การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง

นางสาวนิภาพรรณ ไพจินดา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย
บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)
are the thesis authors' files submitted through the Graduate School.

THE CUT-OFF POINT ESTIMATION OF BINARY LOGISIC REGRESSION MODEL FOR PREDICTIVE CLASSIFICATION USING COMPLEMENTARY LOG-LOG FUNCTION AS A LINK FUNCTION

Miss Nipapan Paijinda

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 2011

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง
โดย	นางสาวนิภาพรรณ ไพจินดา
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะพาณิชยศาสตร์และการบัญชี
(รองศาสตราจารย์ ดร.พสุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

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

.....กรรมการ
(อาจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี)

.....กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.อรุณี กำลัง)

นิภาพรรณ ไพจินดา : การหาจุดแบ่งของตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับการ พยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง (THE CUT–OFF POINT ESTIMATION OF BINARY LOGISTIC REGRESSION MODEL FOR PREDICTIVE CLASSIFICATION USING COMPLEMENTARY LOG-LOG FUNCTION AS A LINK FUNCTION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร.สุพล ดุรงค์วัฒนา , 105 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาจุดแบ่งที่เหมาะสมที่สุดของตัวแบบถดถอยโลจิสติก 2 ประเภทสำหรับการพยากรณ์การจำแนกข้อมูลโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง เพื่อพิจารณาว่าปัจจัยอื่นๆมีผลต่อการคัดเลือกจุดแบ่งที่ 0.5 หรือไม่ ปัจจัยที่สนใจศึกษาในงานวิจัยครั้งนี้ คือ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a) เท่ากับ 0.25,0.5 และ 0.75 จำนวนตัวแปรอิสระ(p) คือ1,2,3,4 และ 5 ขนาดของตัวอย่าง(n) คือ 50,100,150,200 และ 250 ระดับความสัมพันธ์ของตัวแปรอิสระ 3 ระดับ คือ ต่ำ กลาง สูง ทำการจำลองข้อมูลทั้งหมดนี้โดยเทคนิคมอนติคาร์โลด้วยโปรแกรม R และทำการคำนวณหาจุดแบ่งโดยทฤษฎีของ ผลการวิจัยสรุปได้ดังนี้

กรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่ พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งไม่เกินค่า 0.5 สลับสูงต่ำไปมา และที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มขึ้น แต่ไม่มีค่าใดเกินค่า 0.5 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่ พบว่าที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 2 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งไม่เกินค่า 0.5 สลับสูงต่ำไปมา และที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มขึ้น แต่ไม่เกินค่า 0.5 กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่ พบว่า ที่ทุกขนาดตัวอย่างจะให้ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง และสุดท้ายกรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลงแต่ปัจจัยอื่นคงที่พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวจะให้ค่าเฉลี่ยของจุดแบ่งคงที่แต่ไม่เกิน 0.5 และ ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ค่าเฉลี่ยของจุดแบ่งส่วนใหญ่มีค่าเพิ่มสูงขึ้น

ภาควิชา……สถิติ……..…. ลายมือชื่อนิสิต…….……..…..…..….

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

ปีการศึกษา……. 2554……

5281834026 : MAJOR STATISTICS

KEYWORDS : CLASSIFICATION ERROR RATE / CUT-OFF POINT / LOGISTIC REGRESSION / COMPLEMENTARY LOG-LOG FUNCTION

NIPAPAN PAIJINDA : THE CUT-OFF POINT ESTIMATION OF BINARY LOGISIC REGRESSION MODEL FOR PREDICTIVE CLASSIFICATION USING COMPLEMENTARY LOG-LOG FUNCTION AS A LINK FUNCTION. ADVISOR : ASSOC.PROF.SUPOL DURONGWATANA,Ph.D, 105 pp.

The objective of this study is to find out the cut-off point estimation of binary logistic regression model for predictive classification using complementary log-log function as a link function. The interesting factors are failure rate (a) of the values 0.25,0.5, and 0.75 ,number of independent variables (p) are 1,2,3,4, and 5 ,sample size (n) are 50,100,150,200 and 250. Degrees of multicollinearity among independent variables with 3 levels are low, medium, and high level. The data are generated using Monte Carlo technique through R-program. The results are summarized as follow.

As the failure rates change while keeping other factors constant, only one number of independent variable gives the mean value of the cut-off points not exceeding 0.5 in the switching manner. When the number of independent variables increase to 2,3,4, and 5, the mean value of the cut-off points increases as well but none of them exceed 0.5. As the sample size changes and the other factors are kept constant, the number of independent variables of 1 and 2 gives the mean value of the cut-off points not exceeding 0.5; also switching. When the number of independent variables increases to 3,4, and 5, most of the mean values of the cut-off points increase but not exceeding the value of 0.5. As the number of the independent variables change and the other factors are kept constant, the mean value of the cut-off points is found to be increased at every level of the sample sizes. Finally, as the failure rates change when the other factors are kept constant; at the number of independent variable equal to 1, the mean value of the cut-off points is found to be constant not exceeding 0.5. When the number of independent variables increase to 2,3,4, and 5, the mean value of the cut-off points increase.

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

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

Academic Year : 2011

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยความกรุณาและความช่วยเหลืออย่างดียิ่งจาก รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาแนะนำให้คำปรึกษาตลอดจนช่วยตรวจสอบ แก้ไขข้อบกพร่องต่างๆ จนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความเคารพและซาบซึ้งเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอกราบขอบพระคุณ รองศาตราจารย์ ดร. ธีระพร วีระถาวร ประธานกรรมการ และ อาจารย์ ดร. อนุภาพ สมบูรณ์สวัสดี กรรมการ ที่ให้ความช่วยเหลือรวมถึงการให้คำแนะนำตรวจสอบโปรแกรมในการทำการวิจัยครั้งนี้ ซึ่งทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์

ผู้วิจัยขอกราบขอบพระคุณ อาจารย์ ดร.อรุณี กำลัง ที่ท่านเสียสละเวลาอันมีค่ามาเป็นกรรมการภายนอกมหาวิทยาลัย ช่วยตรวจสอบวิทยานิพนธ์ ซึ่งทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลงได้

สุดท้ายนี้ผู้วิจัย ขอขอบพระคุณพระเจ้า ผู้ทรงประทานสติปัญญาและการอวยพรให้กับชีวิตของผู้วิจัย ขอบพระคุณมารดา ผู้ให้การสนับสนุนทั้งทุนการศึกษา ชีวิต กำลังใจ และการอดทนรอคอยจนสำเร็จการศึกษา ขอบคุณเพื่อนร่วมชั้นเรียนทุกท่าน ที่คอยให้ความช่วยเหลือและคำปรึกษาในการเรียนจนจบการศึกษา ขอบคุณพี่น้องในคริสตจักรสำหรับกำลังใจและคำอธิษฐานเผื่อ ขอพระเจ้าอวยพระพรแด่ทุกท่าน

สารบัญ

หน้า

บทคัดย่อภาษาไทย…… ง

บทคัดย่อภาษาอังกฤษ…… จ

กิตติกรรมประกาศ…… ฉ

สารบัญ…… ช

สารบัญตาราง…… ฌ

สารบัญภาพ…… ญ

บทที่

1 บทนำ…… 1

1.1 ความเป็นมาและความสำคัญของปัญหา…… 1

1.2 วัตถุประสงค์ของการวิจัย…… 2

1.3 ขอบเขตของการวิจัย…… 2

1.4 ข้อตกลงเบื้องต้น…… 3

1.5 คำจำกัดความที่ใช้ในการวิจัย…… 4

1.6 ประโยชน์ที่คาดว่าจะได้รับ…… 4

2 เอกสารและงานวิจัยที่เกี่ยวข้อง…… 5

แนวคิดและทฤษฎี…… 5

2.1 ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท…… 5

2.2 ฟังก์ชันคอมพลีเมนทารีล็อก ล็อก…… 7

2.2.1 เปรียบเทียบตัวแบบคอมพลีเมนทารีล็อก ล็อก โลจิท และโพรบิท…. 8

2.3 ฟังก์ชันภาวะน่าจะเป็นของข้อมูลการถดถอยโลจิสติกแบบ 2 ประเภท…… 10

2.4 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด…… 11

2.5 ช่วงความเชื่อมั่น…… 11

2.6 เทคนิคมอนติคาร์โล…… 12

2.7 วิธีการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับตัวแบบถดถอยโลจิสติกแบบ 2 ประเภท…… 12

3 วิธีดำเนินการวิจัย…… 14

3.1 แผนการดำเนินการวิจัย…… 14

3.2 ขั้นตอนในการวิจัย…… 15

3.3 ขั้นตอนการทำงานของโปรแกรม…… 18

บทที่		หน้า
4	ผลการวิเคราะห์ข้อมูล……..	19
	4.1 แสดงค่าเฉลี่ยของจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสมในแต่ละสถานการณ์……..……..	20
	4.1.1 กรณีที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่…….……..	21
	4.1.2 กรณีที่ขนาดของตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ อัตราส่วนความล้มเหลวของลักษณะที่สนใจ และ ระดับความสัมพันธ์ของตัวแปรอิสระคงที่…….....	41
	4.1.3 กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ และ อัตราส่วนความล้มเหลวของลักษณะที่สนใจคงที่……..……...	56
	4.1.4 กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงเมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระและอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาคงที่…….….....	73
5	สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ…….……..	93
	5.1 สรุปผลการวิจัย……	93
	5.1.1 กรณีอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเพิ่มขึ้น.....	94
	5.1.2 กรณีที่ขนาดตัวอย่างเพิ่มขึ้น…….…….	95
	5.1.3 กรณีที่จำนวนของตัวแปรอิสระเพิ่มขึ้น……..….	96
	5.1.4 กรณีที่ระดับความสัมพันธ์ของตัวแปรอิสระเพิ่มขึ้น……..….	97
	5.2 ข้อเสนอแนะ……	99
รายการอ้างอิง……		100
ภาคผนวก…….……..		101
ประวัติผู้เขียนวิทยานิพนธ์……..…...		105

สารบัญตาราง

ตารางที่ หน้า

4.1 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ มีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 21

4.2 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 22

4.3 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 23

4.4 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 24

4.5 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 25

4.6 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 26

4.7 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมี ขนาดตัวอย่าง(n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 27

ตารางที่ หน้า

4.8 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 28

4.9 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 29

4.10 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 30

4.11 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 31

4.12 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 32

4.13 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 33

4.14 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)..... 34

ตารางที่ หน้า

4.15 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)…….. 35

4.16 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 41

4.17 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 43

4.18 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 44

4.19 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 45

4.20 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 47

4.21 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 48

ตารางที่ หน้า

4.22 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 49

4.23 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับต่ำ $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 51

4.24 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิรสะ(ρ)ในระดับกลาง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)…….. 52

4.25 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)…….. 56

4.26 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)…….. 57

4.27 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n) คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p) …….. 59

4.28 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p) …….. 60

ตารางที่ หน้า

4.29 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p) 62

4.30 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)..... 63

4.31 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p) 65

4.32 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง$0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)..... 66

4.33 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p) 68

4.34 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 50 อัตราส่วนความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 73

4.35 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 50 อัตราส่วนความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 74

4.36 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 50 อัตราส่วนความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 75

ตารางที่ หน้า

4.37 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 100 อัตราส่วนความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 76

4.38 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 100 อัตราส่วนความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 77

4.39 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 100 อัตราส่วนความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 78

4.40 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 150 อัตราส่วนความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 79

4.41 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 150 อัตราส่วนความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 80

4.42 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 150 อัตราส่วนความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 81

4.43 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 200 อัตราส่วนความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 82

4.44 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 200 อัตราส่วนความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 83

4.45 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 200 อัตราส่วนความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 84

4.46 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 250 อัตราส่วนความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ) 85

ตารางที่ หน้า

4.47 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 250 อัตราส่วนความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 86

4.48 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง(n) เท่ากับ 250 อัตราส่วนความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)..... 87

สารบัญภาพ

ภาพที่		หน้า
2.1	กราฟของ π_i สำหรับตัวแบบคอมพลีเมนทารีล็อก-ล็อก…….....…….....….	7
3.1	แสดงขั้นตอนการทำงานของโปรแกรม…….....…….	18
4.1	แสดงค่าของจุดแบ่งกรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่…….....……..	36
4.2	แสดงค่าของจุดแบ่งกรณีที่ขนาดของตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจ และระดับความสัมพันธ์ของตัวแปรอิสระคงที่…..……	53
4.3	แสดงค่าของจุดแบ่งกรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง เมื่อ ขนาดของตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ และอัตราส่วนความล้มเหลวของลักษณะที่สนใจคงที่…….....……..	70
4.4	แสดงค่าของจุดแบ่งกรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงเมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาคงที่…….....……	88

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ตัวแบบถดถอยโลจิสติกแบบ 2 กลุ่ม ถูกนำมาใช้ในการพยากรณ์ ตัวแปรตาม เชิงคุณภาพซึ่งมีลักษณะสิ่งที่สนใจเป็นความสำเร็จ และความล้มเหลว โดยใช้ ตัวแปรอิสระเป็นตัวพยากรณ์

ในการทำนายว่าสถานะของบริษัทใกล้จะล้มละลายหรือไม่ หรือ นักวิจัยทางการแพทย์ต้องการพยากรณ์ผู้ป่วยของเขาว่า เสี่ยงต่อการเป็นโรคหรือไม่ โดยการใช้ตัวแปรเชิงปริมาณ บางตัว ตัวอย่างเช่น ความดันโลหิต ระดับคลอเรสเตอรอล ระดับน้ำตาลในเลือด เป็นต้น เป็นตัวพยากรณ์ทางการแพทย์เช่นกัน เกือบทั้งหมดของปัญหาและงานวิจัยจะดำเนินการในการทำนายผลอย่างใดอย่างหนึ่งของตัวแปรตาม โดยใช้กลุ่มของตัวแปรอิสระ โดยปกติมักจะใช้ ระดับที่ 0.5 ในการหาจุดแบ่ง (cut off point) โดยที่ไม่ได้สนใจการแจกแจงของข้อมูลของตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของข้อมูลสองกลุ่ม และความสัมพันธ์ของตัวแปรอิสระ

จุดแบ่งนี้ถูกใช้มาเป็นเวลานาน แต่คำถามคือ เมื่อไหร่ที่จุดแบ่ง(cut off point) ที่ 0.5 นี้จะเหมาะสมที่สุดสำหรับ การจัดหมวดหมู่สำหรับ ข้อมูลหนึ่งในสองกลุ่ม ซึ่งเป็นสิ่งที่ทำให้สนใจศึกษางานวิจัยชิ้นนี้

เกือบทั้งหมดของงานวิจัย จะดำเนินการในการจำแนกบุคคล วิชา หรือ วัตถุ เป็นหนึ่งในสองประเภท โดยการใช้ตัวแปรอิสระ โดยปกติจะใช้จุดแบ่ง (cut off point) หรือ ระดับความน่าจะเป็น ดังนั้นอัตราข้อผิดพลาดในการจัดหมวดหมู่ต้องมีค่าน้อยที่สุด หรือจุดแบ่ง (cut off point) ที่ทำให้ข้อผิดพลาดในการจัดหมวดหมู่มีค่าน้อย เช่นเดียวกับ จุดแบ่งที่ไม่คำนึงถึงการแจกแจงของตัวแปรอิสระ จำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของความสำเร็จและความล้มเหลว ของข้อมูลที่ตั้งไว้ และระดับของ ความสัมพันธ์เชิงพหุระหว่าง ตัวแปรอิสระ

ซึ่งตัวแบบการถดถอยโลจิสติก มีลักษณะความสัมพันธ์ที่ไม่ใช่เส้นตรง ดังนั้นจึงมีการใช้ฟังก์ชันเชื่อมโยง(Link function)เข้ามาช่วยในการพยากรณ์และตีความหมาย ตัวอย่างฟังก์ชันเชื่อมโยง(Link function) เช่น ฟังก์ชันโลจิท ฟังก์ชันโพรบิท และ ฟังก์ชันคอมพลีเมนทารีล็อก-ล็อก เป็นต้น ซึ่งลักษณะการเลือกฟังก์ชันเชื่อมโยงแต่ละฟังกัชันนั้นก็ขึ้นอยู่กับลักษณะการกระจายของข้อมูล ซึ่งข้อมูลที่มีลักษณะการกระจายที่สมมาตร เราจะใช้ฟังก์ชันเชื่อมโยงแบบโลจิท แต่สำหรับข้อมูลที่เราจะทำการศึกษานี้มีลักษณะของข้อมูลที่มีการกระจายที่ไม่สมมาตร คือมีลักษณะที่ล้มเหลวมากกว่าลักษณะที่สำเร็จ เราจึงจำเป็นต้องใช้ฟังก์ชันเชื่อมโยงแบบ คอมพลีเมนทารีล็อก-ล็อกเข้ามาช่วยในการคำนวณจุดแบ่ง (cut off point) แล้วนำไปพยากรณ์หาจุดแบ่งที่ดีที่สุด

1.2 วัตถุประสงค์ของการวิจัย

งานวิจัยนี้ต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลแบบไม่จัดกลุ่มสำหรับข้อมูลที่มีการกระจายที่ไม่สมมาตรซึ่งเป็นข้อมูลที่มีลักษณะของสัดส่วนความล้มเหลวมากกว่าสัดส่วนของความสำเร็จ โดยใช้ตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันเชื่อมโยงแบบคอมพลีเมนทารีล็อก-ล็อกในการประมาณค่าเพื่อช่วยหาจุดแบ่ง เมื่อ

- จำนวนของตัวแปรอิสระเพิ่มขึ้น
- ขนาดตัวอย่างใหญ่ขึ้น
- สัดส่วนของความสำเร็จและความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยน
- ระดับของความสัมพันธ์ระหว่างตัวแปรอิสระด้วยกันมีค่าเพิ่มมากขึ้น

ซึ่งเมื่อพิจารณาจากขอบเขตที่ต้องการศึกษาแล้ว ทำการวิเคราะห์เปรียบเทียบแต่ละขอบเขตร่วมกันว่าขอบเขตแบบไหนให้จุดแบ่งที่เหมาะสมที่สุด

1.3 ขอบเขตของการวิจัย

ต้องการศึกษาตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทเพื่อหาจุดแบ่งที่เหมาะสมที่สุดโดยอาศัยฟังก์ชันเชื่อมโยงคอมพลีเมนทารีล็อก-ล็อกเข้ามาช่วย เราจะทำการจำลองข้อมูลแต่ละชุดโดยการใช้โปรแกรม R โดยเริ่มต้นสร้างจากข้อมูลที่มีการกระจายแบบยูนิฟอร์ม ซึ่งหมายถึงสัดส่วนของความล้มเหลวของข้อมูลนั้นมากกว่าสัดส่วนของความสำเร็จ และทำการวิเคราะห์ข้อมูลตามขอบเขตดังนี้

1. จำนวนของตัวแปรอิสระ (p) แบ่งเป็น 5 ระดับ คือ 1,2,3,4 และ 5
2. ขนาดตัวอย่าง (n)แบ่งเป็น 5 ระดับ คือ 50,100,150,200 และ 250
3. กำหนดให้ตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มี 2 ค่า คือ ความสำเร็จ$(\mathrm{Y}=1)$ และความล้มเหลว $(\mathrm{Y}=0)$ โดยกำหนดอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา คือ a โดยที่ $\mathrm{a}=0.25,0.5$ และ 0.75
4. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) มีเงื่อนไข คือ

   ความสัมพันธ์กันในระดับอย่างต่ำ: $0 < Max\{|\mathrm{r_{ij}}|\} < 0.3$

   ความสัมพันธ์กันในระดับปานกลาง: $0.3 < Max\{|\mathrm{r_{ij}}|\} < 0.6$

   ความสัมพันธ์กันในระดับสูง: $0.6 < Max\{|\mathrm{r_{ij}}|\} < 0.9$

โดยที่ r_{ij} ; $i,j = 1,2,\dots,p$ คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

5. ตัวแปรอิสระมีการแจกแจงเริ่มต้น คือ การแจกแจงแบบยูนิฟอร์ม
6. กำหนดค่าพารามิเตอร์เริ่มต้นของสมการการถดถอยเป็นค่าใดๆ ในการศึกษาครั้งนี้ กำหนดให้ $\beta_i = 15$; $i = 1,2,\dots,p$ และ $\varepsilon_k \sim N(0,100)$; $k = 1,2,\dots,n$ โดย $\sigma^2 = 100$
7. กำหนดระดับนัยสำคัญ (α) ที่ระดับ 0.05
8. ในการศึกษาครั้งนี้ทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation) โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

1.4 ข้อตกลงเบื้องต้น

1. ศึกษาตัวแบบถดถอยโลจิสติกแบบ 2 ประเภท (Binary Logistic Regression) เมื่อมีตัวแปรอิสระมากกว่า 1 ตัว หรือ $(p \geq 2)$ เมื่อ Y มีได้แค่เพียง 2 ค่า จะพบว่าความสัมพันธ์ระหว่าง X กับ Y ไม่ได้อยู่ในรูปเชิงเส้น แต่จะอยู่ในรูป

$$\pi_i = p = \frac{e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}\cdots+\beta_p X_{pi}}}{1+e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}\cdots+\beta_p X_{pi}}} \quad ; i = 1,2,\dots,n$$

เมื่อ π_i คือ ความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สำเร็จ (Y=1) ของลักษณะที่สนใจ

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$ คือ ค่าสัมประสิทธิ์การถดถอยโลจิสติก

$X_{1i}, X_{2i}, \dots, X_{pi}$ คือ ตัวแปรอิสระตัวที่ $i = 1,2,\dots,n$

2. ศึกษาตัวแบบที่มีการแจกแจงเริ่มต้นของตัวแปรอิสระเป็นการแจงแจงแบบยูนิฟอร์ม (Uniform Distribution) ซึ่งมีฟังก์ชันความน่าจะเป็นของตัวแปรอิสระ X คือ

$$f(X;a,b) = \begin{cases} \dfrac{1}{b-a} & ; a < X < b \\ 0 & ; \ X \text{ มีค่าอื่น} \end{cases}$$

