	0.46008487	46.0084868	0.24018782	0.69909629
			0.99	0.44170284	44.1702843	0.18798492	0.70711451

จากตารางที่ 4.49 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ตารางที่ 4.51 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	5	100	0	0.46161	46.1608	0.27983	0.66731
			0.33	0.45621	45.621	0.25078	0.69079
			0.67	0.44935	44.9355	0.20872	0.69572
			0.99	0.40833	40.8332	0.15015	0.71905
		120	0	0.46541	46.5414	0.27372	0.671
			0.33	0.46355	46.355	0.26699	0.70299
			0.67	0.45972	45.9721	0.22517	0.70778
			0.99	0.42575	42.5751	0.09891	0.73452

จากตารางที่ 4.51 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.52 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.5$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.5	6	120	0	0.46217	46.2169	0.24694	0.65623
			0.33	0.46271	46.2708	0.25382	0.71527
			0.67	0.43803	43.8034	0.19376	0.69501
			0.99	0.36773	36.7727	9.58E-09	0.71746

จากตารางที่ 4.52 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.5$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.53 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	2	40	0	0.27727	27.7268	0.12625	0.52265
			0.33	0.28812	28.8119	0.12943	0.5327
			0.67	0.28925	28.925	0.12971	0.54021
			0.99	0.2968	29.6804	0.13289	0.54229
		60	0	0.31754	31.754	0.15127	0.54438
			0.33	0.32304	32.3042	0.16313	0.53465
			0.67	0.3436	34.3597	0.1905	0.56929
			0.99	0.35092	35.0922	0.1819	0.56456
		80	0	0.36052	36.0516	0.20651	0.58619
			0.33	0.37216	37.2157	0.2133	0.58456
			0.67	0.39376	39.3761	0.22817	0.59437
			0.99	0.38854	38.8537	0.22029	0.5746
		100	0	0.40403	40.4034	0.23545	0.61522
			0.33	0.40942	40.9425	0.24483	0.62505
			0.67	0.4243	42.4299	0.24703	0.62776
			0.99	0.41632	41.6321	0.26044	0.60219
		120	0	0.41937	41.9375	0.23791	0.61582
			0.33	0.43912	43.9124	0.25432	0.65165
			0.67	0.44103	44.1031	0.24428	0.64052
			0.99	0.43298	43.2982	0.24341	0.62163

จากตารางที่ 4.53 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $p=2$, $n=40, 60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้นเมื่อระดับความสัมพัทธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มลดลง

ตารางที่ 4.54 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	3	60	0	0.35961	35.961	0.18128	0.59866
			0.33	0.36416	36.4165	0.17854	0.59963
			0.67	0.38339	38.3391	0.16709	0.60909
			0.99	0.39449	39.4486	0.17745	0.61879
		80	0	0.39379	39.3788	0.22074	0.62339
			0.33	0.41612	41.6124	0.21177	0.6522
			0.67	0.42013	42.0127	0.18648	0.66234
			0.99	0.41766	41.766	0.19765	0.63321
		100	0	0.41641	41.6408	0.21114	0.65083
			0.33	0.42704	42.7044	0.20399	0.66039
			0.67	0.42285	42.2854	0.17225	0.65572
			0.99	0.42307	42.307	0.18805	0.6397
		120	0	0.43794	43.7944	0.21962	0.69881
			0.33	0.42929	42.9294	0.21403	0.67132
			0.67	0.42551	42.5512	0.18479	0.67987
			0.99	0.43774	43.774	0.19636	0.67311

จากตารางที่ 4.54 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $p=3$, $n=60, 80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มลดลง

ตารางที่ 4.55 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	4	80	0	0.40506	40.5057	0.18064	0.63361
			0.33	0.41193	41.1925	0.1797	0.66963
			0.67	0.39994	39.9944	0.13297	0.66285
			0.99	0.41054	41.054	8.77E-07	0.68588
		100	0	0.41412	41.4121	0.19606	0.65972
			0.33	0.4189	41.8897	0.15653	0.67981
			0.67	0.4028	40.2804	0.09981	0.69822
			0.99	0.42009	42.0094	0.14641	0.68313
		120	0	0.43084	43.0837	0.21707	0.68736
			0.33	0.42061	42.0607	0.15375	0.69642
			0.67	0.40276	40.2761	0.14374	0.67524
			0.99	0.41959	41.9588	0.15893	0.69205

จากตารางที่ 4.55 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $p=4$, $n=80, 100, 120$ และ M เปลี่ยนแปลงไปเป็น $0, 0.33, 0.67, 0.99$ พบว่าค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.56 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	5	100	0	0.4148	41.4798	0.16203	0.67396
			0.33	0.39466	39.4664	3.09E-08	0.68351
			0.67	0.38456	38.4564	8.98E-09	0.70799
			0.99	0.40809	40.8089	6.83E-08	0.6834
		120	0	0.42674	42.6743	0.18193	0.66697
			0.33	0.40211	40.2111	0.11899	0.70151
			0.67	0.38981	38.981	6.29E-08	0.69617
			0.99	0.39278	39.2782	1.06E-07	0.69536

จากตารางที่ 4.56 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $p=5$, $n=100, 120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.57 แสดงค่าของจุดแบ่ง ค่าร้อยละของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ $a=0.9$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 ตามลำดับ

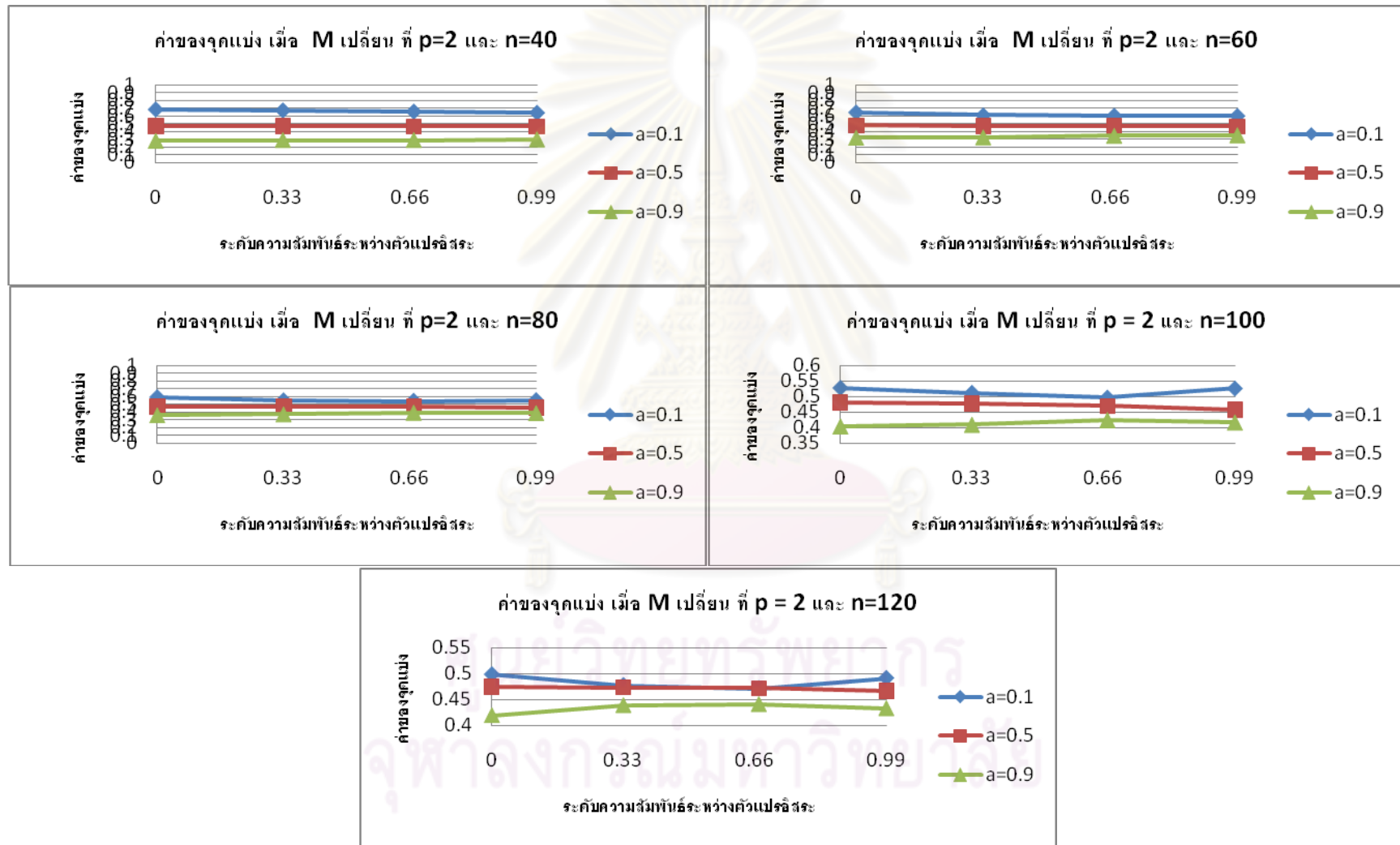
a	p	n	M	\hat{c}	Percent of \hat{c}	CI.Lower of \hat{c}	CI.Upper of \hat{c}
0.9	6	120	0	0.42047	42.0469	0.14876	0.68296
			0.33	0.38475	38.4751	1.14E-08	0.71555
			0.67	0.31568	31.5678	3.73E-09	0.69443
			0.99	0.33698	33.698	8.99E-09	0.6791

จากตารางที่ 4.57 เมื่อพิจารณาค่าของจุดแบ่ง ที่ $a=0.9$, $p=6$, $n=120$ และ M เปลี่ยนแปลงไปเป็น 0, 0.33, 0.67, 0.99 พบว่าค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น แต่เมื่อ $M=0.99$ ค่าของจุดแบ่งจะมีแนวโน้มเพิ่มขึ้น