โดยที่ a และ b เป็นค่าคงที่ และ $a < b$

1.5 คำจำกัดความที่ใช้ในการวิจัย

1. **ตัวแบบถดถอยโลจิสติกแบบ 2 ประเภท** (Binary Logistic Regression Model) หมายถึง ตัวแบบที่ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตามซึ่งเป็นตัวแปรเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า ส่วนตัวแปรอิสระเป็นตัวแปรเชิงปริมาณหรือตัวแปรเชิงคุณภาพ หรืออาจจะมีทั้งตัวแปรเชิงปริมาณหรือตัวแปรเชิงคุณภาพก็ได้ เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรแล้ว จะนำไปใช้ในการประมาณค่าตัวแปรตามหรือการพยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่ง
2. **จุดแบ่ง** (Cut-off point) คือ ค่าความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนกกลุ่มว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มของความสาเร็จและกลุ่มของความล้มเหลวของลักษณะที่สนใจศึกษา
3. **ฟังก์ชันคอมพรีเมนต์ทารี ล็อก-ล็อก** (Complementary Log-Log Function) ขยายมาจากฟังก์ชันโลจิท ที่เป็นค่า Log ของความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สำเร็จที่มีค่าอยู่ในช่วง$(0,1)$ เป็นสมการโลจิทที่อยู่ในช่วง$(-\infty, \infty)$ แต่ต่างกันตรงที่ค่า π_iเพิ่มขึ้นจาก 0 ค่อนข้างช้า แต่ค่าของ π_i จะมีค่าเข้าใกล้ 1 อย่างรวดเร็ว ซึ่งใช้กับข้อมูลที่มีลักษณะการกระจายที่ไม่สมมารตร
4. **การแจกแจงแบบเบอร์นูลลี** สำหรับการทดลองใดๆที่ให้ผลของการทดลองออกมาเป็น 2 กลุ่มเหตุการณ์คือ เหตุการณ์ที่สนใจและกลุ่มเหตุการณ์ที่ไม่สนใจ จะเรียกการทดลองนั้นว่า การทดลองเบอร์นูลลี กล่าวคือ ตัวแปรสุ่ม *Y* เรียกว่า ตัวแปรสุ่มของการแจกแจงแบบเบอร์นูลลี อาจเขียนแทนด้วย $\mathrm{Y}\sim\mathrm{Ber(p)}$ เมื่อ $(\mathrm{Y}=0)$ คือ ไม่เกิดเหตุการณ์ที่สนใจ และ $(\mathrm{Y}=1)$ เมื่อการทดลองเกิดเหตุการณ์ที่สนใจ

   โดยที่ $\mathrm{P(Y=1)=p}$ และ $\mathrm{P(Y=0)=1-p}$, $0<p<1$

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

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถทราบจุดแบ่งในสถานการณ์ต่างๆ ของข้อมูลเพื่อให้ได้จุดแบ่งที่ดีที่สุด เพื่อใช้ประกอบในการตัดสินใจ
2. เพื่อใช้เป็นแนวทางในการวิจัยต่อไป

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยฟังก์ชันคอมพลีเมนทารีล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง สำหรับแต่ละสถานการณ์ที่ต้องการศึกษาต้องการหาค่าเฉลี่ยของจุดแบ่งที่เหมาะสมเพื่อทำการจำแนกประเภทของกลุ่ม

ซึ่งในบทนี้จะกล่าวถึงรายละเอียดเกี่ยวกับตัวแบบถดถอยโลจิสติก แบบ 2 ประเภท ด้วยวิธีภาวะความควรจะเป็นสูงสุด(Maximum Likelihood Estimation) และการหาค่าจุดแบ่งโดยใช้ทฤษฎีของ Hadjicostas P. (2006) ในการหาอัตราความผิดพลาดในการจำแนกกลุ่มหรือสัดส่วนความถูกต้องในการจำแนกกลุ่ม

แนวคิดและทฤษฎี

2.1 ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท (Binary Logistic Regression Model)

ตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทเป็นตัวแบบที่ศึกษาถึงความสัมพันธ์ของตัวแปร 2 กลุ่ม ได้แก่ กลุ่มตัวแปรอิสระ X และกลุ่มของตัวแปรตาม Y ซึ่ง ตัวแปรตามนี้เป็นตัวแปรเชิงคุณภาพมีค่าได้เพียง 2 ค่า โดยพิจารณาในรูปของความสำเร็จและความล้มเหลว โดยที่ $(Y = 1)$ เมื่อพบความสำเร็จ และ $(Y = 0)$ เมื่อพบความล้มเหลว ซึ่ง เมื่อได้ตัวแบบความสัมพันธ์ระหว่างตัวแปรแล้วจะสามารถนำไปใช้พยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่งได้

2.1.1 กรณีที่มีตัวแปรอิสระ 1 ตัว

สมการความถดถอยอย่างง่าย หรือสมการแสดงความสัมพันธ์ระหว่าง Y กับ X จะอยู่ในรูปเชิงเส้นดังนี้

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

โดยที่ $$E(Y) = \beta_0 + \beta_1 X \quad ; -\infty < E(Y) < +\infty$$

สำหรับการวิเคราะห์ความถดถอยโลจิสติกนั้น เมื่อ Y มีเพียง 2 ค่าจะพบว่าความสัมพันธ์ระหว่าง X กับ Y นั้นไม่ได้อยู่ในรูปของเชิงเส้น แต่จะอยู่ในรูปของ

$$E(Y) = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$$

และเรียกสมการนี้ว่า Logistic Response Function โดยที่ $0 \leq E(Y) \leq 1$

2.1.2 กรณีที่มีตัวแปรอิสระมากกว่า 1 ตัว

เมื่อมีตัวแปรอิสระมากกว่า 1 ตัว หรือ ($p \geq 2$) เมื่อ Y มีได้แค่เพียง 2 ค่า จะพบว่าความสัมพันธ์ระหว่าง X กับ Y ไม่ได้อยู่ในรูปเชิงเส้น แต่จะอยู่ในรูป

$$\pi_i = P(\text{เกิดเหตุการณ์}) = \frac{e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}\cdots+\beta_p X_{pi}}}{1+e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}\cdots+\beta_p X_{pi}}}$$
$$P(\text{ไม่เกิดเหตุการณ์}) = 1 - P(\text{เกิดเหตุการณ์}) \; ; i = 1,2,\ldots,n$$

ซึ่งจะพบว่า ทั้งสมการที่ (2.1.1) และ สมการที่ (2.1.2) ต่างก็มีความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระไม่ได้อยู่ในรูปเชิงเส้น จึงมีการปรับให้ความสัมพันธ์อยู่ในรูปเชิงเส้นโดยให้

$$\text{Odds} = \frac{P(\text{เกิดเหตุการณ์})}{P(\text{ไม่เกิดเหตุการณ์})} = \frac{\pi_i}{1-\pi_i}$$

Odds หรือ Odd ratio จะแสดงถึง โอกาสที่จะเกิดเหตุการณ์เป็นกี่เท่าของโอกาสที่จะไม่เกิด เช่น ถ้าได้ Odd ratio = 2.8 แสดงว่า โอกาสที่จะเกิดเหตุการณ์เป็น 2.8 เท่าของเหตุการณ์ที่จะไม่เกิด ถ้าค่า Odd ratio มากกว่า 1 แสดงว่าโอกาสที่จะเกิดเหตุการณ์มากกว่า โอกาสที่จะไม่เกิดเหตุการณ์

Take Log สมการที่ (2.1.3) จะได้

$$\text{Log}(\text{Odds}) = \text{Log}\left[\frac{P(\text{เกิดเหตุการณ์})}{P(\text{ไม่เกิดเหตุการณ์})}\right] = \text{Log}\left[\frac{\pi_i}{1-\pi_i}\right]$$

หรือ

$$\text{Log}(\text{Odds}) = \text{Log}\left[\frac{\pi_i}{1-\pi_i}\right] = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \ldots + \beta_p X_{pi}$$

ซึ่งสมการนี้จะอยู่ในรูปของเชิงเส้น เรียกสมการนี้ว่า Logit Response Function หรือ ฟังก์ชันโลจิท นั่นเอง

2.2 ฟังก์ชันคอมพลีเมนทารี ล็อก-ล็อก

เนื่องจากข้อมูลที่มีลักษณะที่ไม่สมมาตรจึงไม่สามารถใช้ฟังก์ชันโลจิทข้างต้นได้ จึงต้องใช้ฟังก์ชันที่เป็นส่วนขยายจากฟังก์ชันโลจิท คือ ฟังก์ชันคอมพลีเมนทารีล็อก-ล็อก(Complementary log-log function) เป็นส่วนที่ขยายจากตัวแบบโลจิท(Logit model) และตัวแบบโพรบิท(Probit model) เมื่อค่า π_i เพิ่มขึ้นจาก 0 ค่อนข้างช้า แต่ค่าของ π_i จะมีค่าเข้าใกล้ 1 อย่างรวดเร็ว

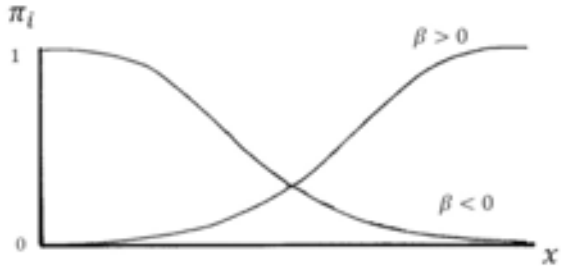
(วีรานันท์ 2544 ; 89-93)

$$\text{Link}[\pi_i] = -\text{link}[1-\pi_i]$$

นั่นคือ
$$\begin{aligned}\text{Logit}[\pi_i] &= \log[\pi_i - \{1-\pi_i\}] \\ &= \log \pi_i - \log[1-\pi_i] \\ &= -\log[\{1-\pi_i\}/\pi_i] \\ &= -\text{logit}[1-\pi_i]\end{aligned}$$

หมายความว่าโค้งของ π_i สำหรับตัวแบบโลจิทและตัวแบบโพรบิท มีรูปแบบสมมาตรรอบจุด $\pi_i = 0.5$ โดยเฉพาะ π_i จะมีค่าเข้าใกล้ 0 ด้วยอัตราที่เท่ากัน กับเมื่อ π_i มีค่าเข้าใกล้ 1 อย่างรวดเร็ว

แต่ถ้าค่าของ π_i เพิ่มขึ้นจาก *0* ค่อนข้างช้าแต่มีค่าเข้าใกล้ 1 อย่างรวดเร็วด้วยตัวแบบโลจิทและตัวแบบโพรบิทจะไม่เหมาะกับข้อมูล จึงควรใช้ตัวแบบอย่างอื่นคือ ตัวแบบคอมพลีเมนทารีล็อก-ล็อก



ภาพที่ 2.1: กราฟของ π_i สำหรับตัวแบบคอมพลีเมนทารีล็อก-ล็อก

ลักษณะกราฟของ π_i เป็นเส้นโค้งของฟังก์ชัน

$$\begin{aligned}\pi_i &= 1-\exp[-\exp(\beta_0+\beta_1 x)] \\ 1-\pi_i &= \exp[-\exp(\beta_0+\beta_1 x)] \quad \text{.....(2.1)}\end{aligned}$$

ซึ่งมีรูปแบบไม่สมมาตรคือ π_i มีค่าลดจาก 1 รวดเร็วกว่าการเข้าใกล้ *0* โดยสอดคล้องกับ *(*กราฟตัวแบบกับ ฟังก์ชันข้างบน) นำไปสู่ตัวแบบคอมพลีเมนทารีล็อก-ล็อก ดังสมการข้างล่าง

$$\begin{aligned}-\log[1-\pi_i] &= \exp(\beta_0+\beta_1 x) \\ \log[-\log\{1-\pi_i\}] &= \beta_0+\beta_1 x \quad \text{.....(2.2)}\end{aligned}$$

การตีความหมายสำหรับตัวแบบคอมพลีเมนทารีล็อก-ล็อก สำหรับ x_1 และ x_2 ใดๆจะพบว่า

$$\begin{aligned}\log[-\log\{1-\pi_2\}] - \log[-\log\{1-\pi_1\}] &= \beta_1(x_2-x_1) \\ \frac{\log[1-\pi_2]}{\log[1-\pi_1]} &= \exp[\beta_1(x_2-x_1)] \\ 1-\pi_2 &= [1-\pi_1]^{\exp[\beta_1(x_2-x_1)]}\end{aligned}$$

ซึ่งหมายความว่า ความน่าจะเป็นที่ไม่สำเร็จ ณ x_2

มีค่าเท่ากับ ความน่าจะเป็นที่ไม่สำเร็จ ณ x_1 ยกกำลัง $\exp[\beta_1(x_2 - x_1)]$

นอกจากตัวแบบ (2.1) กับ (2.2) ยังมีตัวแบบที่น่าสนใจเกี่ยวกับหัวข้อนี้คือตัวแบบที่อยู่ในรูปแบบดังนี้

$$\pi_i = \exp[-\exp(\beta_0 + \beta_1 x)]$$
$$\log[\pi_i] = -\exp(\beta_0 + \beta_1 x)$$
$$-\log[\pi_i] = \exp(\beta_0 + \beta_1 x)$$
$$\log\{-\log[\pi_i]\} = \beta_0 + \beta_1 x \qquad \text{.....(2.3)}$$

สรุปว่า ถ้าเป็นตัวแบบคอมพลีเมนทารีล็อก-ล็อก จะใช้สำหรับความน่าจะเป็นที่สำเร็จ (Probability of success) ส่วน ตัวแบบล็อก-ล็อก จะใช้สำหรับความน่าจะเป็นที่ไม่สำเร็จ

$$\ln\left[-\ln\left[1 - E\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right)\right]\right] = \ln[-\ln[1 - \pi_i]] = \tilde{X}_i^T\tilde{\beta}$$
$$Y_i^* = \tilde{X}_i^T\tilde{\beta} + \varepsilon_i$$
$$Y_i^* = \ln\left[-\ln\left[1 - E\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right)\right]\right] + \varepsilon_i \;;\; \varepsilon_i \sim N(0, \sigma^2), i = 1,2,\dots,n$$

โดย π_i คือ ค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

Y_i คือ ตัวแปรตามเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า คือ 0 และ 1

Y_i^* คือ ตัวแปรแฝง(Latent variable) ที่ไม่สามารถสังเกตได้ จึงไม่ทราบค่าที่แท้จริง ทราบเพียงแต่ผลที่เกิดขึ้น

$\tilde{X}_i$ คือ เวกเตอร์ตัวแปรอิสระหน่วยที่ i

$\tilde{\beta}$ คือ เวกเตอร์พารามิเตอร์ของตัวแบบ

ε_i คือ ตัวแปรค่าคลาดเคลื่อนที่เป็นอิสระกัน หรือค่าคาดหมายอย่างมีเงื่อนไข (Conditional expectation) หน่วยที่ i ตัวแปรสุ่มปกติมาตราฐานโดยที่ $\varepsilon_i \sim N(0,1)$

2.2.1 **เปรียบเทียบตัวแบบคอมพรีเมนทารีล็อก-ล็อก โลจิท และ โพรบิท**

การหาความสัมพันธ์ระหว่างตัวแปรตอบสนองและตัวแปรอิสระเมื่อตัวแปรตอบสนองเป็นตัวแปรเชิงกลุ่มที่มีสองค่า และ ตัวแปรอิสระเป็นตัวแปรเชิงปริมาณ สามารถใช้การวิเคราะห์ที่มีลักษณะไม่เป็นเชิงเส้นของตัวแบบโพรบิทและตัวแบบคอมพรีเมนทารีล็อก-ล็อก ที่มีอยู่บนรากฐานของการแจกแจงสะสมปกติมาตรฐาน หรือตัวแบบโลจิทที่อยู่บนรากฐานของการแจกแจงโลจิสติกได้ ซึ่งตัวแบบทั้งสามมีคุณสมบัติคล้ายกัน การเลือกใช้ตัวแบบใดขึ้นกับพื้นฐานทางทฤษฎีซึ่งมีข้อแตกต่างดังนี้

1. ลักษณะของฟังก์ชันการแจกแจงความน่าจะเป็น กราฟฟังก์ชันการแจกแจงความน่าจะเป็นปกติมาตรฐาน และกราฟการแจกแจงความน่าจะเป็นแบบโลจิสติกจะมีลักษณะเป็นรูประฆังคว่ำ โดยสมมาตรที่ค่าเฉลี่ยตัวแปรตอบสนอง (Y_i) เท่ากับ 0 มีค่ามัธยฐาน ค่าฐานนิยมและความเบ้เท่ากับ 0 ค่าความแปรปรวนมีค่าเท่ากับ 1 ส่วนฟังก์ชันการแจกแจงความน่าจะเป็นโลจิสติกมีค่าเฉลี่ยเท่ากับ 0 ค่าความแปรปรวนเท่ากับ $\frac{\pi^2}{3}$ ในขณะที่ฟังก์ชันการแจกแจงความน่าจะเป็นแบบคอมพรีเมนทารีล็อก-ล็อก มีค่าเฉลี่ยเท่ากับ -0.5772 ค่าความแปรปรวนเท่ากับ $\frac{\pi^2}{6}$
2. ลักษณะฟังก์ชันการแจกแจงสะสม กราฟจะมีลักษณะรูปตัวเอส ทั้งตัวแบบโพรบิทและตัวแบบโลจิสติกมีค่าความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจและความน่าจะเป็นของการไม่เกิดเหตุการณ์ที่สนใจใกล้กันมากจนเกือบจะเท่ากัน ฉะนั้นการใช้ตัวแบบใดก็ไม่ก่อให้เกิดความแตกต่างในการหาค่าความน่าจะเป็น ยกเว้นแต่กรณีที่มีข้อมูลหนาแน่นในช่วงหาง (ชนิเวศรา,2543) ในขณะที่ฟังก์ชันคอมพรีเมนทารีล็อก-ล็อก มีความชันสูงกว่าฟังก์ชันโลจิสติกและฟังก์ชันโพรบิทที่ปรับแล้ว และเมื่อค่าของ $p(x)$ ของตัวแบบคอมพรีเมนทารีล็อก-ล็อกเพิ่มจาก 0 ค่อนข้างช้าแต่มีค่าเข้าใกล้ 1 อย่างรวดเร็ว มากกว่าฟังก์ชันโลจิสติกและโพรบิทที่ปรับแล้ว
3. ในด้านการคำนวณตัวแบบโพรบิทและตัวแบบโลจิสติกจะให้ผลใกล้เคียงกัน โดยตัวแบบโพรบิทอยู่บนรากฐานของการแจกแจงสะสมแบบปกติซึ่งมีแนวคิดและพื้นฐานทางทฤษฎีสนับสนุนอย่างมีเหตุมีผล แต่ตัวแบบโพรบิทมีการคำนวณที่ยุ่งยาก กล่าวคือ ในการกำหนดค่าความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจจะต้องติดอยู่ในรูปแบบของการอินทิเกรตเสมอ จึงทำให้ไม่สะดวกและเวลาในการคำนวณด้วยคอมพิวเตอร์นานกว่า แต่ตัวแบบโลจิสติกคำนวณได้ง่ายไม่ยุ่งยากมาก

2.3 ฟังก์ชันภาวะน่าจะเป็น (Likelihood function) ของข้อมูลการถดถอยโลจิสติกแบบ 2 ประเภท

เมื่อแต่ละค่าสังเกต Y_i คือตัวแปรสุ่มแบบเบอร์นูลลี ซึ่ง

$$\Pr\left(Y_i = 1 \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \pi_i ,$$
$$\Pr\left(Y_i = 0 \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = 1 - \pi_i \quad ; i = 1,2, \dots , n$$

ซึ่งมีการแจกแจงความน่าจะเป็นคือ

$$f_i\left(Y_i = y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \qquad ; y_i = 0,1 \; ; i = 1,2,..,n$$

ฟังก์ชันความน่าจะเป็นร่วมคือ

$$g(Y_1 = y_1, Y_2 = y_2, \dots , Y_n = y_n) = \prod_{i=1}^{n} f_i(Y_i = y_i) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i}$$

ลอการิทึมธรรมชาติของฟังก์ชั่นความน่าจะเป็นร่วมคือ

$$\ln g(Y_1 = y_1, Y_2 = y_2, \dots , Y_n = y_n) = \ln\left\{\prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i}\right\}$$

เมื่อ $\pi_i = E\left(Y_i \middle| \tilde{X}_i^T = \tilde{x}_i^T\right) = 1 - \exp\left[-\exp\left(\tilde{X}_i^T\tilde{\beta}\right)\right]$

$$1 - \pi_i = \exp\left[-\exp\left(\tilde{X}_i^T\tilde{\beta}\right)\right]$$

$$\ln[-\ln[1-\pi_i\,]] = \tilde{X}_i^T\tilde{\beta} \qquad ; i = 1,2,..,n$$

ดังนั้น ลอการิทึมธรรมชาติของฟังก์ชันภาวะน่าจะเป็น (Log-likelihood function) คือ

$$\ln L\left(\tilde{\beta}^T\right) = \sum_{i=1}^{n} \left(y_i \ln\left(1 - \exp\left(-\exp\left(\tilde{X}_i^T\tilde{\beta}\right)\right)\right)\right.$$

$$\left. +(1-y_i)\ln\left(1 - \exp\left(-\exp\left(\tilde{X}_i^T\tilde{\beta}\right)\right)\right)\right)$$

2.4 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation)

เมื่อเราทราบการแจกแจงของ Y เราสามารถสร้างฟังก์ชันภาวะน่าจะเป็นได้ เนื่องจากการประมาณค่าไม่ได้เป็นไปตามรูปแบบ เราจึงต้องใช้วิธีการประมาณเชิงตัวเลข(จำเป็นต้องทำซ้ำเพื่อให้ได้ซึ่งตัวประมาณความน่าจะเป็นสูงสุด (MLE) โดยการทำซ้ำ 1 ครั้งจะได้ตัวประมาณภาวะน่าจะเป็นสูงสุด $(b_0, b_1, b_2, \dots, b_n)$ แล้วทำการคำนวนค่าประมาณของ $\pi_i \ ; i = 1,2, \dots, n$ ดังนี้

$$\hat{\pi}_i = 1 - \exp\left[-\exp\left(\tilde{X}_i^T \tilde{\beta}\right)\right] \quad ; i = 1,2, \dots, n$$

เมื่อประมาณค่าพารามิเตอร์ด้วยวิธีภาวะความน่าเป็นสูงสุดในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภท แล้วจะสามารถนำไปใช้ในการพยากรณ์การจำแนกกลุ่มของตัวแบบ ดังนี้

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

$$\hat{\pi}_i = 1 - \exp\left[-\exp\left(\tilde{X}_i^T \tilde{\beta}\right)\right] > c \quad ; 0 \le c \le 1$$

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

$$\hat{\pi}_i = 1 - \exp\left[-\exp\left(\tilde{X}_i^T \tilde{\beta}\right)\right] \le c \quad ; 0 \le c \le 1$$

เมื่อ C คือ จุดแบ่ง หรือระดับของความน่าจะเป็นที่ใช้ในการพิจารณาการจำแนกกลุ่มว่าแต่ละหน่วยจะอยู่ในกลุ่มใดระหว่างกลุ่มการเกิดเหตุการณ์ที่สนใจ และกลุ่มไม่เกิดเหตุการณ์ที่สนใจ

2.5 ช่วงความเชื่อมั่น (Confidence Interval)

การประมาณค่าแบบช่วง เป็นการประมาณค่าพารามิเตอร์ของประชากรว่าจะอยู่ในช่วงใดช่วงหนึ่ง โดยใช้ข้อมูลตัวอย่างในการประมาณค่า และช่วงของการประมาณค่า จะบอกถึงค่าที่ต่ำที่สุดและค่าที่สูงที่สุดของพารามิเตอร์ และมีโอกาสที่จะคลาดเคลื่อนไปจากค่าจริงน้อยกว่าการประมาณค่าแบบจุด

ในกรณีทั่วไปเราจะให้ระดับความเชื่อมั่น $(1-\propto)100\%$ หมายความว่า โอกาสที่ค่าประมาณจะผิดพลาด $=\propto(100)\%$ ถ้าให้ระดับความเชื่อมั่น$(1-\propto)$ มีค่ามากจะทำให้โอกาสของความผิดพลาดในการประมาณค่า $\propto$ มีค่าน้อย หรือ เรียกโอกาสของความผิดพลาดว่า ระดับนัยสำคัญ$(\propto)$

$$P(L < \mu < U) = (1-\propto)$$

โดยที่ L คือ ขีดจำกัดความเชื่อมั่นล่าง (Lower Confidence Limit)

U คือ ขีดจำกัดความเชื่อมั่นบน (Upper Confidence Limit)

ในงานวิจัยนี้ กำหนดให้ L หาจากตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(\propto/2)$

กำหนดให้ U หาจากตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(1-\propto/2)$

2.6 เทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นการจำลองระบบที่ไม่เปลี่ยนแปลงตามเวลา ซึ่งตัวแบบของการจำลองจะมีลักษณะเป็นตัวแบบทางคณิตศาสตร์ โดยการนำตัวเลขสุ่ม มาประยุกต์ใช้ในการแก้ปัญหาหรือหาคำตอบให้กับระบบที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น ซึ่งมีขั้นตอนที่สำคัญ 3 ขั้น ตอนดังนี้

ขั้นที่ 1 การสร้างเลขสุ่ม (Generate Random Number) จะกำหนดให้มีการแจกแจงแบบยูนิฟอร์มในช่วง [0, 1] และเป็นอิสระซึ่งกันและกัน จากนั้น นำเลขสุ่มนี้ไปสร้างตัวแปรตามลักษณะการแจกแจงที่ต้องการศึกษา เพื่อเป็นข้อมูลของปัญหานั้น ๆ

ขั้นที่ 2 การประยุกต์ใช้เลขสุ่มในการแก้ปัญหา ขั้นตอนนี้เป็นการนำตัวแปรที่ได้จากขั้นตอนแรกมาใช้ในการหาค่าต่างๆ ตามปัญหาที่ต้องการศึกษา

ขั้นที่ 3 การทดลอง ขั้นตอนนี้เป็นการทำวิธีนั้น ซ้ำๆ กัน (Replication) จำนวนหลายครั้ง โดยถือว่าการทำซ้ำๆ กันนั้น เป็นวิธีการเก็บรวบรวมข้อมูลให้มีจำนวนมาก เพื่อลดความไม่แน่นอนของคำตอบในการวิเคราะห์หาค่าต่างๆ ได้

จากหลักการของเทคนิคมอนติคาร์โล จะเห็นว่าการใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีที่จะนำไปสู่แนวคิดในทางทฤษฎีที่เกี่ยวข้องกับการคำนวณโดยเฉพาะทฤษฎีความน่าจะเป็นที่จะนำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริงเพราะไม่มีผลกระทบจากเรื่องอื่นๆ เข้ามาเกี่ยวข้องในการทดลอง เมื่อทำซ้ำๆ กันเป็นจำนวนมากแล้ว ความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่างๆ ในแต่ละครั้งให้หมดไป

2.7 วิธีการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับตัวแบบการถดถอยโลจิสติกแบบ 2ประเภท โดยทฤษฎีของ Hadjicostas P. (2006)

ทฤษฎีของ Hadijicostas P.(2006) เสนอวิธีการที่ใช้ผลลัพธ์ทางคณิตศาสตร์ที่ง่ายแต่มีความถูกต้องแม่นยำในการหาจุดแบ่งที่ทำให้สัดส่วนความถูกต้องในการจำแนกกลุ่มมีค่าสูงสุด

ทฤษฎีบท ให้ $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ สำหรับ $i = 0,1,2,\dots,n$

ให้ I_0 เป็นเซตของ $j\,; j \in \{0,1,2,\dots,n\}$ ที่ซึ่ง $a_j = \max_{0 \le i \le n} a_i$ และให้ C_0 เป็นเซตของ c_0 ทั้งหมด $c_0 \in [0,1]$ที่ซึ่ง $p(C_0) = \max_{c \in [0,1]} p(c)$

แล้ว $C_0 = \cup_{i \in I_0} A_i$

โดยมีวิธีการหาจุดแบ่งตามขั้นตอนดังนี้

ขั้นที่1. เรียงอันดับค่า $\hat{\pi}_i$จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < \cdots < \hat{\pi}_n$ โดยถ้า $\hat{\pi}_i$ คือจุดแบ่งแล้วจะพยากรณ์ให้เป็นกลุ่มของความล้มเหลวของลักษณะที่สนใจศึกษา $(Y = 0)$

ขั้นที่2. สำหรับแต่ละ $i \in \{1,2, \dots, n\}$ ใดๆ หาค่า $M(i)$ ซึ่ง $M(i)$ คือ อันดับของ $\hat{\pi}_i$ ถ้า $\hat{\pi}_i = \hat{\pi}_j$ จะเลือก $M(i)$ที่มีค่ามากที่สุดเป็นอันดับของค่า $\hat{\pi}_i$และ $\hat{\pi}_j$ โดย $M(0) = 0 \; ; i \le M(i) \le n$

ขั้นที่3. การหาค่า $a_i \; ; i \in \{1,2, \dots, n\}$ โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ โดยแบ่งเป็น 2 กรณี คือ

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

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

ขั้นที่4. หา I_0 ซึ่งเป็นเซตของ j ทั้งหมด โดย $j \in \{0,1,2, \dots, n\}$ ที่ซึ่ง $a_j = \max_{0 \le i \le n} a_i$

ขั้นที่5. หา C_0 ซึ่งเป็นเซตของ c_0 ทั้งหมดโดย $c_0 \in [0,1]$ที่ซึ่ง $C_0 = \cup_{i \in I_0} A_i$

โดยพิจารณาจากเงื่อนไขดังนี้

$A_i = [0, \hat{\pi}_1)$ ถ้า $i = 0$

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

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

$A_i = [\hat{\pi}_n, 1)$ ถ้า $i = n$

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

ขั้นที่6. เลือกค่าจุดแบ่ง c ที่เหมาะสมที่สุด ที่ให้สัดส่วนของความถูกต้องในการจำแนกกลุ่มมีค่ามากที่สุดโดยที่ $c \in [0,1]$

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

โดยที่ $p(c)$ คือ สัดส่วนของความถูกต้องในการจำแนกกลุ่มที่จุด c

$N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

$$N(c) = \sum_{j=1}^{M(i)} (1 - y_j) + \sum_{j=M(i)+1}^{n} y_j$$

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทโดยใช้ฟังก์ชันคอมพลีเมนทารี ล็อก ล็อก เป็นฟังก์ชันเชื่อมโยง สำหรับแต่ละสถานการณ์ที่ต้องการศึกษา จะทำการเปรียบเทียบที่ขนาดของตัวอย่างมีค่าเท่ากับ 50,100,150,200 และ 250 จำนวนของตัวแปรอิสระเป็น 1,2,3,4 และ 5 อัตราส่วนของความล้มเหลวของลักษณะที่สนใจ คือ 0.25,0.5 และ 0.75 และแบ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น 3 ระดับ คือ ความสัมพันธ์กันในระดับ ต่ำ กลาง และสูง โดยทำการจำลองข้อมูลและกระทำซ้ำ 500 รอบ โดยเทคนิคมอนติคาร์โล และเขียนด้วยโปรแกรม R

ซึ่งแผนการดำเนินการวิจัย ขั้นตอนในแผนการดำเนินการวิจัย ตลอดจนโปรแกรมที่ใช้ในการวิจัยมีรายละเอียดต่างๆดังนี้

3.1 แผนการดำเนินการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ต่างๆเพื่อเปรียบเทียบดังนี้

1. จำนวนของตัวแปรอิสระ (p) แบ่งเป็น 5 ระดับ คือ 1,2,3,4 และ 5
2. ขนาดตัวอย่าง (n)แบ่งเป็น 5 ระดับ คือ 50,100,150,200 และ 250
3. อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a) คือ 0.25,0.5 และ 0.75
4. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ (ρ) มีเงื่อนไข คือ

   ความสัมพันธ์กันในระดับอย่างต่ำ: $0 < Max\{|\mathrm{r_{ij}}|\} < 0.3$

   ความสัมพันธ์กันในระดับปานกลาง: $0.3 < Max\{|\mathrm{r_{ij}}|\} < 0.6$

   ความสัมพันธ์กันในระดับสูง: $0.6 < Max\{|\mathrm{r_{ij}}|\} < 0.9$

   โดยที่ $\mathrm{r_{ij}}\ ; \mathrm{i,j} = 1,2,\dots,\mathrm{p}$ คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

ดังนั้น จำนวนสถานการณ์ที่ใช้ในการวิจัย มีจำนวน = $5 \times 5 \times 3 \times 3 = 225$ สถานการณ์

5. กำหนดให้ค่าพารามิเตอร์เริ่มต้นของสมการถดถอยใดๆ คือ $\beta_i = 15 ; i = 1,2, \dots, p$
6. กำหนดให้ความคลาดเคลื่อนใดๆ มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย เท่ากับ 0 และ ความแปรปรวนมีค่าเท่ากับ 100 คือ $\varepsilon_k \sim N(0,100); k = 1,2, \dots, n$ โดย $\sigma^2 = 100$
7. กำหนดระดับนัยสำคัญ (α) ที่ระดับ 0.05
8. กำหนดให้การกระทำซ้ำในแต่ละสถานการณ์ คือ 500 รอบ

ซึ่งในแต่ละสถานการณ์จะทำการเปรียบเทียบค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งในแต่ละสถานการณ์

3.2 ขั้นตอนในการวิจัย

ขั้นที่ 1 จำลองข้อมูลตามขอบเขตที่ต้องการศึกษา

ขั้นที่ 2 คำนวณหาจุดแบ่งที่เหมาะสมที่สุดในแต่ละสถานการณ์

ขั้นที่ 3 ทำการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์โดยเทคนิคมอนติคาร์โล

ขั้นที่ 4 คำนวณหาค่าเฉลี่ยร้อยละและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสมที่สุด

ขั้นที่ 5 สรุปผลที่ได้จากการวิจัยในแต่สถานการณ์

ซึ่งรายละเอียดของขั้นตอนมีดังนี้

ขั้นที่ 1 จำลองข้อมูลตามขอบเขตที่ต้องการศึกษา

ขั้นที่1.1 เนื่องจากใช้เทคนิคมอนติคาร์โลในการจำลองข้อมูล จึงต้องทำการเริ่มสร้างตั้งแต่ตัวเลขสุ่มที่มีการแจกแจงยูนิฟอร์มซึ่งเราจะกำหนดให้มีค่าของช่วงบวกและช่วงลบที่เท่ากันดังนี้ $X \sim U(-1,1)$ เพื่อนำไปใช้ในการสร้างตัวแปรสุ่มที่มีการแจกแจง แบร์นูลลีต่อไป

ขั้นที่1.2 สร้างจำนวนของตัวแปรอิสระตามที่กำหนดไว้และให้ตัวแปรอิสระมีความสัมพันธ์กันตามระดับความสัมพันธ์ระหว่างตัวแปรอิสระอยู่สามระดับคือ ต่ำ กลางและสูง และให้มีขนาดตัวอย่างตามขอบเขตของการวิจัย

ขั้นที่1.3 สร้างค่าตัวแปรตาม Y^* ให้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระที่ได้สร้างตอนต้นและมีค่าความคลาดเคลื่อนใดๆ ดังตัวแบบ

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

โดยที่ Y^*คือ เมทริกซ์ของตัวแปรตามที่ทำการพยากรณ์เพื่อกำหนดค่าเบื้องต้น

X คือ เมทริกซ์ของตัวแปรอิสระ

$\tilde{\beta}$ คือ เวกเตอร์พารามิเตอร์ของตัวแบบ กำหนดให้ค่าเริ่มต้นเท่ากับ 15

ε_k คือ ค่าคลาดเคลื่อนที่เป็นอิสระกัน $k = 1,2, \dots, n$ มีการแจกแจงปกติมาตราฐานมีค่า ความแปรปรวนเท่ากับ 100 $\varepsilon_k \sim N(0,100)$

ขั้นที่1.4 แปลงค่าตัวแปร Y^* ให้เป็นตัวแปรตาม Y ให้มีค่าเป็น 0 กับ 1 ตามสัดส่วนความล้มเหลวของลักษณะที่สนใจศึกษาและขนาดตัวอย่างที่กำหนดไว้ในขอบเขต

ขั้นที่1.5 ประมาณค่าพารามิเตอร์โดยใช้ตัวแบบการถดถอยโลจิสติก 2 ประเภท ด้วยวิธีภาวะน่าจะเป็นสูงสุด

ขั้นที่1.6 หาค่าประมาณ $\hat{\pi}_i$ โดยการนำค่าพารามิเตอร์ที่เราได้จากการประมาณข้างต้น และค่าของตัวแปรอิสระ แทนค่ากลับลงไปในตัวแบบการถดถอยโลจิสติก 2 ประเภท

ขั้นที่ 2 คำนวณหาจุดแบ่งที่เหมาะสมที่สุดในแต่ละสถานการณ์

ขั้นที่2.1 เรียงอันดับค่า $\hat{\pi}_i$จากน้อยไปหามาก $\hat{\pi}_1 < \hat{\pi}_2 < \cdots < \hat{\pi}_n$ โดยถ้า $\hat{\pi}_i$ คือจุดแบ่งแล้วจะพยากรณ์ให้เป็นกลุ่มของความล้มเหลวของลักษณะที่สนใจศึกษา $(Y = 0)$

ขั้นที่2.2 สำหรับแต่ละ $i \in \{1,2, \dots, n\}$ ใดๆ หาค่า $M(i)$ ซึ่ง $M(i)$ คือ อันดับของ $\hat{\pi}_i$ ถ้า $\hat{\pi}_i = \hat{\pi}_j$ จะเลือก $M(i)$ที่มีค่ามากที่สุดเป็นอันดับของค่า $\hat{\pi}_i$และ $\hat{\pi}_j$ โดย $M(0) = 0 \; ; i \le M(i) \le n$

ขั้นที่2.3 การหาค่า $a_i \; ; i \in \{1,2, \dots, n\}$ โดย $a_i = \sum_{k=1}^{M(i)} (-1)^{y_k}$ โดยแบ่งเป็น 2 กรณี

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

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

ขั้นที่2.4 หา I_0 ซึ่งเป็นเซตของ j ทั้งหมด โดย $j \in \{0,1,2, \dots, n\}$ ที่ซึ่ง $a_j = \max_{0 \le i \le n} a_i$

ขั้นที่2.5 หา C_0 ซึ่งเป็นเซตของ c_0 ทั้งหมดโดย $c_0 \in [0,1]$ที่ซึ่ง $C_0 = \bigcup_{i \in I_0} A_i$

โดยพิจารณาจากเงื่อนไขดังนี้

$A_i = [0, \hat{\pi}_1)$ ถ้า $i = 0$

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

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

$A_i = [\hat{\pi}_n, 1)$ ถ้า $i = n$

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

ขั้นที่2.6 เลือกค่าจุดแบ่ง c ที่เหมาะสมที่สุด ที่ให้สัดส่วนของความถูกต้องในการจำแนกกลุ่มมีค่ามากที่สุดโดยที่ $c \in [0,1]$

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

โดยที่ $p(c)$ คือ สัดส่วนของความถูกต้องในการจำแนกกลุ่มที่จุด c

$N(c)$ คือ จำนวนของความถูกต้องในการจำแนกกลุ่มที่จุด c

$$N(c) = \sum_{j=1}^{M(i)}(1 - y_j) + \sum_{j=M(i)+1}^{n} y_j$$

ขั้นที่ 3 ทำการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์

ขั้นที่ 4 คำนวณหาค่าเฉลี่ยร้อยละและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสมที่สุด

การค่าเฉลี่ยของจุดแบ่ง $\hat{c}$

ในแต่ละสถานการณ์จากการทดลองซ้ำ 500 รอบในแต่ละสถานการณ์ กำหนัดให้ $\hat{c}$ เป็นตัวประมาณค่าของพารามิเตอร์ c

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

โดยที่ N คือ จำนวนรอบในการกระทำซ้ำ ($N = 500$)

ช่วงความเชื่อมั่นของจุดแบ่ง

ช่วงความเชื่อมั่นนี้จะให้ค่าของจุดแบ่งต่ำสุดและสูงสุดที่เป็นไปได้ในแต่ละสถานการณ์ที่ระดับนัยสำคัญเท่ากับ 0.05

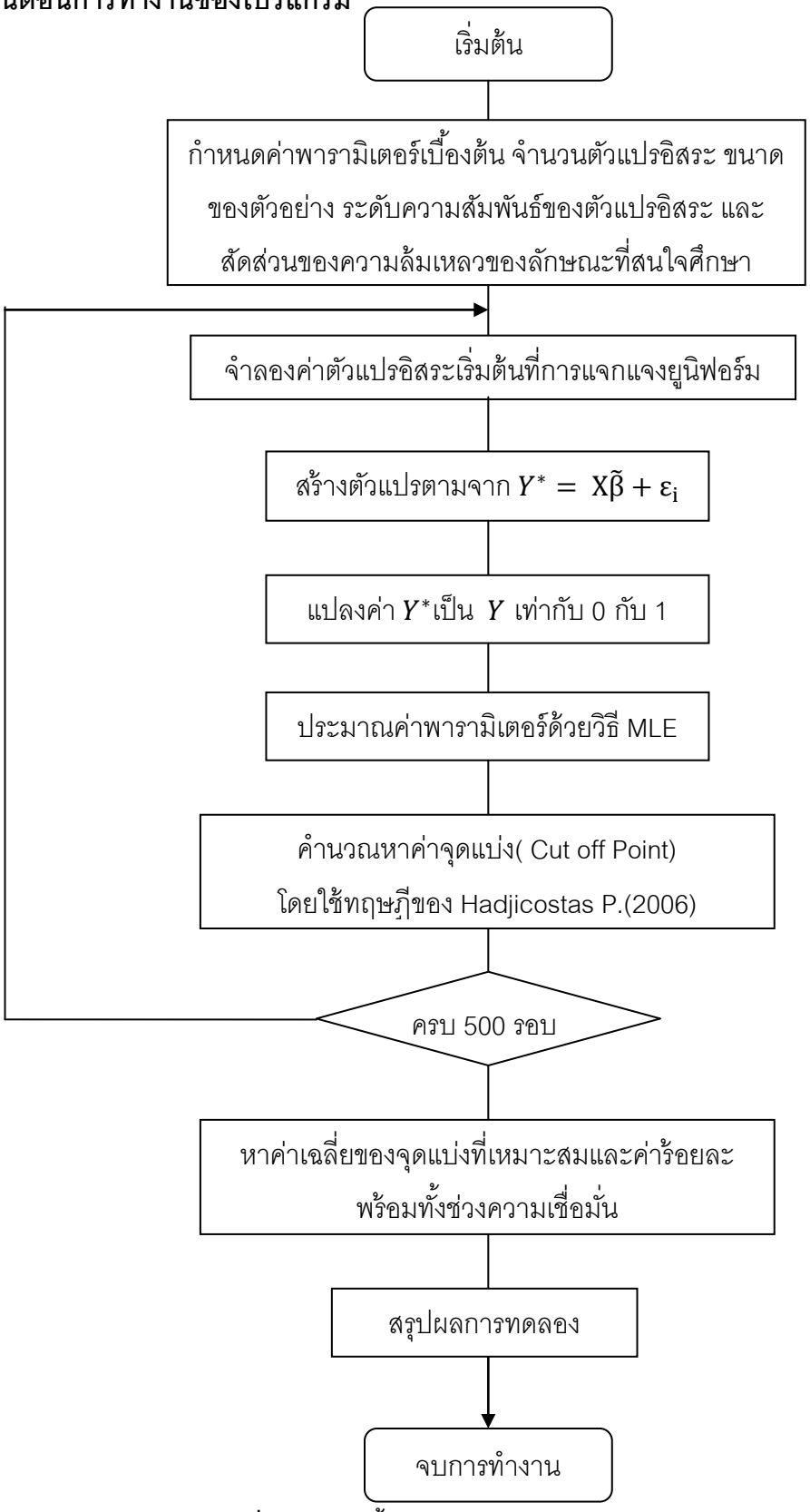
$$P(L < \hat{c} < U) = (1 - \alpha)$$

โดยที่ L คือ ขีดจำกัดความเชื่อมั่นล่าง (Lower Confidence Limit) เปอร์เซ็นต์ไทล์ที่ 2.5

U คือ ขีดจำกัดความเชื่อมั่นบน (Upper Confidence Limit) เปอร์เซ็นต์ไทล์ที่ 9.75

ขั้นที่ 5 สรุปผลที่ได้จากการวิจัยในแต่สถานการณ์

3.3 ขั้นตอนการทำงานของโปรแกรม



รูปที่ 3.1 **แสดงขั้นตอนของการทำงานของโปรแกรม**

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ต้องการหาจุดแบ่งที่เหมาะสมสำหรับการจำแนกกลุ่มของข้อมูลในตัวแบบถดถอยโลจิสติกแบบ 2 ประเภทโดยมีฟังก์ชันคอมพลีเมนทารีล็อก-ล็อกเป็นฟังก์ชันเชื่อมโยง โดยตัวแปรอิสระนั้นมีการแจกแจงเริ่มต้นเป็นการแจกแจงแบบยูนิฟอร์ม จากนั้นพิจารณาปัจจัยอื่นๆ ว่ามีผลต่อการหาจุดแบ่งหรือไม่ เช่น จำนวนตัวแปรอิสระที่เพิ่มขึ้น ขนาดตัวอย่างที่เปลี่ยนไป อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเพิ่มขึ้นจากระดับต่ำ กลาง และ สูง รวมถึงระดับความสัมพันธ์ของตัวแปรอิสระที่มีค่าสูงขึ้น ซึ่งผลการวิเคราะห์ข้อมูลอยู่ภายใต้สถานการณ์ดังต่อไปนี้

1. กำหนดให้อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา(a) เพิ่มขึ้นจาก 0.25,0.5 และ 0.75 ตามลำดับ โดยที่ จำนวนตัวแปรอิสระ(p) ขนาดตัวอย่าง(n) และระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)มีค่าคงที่
2. กำหนดให้ขนาดตัวอย่าง(n)เพิ่มขึ้น จาก 50,100,150,200 ไปจนถึง 250 ตามลำดับ โดยที่ จำนวนตัวแปรอิสระ(p) อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) และระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)มีค่าคงที่
3. กำหนดให้จำนวนของตัวแปรอิสระ(p)เพิ่มขึ้นจาก 1,2,3,4 และ 5 ตามลำดับ โดยที่ขนาดของตัวอย่าง(n) อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) และระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)มีค่าคงที่
4. กำหนดให้ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)เพิ่มขึ้นจากระดับ ต่ำ กลาง และสูง ตามลำดับ โดยที่ จำนวนตัวแปรอิสระ(p) ขนาดตัวอย่าง(n) และอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a) คงที่

งานวิจัยนี้นำเสนอผลของการวิเคราะห์ข้อมูลในรูปแบบตาราง ซึ่งกำหนดสัญลักษณ์แทนความหมายต่างๆดังนี้

n	คือ ขนาดตัวอย่าง
p	คือ จำนวนตัวแปรอิสระ
ρ	คือ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ
a	คือ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา
$\bar{c}$	คือ ค่าเฉลี่ยของจุดแบ่ง