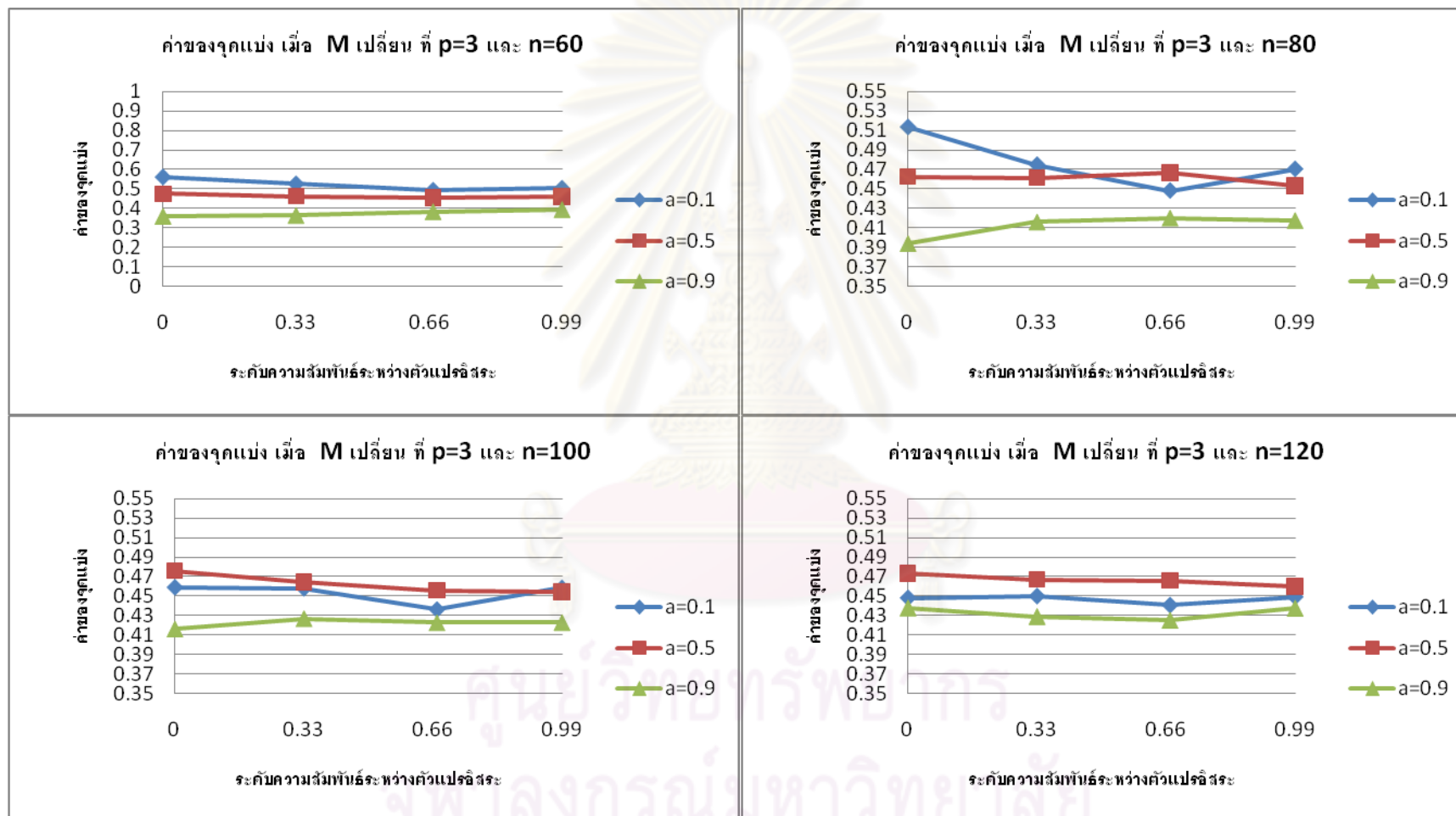
จากตารางที่ 4.43- 4.57 สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่ ได้ดังรูปที่ 4.15-4.19 ต่อไปนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

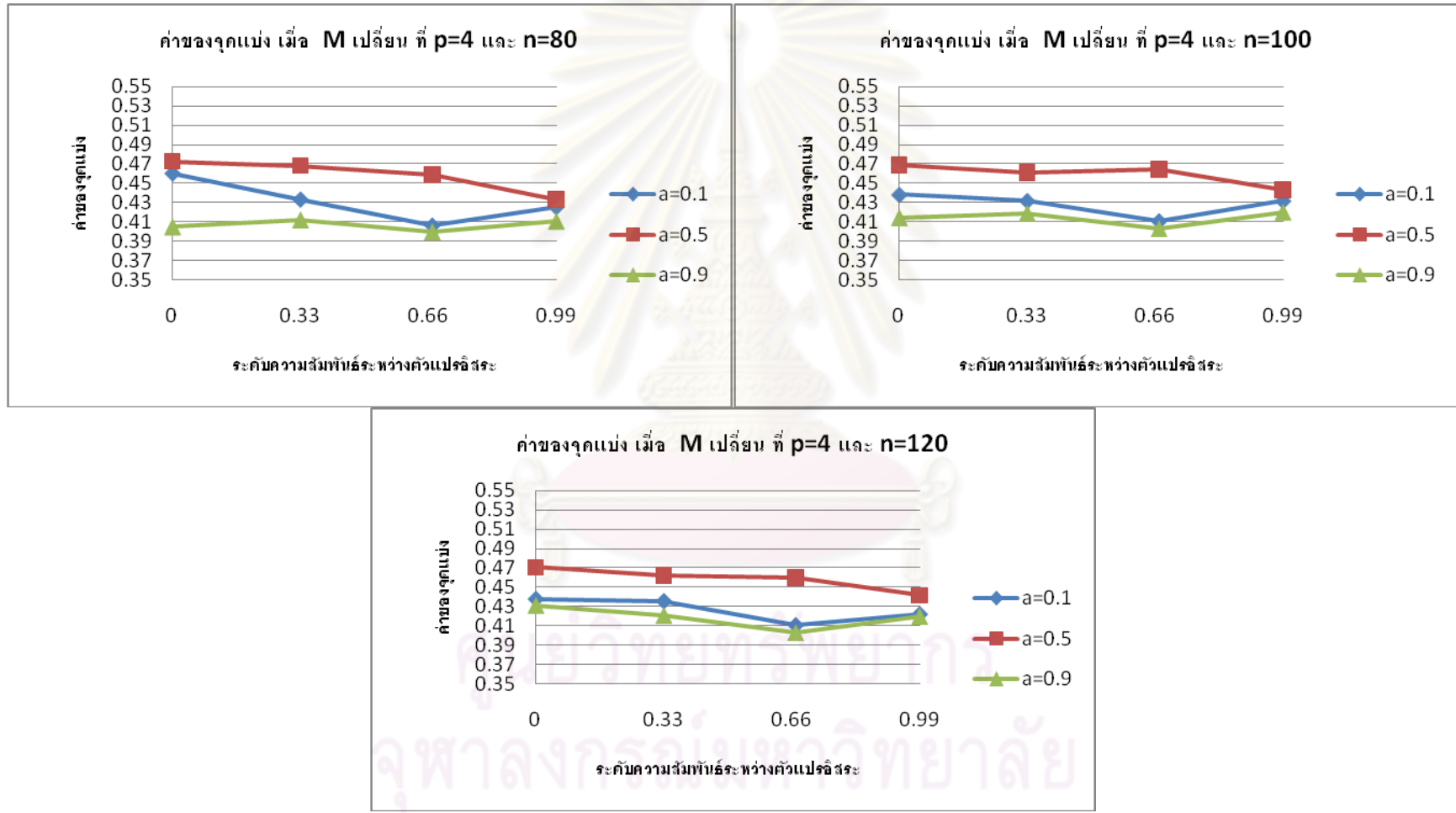
รูปที่ 4.15 แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป ที่ $p = 2$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 40, 60, 80, 100$ และ 120



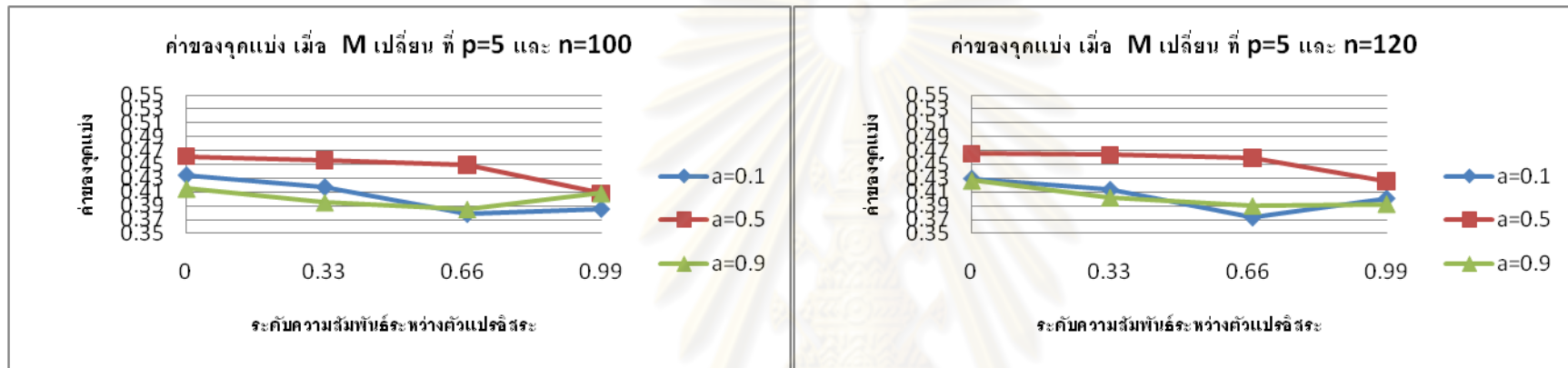
รูปที่ 4.16 แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไปที่ $p = 3$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 60, 80, 100$ และ 120



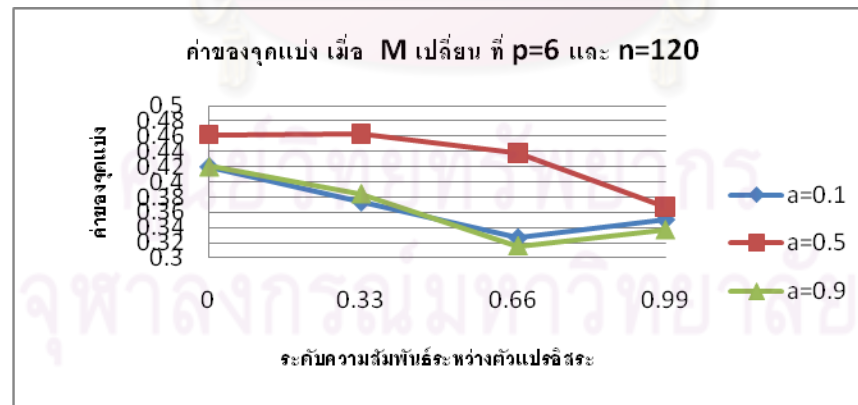
รูปที่ 4.17 แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป ที่ $p = 4$, $a = 0.1, 0.5, 0.9$ โดยแยกตาม $n = 80, 100$ และ 120



รูปที่ 4.18 แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป ที่ $p = 5$, $a=0.1, 0.5, 0.9$ โดยแยกตาม $n = 100$ และ 120



รูปที่ 4.19 แสดงค่าของจุดแบ่ง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป ที่ $p = 6$, $a=0.1, 0.5, 0.9$ โดยแยกตาม $n = 120$



4.2 สมการถดถอยพหุคูณ เพื่อใช้ประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่น ๆ

เมื่อได้ค่าร้อยละของจุดแบ่งที่เหมาะสมในแต่ละสถานการณ์แล้ว จะนำค่าของจุดแบ่งนั้นไปเข้าสมการถดถอยพหุคูณที่มีผลอันตรกิริยา เพื่อประมาณค่าพารามิเตอร์ แล้วนำค่าพารามิเตอร์นั้นไปใช้สำหรับการประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่น ๆ ต่อไป ซึ่งผลการวิจัยได้นำเสนอในตาราง 4.58 ดังต่อไปนี้

ตารางที่ 4.58 แสดงการประมาณค่าพารามิเตอร์จากตัวแบบการถดถอยพหุคูณ

	Estimate	Std.Error	t value	Pr(> t)
(Intercept)	114.638	3.201	35.808	0.000
n	-0.004	0.000	-13.147	0.000
p	-0.176	0.016	-10.884	0.000
M	-0.142	0.053	-2.709	0.007
a	-1.279	0.051	-25.181	0.000
ap	0.003	0.000	12.163	0.000
an	8.46E-05	0.000	16.266	0.000
aM	0.003	0.000	5.86	0.000
np	1.15E-05	0.000	7.784	0.000
nM	3.18E-06	0.000	0.676	0.500
pM	0	0.000	1.605	0.110
apn	-2.21E-07	0.000	-9.313	0.000
apM	-8.05E-06	0.000	-2.356	0.020
pnM	-2.44E-08	0.000	-1.113	0.267
apnM	2.13E-10	0.000	0.842	0.401

จากตารางที่ 4.58 ผลวิเคราะห์โดยใช้ตัวแบบการถดถอยพหุคูณที่มีผลอันตรกิริยา พบว่า ปัจจัยที่มีผลต่อค่าของจุดแบ่ง คือ ขนาดตัวอย่าง, จำนวนตัวแปรอิสระ สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ และอันตรกิริยามีผลต่อค่าร้อยละของจุดแบ่ง คือ ap, an, aM, np, apn และ apM ที่ระดับนัยสำคัญเท่ากับ 0.05 โดยถ้าขนาดตัวอย่าง

เพิ่มขึ้น 1% ค่าร้อยละของจุดแบ่งจะลดลง 0.004% เมื่อปัจจัยอื่นๆ คงที่, จำนวนตัวแปรอิสระเพิ่มขึ้น 1% ค่าร้อยละของจุดแบ่งจะลดลง 0.176% เมื่อปัจจัยอื่นๆ คงที่, ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น 1% ค่าร้อยละของจุดแบ่งจะลดลง 0.142% เมื่อปัจจัยอื่นๆ คงที่และสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น 1% ค่าร้อยละของจุดแบ่งจะลดลง 1.279% เมื่อปัจจัยอื่นๆ คงที่

ตารางที่ 4.59 แสดงค่าสัมประสิทธิ์การตัดสินใจเชิงซ้อน(R Square)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.956 ^a	.914	.908	2.62894

a. Predictors: (Constant), apnM, n, a, p, M, pnM, apn, aM, an, nM, np, apM, ap, pM

จากตารางที่ 4.59 แสดงค่าความผันแปรทั้งหมดของจุดแบ่งมีสาเหตุจากความผันแปรของ n, a, p, M, aM, an, nM, np, ap, pM, apn, apM, pnM, apnM, คิดเป็น 91.4% (R^2) ส่วนความผันแปรของค่าของจุดแบ่งที่เหลืออีก 8.6% เกิดจากสาเหตุอื่นๆ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบโพรบิตแบบ 2 ประเภทสำหรับแต่ละสถานการณ์ที่ต้องการศึกษา และทำการวิเคราะห์ด้วยตัวแบบการถดถอยพหุคูณที่มีผลอันตรกิริยา เพื่อประมาณค่าพารามิเตอร์สำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่นๆ ต่อไป การจำลองข้อมูลในแต่ละสถานการณ์จะจำลองขึ้น โดยใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรม R ซึ่งกำหนดระดับนัยสำคัญ (α) ในการวิจัยครั้งนี้ที่ระดับ 0.05 โดยจะแสดงข้อสรุปผลการวิจัยตามวัตถุประสงค์หลักของการวิจัย ดังต่อไปนี้

1. แสดงผลสรุปของค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท เมื่อ จำนวนตัวแปรอิสระเปลี่ยนแปลงไป, ขนาดตัวอย่างเปลี่ยนแปลงไป, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดของข้อมูลเปลี่ยนแปลงไป, ระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนแปลงไป
2. แสดงผลสรุปของสมการถดถอยพหุคูณ สำหรับใช้ในการประมาณค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ในสถานการณ์อื่นๆ ต่อไป

5.1 สรุปของค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบโพรบิตแบบ 2 ประเภท ภายใต้สถานการณ์ต่าง ๆ ที่สนใจศึกษา

พบว่าค่าของจุดแบ่งที่เหมาะสมมีแนวโน้มเพิ่มขึ้นหรือลดลงต่างกันภายใต้สถานการณ์ต่าง ๆ ที่สนใจศึกษาแตกต่างกัน โดยสรุปแยกตามสถานการณ์ต่าง ๆ ที่สนใจศึกษา ดังต่อไปนี้

จุฬาลงกรณ์มหาวิทยาลัย

เมื่อจำนวนตัวแปรอิสระ (p) เปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่

ตารางที่ 5.1 แสดงผลสรุปค่าของจุดแบ่งกรณีจำนวนตัวแปรอิสระ (p) เปลี่ยนแปลงไป แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M), สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a) และขนาดตัวอย่าง (n) คงที่

สถานการณ์	ค่าของจุดแบ่ง
$a = 0.1, M = 0, n = 40, 60, 80, 100, 120$	มีแนวโน้มลดลง
$a = 0.1, M = 0.33, n = 60, 80, 100, 120$	
$a = 0.1, M = 0.67, n = 60, 80, 100, 120$	
$a = 0.1, M = 0.99, n = 60, 80, 100, 120$	
$a = 0.5, M = 0, n = 40, 60, 80, 100, 120$	
$a = 0.5, M = 0.33, n = 60, 80, 100, 120$	
$a = 0.5, M = 0.67, n = 60, 80, 100, 120$	
$a = 0.5, M = 0.99, n = 60, 80, 100, 120$	
$a = 0.9, M = 0, n = 40, 60, 80, 100, 120$	มีแนวโน้มเพิ่มขึ้น
$a = 0.9, M = 0.33, n = 60, 80$	
$a = 0.9, M = 0.67, n = 60, 80$	
$a = 0.9, M = 0.99, n = 60, 80$	
$a = 0.9, M = 0.33, n = 100, 120$	มีแนวโน้มลดลง
$a = 0.9, M = 0.67, n = 100, 120$	
$a = 0.9, M = 0.99, n = 100, 120$	

จากตารางที่ 5.1 ค่าของจุดแบ่งเมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลงไป ที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและขนาดตัวอย่าง คงที่ สรุปผลได้ว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 และ 0.5 ที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระและทุกระดับของขนาดตัวอย่าง พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ตัวแปรอิสระไม่มีความสัมพันธ์กัน ($M = 0$) และที่ทุกระดับของขนาดตัวอย่าง พบว่า ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ตัวแปรอิสระมีความสัมพันธ์กัน ($M = 0.33, 0.67, 0.99$) และขนาดตัวอย่างปานกลาง ($n = 60, 80$) พบว่า ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ตัวแปรอิสระมีความสัมพันธ์กัน ($M = 0.33, 0.67, 0.99$) และขนาดตัวอย่างใหญ่ ($n = 100, 120$) พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

จากผลลัพธ์ของค่าจุดแบ่งจะเห็นได้ว่าเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น เมื่อขนาดตัวอย่างเล็กและปานกลาง ($n = 20, 40, 60, 80$) ค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 และ 0.9 จะมีแนวโน้มในการลู่ออกค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 ซึ่งค่านี้เป็นค่าของจุดแบ่งที่ถูกกำหนดให้ใช้กันอยู่ในปัจจุบัน แต่เมื่อขนาดตัวอย่างใหญ่ ($n = 100, 120$) ค่าจุดแบ่งของทั้งสองสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจที่ 0.1 และ 0.9 จะมีค่าต่ำกว่าค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ 0.5