สัญลักษณ์แทนความหมาย (ต่อ)

Cl. Lower of $\hat{c}$ แทน ค่าต่ำสุดของช่วงความเชื่อมั่นของจุดแบ่ง

Cl. Upper of $\hat{c}$ แทน ค่าสูงสุดของช่วงความเชื่อมั่นของจุดแบ่ง

4.1 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสมในแต่ละสถานการณ์

4.1.1 แสดงค่าเฉลี่ยจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง โดยที่ ขนาดของตัวอย่าง จำนวนตัวแปรอิสระ และระดับความสัมพันธ์ของตัวแปรอิสระคงที่ ดังตารางที่ 4.1-4.15

4.1.2 แสดงค่าเฉลี่ยจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดของตัวอย่างเปลี่ยนแปลง โดยที่ จำนวนตัวแปรอิสระ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา และระดับความสัมพันธ์ของตัวแปรอิสระคงที่ นำเสนอในตารางที่ 4.16-4.24

4.1.3 แสดงค่าเฉลี่ยจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อจำนวนตัวแปรอิสระเปลี่ยนแปลง โดยที่ ขนาดตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ และอัตราส่วนความล้มเหลวของลักษณะที่สนใจคงที่ นำเสนอในตารางที่ 4.25-4.33

4.1.4 แสดงค่าเฉลี่ยจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง โดยที่ ขนาดตัวอย่าง จำนวนตัวแปรอิสระและอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาคงที่ นำเสนอในตารางที่ 4.34-4.48

4.1.1 กรณีที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.1-4.15

ตารางที่ 4.1 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	0.25	0.4882	0.2628	0.7205
			0.5	0.4848	0.2611	0.7578
			0.75	0.4883	0.2799	0.7150
		2	0.25	0.4085	0.0853	0.7163
			0.5	0.4334	0.1505	0.7024
			0.75	0.4434	0.1472	0.7411
		3	0.25	0.2238	0.0000	0.6755
			0.5	0.2220	0.0000	0.6752
			0.75	0.2637	0.0000	0.6488
		4	0.25	0.0469	0.0000	0.5724
			0.5	0.0515	0.0000	0.4970
			0.75	0.0477	0.0000	0.5043
		5	0.25	0.0058	0.0000	0.0948
			0.5	0.0075	0.0000	0.0539
			0.75	0.0090	0.0000	0.0207

จากตารางที่ 4.1 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นลงสลับกัน ตัวแปรอิสระเท่ากับ 2 ตัวพบว่าค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ตัวแปรอิสระเท่ากับ 3 และ 4 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าขึ้นลงสลับกัน และที่ตัวแปรอิสระเท่ากับ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.2 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	0.25	0.4255	0.1468	0.7091
			0.5	0.4356	0.1141	0.7127
			0.75	0.4425	0.1807	0.7462
		3	0.25	0.2478	0.0000	0.6551
			0.5	0.2496	0.0000	0.6268
			0.75	0.3004	0.0000	0.6952
		4	0.25	0.0641	0.0000	0.5237
			0.5	0.0837	0.0000	0.5649
			0.75	0.1038	0.0000	0.6475
		5	0.25	0.0208	0.0000	0.3755
			0.5	0.0242	0.0000	0.4876
			0.75	0.0192	0.0000	0.3882

จากตารางที่ 4.2 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 2,3 และ 4 ตัว พบว่าค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และ ที่ตัวแปรอิสระเท่ากับ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา

ตารางที่ 4.3 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	0.25	0.4186	0.0916	0.7224
			0.5	0.4391	0.1590	0.7346
			0.75	0.4460	0.1555	0.7423
		3	0.25	0.3070	0.0000	0.6978
			0.5	0.3104	0.0000	0.7046
			0.75	0.3379	0.0000	0.7035
		4	0.25	0.1061	0.0000	0.6467
			0.5	0.1149	0.0000	0.6357
			0.75	0.1261	0.0000	0.6403
		5	0.25	0.0265	0.0000	0.4866
			0.5	0.0366	0.0000	0.5326
			0.75	0.0358	0.0000	0.5545

จากตารางที่ 4.3 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 50 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 2,3 และ 4 ตัว พบว่าค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และ ที่ตัวแปรอิสระเท่ากับ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา

ตารางที่ 4.4 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	0.25	0.4846	0.3125	0.6700
			0.5	0.4807	0.3000	0.6825
			0.75	0.4833	0.3211	0.6935
		2	0.25	0.4281	0.1575	0.6860
			0.5	0.4355	0.1941	0.6971
			0.75	0.4464	0.2160	0.6711
		3	0.25	0.3334	0.0000	0.6609
			0.5	0.3326	0.0000	0.6423
			0.75	0.3522	0.0000	0.6709
		4	0.25	0.0973	0.0000	0.5606
			0.5	0.1058	0.0000	0.5766
			0.75	0.1546	0.0000	0.5825
		5	0.25	0.0157	0.0000	0.3200
			0.5	0.0158	0.0000	0.2717
			0.75	0.0171	0.0000	0.3238

จากตารางที่ 4.4 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 3 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระ เท่ากับ 2,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.5 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	0.25	0.4269	0.1931	0.6639
			0.5	0.4472	0.2250	0.6816
			0.75	0.4563	0.2390	0.7011
		3	0.25	0.3626	0.0000	0.6629
			0.5	0.3709	0.0000	0.6790
			0.75	0.3878	0.0000	0.6790
		4	0.25	0.1465	0.0000	0.6015
			0.5	0.1700	0.0000	0.6228
			0.75	0.1939	0.0000	0.6059
		5	0.25	0.0277	0.0000	0.3656
			0.5	0.0332	0.0000	0.4275
			0.75	0.0374	0.0000	0.4978

จากตารางที่ 4.5 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.6 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง$0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	0.25	0.4464	0.2421	0.6954
			0.5	0.4578	0.2470	0.7052
			0.75	0.4660	0.2688	0.6818
		3	0.25	0.3955	0.0000	0.6862
			0.5	0.4096	0.0971	0.6791
			0.75	0.4318	0.1307	0.6791
		4	0.25	0.2268	0.0000	0.6265
			0.5	0.2494	0.0000	0.6806
			0.75	0.2730	0.0000	0.6256
		5	0.25	0.0720	0.0000	0.5657
			0.5	0.0848	0.0000	0.5780
			0.75	0.0793	0.0000	0.5418

จากตารางที่ 4.6 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 100 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2,3 และ 4 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่จำนวนตัวแปรอิสระเท่ากับ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา

ตารางที่ 4.7 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	0.25	0.4782	0.3179	0.6653
			0.5	0.4744	0.3164	0.6450
			0.75	0.4831	0.3365	0.6528
		2	0.25	0.4450	0.2241	0.6509
			0.5	0.4423	0.2468	0.6541
			0.75	0.4546	0.2717	0.6694
		3	0.25	0.3656	0.0697	0.6509
			0.5	0.3828	0.1051	0.6441
			0.75	0.4001	0.1057	0.6547
		4	0.25	0.1625	0.0000	0.5854
			0.5	0.1758	0.0000	0.5918
			0.75	0.2165	0.0000	0.6159
		5	0.25	0.0231	0.0000	0.3867
			0.5	0.0267	0.0000	0.4096
			0.75	0.0281	0.0000	0.3847

จากตารางที่ 4.7 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0 < Max\{|r_{ij}|\} < 0.3$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระ เท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.8 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	0.25	0.4327	0.2377	0.6439
			0.5	0.4502	0.2572	0.6638
			0.75	0.4635	0.2803	0.6672
		3	0.25	0.3794	0.0935	0.6534
			0.5	0.3968	0.1464	0.6673
			0.75	0.4052	0.1700	0.6470
		4	0.25	0.2277	0.0000	0.6299
			0.5	0.2473	0.0000	0.6369
			0.75	0.2594	0.0000	0.6132
		5	0.25	0.0501	0.0000	0.4599
			0.5	0.0491	0.0000	0.5371
			0.75	0.0735	0.0000	0.5163

จากตารางที่ 4.8 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.9 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง (n) เท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	0.25	0.4427	0.2438	0.6552
			0.5	0.4567	0.2836	0.6672
			0.75	0.4616	0.2996	0.6569
		3	0.25	0.4064	0.1513	0.6772
			0.5	0.4222	0.1852	0.6725
			0.75	0.4299	0.2087	0.6771
		4	0.25	0.3141	0.0000	0.6569
			0.5	0.3454	0.0000	0.6483
			0.75	0.3711	0.0000	0.6802
		5	0.25	0.1194	0.0000	0.6014
			0.5	0.1440	0.0000	0.5683
			0.75	0.1785	0.0000	0.6243

จากตารางที่ 4.9 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 150 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.10 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	0.25	0.4726	0.3281	0.6353
			0.5	0.4766	0.3332	0.6434
			0.75	0.4718	0.3304	0.6445
		2	0.25	0.4416	0.2586	0.6239
			0.5	0.4546	0.2675	0.6411
			0.75	0.4569	0.2774	0.6419
		3	0.25	0.3811	0.0905	0.6452
			0.5	0.3842	0.1469	0.6287
			0.75	0.3972	0.1181	0.6423
		4	0.25	0.1940	0.0000	0.5723
			0.5	0.2255	0.0000	0.6292
			0.75	0.2620	0.0000	0.6264
		5	0.25	0.0367	0.0000	0.4338
			0.5	0.0415	0.0000	0.4054
			0.75	0.0567	0.0000	0.4800

จากตารางที่4.10 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.11 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	0.25	0.4407	0.2579	0.6295
			0.5	0.4528	0.2751	0.6386
			0.75	0.4573	0.2911	0.6273
		3	0.25	0.3937	0.1468	0.6349
			0.5	0.4016	0.1586	0.6756
			0.75	0.4178	0.1960	0.6579
		4	0.25	0.2691	0.0000	0.6273
			0.5	0.2819	0.0000	0.5835
			0.75	0.3232	0.0000	0.6466
		5	0.25	0.0642	0.0000	0.5359
			0.5	0.0735	0.0000	0.5260
			0.75	0.1001	0.0000	0.5503

จากตารางที่4.11 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.12 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	0.25	0.4519	0.2821	0.6529
			0.5	0.4587	0.2921	0.6567
			0.75	0.4585	0.2924	0.6460
		3	0.25	0.4090	0.1922	0.6500
			0.5	0.4331	0.2167	0.6711
			0.75	0.4510	0.2386	0.6881
		4	0.25	0.3626	0.0000	0.6596
			0.5	0.3744	0.0000	0.6341
			0.75	0.3985	0.0427	0.6676
		5	0.25	0.1856	0.0000	0.5878
			0.5	0.2043	0.0000	0.5986
			0.75	0.2565	0.0000	0.6455

จากตารางที่4.12 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 200 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระ เท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.13 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	0.25	0.4730	0.3336	0.6254
			0.5	0.4766	0.3389	0.6178
			0.75	0.4752	0.3470	0.6231
		2	0.25	0.4523	0.2875	0.6420
			0.5	0.4566	0.2799	0.6209
			0.75	0.4558	0.3161	0.6153
		3	0.25	0.3706	0.1130	0.6354
			0.5	0.3884	0.1457	0.6325
			0.75	0.4171	0.1702	0.6704
		4	0.25	0.2537	0.0000	0.6000
			0.5	0.2570	0.0000	0.6283
			0.75	0.2982	0.0000	0.6119
		5	0.25	0.0489	0.0000	0.4587
			0.5	0.0508	0.0000	0.4189
			0.75	0.0631	0.0000	0.4650

จากตารางที่4.13 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระ เท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.14 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a)

n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	0.25	0.4428	0.2679	0.6436
			0.5	0.4508	0.2852	0.6208
			0.75	0.4632	0.3139	0.6458
		3	0.25	0.3849	0.1523	0.6406
			0.5	0.4086	0.1656	0.6619
			0.75	0.4311	0.1893	0.6484
		4	0.25	0.2899	0.0000	0.6102
			0.5	0.3188	0.0000	0.6258
			0.75	0.3516	0.0000	0.6493
		5	0.25	0.0979	0.0000	0.4979
			0.5	0.1148	0.0000	0.5400
			0.75	0.1342	0.0000	0.5247

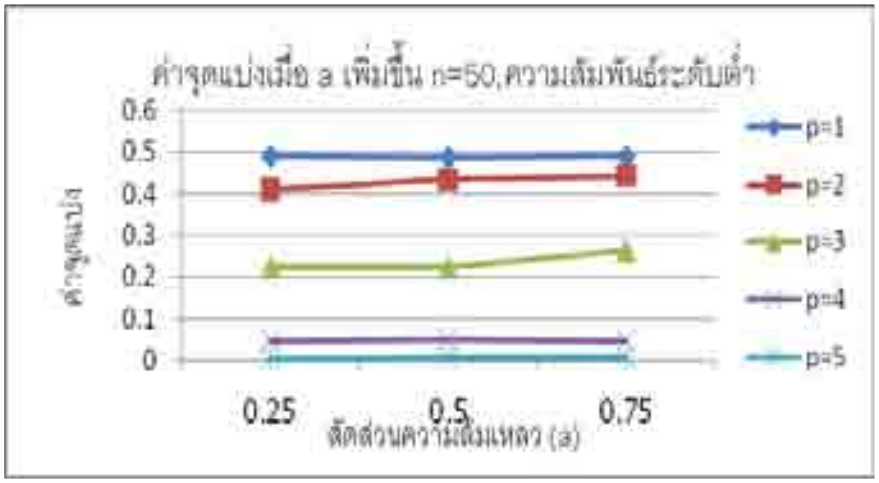
จากตารางที่4.14 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

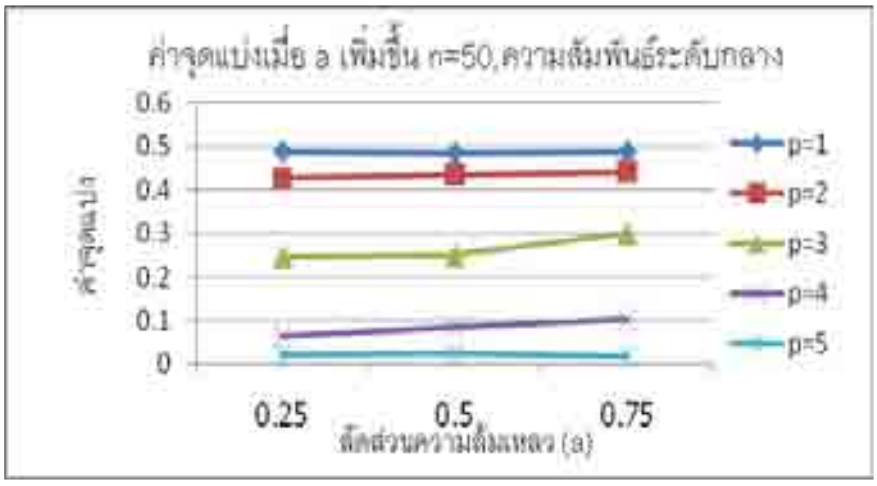
ตารางที่ 4.15 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อมีขนาดตัวอย่าง(n) เท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา (a)

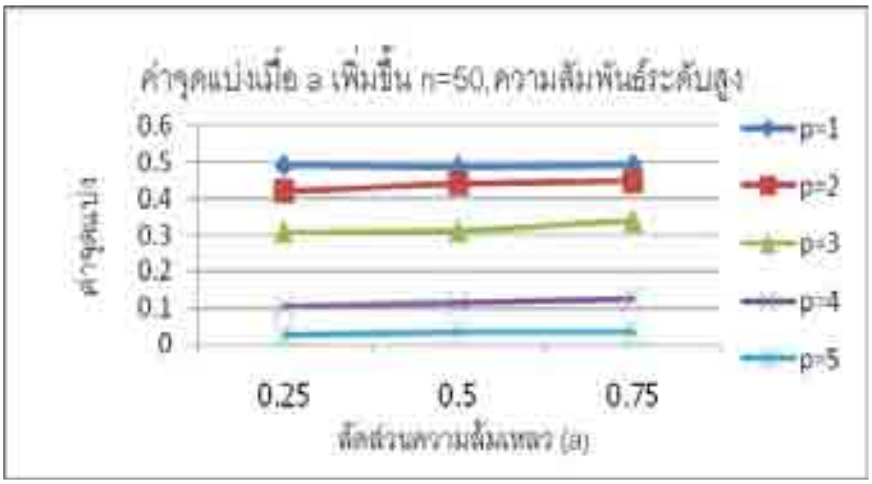
n	ρ	p	a	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	0.25	0.4466	0.2980	0.6263
			0.5	0.4592	0.3043	0.6264
			0.75	0.4694	0.3114	0.6311
		3	0.25	0.4179	0.2069	0.6270
			0.5	0.4250	0.2376	0.6538
			0.75	0.4465	0.2475	0.6851
		4	0.25	0.3769	0.0000	0.6391
			0.5	0.3866	0.1291	0.6555
			0.75	0.4075	0.1522	0.6429
		5	0.25	0.2302	0.0000	0.6221
			0.5	0.2603	0.0000	0.6125
			0.75	0.3047	0.0000	0.6385

จากตารางที่4.15 เมื่อพิจารณาค่าจุดแบ่งที่ขนาดตัวอย่างเท่ากับ 250 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ โดยที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง พบว่า ที่จำนวนตัวแปรอิสระ เท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

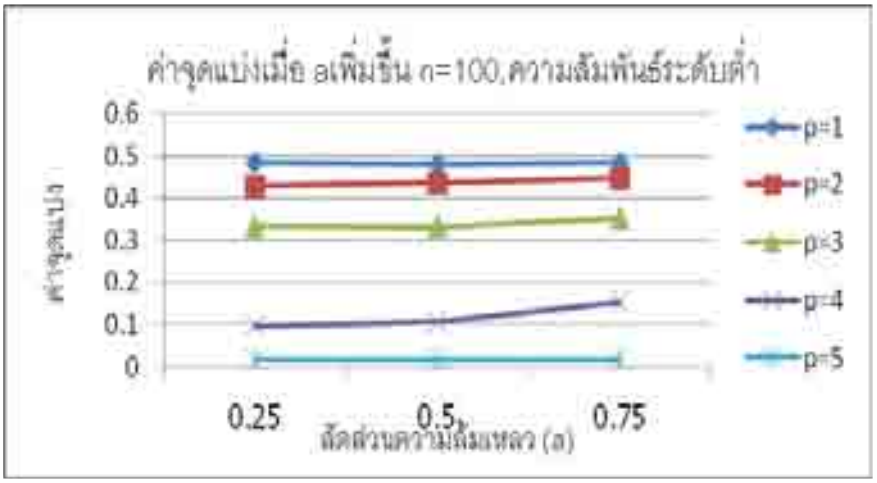
ภาพที่ 4.1 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

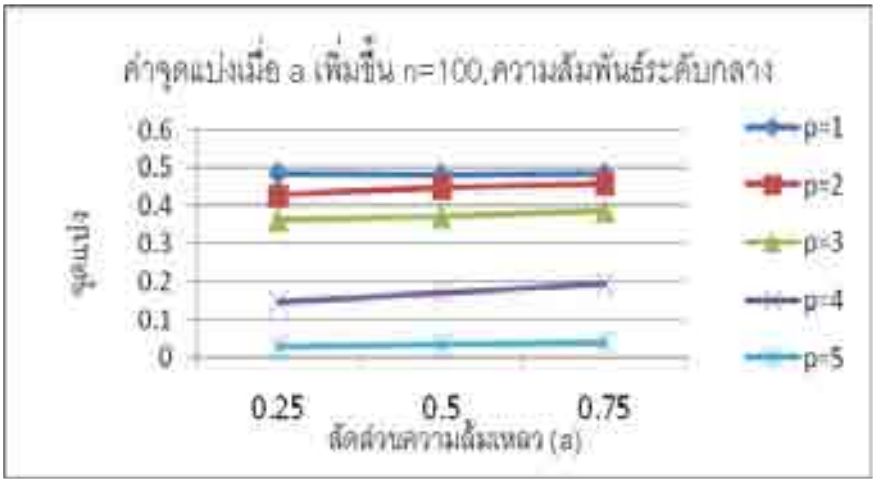


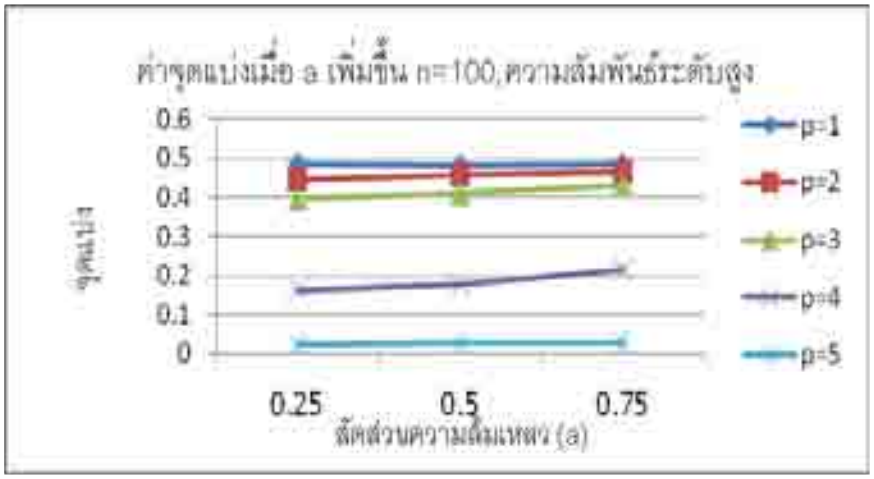




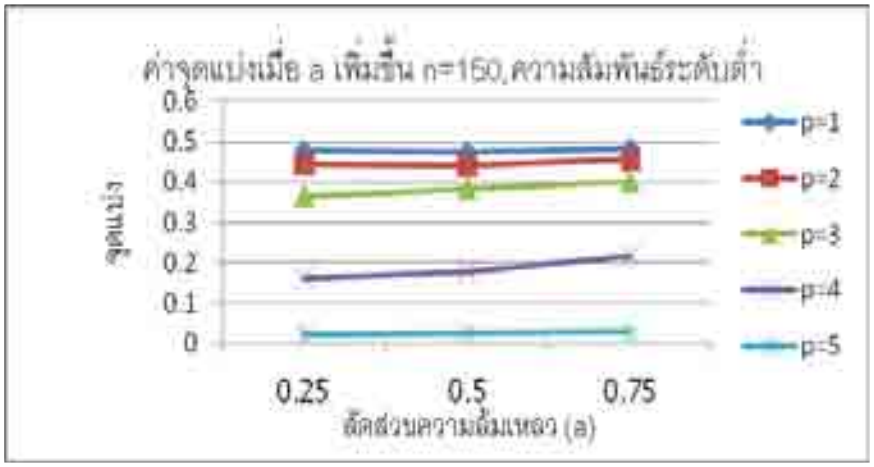
ภาพที่ 4.1 (ต่อ)