เมื่อขนาดตัวอย่างเปลี่ยนแปลงไป (n) แต่ปัจจัยอื่น ๆ คงที่

ตารางที่ 5.2 แสดงผลสรุปค่าของจุดแบ่งกรณีขนาดตัวอย่าง (n) เปลี่ยนแปลงไป แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M), สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a) และจำนวนตัวแปรอิสระ (p) คงที่

สถานการณ์	ค่าของจุดแบ่ง
$a = 0.1, M = 0, p = 1, 2, 3, 4, 5, 6$	มีแนวโน้มลดลง
$a = 0.1, M = 0.33, p = 2, 3, 4, 5, 6$	
$a = 0.1, M = 0.67, p = 2, 3, 4, 5, 6$	
$a = 0.1, M = 0.99, p = 2, 3, 4, 5, 6$	
$a = 0.5, M = 0, p = 1, 2, 3, 4, 5, 6$	มีค่าใกล้เคียงกันมาก ไม่สามารถแยกได้ชัดเจน
$a = 0.5, M = 0.33, p = 2, 3, 4, 5, 6$	
$a = 0.5, M = 0.67, p = 2, 3, 4, 5, 6$	
$a = 0.5, M = 0.99, p = 2, 3, 4, 5, 6$	
$a = 0.9, M = 0, p = 1, 2, 3, 4, 5, 6$	มีแนวโน้มเพิ่มขึ้น
$a = 0.9, M = 0.33, p = 2, 3, 4, 5, 6$	
$a = 0.9, M = 0.67, p = 2, 3, 4, 5, 6$	
$a = 0.9, M = 0.99, p = 2, 3, 4, 5, 6$	

จากตารางที่ 5.2 ค่าของจุดแบ่งเมื่อขนาดตัวอย่างเปลี่ยนแปลงไป แต่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจและจำนวนตัวแปรอิสระคงที่ สรุปผลได้ว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 ที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระและทุกระดับของจำนวนตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระและทุกระดับของจำนวนตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีค่าใกล้เคียงกันมาก ไม่สามารถแยกได้ชัดเจน เมื่อขนาดตัวอย่างเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระและทุกระดับของจำนวนตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

จากผลลัพธ์ของค่าของจุดแบ่งจะเห็นได้ว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นและจำนวนตัวแปรอิสระอยู่ในระดับน้อย ($p = 1, 2$) ค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 และ 0.9 จะมีแนวโน้มในการลู่อู่เข้าหาค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ซึ่งค่านี้เป็นค่าของจุดแบ่งที่ถูกกำหนดให้ใช้กันอยู่ในปัจจุบัน แต่เมื่อจำนวนตัวแปรอิสระอยู่ในระดับปานกลางและมาก ($p = 3, 4, 5, 6$) ค่าจุดแบ่งของทั้งสองสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจที่ 0.1 และ 0.9 จะมีค่าต่ำกว่าค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ 0.5



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a) เปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่

ตารางที่ 5.3 แสดงผลสรุปค่าของจุดแบ่งกรณีสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a) เปลี่ยนแปลงไป แต่ขนาดตัวอย่าง (n), จำนวนตัวแปรอิสระ (p) และระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M) คงที่

สถานการณ์	ค่าของจุดแบ่ง
p = 1, M = 0, n = 20, 40, 60, 80, 100, 120	มีแนวโน้มลดลง
p = 2, M = 0, n = 40, 60, 80, 100, 120	
p = 2, M = 0.33, n = 40, 60, 80, 100, 120	
p = 2, M = 0.67, n = 40, 60, 80, 100, 120	
p = 2, M = 0.99, n = 40, 60, 80, 100, 120	
p = 3, M = 0, 0.33, 0.67, 0.99, n = 60, 80	
p = 3, M = 0, 0.33, 0.67, 0.99, n = 100, 120	มีค่าเพิ่มขึ้นสูงสุดที่ a = 0.5 และหลังจากนั้นจะมีค่าลดลง
p = 4, M = 0, n = 80, 100, 120	
p = 4, M = 0.33, n = 80, 100, 120	
p = 4, M = 0.67, n = 80, 100, 120	
p = 4, M = 0.99, n = 80, 100, 120	
p = 5, M = 0, n = 100, 120	
p = 5, M = 0.33, n = 100, 120	
p = 5, M = 0.67, n = 100, 120	
p = 5, M = 0.99, n = 100, 120	
p = 6, M = 0, 0.33, 0.67, 0.99, n = 120	

จากตารางที่ 5.3 ค่าของจุดแบ่งเมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป แต่ขนาดตัวอย่าง, จำนวนตัวแปรอิสระและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่สรุปผลได้ว่า

เมื่อจำนวนตัวแปรอิสระน้อย ($p = 1, 2$) ที่ทุกระดับของขนาดตัวอย่างและทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระปานกลาง ($p = 3$) ที่ขนาดตัวอย่างปานกลาง ($n = 60, 80$) และที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระปานกลาง ($p = 3$) ที่ขนาดตัวอย่างใหญ่ ($n = 100, 120$) และที่ทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีค่าเพิ่มขึ้นสูงสุด ที่ $a = 0.5$ และหลังจากนั้นจะมีค่าลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระปานกลางและมาก ($p = 4, 5, 6$) ที่ทุกระดับของขนาดตัวอย่าง และทุกระดับความสัมพันธ์ระหว่างตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีค่าเพิ่มขึ้นสูงสุด ที่ $a = 0.5$ และหลังจากนั้นจะมีค่าลดลง เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

จากผลลัพธ์ของค่าของจุดแบ่งจะเห็นได้ว่าเมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเปลี่ยนแปลงไป ค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 จะมีค่าสูงที่สุด เมื่อเทียบกับค่าของจุดแบ่งที่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.1 และ 0.9 ซึ่งค่าสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5 นี้เป็นค่าของจุดแบ่งที่ถูกกำหนดให้ใช้กันอยู่ในปัจจุบัน เป็นค่าจุดแบ่งที่ถูกกำหนดให้ใช้กันในปัจจุบันเพราะเหตุผลที่ว่ากลุ่มของการเกิดเหตุการณ์ที่สนใจมีโอกาสเกิดขึ้นเท่ากับกลุ่มของการไม่เกิดเหตุการณ์ที่สนใจ

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M) เปลี่ยนแปลงไป แต่ปัจจัยอื่น ๆ คงที่

ตารางที่ 5.4 แสดงผลสรุปค่าของจุดแบ่งกรณีระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (M) เปลี่ยนแปลงไป แต่สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล (a), ขนาดตัวอย่าง (n) และจำนวนตัวแปรอิสระ (p) คงที่

สถานการณ์	ค่าของจุดแบ่ง
$a = 0.1, p = 2, n = 40, 60$	มีแนวโน้มลดลง
$a = 0.1, p = 2, n = 80, 100, 120$	มีค่าลดลงต่ำสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าเพิ่มขึ้น
$a = 0.1, p = 3, n = 60, 80, 100, 120$	
$a = 0.1, p = 4, n = 80, 100, 120$	
$a = 0.1, p = 5, n = 100, 120$	
$a = 0.1, p = 6, n = 120$	
$a = 0.5, p = 2, n = 40, 60, 80, 100, 120$	มีแนวโน้มลดลง
$a = 0.5, p = 3, n = 60, 80, 100, 120$	
$a = 0.5, p = 4, n = 80, 100, 120$	
$a = 0.5, p = 5, n = 100, 120$	
$a = 0.5, p = 6, n = 120$	
$a = 0.9, p = 2, n = 40, 60$	มีแนวโน้มเพิ่มขึ้น
$a = 0.9, p = 2, n = 80, 100, 120$	มีค่าเพิ่มขึ้นสูงสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าลดลง
$a = 0.9, p = 3, n = 60$	มีแนวโน้มเพิ่มขึ้น
$a = 0.9, p = 3, n = 80, 100$	มีค่าเพิ่มขึ้นสูงสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าลดลง
$a = 0.9, p = 3, n = 120$	มีค่าลดลงต่ำสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าเพิ่มขึ้น
$a = 0.9, p = 4, n = 80, 100, 120$	
$a = 0.9, p = 5, n = 100, 120$	
$a = 0.9, p = 6, n = 120$	

จากตารางที่ 5.4 ค่าของจุดแบ่งเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจ, ขนาดตัวอย่างและจำนวนตัวแปรอิสระคงที่ สรุปผลได้ว่า

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.1 ที่ทุกระดับของจำนวนตัวแปรอิสระและทุกระดับของขนาดตัวอย่าง พบว่า ค่าของจุดแบ่งมีค่าลดลงต่ำสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าเพิ่มขึ้น เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น ยกเว้น ที่จำนวนตัวแปรอิสระเท่ากับ 2 และขนาดตัวอย่างเท่ากับ 40 และ 60 ที่ค่าของจุดแบ่งมีแนวโน้มลดลง

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.5 ที่ทุกระดับของขนาดตัวอย่างและทุกระดับของจำนวนตัวแปรอิสระ พบว่า ค่าของจุดแบ่งมีแนวโน้มลดลง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่จำนวนตัวแปรอิสระเท่ากับ 2 และขนาดตัวอย่างเท่ากับ 40 และ 60 พบว่า ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น ยกเว้น ที่ขนาดตัวอย่างเท่ากับ 80, 100 และ 120 ที่ค่าของจุดแบ่งมีค่าเพิ่มขึ้นสูงสุดที่ $M = 0.67$ และหลังจากนั้นจะมีค่าลดลง

เมื่อสัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจมีค่าเท่ากับ 0.9 ที่ $p = 3, n = 120$ ที่ $p = 4, n = 80, 100, 120$ ที่ $p = 5, n = 100, 120$ ที่ $p = 6, n = 120$ ค่าของจุดแบ่งมีค่าลดลงต่ำสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าเพิ่มขึ้น เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น ยกเว้นที่ $p = 2, n = 40, 60$ และที่ $p = 3, n = 60$ พบว่า ค่าของจุดแบ่งมีแนวโน้มเพิ่มขึ้น เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระเพิ่มขึ้น และที่ $p = 2, n = 80, 100, 120$ และ $p = 3, n = 80, 100$ ค่าของจุดแบ่งมีค่าเพิ่มขึ้นสูงสุด ที่ $M = 0.67$ และหลังจากนั้นจะมีค่าลดลง

จากทฤษฎีการหาค่าของจุดแบ่ง Hadjicostas P. (2006) จะสามารถสรุปได้ว่า จำนวนตัวแปรอิสระ, ขนาดตัวอย่าง, สัดส่วนของการไม่เกิดเหตุการณ์ที่สนใจในแต่ละชุดข้อมูล และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็นปัจจัยที่มีผลต่อค่าของจุดแบ่งที่เหมาะสมสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบโพรบิตแบบ 2 ประเภท

5.2 สรุปผลการวิเคราะห์การถดถอยพหุคูณ

ผลการวิเคราะห์ตัวแบบการถดถอยพหุคูณที่มีผลอันตรกิริยาจากผลลัพธ์ของทุกสถานการณ์ พบว่าค่าของสัมประสิทธิ์การตัดสินใจ (R^2) มีค่าสูง คือ ซึ่งสามารถนำสมการของตัวแบบการถดถอยพหุคูณที่มีผลอันตรกิริยานี้ไปใช้ประมาณหาค่าจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่น ๆ ได้ สมการการถดถอยพหุคูณ แสดงได้ดังนี้

Percent of $\hat{c} = 114.638 - (0.176)(p) - (0.004)(n) - (1.279)(a) - (0.142)(M) + (0.003)(ap) + (8.46E-05)(an) + (0.003)(aM) + (1.15E-05)(np) + (3.18E-06)(nM) + (0)(pM) - (2.21E-07)(apn) - (8.05E-06)(apM) - (2.44E-08)(pnM) + (2.13E-10)(apnM)$

5.3 ข้อเสนอแนะ

ผลการวิจัยในครั้งนี้มีข้อเสนอแนะ 2 ด้าน คือ

5.3.1 ด้านการนำไปใช้ประโยชน์

1. เมื่อต้องการหาค่าจุดแบ่งที่เหมาะสมสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบโพรบิตแบบ 2 ประเภท สามารถนำค่าจุดแบ่งนี้ไปใช้ได้ ตามแต่ละสถานการณ์ที่ต้องการศึกษา

2. เมื่อต้องการประมาณหาค่าจุดแบ่งที่เหมาะสมที่สุดในสถานการณ์อื่น ๆ นอกเหนือจากการวิจัยนี้ สามารถนำสมการการถดถอยพหุคูณนี้ไปใช้ประมาณหาค่าจุดแบ่งที่เหมาะสมได้

5.3.2 ด้านการศึกษาวิจัย

1. ในการวิจัยครั้งนี้ได้ทำการศึกษาค่าจุดแบ่งสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบโพรบิตแบบ 2 ประเภทเท่านั้น ในการวิจัยครั้งต่อไปอาจทำการศึกษาค่าจุดแบ่งสำหรับตัวแบบโพรบิตแบบหลายกลุ่ม (Multinomial Probit Model)

2. ในการวิจัยครั้งนี้ได้ทำการศึกษาค่าจุดแบ่งเมื่อข้อมูลตัวแปรอิสระเริ่มต้นมีการแจกแจงแบบยูนิฟอร์มเท่านั้น ในการวิจัยครั้งต่อไปอาจทำการศึกษาค่าจุดแบ่งเมื่อข้อมูลตัวแปรอิสระมีการแจกแจงแบบอื่นๆ เช่น การแจกแจงปกติ เป็นต้น และทำการทดสอบการแจกแจงของตัวแปรอิสระว่ามีการแจกแจงตามเดิมตามที่สนใจศึกษาหรือไม่ เมื่อทำการจำลองข้อมูลแล้ว

รายการอ้างอิง

ภาษาไทย

กัลยา วานิชย์บัญชา. การวิเคราะห์ข้อมูลหลายตัวแปร. กรุงเทพมหานคร : สำนักพิมพ์
จุฬาลงกรณ์มหาวิทยาลัย, 2548.

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

ชนิศวรา ฉัตรแก้ว. การการถดถอยเมื่อตัวแปรตามมีสองลักษณะโดยใช้ตัวแบบความน่าจะเป็น
เชิงเส้น ตัวแบบโพรบิตแบบ 2 ประเภทและตัวแบบโลจิท. วิทยานิพนธ์ปริญญา
มหาบัณฑิต, ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2543.

ภาษาอังกฤษ

Hadjicostas, P. Maximizing Proportions of Correct Classifications in Binary Logistic
Regression. Journal of Applied Statistics 33(2006): 629-640.

Robert, S., and Daniel, L. Econometric Model and Economic Forecasts, New York:
McGraw-Hill Book Company, 1981.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรมการหาค่าของจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลใน ตัวแบบโพรบิตแบบ 2 ประเภท

- Main Program

```
options(digits=20)
```

```
dataD0 <- new("list")
```

```
dataD1 <- new("list")
```

```
dataD2 <- new("list")
```

```
dataD3 <- new("list")
```

```
errorTerm <- new("list") #create error term
```

```
Y0est <- new("list")
```

```
Y1est <- new("list")
```

```
Y2est <- new("list")
```

```
Y3est <- new("list")
```

```
MD0 <- new("list")
```

```
MD1 <- new("list")
```

```
MD2 <- new("list")
```

```
MD3 <- new("list")
```

```
MD0A01 <- new("list")
```

```
MD0A05 <- new("list")
```

```
MD0A09 <- new("list")
```

```
MD1A01 <- new("list")
```

```
MD1A05 <- new("list")
```

```
MD1A09 <- new("list")
```

```
MD2A01 <- new("list")
```

```
MD2A05 <- new("list")
```

```
MD2A09 <- new("list")
```

```
MD3A01 <- new("list")
```

```
MD3A05 <- new("list")
```

```
MD3A09 <- new("list")
```

#Probit

MD0A01P <- new("list")

MD0A05P <- new("list")

MD0A09P <- new("list")

MD1A01P <- new("list")

MD1A05P <- new("list")

MD1A09P <- new("list")

MD2A01P <- new("list")

MD2A05P <- new("list")

MD2A09P <- new("list")

MD3A01P <- new("list")

MD3A05P <- new("list")

MD3A09P <- new("list")

#Probit

MD0A01PMI <- new("list")

MD0A05PMI <- new("list")

MD0A09PMI <- new("list")

MD1A01PMI <- new("list")

MD1A05PMI <- new("list")

MD1A09PMI <- new("list")

MD2A01PMI <- new("list")

MD2A05PMI <- new("list")

MD2A09PMI <- new("list")

MD3A01PMI <- new("list")

MD3A05PMI <- new("list")

MD3A09PMI <- new("list")



คณะสัตวแพทย์
จุฬาลงกรณ์มหาวิทยาลัย

```
#Probit
```

```
MD0A01PAI <- new("list")
```

```
MD0A05PAI <- new("list")
```

```
MD0A09PAI <- new("list")
```

```
MD1A01PAI <- new("list")
```

```
MD1A05PAI <- new("list")
```

```
MD1A09PAI <- new("list")
```

```
MD2A01PAI <- new("list")
```

```
MD2A05PAI <- new("list")
```

```
MD2A09PAI <- new("list")
```

```
MD3A01PAI <- new("list")
```

```
MD3A05PAI <- new("list")
```

```
MD3A09PAI <- new("list")
```

```
#Probit
```

```
MD0A01PCP <- new("list")
```

```
MD0A05PCP <- new("list")
```

```
MD0A09PCP <- new("list")
```

```
MD1A01PCP <- new("list")
```

```
MD1A05PCP <- new("list")
```

```
MD1A09PCP <- new("list")
```

```
MD2A01PCP <- new("list")
```

```
MD2A05PCP <- new("list")
```

```
MD2A09PCP <- new("list")
```

```
MD3A01PCP <- new("list")
```

```
MD3A05PCP <- new("list")
```

```
MD3A09PCP <- new("list")
```

```

#Probit
MD0A01PCER <- new("list")
MD0A05PCER <- new("list")
MD0A09PCER <- new("list")
MD1A01PCER <- new("list")
MD1A05PCER <- new("list")
MD1A09PCER <- new("list")
MD2A01PCER <- new("list")
MD2A05PCER <- new("list")
MD2A09PCER <- new("list")
MD3A01PCER <- new("list")
MD3A05PCER <- new("list")
MD3A09PCER <- new("list")

#INPUT
n <- (ใส่ค่า n ตามต้องการ)
p <- c(1, 2, 3, 4, 5, 6)
b <- 0.1
loop <- 500
seed <- 1
i <- 1 #start loop, check condition
set.seed(seed)
errorTerm <- numError(n, 0, 5, loop)
while(n/p[i]>=20 & i<=length(p)){

#Generate random number
set.seed(seed)
dataD0[[i]] <- numGen(n, p[i], 0, 0.1, loop)
set.seed(seed)
dataD1[[i]] <- numGen(n, p[i], (1/3), 0.1, loop)
set.seed(seed)