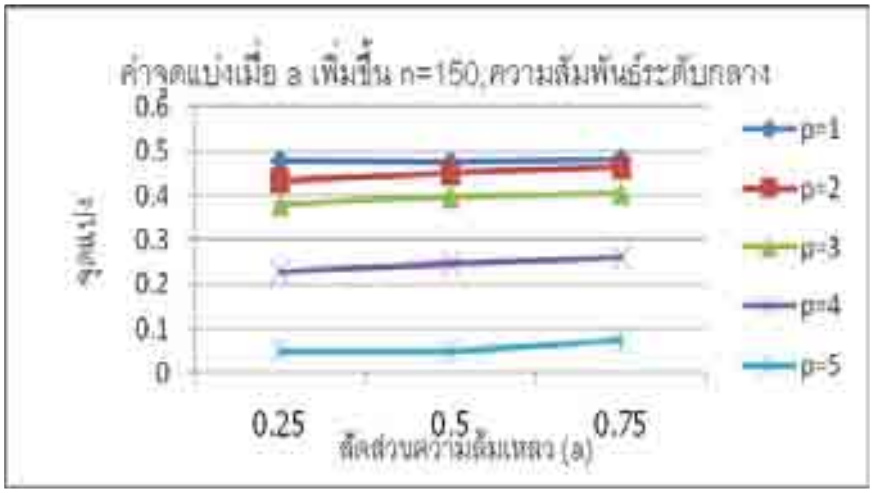

ค่าจุดแบ่งเมื่อ aเพิ่มขึ้น n=100,ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5

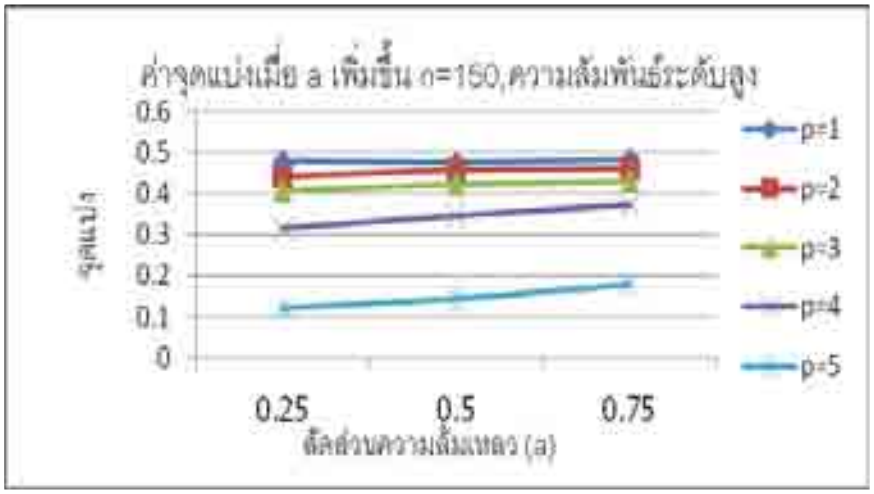

ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=100,ความสัมพันธ์ระดับกลาง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=100,ความสัมพันธ์ระดับสูง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5

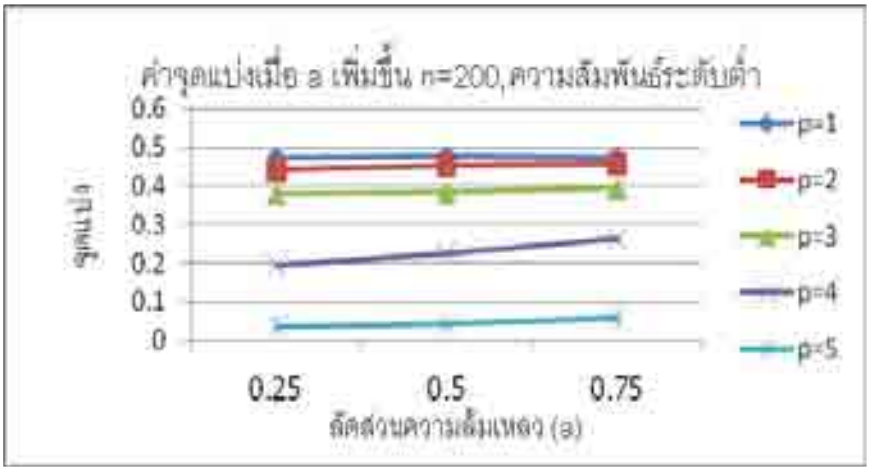
ภาพที่ 4.1 (ต่อ)

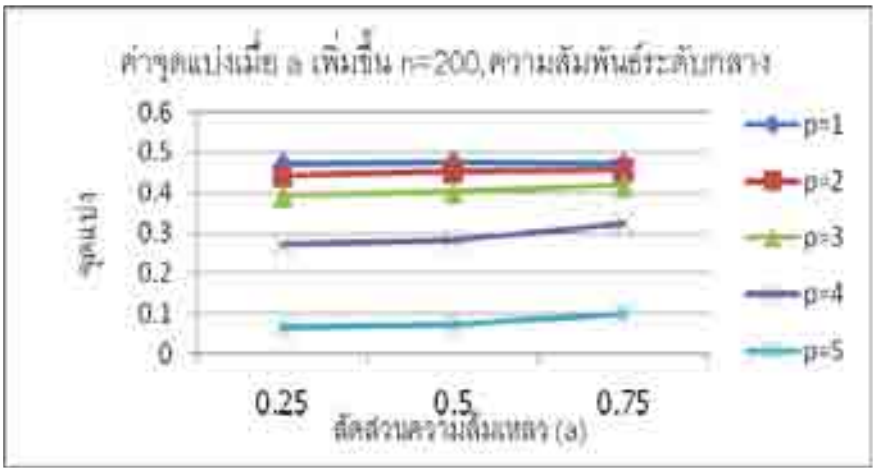

ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=150, ความสัมพันธ์ระดับต่ำ
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
อัตส่วนความสัมเหลว (a)
p=1
p=2
p=3
p=4
p=5

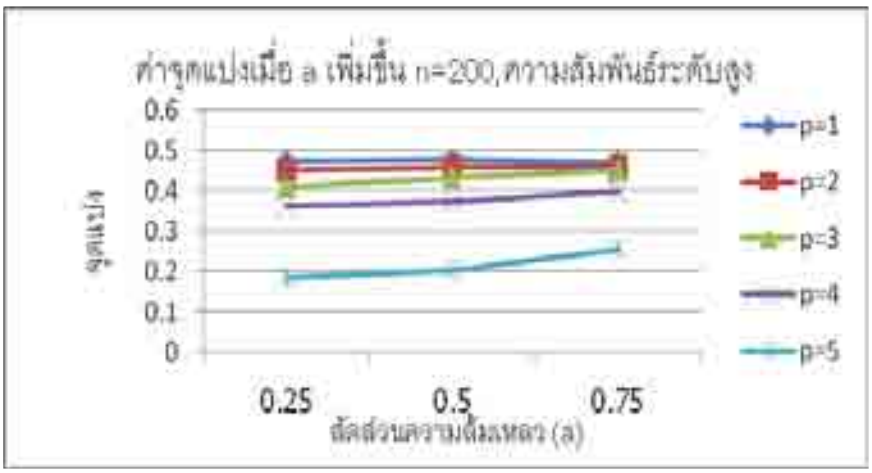

ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=150, ความสัมพันธ์ระดับกลาง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
อัตส่วนความสัมเหลว (a)
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=150, ความสัมพันธ์ระดับสูง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
อัตส่วนความสัมเหลว (a)
p=1
p=2
p=3
p=4
p=5

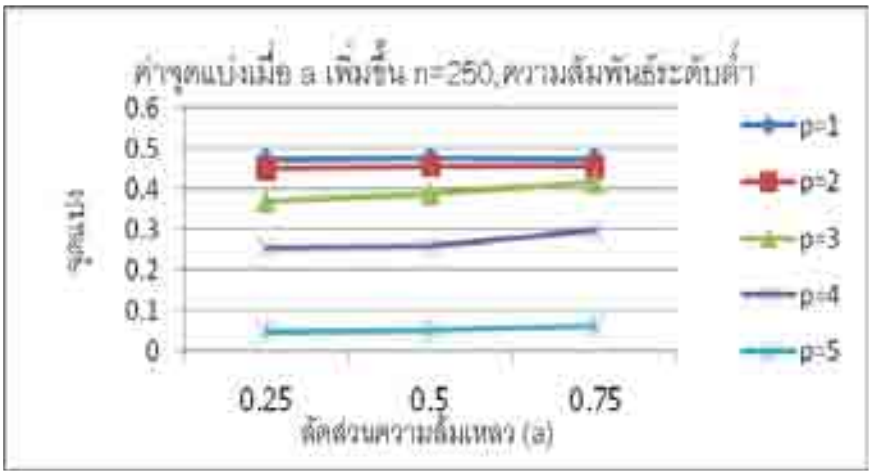
ภาพที่ 4.1 (ต่อ)

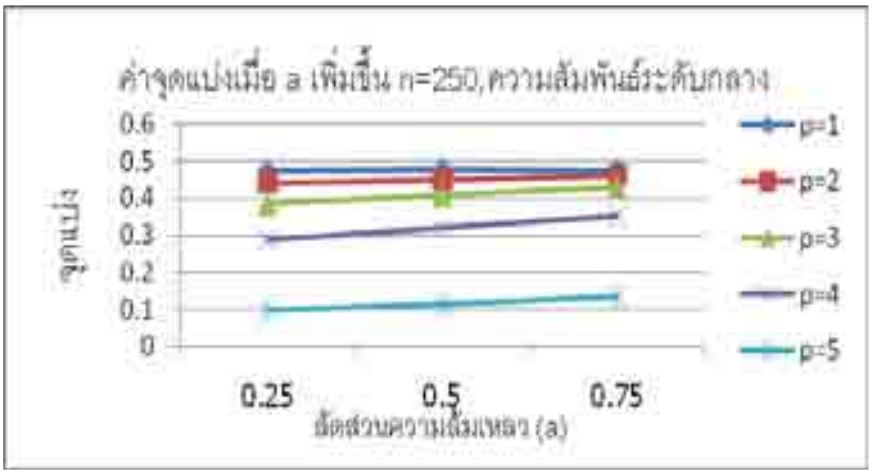

ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=200, ความสัมพันธ์ระดับต่ำ
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5

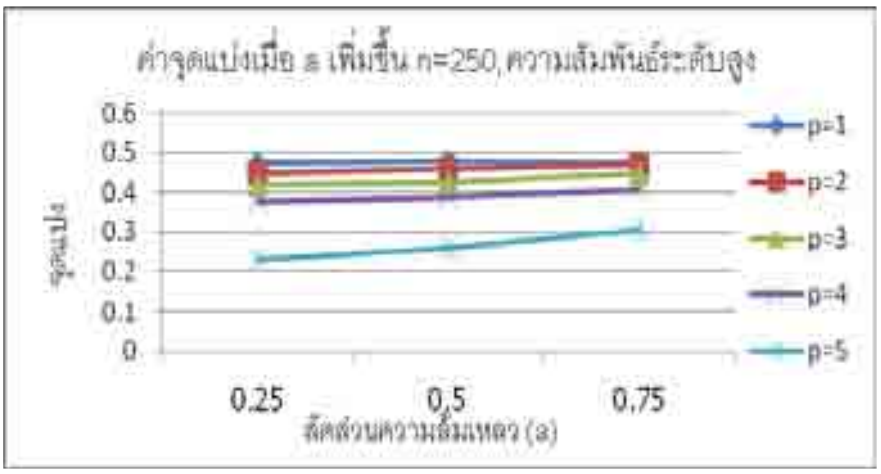

ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=200, ความสัมพันธ์ระดับกลาง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=200, ความสัมพันธ์ระดับสูง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความเข้มแหลว (a)
p=1
p=2
p=3
p=4
p=5

ภาพที่ 4.1 (ต่อ)


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=250,ความสัมพันธ์ระดับต่ำ
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความล้มเหลว (a)
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=250,ความสัมพันธ์ระดับกลาง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความล้มเหลว (a)
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ a เพิ่มขึ้น n=250,ความสัมพันธ์ระดับสูง
จุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.5
0.75
สัดส่วนความล้มเหลว (a)
p=1
p=2
p=3
p=4
p=5

4.1.2 กรณีที่ขนาดของตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจ และระดับความสัมพันธ์ของตัวแปรอิสระคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.16 -4.24 ดังนี้

ตารางที่ 4.16 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	50	0.4882	0.2628	0.7205
			100	0.4846	0.3125	0.6700
			150	0.4782	0.3179	0.6653
			200	0.4726	0.3281	0.6353
			250	0.4730	0.3336	0.6254
		2	50	0.2238	0.0000	0.6755
			100	0.4281	0.1575	0.6860
			150	0.4450	0.2241	0.6509
			200	0.4416	0.2586	0.6239
			250	0.4523	0.2875	0.6420
		3	50	0.4883	0.2799	0.7150
			100	0.3334	0.0000	0.6609
			150	0.3656	0.0697	0.6509
			200	0.3811	0.0905	0.6452
			250	0.3706	0.1130	0.6354

ตารางที่ 4.16 (ต่อ)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	4	50	0.0469	0.0000	0.5724
			100	0.0973	0.0000	0.5606
			150	0.1625	0.0000	0.5854
			200	0.1940	0.0000	0.5723
			250	0.2537	0.0000	0.6000
		5	50	0.0058	0.0000	0.0948
			100	0.0157	0.0000	0.3200
			150	0.0231	0.0000	0.3867
			200	0.0367	0.0000	0.4338
			250	0.0489	0.0000	0.4587

จากตารางที่4.16 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างต่ำ $0 < Max\{\lvert r_{ij}\rvert\} < 0.3$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 1,2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และ จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.17 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	50	0.4255	0.1468	0.7091
			100	0.4269	0.1931	0.6639
			150	0.4327	0.2377	0.6439
			200	0.4407	0.2579	0.6295
			250	0.4428	0.2679	0.6436
		3	50	0.2478	0.0000	0.6551
			100	0.3626	0.0000	0.6629
			150	0.3794	0.0935	0.6534
			200	0.3937	0.1468	0.6349
			250	0.3849	0.1523	0.6406
		4	50	0.0641	0.0000	0.5237
			100	0.1465	0.0000	0.6015
			150	0.2277	0.0000	0.6299
			200	0.2691	0.0000	0.6273
			250	0.2899	0.0000	0.6102
		5	50	0.0208	0.0000	0.3755
			100	0.0277	0.0000	0.3656
			150	0.0501	0.0000	0.4599
			200	0.0642	0.0000	0.5359
			250	0.0979	0.0000	0.4979

จากตารางที่4.17 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และจำนวนตัวแปรอิสระเท่ากับ 2,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.18 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัวโดยจำแนกตามขนาดตัวอย่าง

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	50	0.4186	0.0916	0.7224
			100	0.4464	0.2421	0.6954
			150	0.4427	0.2438	0.6552
			200	0.4519	0.2821	0.6529
			250	0.4466	0.2980	0.6263
		3	50	0.3070	0.0000	0.6978
			100	0.3955	0.0000	0.6862
			150	0.4064	0.1513	0.6772
			200	0.4090	0.1922	0.6500
			250	0.4179	0.2069	0.6270
		4	50	0.1061	0.0000	0.6467
			100	0.2268	0.0000	0.6265
			150	0.3141	0.0000	0.6569
			200	0.3626	0.0000	0.6596
			250	0.3769	0.0000	0.6391
		5	50	0.0265	0.0000	0.4866
			100	0.0720	0.0000	0.5657
			150	0.1194	0.0000	0.6014
			200	0.1856	0.0000	0.5878
			250	0.2302	0.0000	0.6221

จากตารางที่4.18 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับอย่างสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และ จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.19 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	50	0.4848	0.2611	0.7578
			100	0.4807	0.3000	0.6825
			150	0.4567	0.2836	0.6672
			200	0.4766	0.3332	0.6434
			250	0.4766	0.3389	0.6178
		2	50	0.4334	0.1505	0.7024
			100	0.4355	0.1941	0.6971
			150	0.4423	0.2468	0.6541
			200	0.4546	0.2675	0.6411
			250	0.4566	0.2799	0.6209
		3	50	0.2220	0.0000	0.6752
			100	0.3326	0.0000	0.6423
			150	0.3828	0.1051	0.6441
			200	0.3842	0.1469	0.6287
			250	0.3884	0.1457	0.6325
		4	50	0.0515	0.0000	0.4970
			100	0.1058	0.0000	0.5766
			150	0.1758	0.0000	0.5918
			200	0.2255	0.0000	0.6292
			250	0.2570	0.0000	0.6283

ตารางที่ 4.19 (ต่อ)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0. < Max\{\lvert r_{ij}\rvert\} < 0.3$	5	50	0.0075	0.0000	0.0539
			100	0.0158	0.0000	0.2717
			150	0.0267	0.0000	0.4096
			200	0.0415	0.0000	0.4054
			250	0.0508	0.0000	0.4189

จากตารางที่4.19 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และ จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.20 แสดงค่าจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	50	0.4356	0.1141	0.7127
			100	0.4472	0.2250	0.6816
			150	0.4502	0.2572	0.6638
			200	0.4528	0.2751	0.6386
			250	0.4508	0.2852	0.6208
		3	50	0.2496	0.0000	0.6268
			100	0.3709	0.0000	0.6790
			150	0.3968	0.1464	0.6673
			200	0.4016	0.1586	0.6756
			250	0.4086	0.1656	0.6619
		4	50	0.0837	0.0000	0.5649
			100	0.1700	0.0000	0.6228
			150	0.2473	0.0000	0.6369
			200	0.2819	0.0000	0.5835
			250	0.3188	0.0000	0.6258
		5	50	0.0242	0.0000	0.4876
			100	0.0332	0.0000	0.4275
			150	0.0491	0.0000	0.5371
			200	0.0735	0.0000	0.5260
			250	0.1148	0.0000	0.5400

จากตารางที่4.20 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และ จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.21 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	50	0.4391	0.1590	0.7346
			100	0.4578	0.2470	0.7052
			150	0.4567	0.2836	0.6672
			200	0.4587	0.2921	0.6567
			250	0.4592	0.3043	0.6264
		3	50	0.3104	0.0000	0.7046
			100	0.4096	0.0971	0.6791
			150	0.4222	0.1852	0.6725
			200	0.4331	0.2167	0.6711
			250	0.4250	0.2376	0.6538
		4	50	0.1149	0.0000	0.6357
			100	0.2494	0.0000	0.6806
			150	0.3454	0.0000	0.6483
			200	0.3744	0.0000	0.6341
			250	0.3866	0.1291	0.6555
		5	50	0.0366	0.0000	0.5326
			100	0.0848	0.0000	0.5780
			150	0.1440	0.0000	0.5683
			200	0.2043	0.0000	0.5986
			250	0.2603	0.0000	0.6125

จากตารางที่4.21 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และ จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.22 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ตัวแปรอิสระ(p) คือ 1,2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0 < Max\{\|r_{ij}\|\} < 0.3$	1	50	0.4883	0.2799	0.7150
			100	0.4833	0.3211	0.6935
			150	0.4831	0.3365	0.6528
			200	0.4718	0.3304	0.6445
			250	0.4752	0.3470	0.6231
		2	50	0.4434	0.1472	0.7411
			100	0.4464	0.2160	0.6711
			150	0.4546	0.2717	0.6694
			200	0.4569	0.2774	0.6419
			250	0.4558	0.3161	0.6153
		3	50	0.2637	0.0000	0.6488
			100	0.3522	0.0000	0.6709
			150	0.4001	0.1057	0.6547
			200	0.3972	0.1181	0.6423
			250	0.4171	0.1702	0.6704
		4	50	0.0477	0.0000	0.5043
			100	0.1546	0.0000	0.5825
			150	0.2165	0.0000	0.6159
			200	0.2620	0.0000	0.6264
			250	0.2982	0.0000	0.6119

ตารางที่ 4.22 (ต่อ)

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	5	50	0.0090	0.0000	0.0207
			100	0.0171	0.0000	0.3238
			150	0.0281	0.0000	0.3847
			200	0.0567	0.0000	0.4800
			250	0.0631	0.0000	0.4650

จากตารางที่4.22 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 1,2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และจำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.23 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับต่ำ $0.3 < Max\{|r_{ij}|\} < 0.6$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง

a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	2	50	0.4425	0.1807	0.7462
			100	0.4563	0.2390	0.7011
			150	0.4635	0.2803	0.6672
			200	0.4573	0.2911	0.6273
			250	0.4632	0.3139	0.6458
		3	50	0.3004	0.0000	0.6952
			100	0.3878	0.0000	0.6790
			150	0.4052	0.1700	0.6470
			200	0.4178	0.1960	0.6579
			250	0.4311	0.1893	0.6484
		4	50	0.1038	0.0000	0.6475
			100	0.1939	0.0000	0.6059
			150	0.2594	0.0000	0.6132
			200	0.3232	0.0000	0.6466
			250	0.2982	0.0000	0.6119
		5	50	0.0192	0.0000	0.3882
			100	0.0374	0.0000	0.4978
			150	0.0735	0.0000	0.5163
			200	0.1001	0.0000	0.5503
			250	0.1342	0.0000	0.5247

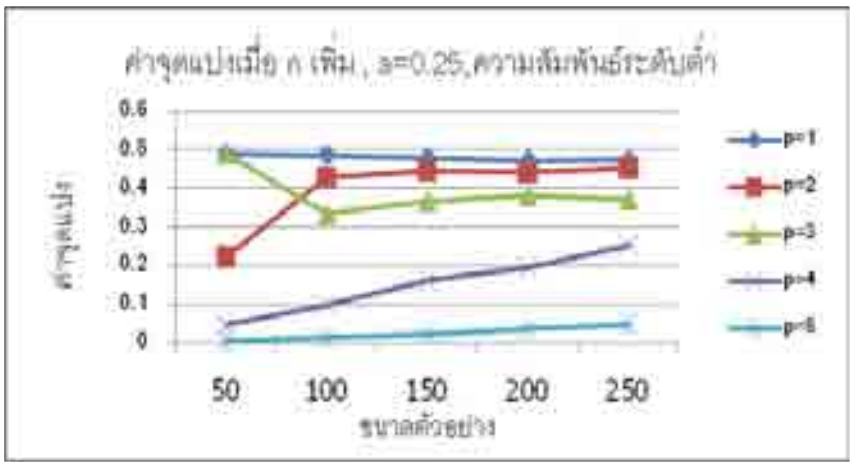
จากตารางที่4.23 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 2 และ 4 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และจำนวนตัวแปรอิสระเท่ากับ 3 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

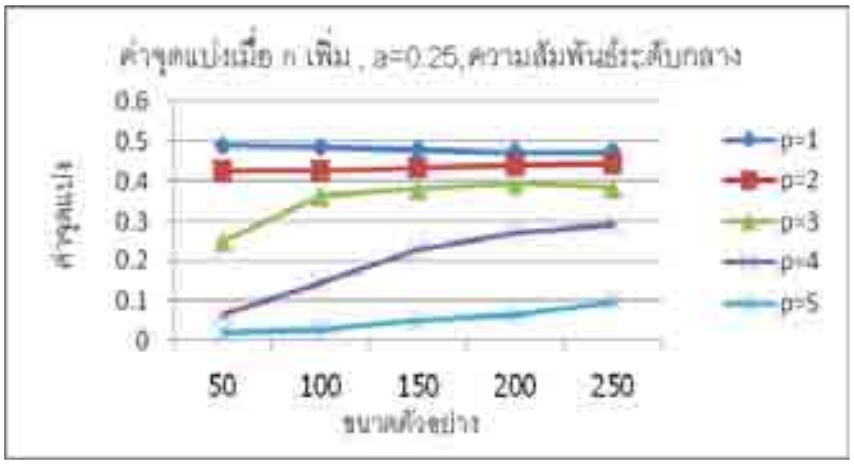
ตารางที่ 4.24 แสดงค่าเฉลี่ยจุดแบ่ง และช่วงความเชื่อมั่นของจุดแบ่ง เมื่ออัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)ในระดับกลาง $0.6 < Max\{|r_{ij}|\} < 0.9$ ตัวแปรอิสระ(p) คือ 2,3,4 และ 5 ตัว โดยจำแนกตามขนาดตัวอย่าง(n)

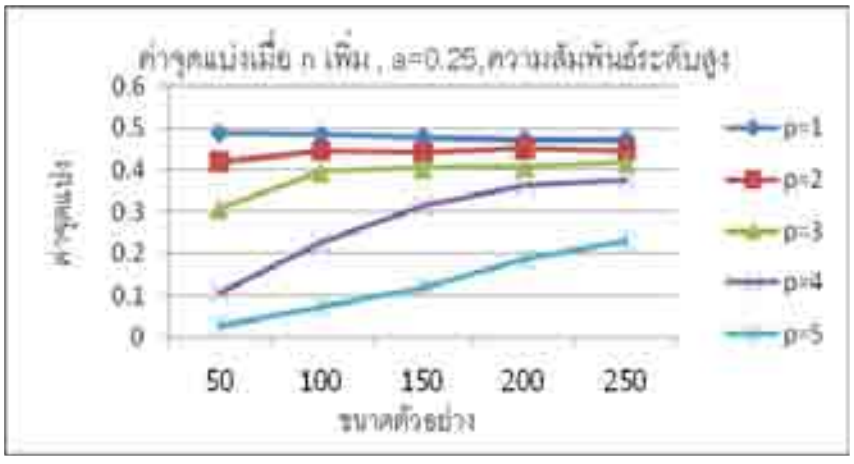
a	ρ	p	n	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	2	50	0.4460	0.1555	0.7423
			100	0.4660	0.2688	0.6818
			150	0.4616	0.2996	0.6569
			200	0.4585	0.2924	0.6460
			250	0.4694	0.3114	0.6311
		3	50	0.3379	0.0000	0.7035
			100	0.4318	0.1307	0.6791
			150	0.4299	0.2087	0.6771
			200	0.4510	0.2386	0.6881
			250	0.4465	0.2475	0.6851
		4	50	0.1261	0.0000	0.6403
			100	0.2730	0.0000	0.6256
			150	0.3711	0.0000	0.6802
			200	0.3985	0.0427	0.6676
			250	0.4075	0.1522	0.6429
		5	50	0.0358	0.0000	0.5545
			100	0.0793	0.0000	0.5418
			150	0.1785	0.0000	0.6243
			200	0.2565	0.0000	0.6455
			250	0.3047	0.0000	0.6385

จากตารางที่4.24 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดของตัวอย่างเพิ่มขึ้น พบว่า จำนวนตัวแปรอิสระเท่ากับ 2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา และจำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

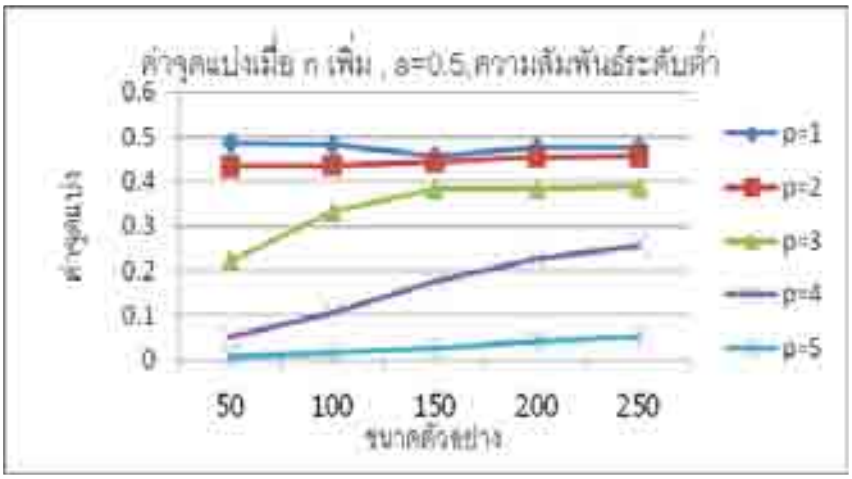
ภาพที่ 4.2 แสดงค่าเฉลี่ยของจุดแบ่งกรณีที่ขนาดของตัวอย่างเปลี่ยนแปลง เมื่อจำนวนตัวแปรอิสระ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจ และระดับความสัมพันธ์ของตัวแปรอิสระคงที่



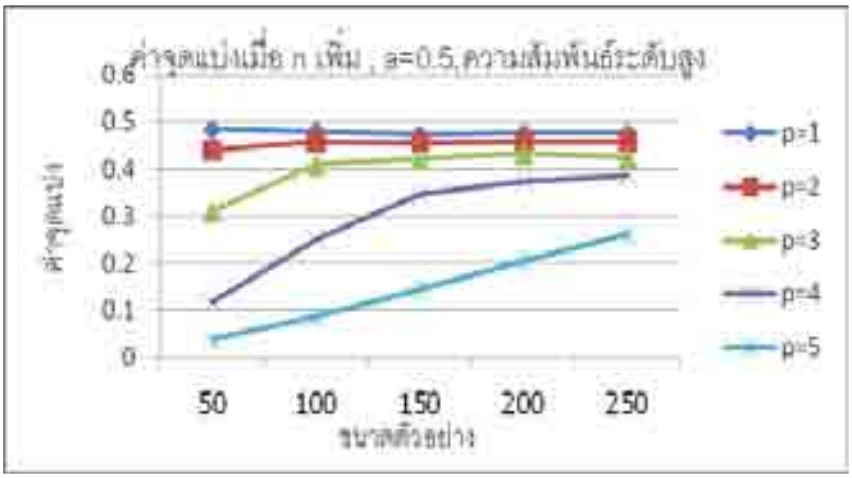




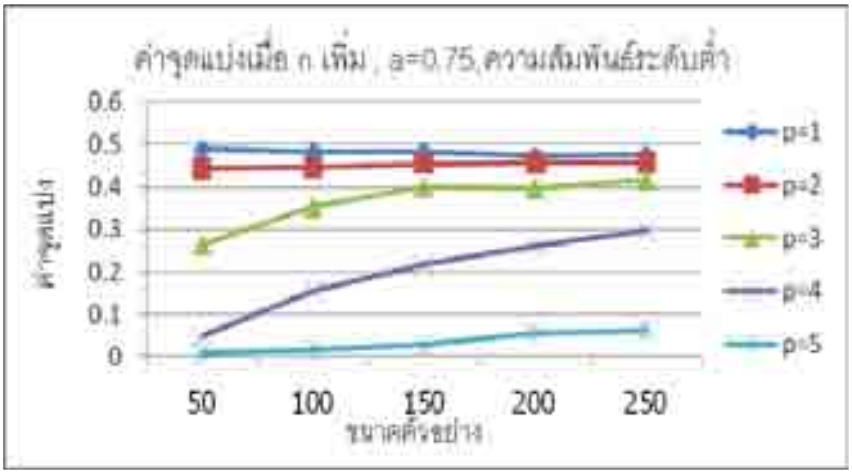
ภาพที่ 4.2 (ต่อ)

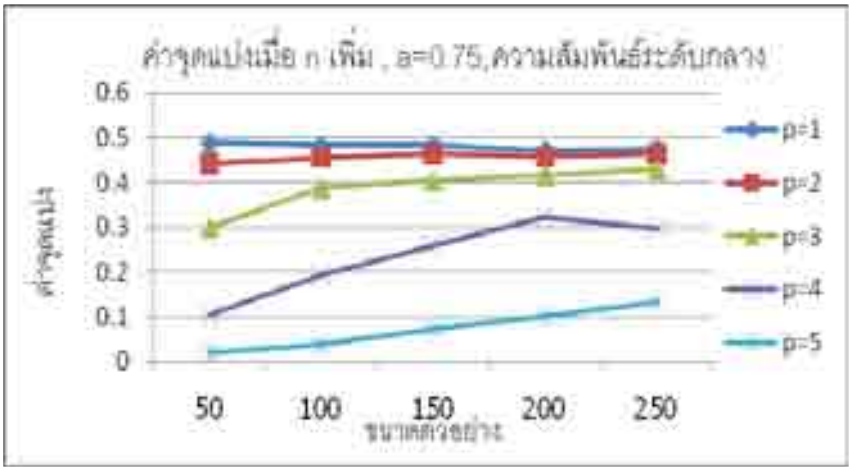

ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.5,ความสัมพันธ์ระดับต่ำ
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
ขนาดตัวอย่าง
p=1
p=2
p=3
p=4
p=5

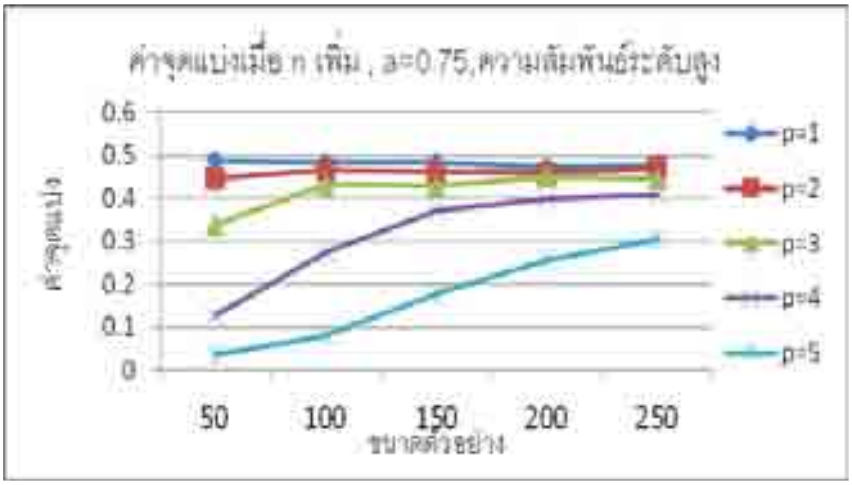
ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.5,ความสัมพันธ์ระดับกลาง
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
ขนาดตัวอย่าง
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.5,ความสัมพันธ์ระดับสูง
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
ขนาดตัวอย่าง
p=1
p=2
p=3
p=4
p=5

ภาพที่ 4.2 (ต่อ)


ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.75,ความสัมพันธ์ระดับต่ำ
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
ขนาดตัวอย่าง
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.75,ความสัมพันธ์ระดับกลาง
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
p=1
p=2
p=3
p=4
p=5


ค่าจุดแบ่งเมื่อ n เพิ่ม , a=0.75,ความสัมพันธ์ระดับสูง
0.6
0.5
0.4
0.3
0.2
0.1
0
50
100
150
200
250
ขนาดตัวอย่าง
p=1
p=2
p=3
p=4
p=5

4.1.3 กรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ และอัตราส่วนความล้มเหลวของลักษณะที่สนใจคงที่

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 4.25-4.33 ดังนี้

ตารางที่ 4.25 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0 < Max\{\|r_{ij}\|\} < 0.3$	50	1	0.4882	0.2628	0.7205
			2	0.4085	0.0853	0.7163
			3	0.2238	0.0000	0.6755
			4	0.0469	0.0000	0.5724
			5	0.0058	0.0000	0.0948
		100	1	0.4846	0.3125	0.6700
			2	0.4281	0.1575	0.6860
			3	0.3334	0.0000	0.6609
			4	0.0973	0.0000	0.5606
			5	0.0157	0.0000	0.3200
		150	1	0.4782	0.3179	0.6653
			2	0.4450	0.2241	0.6509
			3	0.3656	0.0697	0.6509
			4	0.1625	0.0000	0.5854
			5	0.0231	0.0000	0.3867

ตารางที่ 4.25 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	200	1	0.4726	0.3281	0.6353
			2	0.4416	0.2586	0.6239
			3	0.3811	0.0905	0.6452
			4	0.1940	0.0000	0.5723
			5	0.0367	0.0000	0.4338
		250	1	0.4730	0.3336	0.6254
			2	0.4523	0.2875	0.6420
			3	0.3706	0.1130	0.6354
			4	0.2537	0.0000	0.6000
			5	0.0489	0.0000	0.4587

จากตารางที่ 4.25 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.26 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.3 < Max\{\lvert r_{ij}\rvert\} < 0.6$	50	1	0.4882	0.2628	0.7205
			2	0.4255	0.1468	0.7091
			3	0.2478	0.0000	0.6551
			4	0.0641	0.0000	0.5237
			5	0.0208	0.0000	0.3755

ตารางที่ 4.26 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.3 < Max\{\lvert r_{ij}\rvert\} < 0.6$	100	1	0.4846	0.3125	0.6700
			2	0.4269	0.1931	0.6639
			3	0.3626	0.0000	0.6629
			4	0.1465	0.0000	0.6015
			5	0.0277	0.0000	0.3656
		150	1	0.4782	0.3179	0.6653
			2	0.4327	0.2377	0.6439
			3	0.3794	0.0935	0.6534
			4	0.2277	0.0000	0.6299
			5	0.0501	0.0000	0.4599
		200	1	0.4726	0.3281	0.6353
			2	0.4416	0.2586	0.6239
			3	0.3937	0.1468	0.6349
			4	0.2691	0.0000	0.6273
			5	0.0642	0.0000	0.5359
		250	1	0.4730	0.3336	0.6254
			2	0.4428	0.2679	0.6436
			3	0.3849	0.1523	0.6406
			4	0.2899	0.0000	0.6102
			5	0.0979	0.0000	0.4979

จากตารางที่4.26 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.27 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.25 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n) คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	50	1	0.4882	0.2628	0.7205
			2	0.4186	0.0916	0.7224
			3	0.3070	0.0000	0.6978
			4	0.1061	0.0000	0.6467
			5	0.0265	0.0000	0.4866
		100	1	0.4846	0.3125	0.6700
			2	0.4464	0.2421	0.6954
			3	0.3955	0.0000	0.6862
			4	0.2268	0.0000	0.6265
			5	0.0720	0.0000	0.5657
		150	1	0.4782	0.3179	0.6653
			2	0.4427	0.2438	0.6552
			3	0.4064	0.1513	0.6772
			4	0.3141	0.0000	0.6569
			5	0.1194	0.0000	0.6014

ตารางที่ 4.27 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.25	$0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$	200	1	0.4726	0.3281	0.6353
			2	0.4519	0.2821	0.6529
			3	0.4090	0.1922	0.6500
			4	0.3626	0.0000	0.6596
			5	0.1856	0.0000	0.5878
		250	1	0.4730	0.3336	0.6254
			2	0.4466	0.2980	0.6263
			3	0.4179	0.2069	0.6270
			4	0.3769	0.0000	0.6391
			5	0.2302	0.0000	0.6221

จากตารางที่4.27 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.28 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{\lvert r_{ij}\rvert\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	50	1	0.4848	0.2611	0.7578
			2	0.4334	0.1505	0.7024
			3	0.2220	0.0000	0.6752
			4	0.0515	0.0000	0.4970
			5	0.0075	0.0000	0.0539

ตารางที่ 4.28 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	100	1	0.4807	0.3000	0.6825
			2	0.4355	0.1941	0.6971
			3	0.3326	0.0000	0.6423
			4	0.1058	0.0000	0.5766
			5	0.0158	0.0000	0.2717
		150	1	0.4744	0.3164	0.6450
			2	0.4423	0.2468	0.6541
			3	0.3828	0.1051	0.6441
			4	0.1758	0.0000	0.5918
			5	0.0267	0.0000	0.4096
		200	1	0.4766	0.3332	0.6434
			2	0.4546	0.2675	0.6411
			3	0.3842	0.1469	0.6287
			4	0.2255	0.0000	0.6292
			5	0.0415	0.0000	0.4054
		250	1	0.4766	0.3389	0.6178
			2	0.4566	0.2799	0.6209
			3	0.3884	0.1457	0.6325
			4	0.2570	0.0000	0.6283
			5	0.0508	0.0000	0.4189

จากตารางที่4.28 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ $0 < Max\{\lvert r_{ij}\rvert\} < 0.3$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.29 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.3 < Max\{\|r_{ij}\|\} < 0.6$	50	1	0.4848	0.2611	0.7578
			2	0.4356	0.1141	0.7127
			3	0.2496	0.0000	0.6268
			4	0.0837	0.0000	0.5649
			5	0.0242	0.0000	0.4876
		100	1	0.4807	0.3000	0.6825
			2	0.4472	0.2250	0.6816
			3	0.3709	0.0000	0.6790
			4	0.1700	0.0000	0.6228
			5	0.0332	0.0000	0.4275
		150	1	0.4744	0.3164	0.6450
			2	0.4502	0.2572	0.6638
			3	0.3968	0.1464	0.6673
			4	0.2473	0.0000	0.6369
			5	0.0491	0.0000	0.5371

ตารางที่ 4.29 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.3 < Max\{\lvert r_{ij}\rvert\} < 0.6$	200	1	0.4766	0.3332	0.6434
			2	0.4528	0.2751	0.6386
			3	0.4016	0.1586	0.6756
			4	0.2819	0.0000	0.5835
			5	0.0735	0.0000	0.5260
		250	1	0.4766	0.3389	0.6178
			2	0.4508	0.2852	0.6208
			3	0.4086	0.1656	0.6619
			4	0.3188	0.0000	0.6258
			5	0.1148	0.0000	0.5400

จากตารางที่4.29 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.30 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.5 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$	50	1	0.4848	0.2611	0.7578
			2	0.4391	0.1590	0.7346
			3	0.3104	0.0000	0.7046
			4	0.1149	0.0000	0.6357
			5	0.0366	0.0000	0.5326

ตารางที่ 4.30 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.5	$0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$	100	1	0.4807	0.3000	0.6825
			2	0.4578	0.2470	0.7052
			3	0.4096	0.0971	0.6791
			4	0.2494	0.0000	0.6806
			5	0.0848	0.0000	0.5780
		150	1	0.4744	0.3164	0.6450
			2	0.4567	0.2836	0.6672
			3	0.4222	0.1852	0.6725
			4	0.3454	0.0000	0.6483
			5	0.1440	0.0000	0.5683
		200	1	0.4766	0.3332	0.6434
			2	0.4587	0.2921	0.6567
			3	0.4331	0.2167	0.6711
			4	0.3744	0.0000	0.6341
			5	0.2043	0.0000	0.5986
		250	1	0.4766	0.3389	0.6178
			2	0.4592	0.3043	0.6264
			3	0.4250	0.2376	0.6538
			4	0.3866	0.1291	0.6555
			5	0.2603	0.0000	0.6125

จากตารางที่4.30 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับสูง $0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.31 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0 < Max\{\|r_{ij}\|\} < 0.3$	50	1	0.4883	0.2799	0.7150
			2	0.4434	0.1472	0.7411
			3	0.2637	0.0000	0.6488
			4	0.0477	0.0000	0.5043
			5	0.0090	0.0000	0.0207
		100	1	0.4833	0.3211	0.6935
			2	0.4464	0.2160	0.6711
			3	0.3522	0.0000	0.6709
			4	0.1546	0.0000	0.5825
			5	0.0171	0.0000	0.3238
		150	1	0.4831	0.3365	0.6528
			2	0.4546	0.2717	0.6694
			3	0.4001	0.1057	0.6547
			4	0.2165	0.0000	0.6159
			5	0.0281	0.0000	0.3847

ตารางที่ 4.31 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0 < Max\{\lvert r_{ij}\rvert\} < 0.3$	200	1	0.4718	0.3304	0.6445
			2	0.4569	0.2774	0.6419
			3	0.3972	0.1181	0.6423
			4	0.2620	0.0000	0.6264
			5	0.0567	0.0000	0.4800
		250	1	0.4752	0.3470	0.6231
			2	0.4558	0.3161	0.6153
			3	0.4171	0.1702	0.6704
			4	0.2982	0.0000	0.6119
			5	0.0631	0.0000	0.4650

จากตารางที่4.31 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระในระดับต่ำ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.32 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตาม จำนวนตัวแปรอิสระ(p)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.3 < Max\{\lvert r_{ij}\rvert\} < 0.6$	50	1	0.4883	0.2799	0.7150
			2	0.4425	0.1807	0.7462
			3	0.3004	0.0000	0.6952
			4	0.1038	0.0000	0.6475
			5	0.0192	0.0000	0.3882

ตารางที่ 4.32 (ต่อ)

a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.3 < Max\{\lvert r_{ij}\rvert\} < 0.6$	100	1	0.4833	0.3211	0.6935
			2	0.4563	0.2390	0.7011
			3	0.3878	0.0000	0.6790
			4	0.1939	0.0000	0.6059
			5	0.0374	0.0000	0.4978
		150	1	0.4831	0.3365	0.6528
			2	0.4635	0.2803	0.6672
			3	0.4052	0.1700	0.6470
			4	0.2594	0.0000	0.6132
			5	0.0735	0.0000	0.5163
		200	1	0.4718	0.3304	0.6445
			2	0.4573	0.2911	0.6273
			3	0.4178	0.1960	0.6579
			4	0.3232	0.0000	0.6466
			5	0.1001	0.0000	0.5503
		250	1	0.4752	0.3470	0.6231
			2	0.4632	0.3139	0.6458
			3	0.4311	0.1893	0.6484
			4	0.3516	0.0000	0.6493
			5	0.1342	0.0000	0.5247

จากตารางที่4.32 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับกลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

ตารางที่ 4.33 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) เท่ากับ 0.75 ความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)อยู่ในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ ขนาดตัวอย่าง(n)คือ 50,100,150,200 และ 250 โดยจำแนกตามจำนวนตัวแปรอิสระ(p)