```

```

dataD2[[i]] <- numGen(n, p[i], (2/3), 0.1, loop)
set.seed(seed)
dataD3[[i]] <- numGen(n, p[i], 0.99, 0.1, loop)

#Calculation Y (dependent variable)
Y0est[[i]] <- lapply(dataD0[[i]], FUN=tBeta, b)
Y0est[[i]] <- CalcY(Y0est[[i]], errorTerm)
Y1est[[i]] <- lapply(dataD1[[i]], FUN=tBeta, b)
Y1est[[i]] <- CalcY(Y1est[[i]], errorTerm)
Y2est[[i]] <- lapply(dataD2[[i]], FUN=tBeta, b)
Y2est[[i]] <- CalcY(Y2est[[i]], errorTerm)
Y3est[[i]] <- lapply(dataD3[[i]], FUN=tBeta, b)
Y3est[[i]] <- CalcY(Y3est[[i]], errorTerm)

#Merge data and transform to data frame
MD0[[i]] <- JoinData(dataD0[[i]], errorTerm, Y0est[[i]])
MD1[[i]] <- JoinData(dataD1[[i]], errorTerm, Y1est[[i]])
MD2[[i]] <- JoinData(dataD2[[i]], errorTerm, Y2est[[i]])
MD3[[i]] <- JoinData(dataD3[[i]], errorTerm, Y3est[[i]])

#Calculation the Ynew in binary(0,1)
MD0A01[[i]] <- CalcYnew(MD0[[i]], 0.1)
MD0A05[[i]] <- CalcYnew(MD0[[i]], 0.5)
MD0A09[[i]] <- CalcYnew(MD0[[i]], 0.9)
MD1A01[[i]] <- CalcYnew(MD1[[i]], 0.1)
MD1A05[[i]] <- CalcYnew(MD1[[i]], 0.5)
MD1A09[[i]] <- CalcYnew(MD1[[i]], 0.9)
MD2A01[[i]] <- CalcYnew(MD2[[i]], 0.1)
MD2A05[[i]] <- CalcYnew(MD2[[i]], 0.5)
MD2A09[[i]] <- CalcYnew(MD2[[i]], 0.9)
MD3A01[[i]] <- CalcYnew(MD3[[i]], 0.1)

```

```

MD3A05[[i]] <- CalcYnew(MD3[[i]], 0.5)
MD3A09[[i]] <- CalcYnew(MD3[[i]], 0.9)

#Calculation the Ypred from x independent variables
if(i==1){
  #Probit
  MD0A01P[[i]] <- CalcYpred1x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred1x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred1x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred1x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred1x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred1x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred1x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred1x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred1x(MD2A09[[i]], "probit")
  MD3A01P[[i]] <- CalcYpred1x(MD3A01[[i]], "probit")
  MD3A05P[[i]] <- CalcYpred1x(MD3A05[[i]], "probit")
  MD3A09P[[i]] <- CalcYpred1x(MD3A09[[i]], "probit")
}
if(i==2){

  #Probit
  MD0A01P[[i]] <- CalcYpred2x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred2x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred2x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred2x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred2x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred2x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred2x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred2x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred2x(MD2A09[[i]], "probit")
}

```

```

MD3A01P[[i]] <- CalcYpred2x(MD3A01[[i]], "probit")
MD3A05P[[i]] <- CalcYpred2x(MD3A05[[i]], "probit")
MD3A09P[[i]] <- CalcYpred2x(MD3A09[[i]], "probit")
}
if(i==3){
  #Probit
  MD0A01P[[i]] <- CalcYpred3x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred3x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred3x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred3x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred3x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred3x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred3x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred3x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred3x(MD2A09[[i]], "probit")
  MD3A01P[[i]] <- CalcYpred3x(MD3A01[[i]], "probit")
  MD3A05P[[i]] <- CalcYpred3x(MD3A05[[i]], "probit")
  MD3A09P[[i]] <- CalcYpred3x(MD3A09[[i]], "probit")
}
if(i==4){
  #Probit
  MD0A01P[[i]] <- CalcYpred4x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred4x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred4x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred4x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred4x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred4x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred4x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred4x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred4x(MD2A09[[i]], "probit")
  MD3A01P[[i]] <- CalcYpred4x(MD3A01[[i]], "probit")
}

```



```

MD3A05P[[i]] <- CalcYpred4x(MD3A05[[i]], "probit")
MD3A09P[[i]] <- CalcYpred4x(MD3A09[[i]], "probit")
}
if(i==5){
  #Probit
  MD0A01P[[i]] <- CalcYpred5x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred5x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred5x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred5x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred5x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred5x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred5x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred5x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred5x(MD2A09[[i]], "probit")
  MD3A01P[[i]] <- CalcYpred5x(MD3A01[[i]], "probit")
  MD3A05P[[i]] <- CalcYpred5x(MD3A05[[i]], "probit")
  MD3A09P[[i]] <- CalcYpred5x(MD3A09[[i]], "probit")
}
if(i==6){
  #Probit
  MD0A01P[[i]] <- CalcYpred6x(MD0A01[[i]], "probit")
  MD0A05P[[i]] <- CalcYpred6x(MD0A05[[i]], "probit")
  MD0A09P[[i]] <- CalcYpred6x(MD0A09[[i]], "probit")
  MD1A01P[[i]] <- CalcYpred6x(MD1A01[[i]], "probit")
  MD1A05P[[i]] <- CalcYpred6x(MD1A05[[i]], "probit")
  MD1A09P[[i]] <- CalcYpred6x(MD1A09[[i]], "probit")
  MD2A01P[[i]] <- CalcYpred6x(MD2A01[[i]], "probit")
  MD2A05P[[i]] <- CalcYpred6x(MD2A05[[i]], "probit")
  MD2A09P[[i]] <- CalcYpred6x(MD2A09[[i]], "probit")
  MD3A01P[[i]] <- CalcYpred6x(MD3A01[[i]], "probit")
  MD3A05P[[i]] <- CalcYpred6x(MD3A05[[i]], "probit")
}

```

```

MD3A09P[[i]] <- CalcYpred6x(MD3A09[[i]], "probit")
}

```

#Probit: Calculation the M(i)

```

MD0A01PMI[[i]] <- CalcMI(MD0A01P[[i]])
MD0A05PMI[[i]] <- CalcMI(MD0A05P[[i]])
MD0A09PMI[[i]] <- CalcMI(MD0A09P[[i]])
MD1A01PMI[[i]] <- CalcMI(MD1A01P[[i]])
MD1A05PMI[[i]] <- CalcMI(MD1A05P[[i]])
MD1A09PMI[[i]] <- CalcMI(MD1A09P[[i]])
MD2A01PMI[[i]] <- CalcMI(MD2A01P[[i]])
MD2A05PMI[[i]] <- CalcMI(MD2A05P[[i]])
MD2A09PMI[[i]] <- CalcMI(MD2A09P[[i]])
MD3A01PMI[[i]] <- CalcMI(MD3A01P[[i]])
MD3A05PMI[[i]] <- CalcMI(MD3A05P[[i]])
MD3A09PMI[[i]] <- CalcMI(MD3A09P[[i]])

```

#Probit: Calculation the a(i)

```

MD0A01PAI[[i]] <- CalcAI(MD0A01PMI[[i]])
MD0A05PAI[[i]] <- CalcAI(MD0A05PMI[[i]])
MD0A09PAI[[i]] <- CalcAI(MD0A09PMI[[i]])
MD1A01PAI[[i]] <- CalcAI(MD1A01PMI[[i]])
MD1A05PAI[[i]] <- CalcAI(MD1A05PMI[[i]])
MD1A09PAI[[i]] <- CalcAI(MD1A09PMI[[i]])
MD2A01PAI[[i]] <- CalcAI(MD2A01PMI[[i]])
MD2A05PAI[[i]] <- CalcAI(MD2A05PMI[[i]])
MD2A09PAI[[i]] <- CalcAI(MD2A09PMI[[i]])
MD3A01PAI[[i]] <- CalcAI(MD3A01PMI[[i]])
MD3A05PAI[[i]] <- CalcAI(MD3A05PMI[[i]])
MD3A09PAI[[i]] <- CalcAI(MD3A09PMI[[i]])

```

```

#Probit: Calculation the cp
MD0A01PCP[[i]] <- cut.of.point(MD0A01PAI[[i]])
MD0A05PCP[[i]] <- cut.of.point(MD0A05PAI[[i]])
MD0A09PCP[[i]] <- cut.of.point(MD0A09PAI[[i]])
MD1A01PCP[[i]] <- cut.of.point(MD1A01PAI[[i]])
MD1A05PCP[[i]] <- cut.of.point(MD1A05PAI[[i]])
MD1A09PCP[[i]] <- cut.of.point(MD1A09PAI[[i]])
MD2A01PCP[[i]] <- cut.of.point(MD2A01PAI[[i]])
MD2A05PCP[[i]] <- cut.of.point(MD2A05PAI[[i]])
MD2A09PCP[[i]] <- cut.of.point(MD2A09PAI[[i]])
MD3A01PCP[[i]] <- cut.of.point(MD3A01PAI[[i]])
MD3A05PCP[[i]] <- cut.of.point(MD3A05PAI[[i]])
MD3A09PCP[[i]] <- cut.of.point(MD3A09PAI[[i]])
print(paste(n, p[i])) #for checking my program
i <- i+1
}

# Generalized Linear Models, Link: Probit
slcList <- length(MD1A01PCP)
slcList <- ifelse(slcList<2, 2, slcList)
MD0A01Pdt <- unlistMatrix(MD0A01PCP, loop)
MD0A05Pdt <- unlistMatrix(MD0A05PCP, loop)
MD0A09Pdt <- unlistMatrix(MD0A09PCP, loop)
MD1A01Pdt <- unlistMatrix(MD1A01PCP, loop)[2:slcList]
MD1A05Pdt <- unlistMatrix(MD1A05PCP, loop)[2:slcList]
MD1A09Pdt <- unlistMatrix(MD1A09PCP, loop)[2:slcList]
MD2A01Pdt <- unlistMatrix(MD2A01PCP, loop)[2:slcList]
MD2A05Pdt <- unlistMatrix(MD2A05PCP, loop)[2:slcList]
MD2A09Pdt <- unlistMatrix(MD2A09PCP, loop)[2:slcList]
MD3A01Pdt <- unlistMatrix(MD3A01PCP, loop)[2:slcList]
MD3A05Pdt <- unlistMatrix(MD3A05PCP, loop)[2:slcList]