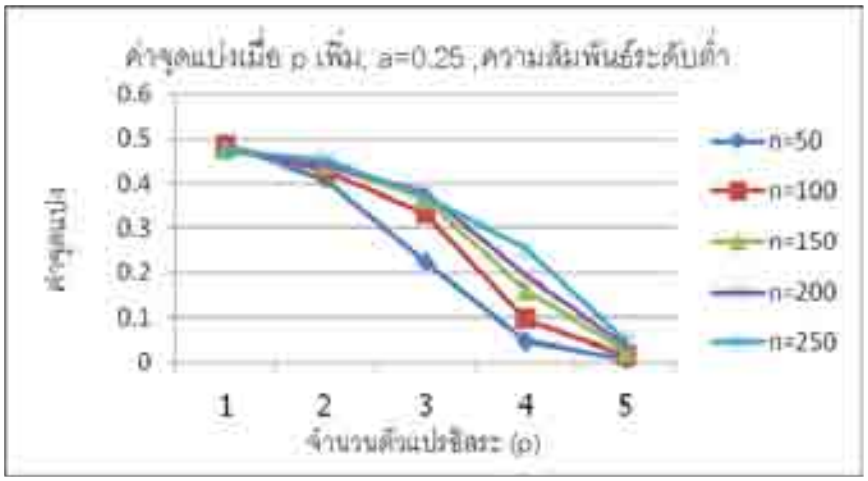
a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.6 < Max\{\|r_{ij}\|\} < 0.9$	50	1	0.4883	0.2799	0.7150
			2	0.4460	0.1555	0.7423
			3	0.3379	0.0000	0.7035
			4	0.1261	0.0000	0.6403
			5	0.0358	0.0000	0.5545
		100	1	0.4833	0.3211	0.6935
			2	0.4660	0.2688	0.6818
			3	0.4318	0.1307	0.6791
			4	0.2730	0.0000	0.6256
			5	0.0793	0.0000	0.5418
		150	1	0.4831	0.3365	0.6528
			2	0.4616	0.2996	0.6569
			3	0.4299	0.2087	0.6771
			4	0.3711	0.0000	0.6802
			5	0.1785	0.0000	0.6243

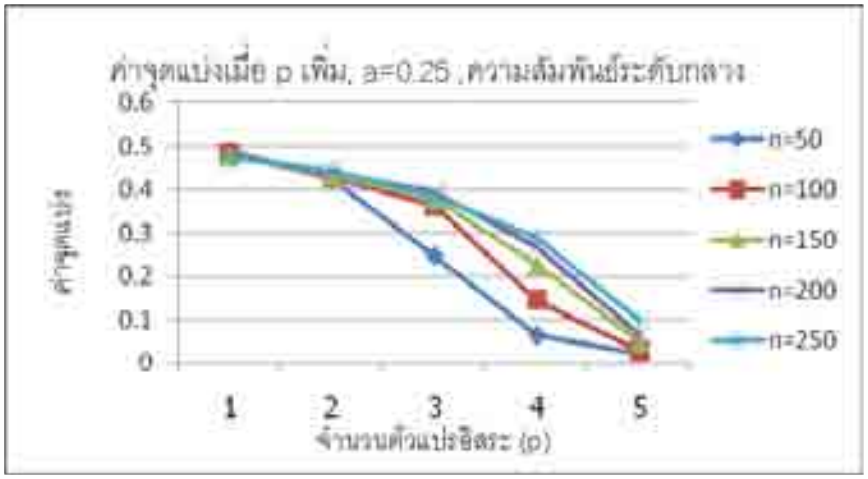
ตารางที่ 4.33 (ต่อ)

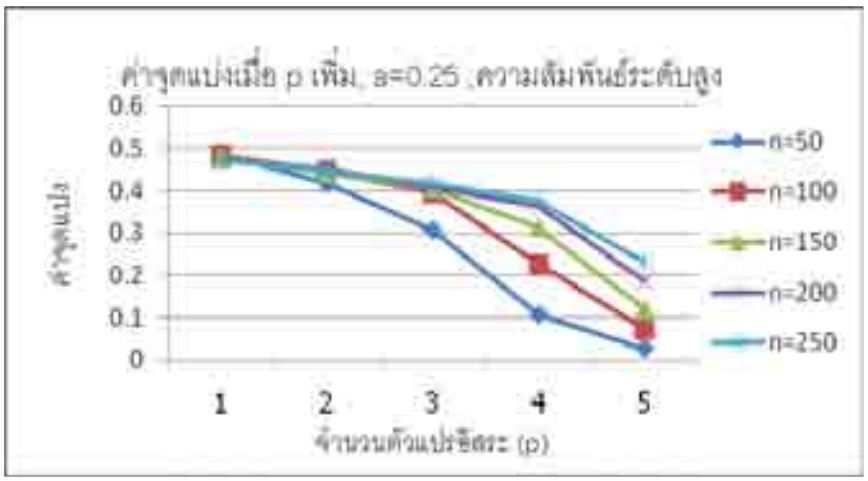
a	ρ	n	p	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
0.75	$0.6 < Max\{\lvert r_{ij}\rvert\} < 0.9$	200	1	0.4718	0.3304	0.6445
			2	0.4585	0.2924	0.6460
			3	0.4510	0.2386	0.6881
			4	0.3985	0.0427	0.6676
			5	0.2565	0.0000	0.6455
		250	1	0.4752	0.3470	0.6231
			2	0.4694	0.3114	0.6311
			3	0.4465	0.2475	0.6851
			4	0.4075	0.1522	0.6429
			5	0.3047	0.0000	0.6385

จากตารางที่4.33 เมื่อพิจารณาค่าจุดแบ่งที่มีอัตราส่วนของความล้มเหลวเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระในระดับสูง $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น พบว่า ที่ขนาดของตัวอย่าง คือ 50,100,150,200 และ 250 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าลดลง

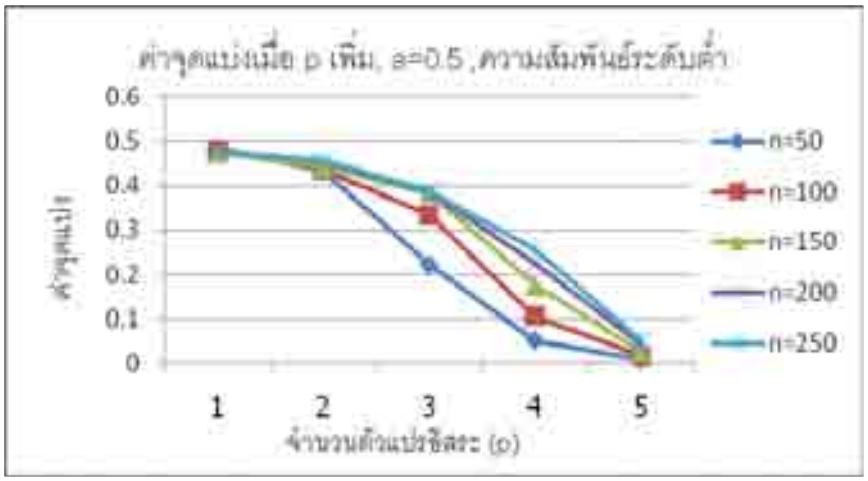
ภาพที่ 4.3 แสดงค่าเฉลี่ยของจุดแบ่งกรณีที่จำนวนตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง ระดับความสัมพันธ์ของตัวแปรอิสระ และอัตราส่วนความล้มเหลวของลักษณะที่สนใจคงที่

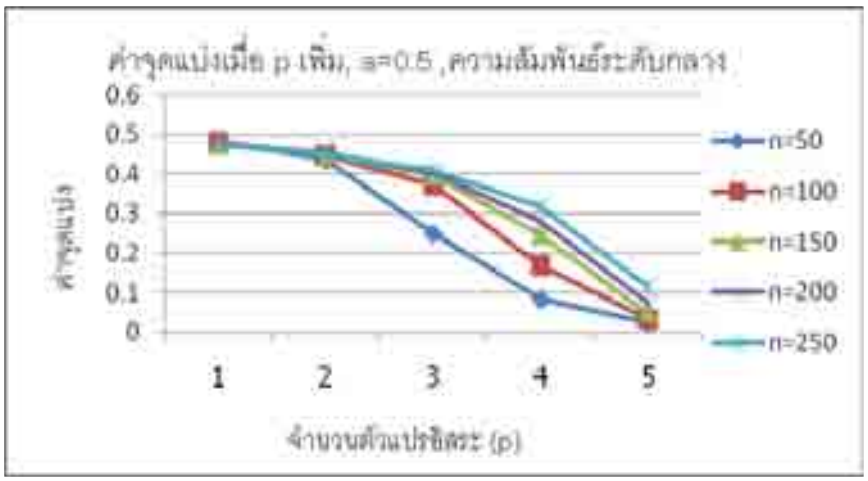


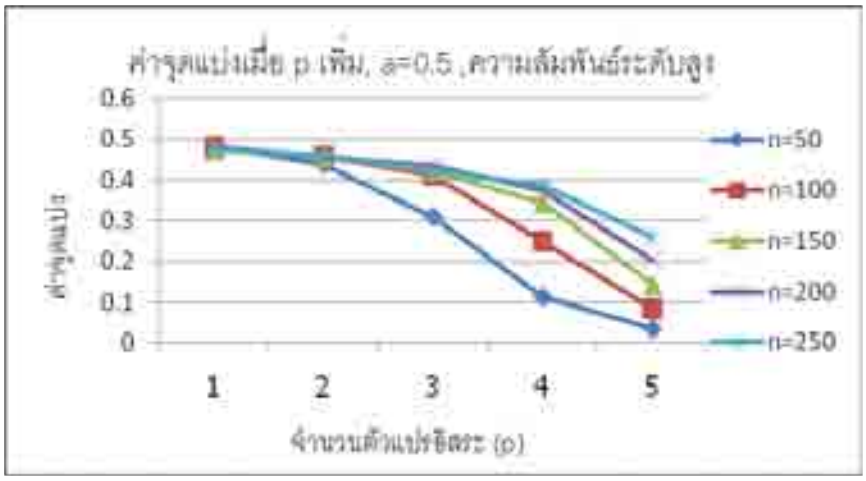




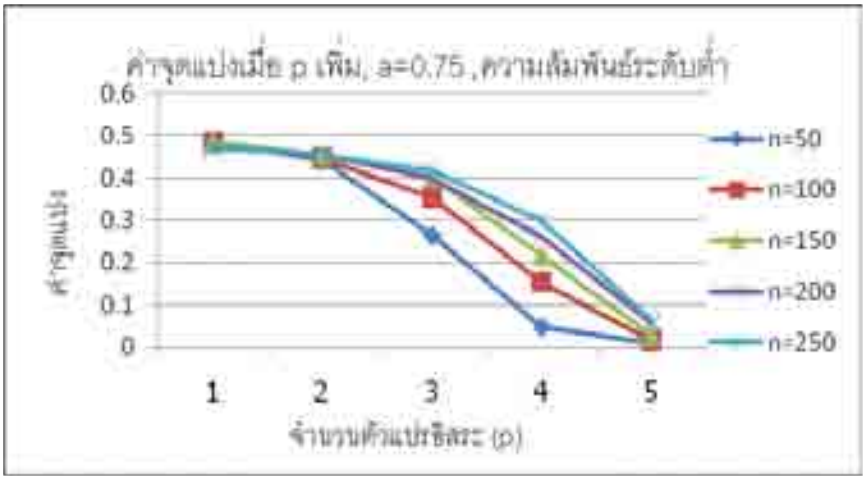
ภาพที่ 4.3 (ต่อ)

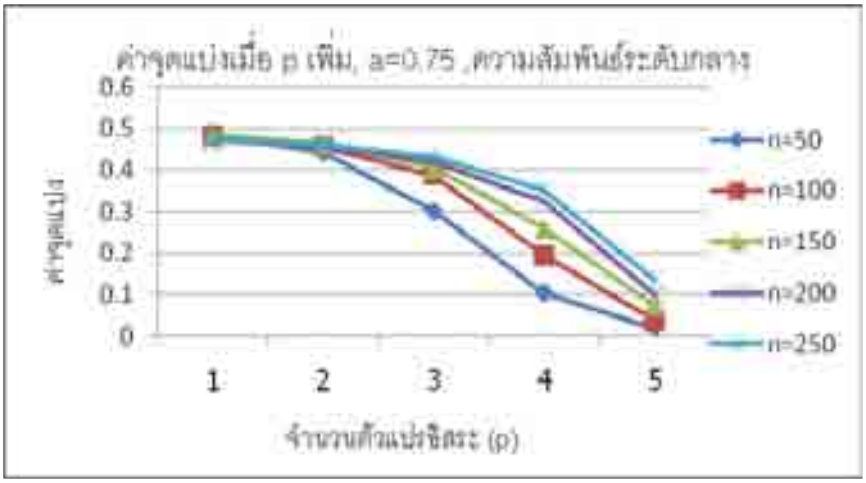

ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.5 ,ความสัมพันธ์ระดับต่ำ
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250

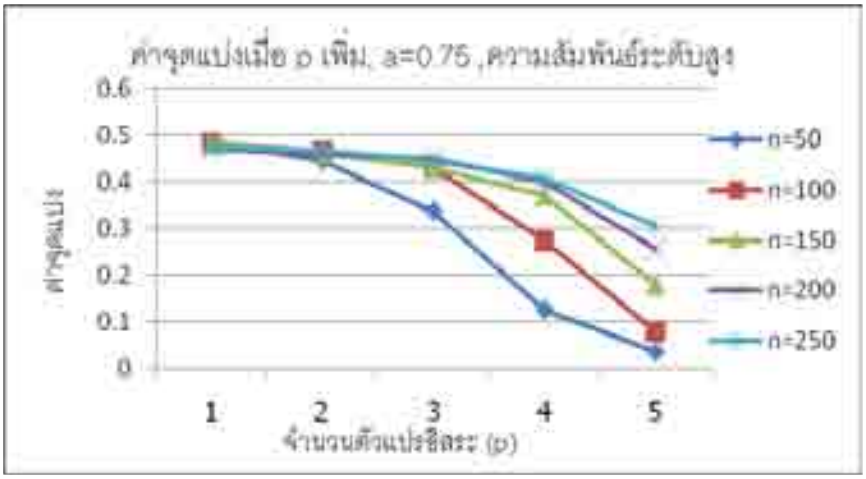

ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.5 ,ความสัมพันธ์ระดับกลาง
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.5 ,ความสัมพันธ์ระดับสูง
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250

ภาพที่ 4.3 (ต่อ)


ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.75 ,ความสัมพันธ์ระดับต่ำ
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.75 ,ความสัมพันธ์ระดับกลาง
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250


ค่าจุดแบ่งเมื่อ p เพิ่ม, a=0.75 ,ความสัมพันธ์ระดับสูง
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
จำนวนตัวแปรอิสระ (p)
n=50
n=100
n=150
n=200
n=250

4.1.4 กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระและอัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาคงที่

ซึ่งผลการวิจัยนำเสนอในตารางที่ 4.34-4.48

ตารางที่ 4.34 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 50 อัตราส่วนของความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	0.25	1	ต่ำ	0.4882	0.2628	0.7205
			กลาง	0.4882	0.2628	0.7205
			สูง	0.4882	0.2628	0.7205
		2	ต่ำ	0.4085	0.0853	0.7163
			กลาง	0.4255	0.1468	0.7091
			สูง	0.4186	0.0916	0.7224
		3	ต่ำ	0.2238	0.0000	0.6755
			กลาง	0.2478	0.0000	0.6551
			สูง	0.3070	0.0000	0.6978
		4	ต่ำ	0.0469	0.0000	0.5724
			กลาง	0.0641	0.0000	0.5237
			สูง	0.1061	0.0000	0.6467
		5	ต่ำ	0.0058	0.0000	0.0948
			กลาง	0.0208	0.0000	0.3755
			สูง	0.0265	0.0000	0.4866

จากตารางที่4.34 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 50 อัตราส่วนของความล้มเหลวเท่ากับ 0.25 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4882 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.35 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 50 อัตราส่วนของความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	0.5	1	ต่ำ	0.4848	0.2611	0.7578
			กลาง	0.4848	0.2611	0.7578
			สูง	0.4848	0.2611	0.7578
		2	ต่ำ	0.4334	0.1505	0.7024
			กลาง	0.4356	0.1141	0.7127
			สูง	0.4391	0.1590	0.7346
		3	ต่ำ	0.2220	0.0000	0.6752
			กลาง	0.2496	0.0000	0.6268
			สูง	0.3104	0.0000	0.7046
		4	ต่ำ	0.0515	0.0000	0.4970
			กลาง	0.0837	0.0000	0.5649
			สูง	0.1149	0.0000	0.6357
		5	ต่ำ	0.0075	0.0000	0.0539
			กลาง	0.0242	0.0000	0.4876
			สูง	0.0366	0.0000	0.5326

จากตารางที่4.35 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 50 อัตราส่วนของความล้มเหลวเท่ากับ 0.5 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4848 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.36 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 50 อัตราส่วนของความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
50	0.75	1	ต่ำ	0.4883	0.2799	0.7150
			กลาง	0.4883	0.2799	0.7150
			สูง	0.4883	0.2799	0.7150
		2	ต่ำ	0.4434	0.1472	0.7411
			กลาง	0.4425	0.1807	0.7462
			สูง	0.4460	0.1555	0.7423
		3	ต่ำ	0.2637	0.0000	0.6488
			กลาง	0.3004	0.0000	0.6952
			สูง	0.3379	0.0000	0.7035
		4	ต่ำ	0.0477	0.0000	0.5043
			กลาง	0.1038	0.0000	0.6475
			สูง	0.1261	0.0000	0.6403
		5	ต่ำ	0.0090	0.0000	0.0207
			กลาง	0.0192	0.0000	0.3882
			สูง	0.0358	0.0000	0.5545

จากตารางที่4.36 เมื่อพิจารณาจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 50 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4883 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.37 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 100 อัตราส่วนของความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	0.25	1	ต่ำ	0.4846	0.3125	0.6700
			กลาง	0.4846	0.3125	0.6700
			สูง	0.4846	0.3125	0.6700
		2	ต่ำ	0.4281	0.1575	0.6860
			กลาง	0.4269	0.1931	0.6639
			สูง	0.4464	0.2421	0.6954
		3	ต่ำ	0.3334	0.0000	0.6609
			กลาง	0.3626	0.0000	0.6629
			สูง	0.3955	0.0000	0.6862
		4	ต่ำ	0.0973	0.0000	0.5606
			กลาง	0.1465	0.0000	0.6015
			สูง	0.2268	0.0000	0.6265
		5	ต่ำ	0.0157	0.0000	0.3200
			กลาง	0.0277	0.0000	0.3656
			สูง	0.0720	0.0000	0.5657

จากตารางที่4.37 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 100 อัตราส่วนของความล้มเหลวเท่ากับ 0.25 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4846 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.38 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 100 อัตราส่วนของความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	0.5	1	ต่ำ	0.4807	0.3000	0.6825
			กลาง	0.4807	0.3000	0.6825
			สูง	0.4807	0.3000	0.6825
		2	ต่ำ	0.4355	0.1941	0.6971
			กลาง	0.4472	0.2250	0.6816
			สูง	0.4578	0.2470	0.7052
		3	ต่ำ	0.3326	0.0000	0.6423
			กลาง	0.3709	0.0000	0.6790
			สูง	0.4096	0.0971	0.6791
		4	ต่ำ	0.1058	0.0000	0.5766
			กลาง	0.1700	0.0000	0.6228
			สูง	0.2494	0.0000	0.6806
		5	ต่ำ	0.0158	0.0000	0.2717
			กลาง	0.0332	0.0000	0.4275
			สูง	0.0848	0.0000	0.5780

จากตารางที่4.38 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 100 อัตราส่วนของความล้มเหลวเท่ากับ 0.5 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4807 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.39 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 100 อัตราส่วนของความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
100	0.75	1	ต่ำ	0.4833	0.3211	0.6935
			กลาง	0.4833	0.3211	0.6935
			สูง	0.4833	0.3211	0.6935
		2	ต่ำ	0.4464	0.2160	0.6711
			กลาง	0.4563	0.2390	0.7011
			สูง	0.4660	0.2688	0.6818
		3	ต่ำ	0.3522	0.0000	0.6709
			กลาง	0.3878	0.0000	0.6790
			สูง	0.4318	0.1307	0.6791
		4	ต่ำ	0.1546	0.0000	0.5825
			กลาง	0.1939	0.0000	0.6059
			สูง	0.2730	0.0000	0.6256
		5	ต่ำ	0.0171	0.0000	0.3238
			กลาง	0.0374	0.0000	0.4978
			สูง	0.0793	0.0000	0.5418

จากตารางที่4.39 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 100 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4833 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.40 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อขนาดตัวอย่าง (n) เท่ากับ 150 สัดส่วนของความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	0.25	1	ต่ำ	0.4782	0.3179	0.6653
			กลาง	0.4782	0.3179	0.6653
			สูง	0.4782	0.3179	0.6653
		2	ต่ำ	0.4450	0.2241	0.6509
			กลาง	0.4327	0.2377	0.6439
			สูง	0.4427	0.2438	0.6552
		3	ต่ำ	0.3656	0.0697	0.6509
			กลาง	0.3794	0.0935	0.6534
			สูง	0.4064	0.1513	0.6772
		4	ต่ำ	0.1625	0.0000	0.5854
			กลาง	0.2277	0.0000	0.6299
			สูง	0.3141	0.0000	0.6569
		5	ต่ำ	0.0231	0.0000	0.3867
			กลาง	0.0501	0.0000	0.4599
			สูง	0.1194	0.0000	0.6014

จากตารางที่4.40 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 150 อัตราส่วนของความล้มเหลวเท่ากับ 0.25 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4782 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.41 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 150 อัตราส่วนของความล้มเหลว (a) เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	0.5	1	ต่ำ	0.4744	0.3164	0.6450
			กลาง	0.4744	0.3164	0.6450
			สูง	0.4744	0.3164	0.6450
		2	ต่ำ	0.4423	0.2468	0.6541
			กลาง	0.4502	0.2572	0.6638
			สูง	0.4567	0.2836	0.6672
		3	ต่ำ	0.3828	0.1051	0.6441
			กลาง	0.3968	0.1464	0.6673
			สูง	0.4222	0.1852	0.6725
		4	ต่ำ	0.1758	0.0000	0.5918
			กลาง	0.2473	0.0000	0.6369
			สูง	0.3454	0.0000	0.6483
		5	ต่ำ	0.0267	0.0000	0.4096
			กลาง	0.0491	0.0000	0.5371
			สูง	0.1440	0.0000	0.5683

จากตารางที่4.41 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 150 อัตราส่วนของความล้มเหลวเท่ากับ 0.5 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4744 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.42 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 150 สัดส่วนของความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
150	0.75	1	ต่ำ	0.4831	0.3365	0.6528
			กลาง	0.4831	0.3365	0.6528
			สูง	0.4831	0.3365	0.6528
		2	ต่ำ	0.4546	0.2717	0.6694
			กลาง	0.4635	0.2803	0.6672
			สูง	0.4616	0.2996	0.6569
		3	ต่ำ	0.4001	0.1057	0.6547
			กลาง	0.4052	0.1700	0.6470
			สูง	0.4299	0.2087	0.6771
		4	ต่ำ	0.2165	0.0000	0.6159
			กลาง	0.2594	0.0000	0.6132
			สูง	0.3711	0.0000	0.6802
		5	ต่ำ	0.0281	0.0000	0.3847
			กลาง	0.0735	0.0000	0.5163
			สูง	0.1785	0.0000	0.6243

จากตารางที่4.42 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 150 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4831 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.43 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 200 อัตราส่วนของความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	0.25	1	ต่ำ	0.4726	0.3281	0.6353
			กลาง	0.4726	0.3281	0.6353
			สูง	0.4726	0.3281	0.6353
		2	ต่ำ	0.4416	0.2586	0.6239
			กลาง	0.4407	0.2579	0.6295
			สูง	0.4519	0.2821	0.6529
		3	ต่ำ	0.3811	0.0905	0.6452
			กลาง	0.3937	0.1468	0.6349
			สูง	0.4090	0.1922	0.6500
		4	ต่ำ	0.1940	0.0000	0.5723
			กลาง	0.2691	0.0000	0.6273
			สูง	0.3626	0.0000	0.6596
		5	ต่ำ	0.0367	0.0000	0.4338
			กลาง	0.0642	0.0000	0.5359
			สูง	0.1856	0.0000	0.5878

จากตารางที่4.43 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 200 อัตราส่วนของความล้มเหลวเท่ากับ 0.25 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4726 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.44 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 200 อัตราส่วนของความล้มเหลวเท่ากับ 0.5 ตัวแปรอิสระ(p)เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	0.5	1	ต่ำ	0.4766	0.3332	0.6434
			กลาง	0.4766	0.3332	0.6434
			สูง	0.4766	0.3332	0.6434
		2	ต่ำ	0.4546	0.2675	0.6411
			กลาง	0.4528	0.2751	0.6386
			สูง	0.4587	0.2921	0.6567
		3	ต่ำ	0.3842	0.1469	0.6287
			กลาง	0.4016	0.1586	0.6756
			สูง	0.4331	0.2167	0.6711
		4	ต่ำ	0.2255	0.0000	0.6292
			กลาง	0.2819	0.0000	0.5835
			สูง	0.3744	0.0000	0.6341
		5	ต่ำ	0.0415	0.0000	0.4054
			กลาง	0.0735	0.0000	0.5260
			สูง	0.2043	0.0000	0.5986

จากตารางที่ 4.44 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 200 สัดส่วนของความล้มเหลวเท่ากับ 0.5 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4766 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.45 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 200 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 ตัวแปรอิสระ(p)เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
200	0.75	1	ต่ำ	0.4718	0.3304	0.6445
			กลาง	0.4718	0.3304	0.6445
			สูง	0.4718	0.3304	0.6445
		2	ต่ำ	0.4569	0.2774	0.6419
			กลาง	0.4573	0.2911	0.6273
			สูง	0.4585	0.2924	0.6460
		3	ต่ำ	0.3972	0.1181	0.6423
			กลาง	0.4178	0.1960	0.6579
			สูง	0.4510	0.2386	0.6881
		4	ต่ำ	0.2620	0.0000	0.6264
			กลาง	0.3232	0.0000	0.6466
			สูง	0.3985	0.0427	0.6676
		5	ต่ำ	0.0567	0.0000	0.4800
			กลาง	0.1001	0.0000	0.5503
			สูง	0.2565	0.0000	0.6455

จากตารางที่4.45 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 200 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4718 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.46 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 250 อัตราส่วนของความล้มเหลว เท่ากับ 0.25 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	0.25	1	ต่ำ	0.4730	0.3336	0.6254
			กลาง	0.4730	0.3336	0.6254
			สูง	0.4730	0.3336	0.6254
		2	ต่ำ	0.4523	0.2875	0.6420
			กลาง	0.4428	0.2679	0.6436
			สูง	0.4466	0.2980	0.6263
		3	ต่ำ	0.3706	0.1130	0.6354
			กลาง	0.3849	0.1523	0.6406
			สูง	0.4179	0.2069	0.6270
		4	ต่ำ	0.2537	0.0000	0.6000
			กลาง	0.2899	0.0000	0.6102
			สูง	0.3769	0.0000	0.6391
		5	ต่ำ	0.0489	0.0000	0.4587
			กลาง	0.0979	0.0000	0.4979
			สูง	0.2302	0.0000	0.6221

จากตารางที่4.46 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 250 อัตราส่วนของความล้มเหลวเท่ากับ 0.25 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4730 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

ตารางที่ 4.47 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 250 อัตราส่วนของความล้มเหลว เท่ากับ 0.5 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	0.5	1	ต่ำ	0.4766	0.3389	0.6178
			กลาง	0.4766	0.3389	0.6178
			สูง	0.4766	0.3389	0.6178
		2	ต่ำ	0.4566	0.2799	0.6209
			กลาง	0.4508	0.2852	0.6208
			สูง	0.4592	0.3043	0.6264
		3	ต่ำ	0.3884	0.1457	0.6325
			กลาง	0.4086	0.1656	0.6619
			สูง	0.4250	0.2376	0.6538
		4	ต่ำ	0.2570	0.0000	0.6283
			กลาง	0.3188	0.0000	0.6258
			สูง	0.3866	0.1291	0.6555
		5	ต่ำ	0.0508	0.0000	0.4189
			กลาง	0.1148	0.0000	0.5400
			สูง	0.2603	0.0000	0.6125

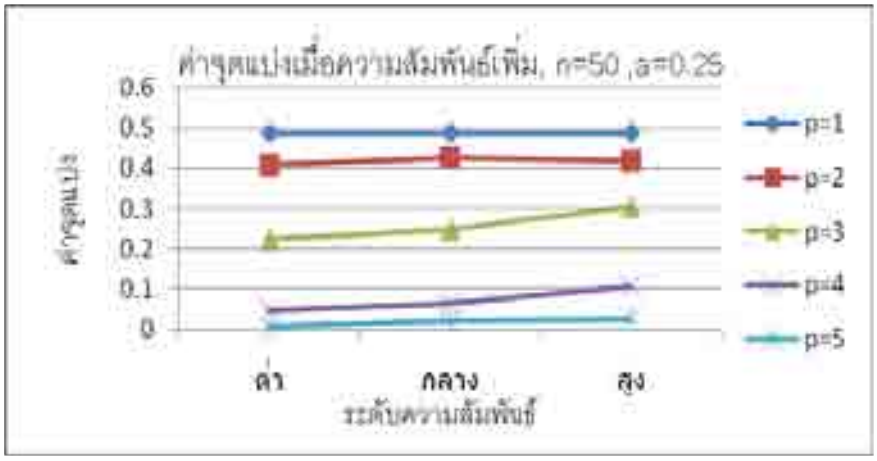
จากตารางที่4.47 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 250 อัตราส่วนของความล้มเหลวเท่ากับ 0.5 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4766 ที่ตัวแปรอิสระเท่ากับ 2 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับกัน ที่ตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

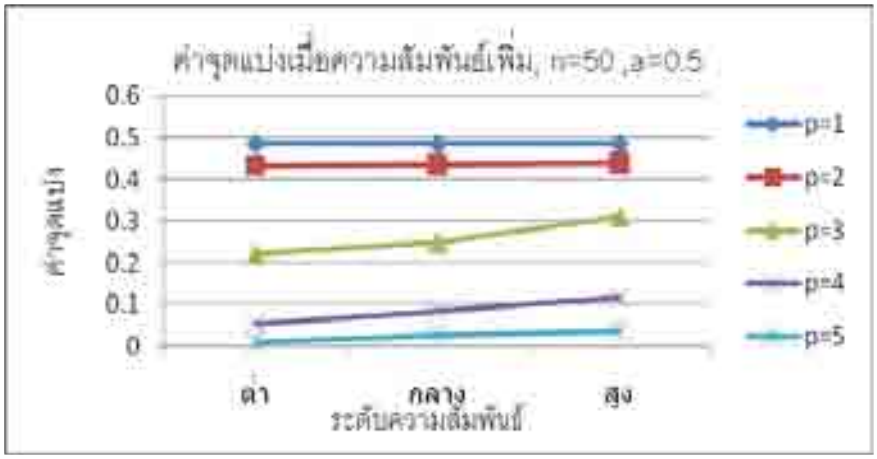
ตารางที่ 4.48 แสดงค่าเฉลี่ยของจุดแบ่งและช่วงความเชื่อมั่นของจุดแบ่งที่เหมาะสม เมื่อ ขนาดตัวอย่าง (n) เท่ากับ 250 อัตราส่วนของความล้มเหลว เท่ากับ 0.75 ตัวแปรอิสระ(p) เท่ากับ 1,2,3,4 และ 5 ตัว โดยจำแนกตามความสัมพันธ์ระหว่างตัวแปรอิสระ(ρ)

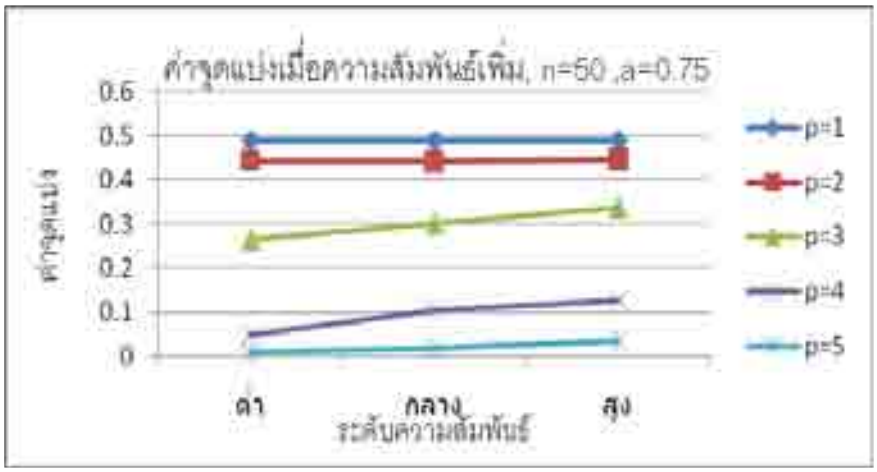
n	a	p	ρ	$\bar{\hat{c}}$	Cl. Lower of $\hat{c}$	Cl. Upper of $\hat{c}$
250	0.75	1	ต่ำ	0.4752	0.3470	0.6231
			กลาง	0.4752	0.3470	0.6231
			สูง	0.4752	0.3470	0.6231
		2	ต่ำ	0.4558	0.3161	0.6153
			กลาง	0.4632	0.3139	0.6458
			สูง	0.4694	0.3114	0.6311
		3	ต่ำ	0.4171	0.1702	0.6704
			กลาง	0.4311	0.1893	0.6484
			สูง	0.4465	0.2475	0.6851
		4	ต่ำ	0.2982	0.0000	0.6119
			กลาง	0.3516	0.0000	0.6493
			สูง	0.4075	0.1522	0.6429
		5	ต่ำ	0.0631	0.0000	0.4650
			กลาง	0.1342	0.0000	0.5247
			สูง	0.3047	0.0000	0.6385

จากตารางที่4.48 เมื่อพิจารณาค่าจุดแบ่งที่มีขนาดตัวอย่างเท่ากับ 250 อัตราส่วนของความล้มเหลวเท่ากับ 0.75 โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง พบว่า ที่ตัวแปรอิสระเท่ากับ 1 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่ คือ 0.4882 ที่ตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

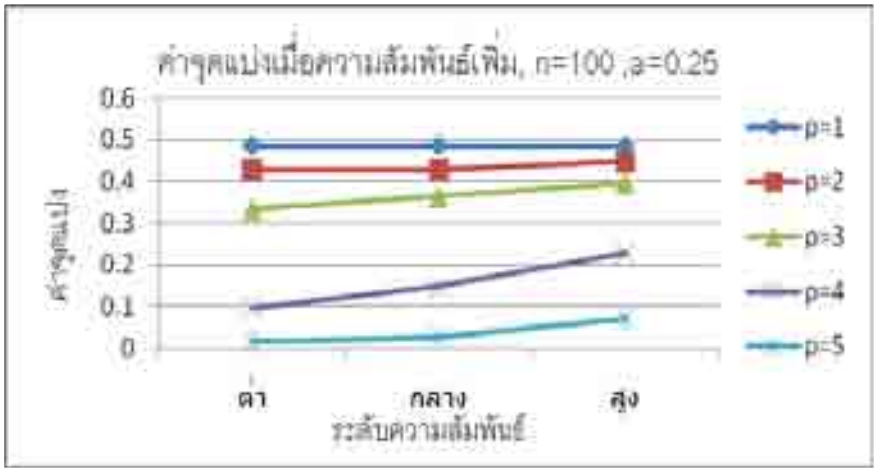
ภาพที่ 4.4 แสดงค่าเฉลี่ยของจุดแบ่ง กรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลง เมื่อขนาดตัวอย่าง จำนวนตัวแปรอิสระ และอัตราส่วนของความล้มเหลว ของลักษณะที่สนใจศึกษาคงที่

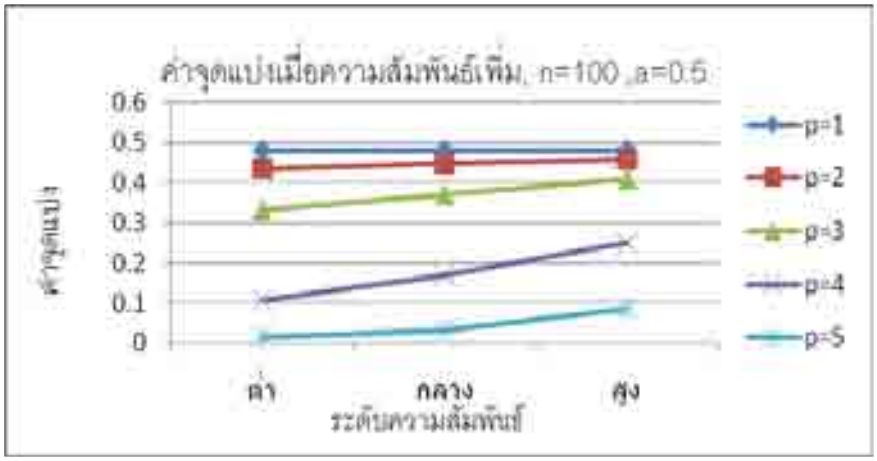


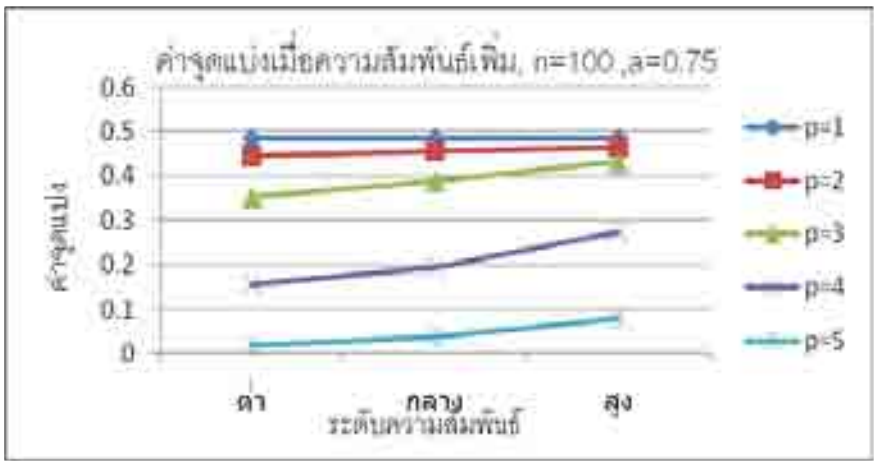




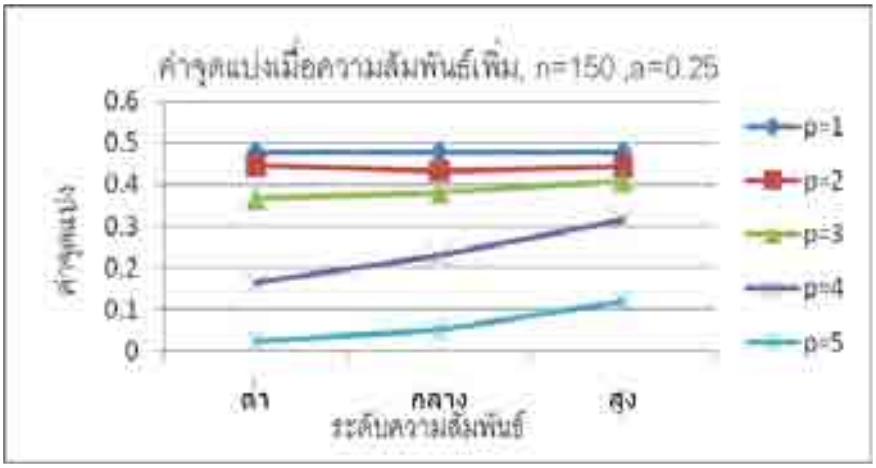
ภาพที่ 4.4 (ต่อ)

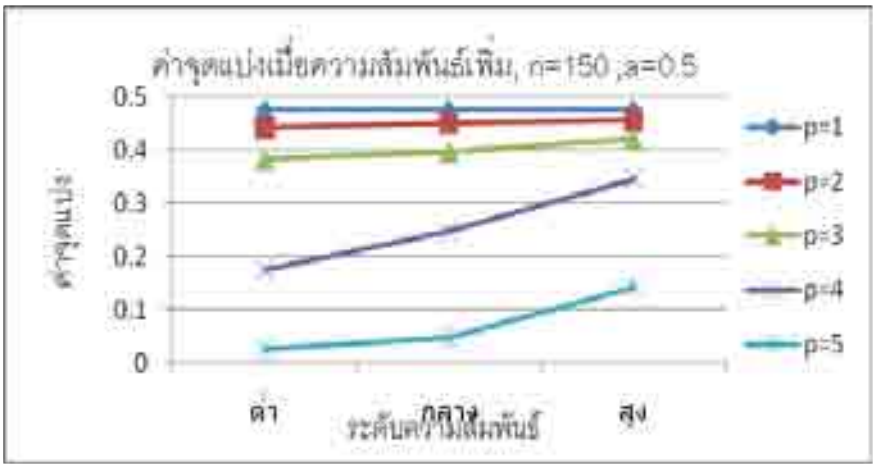

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=100 ,a=0.25
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

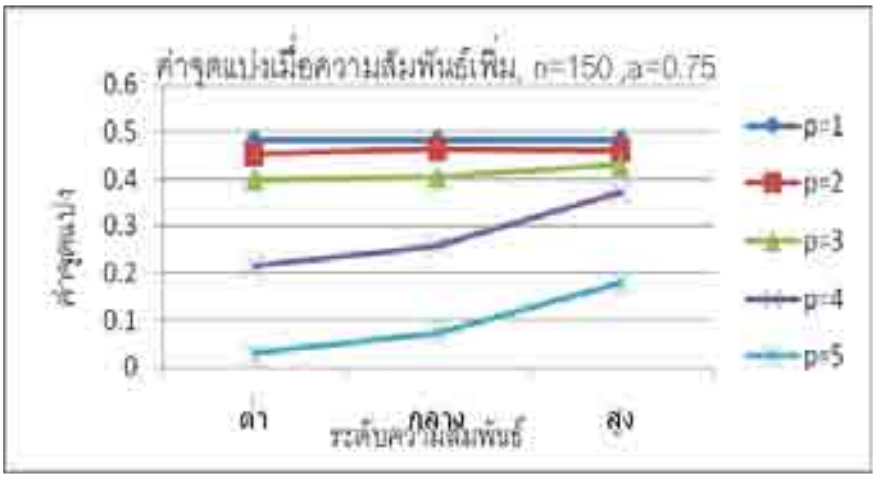

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=100 ,a=0.5
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=100 ,a=0.75
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

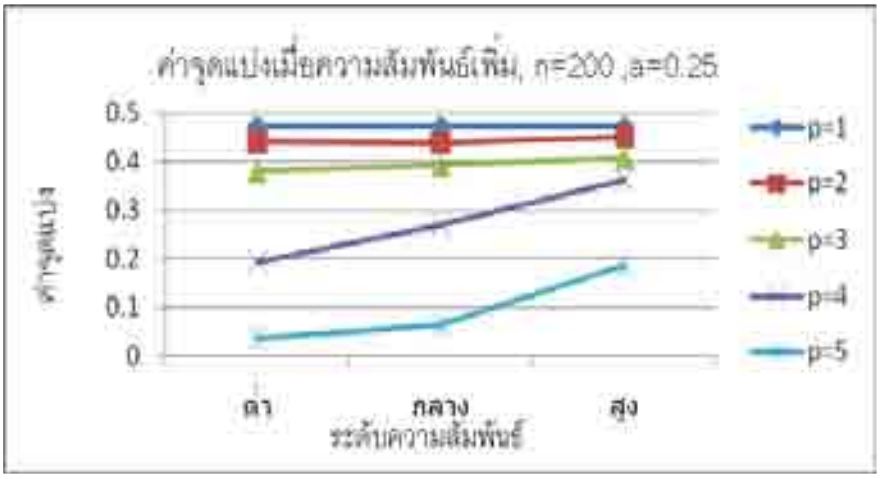
ภาพที่ 4.4 (ต่อ)

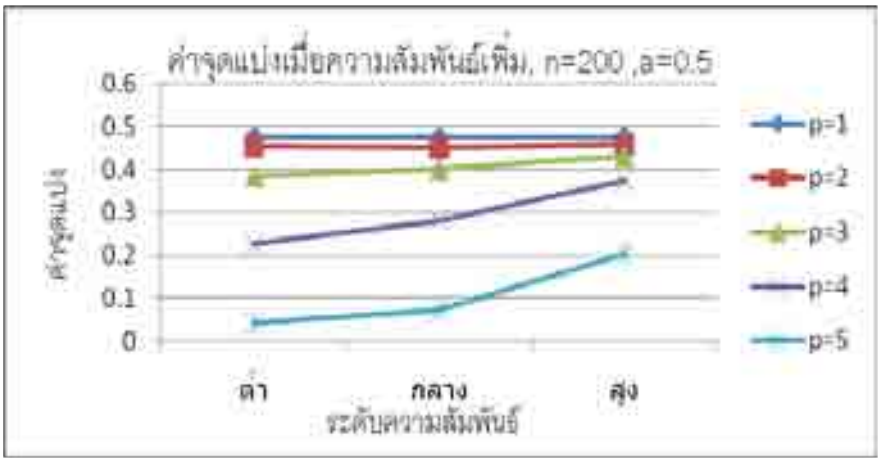

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=150 ,a=0.25
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

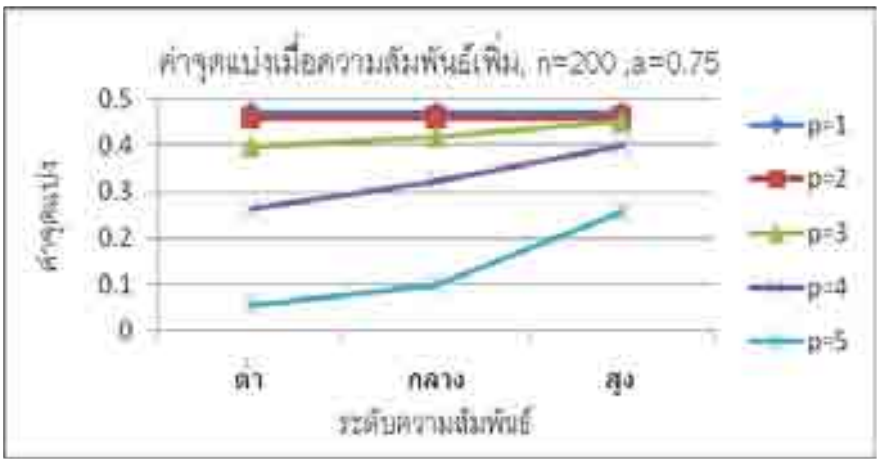

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=150 ,a=0.5
ค่าจุดแบ่ง
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=150 ,a=0.75
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