```

```
MD3A09Pdt <- unlistMatrix(MD3A09PCP, loop)[2:slcList]
```

```
if(n!=20){
```

```
  #Probit: Calculation percent of C0
```

```
  MD0A01PC0 <- lapply(MD0A01Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD0A05PC0 <- lapply(MD0A05Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD0A09PC0 <- lapply(MD0A09Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD1A01PC0 <- lapply(MD1A01Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD1A05PC0 <- lapply(MD1A05Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD1A09PC0 <- lapply(MD1A09Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD2A01PC0 <- lapply(MD2A01Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD2A05PC0 <- lapply(MD2A05Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD2A09PC0 <- lapply(MD2A09Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD3A01PC0 <- lapply(MD3A01Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD3A05PC0 <- lapply(MD3A05Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  MD3A09PC0 <- lapply(MD3A09Pdt,FUN=function(x){mean(x[,1])*100})
```

```
  #Probit: Confidence Interval
```

```
  MD0A01PCI <- lapply(MD0A01Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD0A05PCI <- lapply(MD0A05Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD0A09PCI <- lapply(MD0A09Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD1A01PCI <- lapply(MD1A01Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD1A05PCI <- lapply(MD1A05Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD1A09PCI <- lapply(MD1A09Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD2A01PCI <- lapply(MD2A01Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD2A05PCI <- lapply(MD2A05Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD2A09PCI <- lapply(MD2A09Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD3A01PCI <- lapply(MD3A01Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD3A05PCI <- lapply(MD3A05Pdt,FUN=function(x){percentile(x[,1])})
```

```
  MD3A09PCI <- lapply(MD3A09Pdt,FUN=function(x){percentile(x[,1])})
```

```
pTemp <- c(rep(c(1:(n/20)),3), rep(c(2:(n/20)),9))
```

```

nTemp <- rep(n, length(pTemp))

rep(c(rep(0.1,(n/20-1)), rep(0.5,(n/20-1)), rep(0.9,(n/20-1))),3))
mTemp <- c(rep(0, (3*(n/20))), rep(0.33, (3*(n/20-1))),
           rep(0.67, (3*(n/20-1))), rep(0.99, (3*(n/20-1))))

#Probit
perTempP <- c(unlist(MD0A01PC0),unlist(MD0A05PC0),unlist(MD0A09PC0),
             unlist(MD1A01PC0),unlist(MD1A05PC0),unlist(MD1A09PC0),
             unlist(MD2A01PC0),unlist(MD2A05PC0),unlist(MD2A09PC0),
             unlist(MD3A01PC0),unlist(MD3A05PC0),unlist(MD3A09PC0))

DataFinalP <- matrix(NA, length(pTemp), 5)
colnames(DataFinalP) <- c("n","p","m","a","percent")
DataFinalP[,1] <- nTemp
DataFinalP[,2] <- pTemp
DataFinalP[,3] <- mTemp
DataFinalP[,4] <- aTemp
DataFinalP[,5] <- perTempP

#Confidence Interval
CIPMatrix <- matrix(c(unlist(MD0A01PCI),unlist(MD0A05PCI),unlist(MD0A09PCI),
                    unlist(MD1A01PCI),unlist(MD1A05PCI),unlist(MD1A09PCI),
                    unlist(MD2A01PCI),unlist(MD2A05PCI),unlist(MD2A09PCI),
                    unlist(MD3A01PCI),unlist(MD3A05PCI),unlist(MD3A09PCI)),
                    length(pTemp), 2, byrow=T)
colnames(CIPMatrix) <- c("CI.Lower","CI.Upper")
allInfProbit <- cbind(DataFinalP, CIPMatrix)
}

if(n==20){

#Probit: Calculation percent of C0
MD0A01PC0 <- lapply(MD0A01Pdt,FUN=function(x){mean(x[,1])*100})

```

```

MD0A05PC0 <- lapply(MD0A05Pdt,FUN=function(x){mean(x[,1])*100})
MD0A09PC0 <- lapply(MD0A09Pdt,FUN=function(x){mean(x[,1])*100})
#Probit: Confidence Interval
MD0A01PCI <- lapply(MD0A01Pdt,FUN=function(x){percentile(x[,1])})
MD0A05PCI <- lapply(MD0A05Pdt,FUN=function(x){percentile(x[,1])})
MD0A09PCI <- lapply(MD0A09Pdt,FUN=function(x){percentile(x[,1])})

pTemp <- rep(1,3)
nTemp <- rep(n, 3)
aTemp <- c(0.1,0.5,0.9)
mTemp <- rep(0,3)

#Probit
perTempP <- c(unlist(MD0A01PC0),unlist(MD0A05PC0),unlist(MD0A09PC0))
DataFinalP <- matrix(NA, length(pTemp), 5)
colnames(DataFinalP) <- c("n","p","m","a","percent")
DataFinalP[,1] <- nTemp
DataFinalP[,2] <- pTemp
DataFinalP[,3] <- mTemp
DataFinalP[,4] <- aTemp
DataFinalP[,5] <- perTempP
#Confidence Interval
CIPMatrix <- matrix(c(unlist(MD0A01PCI),unlist(MD0A05PCI),unlist(MD0A09PCI)),
  length(pTemp), 2, byrow=T)
colnames(CIPMatrix) <- c("CI.Lower","CI.Upper")
allInfProbit <- cbind(DataFinalP, CIPMatrix)
}

#Export data into Excel
write.csv(allInfProbit, paste("C:/ProbitALL","_n",n,".csv",sep=""), row.names=F)
write.csv(DataFinalP, paste("C:/ProbitDataInf","_n",n,".csv",sep=""), row.names=F)

```

- Create matrix with the correlation value

```

cor.matrix <- function(p=1, corv=0){
  m1 <- 1
  m2 <- matrix(c(1, corv, corv, 1), 2, 2)
  dimnames(m2) <- list(c("x1","x2"), c("x1","x2"))
  m3 <- matrix(c(1, corv, corv^2, corv, 1, corv, corv^2, corv, 1), 3, 3)
  dimnames(m3) <- list(c("x1","x2","x3"), c("x1","x2","x3"))
  m4 <- matrix(c(1, corv, corv^2, corv^3, corv, 1, corv, corv^2, corv^2,
    corv, 1, corv, corv^3, corv^2, corv, 1), 4, 4)
  dimnames(m4) <- list(c("x1","x2","x3","x4"), c("x1","x2","x3","x4"))
  m5 <- matrix(c(1, corv, corv^2, corv^3, corv^4, corv, 1, corv, corv^2,
    corv^3, corv^2, corv, 1, corv, corv^2, corv^3, corv^2, corv,
    1, corv, corv^4, corv^3, corv^2, corv, 1), 5, 5)
  dimnames(m5) <- list(c("x1","x2","x3","x4","x5"), c("x1","x2","x3",
    "x4","x5"))
  m6 <- matrix(c(1, corv, corv^2, corv^3, corv^4, corv^5, corv, 1, corv,
    corv^2, corv^3, corv^4, corv^2, corv, 1, corv, corv^2, corv^3,
    corv^3, corv^2, corv, 1, corv, corv^2, corv^4, corv^3, corv^2,
    corv, 1, corv, corv^5, corv^4, corv^3, corv^2, corv, 1), 6, 6)
  dimnames(m6) <- list(c("x1","x2","x3","x4","x5","x6"), c("x1","x2",
    "x3","x4","x5","x6"))
  cor.mtx <- list(m1, m2, m3, m4, m5, m6)
  output <- cor.mtx[[p]]
  return(output)
}

```

- Number generator with correlation matrix

```

numGen <- function(n, p=1, cor.value, error=0.01, nloop=1){
  np <- p
  cv <- cor.value

```



```

MC <- cor.matrix(np, cv)
chol.matr <- chol(MC)
num <- new("list")
for(i in 1:nloop){
  stop.loop <- 0
  while(stop.loop<1){
    num[[i]] <- matrix(runif(n*p, -(n/2), n/2), n)
    num[[i]] <- num[[i]]%*%chol.matr
    if(max(abs(cor(num[[i]])-MC))<=error){
      stop.loop <- stop.loop+1
    }
    stop.loop
  }
}
output <- num
return(output)
}

```

- Generator the error term

```

numError <- function(n, Mean=0, Var=1, nloop=1){
  num <- new("list")
  for(i in 1:nloop){
    num[[i]] <- rnorm(n, Mean, Var)
  }
  output <- num
  return(output)
}

```

- Calculation Y, $Y=b_0+b_1X_1+b_2X_2+\dots+b_nX_n$

```

#tBeta <- function(x, b){
#  b+apply(rep(b, ncol(x))*x, 1, sum)

```

```
#}
tBeta <- function(x, b0=1, b1=0.15, By=0.15){
  b0+apply(seq(b1, By, length.out=ncol(x))*x, 1, sum)
}
```

- Calculation dependent variable

```
CalcY <- function(x, y){
  Length <- length(x)
  Y.calc <- new("list")
  for(i in 1:Length){
    Y.calc[[i]] <- x[[i]]+y[[i]]
  }
  output <- Y.calc
  return(output)
}
```

- Merge data and transform to data frame

```
JoinData <- function(x, y, z){
  Length <- length(x)
  DataTemp <- new("list")
  for(i in 1:Length){
    DataTemp[[i]] <- cbind(x[[i]], y[[i]], z[[i]])
    DataTemp[[i]] <- as.data.frame(DataTemp[[i]][order(DataTemp[[i]]
      [,ncol(DataTemp[[i]])]),])
  }
  output <- DataTemp
  return(output)
}
```

- Calculation the y in binary

```
CalcYnew <- function(dataProg, Prob){
```

```

Length <- length(dataProg)
Ynew <- new("list")
for(i in 1:Length){
  Ynew <- rep(1, nrow(dataProg[[i]]))
  dataProg[[i]]$Ynew <- Ynew
  dataProg[[i]]$Ynew[1:(Prob*nrow(dataProg[[i]]))] <- 0
}
output <- dataProg
return(output)
}

```

- Calculation the Ypred from 1 independent variables

```

CalcYpred1x <- function(dataProg, Link="logit"){
  Length <- length(dataProg)
  Ypred <- new("list")
  for(i in 1:Length){
    dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
      (ncol(dataProg[[i]])-2))]
    Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
      dataProgTr[,1], family=binomial(link=Link),
      data=dataProgTr)))
    dataProg[[i]]$Ypred <- Ypred[[i]]
  }
  output <- dataProg
  return(output)
}

```

- Calculation the Ypred from 2 independent variables

```

CalcYpred2x <- function(dataProg, Link="logit"){
  Length <- length(dataProg)

```

```

Ypred <- new("list")
for(i in 1:Length){
  dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
    (ncol(dataProg[[i]])-2))]
  Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
    dataProgTr[,1]+dataProgTr[,2], family=binomial(link=Link),
    data=dataProgTr)))
  dataProg[[i]]$Ypred <- Ypred[[i]]
}
output <- dataProg
return(output)
}

```

- Calculation the Ypred from 3 independent variables

```

CalcYpred3x <- function(dataProg, Link="logit"){
  Length <- length(dataProg)
  Ypred <- new("list")
  for(i in 1:Length){
    dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
      (ncol(dataProg[[i]])-2))]
    Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
      dataProgTr[,1]+dataProgTr[,2]+dataProgTr[,3], family=binomial(link=Link),
      data=dataProgTr)))
    dataProg[[i]]$Ypred <- Ypred[[i]]
  }
  output <- dataProg
  return(output)
}

```

- Calculation the Ypred from 4 independent variables

```

CalcYpred4x <- function(dataProg, Link="logit"){

```

```

Length <- length(dataProg)
Ypred <- new("list")
for(i in 1:Length){
  dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
    (ncol(dataProg[[i]])-2))]
  Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
    dataProgTr[,1]+dataProgTr[,2]+dataProgTr[,3]+dataProgTr[,4],
    family=binomial(link=Link), data=dataProgTr)))
  dataProg[[i]]$Ypred <- Ypred[[i]]
}
output <- dataProg
return(output)
}

```

- Calculation the Ypred from 5 independent variables

```

CalcYpred5x <- function(dataProg, Link="logit"){
  Length <- length(dataProg)
  Ypred <- new("list")
  for(i in 1:Length){
    dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
      (ncol(dataProg[[i]])-2))]
    Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
      dataProgTr[,1]+dataProgTr[,2]+dataProgTr[,3]+dataProgTr[,4]+
      dataProgTr[,5], family=binomial(link=Link), data=dataProgTr)))
    dataProg[[i]]$Ypred <- Ypred[[i]]
  }
  output <- dataProg
  return(output)
}

```

- Calculation the Ypred from 6 independent variables

```

CalcYpred6x <- function(dataProg, Link="logit"){

```

```

Length <- length(dataProg)
Ypred <- new("list")
for(i in 1:Length){
  dataProgTr <- dataProg[[i]][,-c((ncol(dataProg[[i]])-1),
    (ncol(dataProg[[i]])-2))]
  Ypred[[i]] <- as.vector(fitted(glm(dataProgTr[,ncol(dataProgTr)]~
    dataProgTr[,1]+dataProgTr[,2]+dataProgTr[,3]+dataProgTr[,4]+
    dataProgTr[,5]+dataProgTr[,6], family=binomial(link=Link),
    data=dataProgTr)))
  dataProg[[i]]$Ypred <- Ypred[[i]]
}
output <- dataProg
return(output)
}

```

- Calculation the M(i)

```

CalcMI <- function(dataProg){
  Length <- length(dataProg)
  mRank <- new("list")
  for(i in 1:Length){
    dataProg[[i]] <- dataProg[[i]][order(dataProg[[i]]$Ypred),]
    mRank[[i]] <- rank(dataProg[[i]]$Ypred)
    mTemp1 <- cbind(as.vector(unlist(mRank[[i]])), order(order(unlist(mRank[[i]]))))
    rewRank <-
as.numeric(names(table(unlist(mRank[[i]]))[table(unlist(mRank[[i]])!=1]))
    if(length(rewRank)!=0){
      mTemp2 <- mTemp1[!is.na(match(mTemp1[,1], rewRank)),]
      newRank <- aggregate(mTemp2[,2], list(mTemp2[,1]), FUN=max)
      newRank <- newRank[match(mTemp1[,1], newRank[,1]), 2]
      mRank[[i]][!is.na(newRank)] <- newRank[!is.na(newRank)]
      dataProg[[i]]$MI <- mRank[[i]]
    }
  }
}

```

```

    }
    else{
      dataProg[[i]]$MI <- mRank[[i]]
    }
  }
  output <- dataProg
  return(output)
}

```

- Calculation the a(i)

```

CalcAI <- function(dataProg){
  Length <- length(dataProg)
  AI <- new("list")
  for(i in 1:Length){
    MI <- c(0, dataProg[[i]]$MI)
    Yn <- c(dataProg[[i]]$Ynew, NA)
    iTemp <- c(0:(length(MI)-1))
    Co <- rep(-1, length(MI))

    Case1st <- c(iTemp[-length(MI)]+1<=MI[-length(MI)], NA)
    Case2nd <- c(MI[-length(MI)]<iTemp[-length(MI)]+1, NA)
    Pos1st <- rep(NA, length(MI))
    Pos2nd <- rep(NA, length(MI))
    Pos1st[which(!is.na(Case2nd) & Case2nd==1)] <- MI[which(!is.na(Case2nd) &
    Case2nd==1)]+1
    Pos2nd[which(!is.na(Case2nd) & Case2nd==1)+1] <- MI[which(!is.na(Case2nd) &
    Case2nd==1)+1]

    dataTemp <- as.data.frame(cbind(Yn, MI, iTemp, Co, Case1st, Case2nd, Pos1st,
    Pos2nd))

    AI[[i]] <- c(0, rep(NA, length(MI)))
    for(j in 2:(length(AI[[i]])-1)){

```



```

if(Case2nd[j-1]==1){
  AI[[i]][j] <- AI[[i]][j-1]+sum(Co[Pos1st[j-1]:Pos2nd[j+1-1]]^Yn[Pos1st[j-
1]:Pos2nd[j+1-1]])
}
else{
  AI[[i]][j] <- AI[[i]][j-1]
}
}
AI[[i]] <- AI[[i]][2:(length(AI[[i]))-1)]
dataProg[[i]]$AI <- AI[[i]]
}
output <- dataProg
return(output)
}

```

- Calculation the cut of point

```

cut.of.point <- function(dataProg){
  Length <- length(dataProg)
  lower <- new("list")
  upper <- new("list")
  for(i in 1:Length){
    ai <- dataProg[[i]]$AI
    y.pred <- dataProg[[i]]$Ypred
    lower[[i]] <- ifelse(which(ai==max(ai))[1]==length(ai),
      y.pred[which(ai==max(ai))[1]],
      ifelse(which(ai==max(ai))==0, 0,
        y.pred[which(ai==max(ai))[1]]))
    upper[[i]] <- ifelse(which(ai==max(ai))[1]==length(ai) | max(which(
      ai==max(ai))==length(ai), 1,
      ifelse(which(ai==max(ai))[1]==0, y.pred[1],
        y.pred[max(which(ai==max(ai))+1]))

```

```

}
output <- list(lower, upper)
names(output) <- c("lower","upper")
return(output)
}
cer.value <- function(dataProg){
  Length <- length(dataProg)
  CER <- new("list")
  for(i in 1:Length){
    nDt <- length(dataProg[[i]]$AI)
    posLow <- which(dataProg[[i]]$AI==max(dataProg[[i]]$AI))[1]
    nX1 <- sum(dataProg[[i]]$Ynew[1:posLow]==0)
    nX2 <- sum(dataProg[[i]]$Ynew[posLow:nDt]==1)
    CER[[i]] <- (nX1+nX2)/nDt
  }
  output <- CER
  return(output)
}

```

- Create new dataset, unlist to matrix

```

unlistMatrix <- function(dataProg, Loop){
  Length <- length(dataProg)
  dataMt <- new("list")
  for(i in 1:Length){
    dataMt[[i]] <- matrix(unlist(dataProg[[i]]), Loop, length(dataProg[[i]]))
    colnames(dataMt[[i]]) <- c("LOWER", "UPPER")
  }
  output <- dataMt
  return(output)
}

```

- Get data from percentile rank

```
#ROUND(((500+1)*2.5)/100,0)
percentile <- function(dataProg, alpha=0.05){
  dataProg <- dataProg[order(dataProg)]
  posL <- round((((length(dataProg)+1)*((alpha/2)*100))/100),0)
  posL <- ifelse(posL<1, 1, posL)
  posU <- round((((length(dataProg)+1)*((1-(alpha/2))*100))/100),0)
  posU <- ifelse(posU>length(dataProg), length(dataProg), posU)
  LCI <- dataProg[posL]
  UCI <- dataProg[posU]
  output <- matrix(c(LCI, UCI),1,2)
  colnames(output) <- c("LOWER", "UPPER")
  return(output)
}
```



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวชลลดา กล้วยไม้ เกิดเมื่อวันที่ 11 เมษายน พุทธศักราช 2528 สำเร็จการศึกษา
ระดับปริญญาวิทยาศาสตรบัณฑิต สาขาสถิติ จากภาควิชาสถิติ คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยธรรมศาสตร์ ในปีการศึกษา 2550 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตร
มหาบัณฑิต สาขาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์
มหาวิทยาลัย ในปีการศึกษา 2551



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย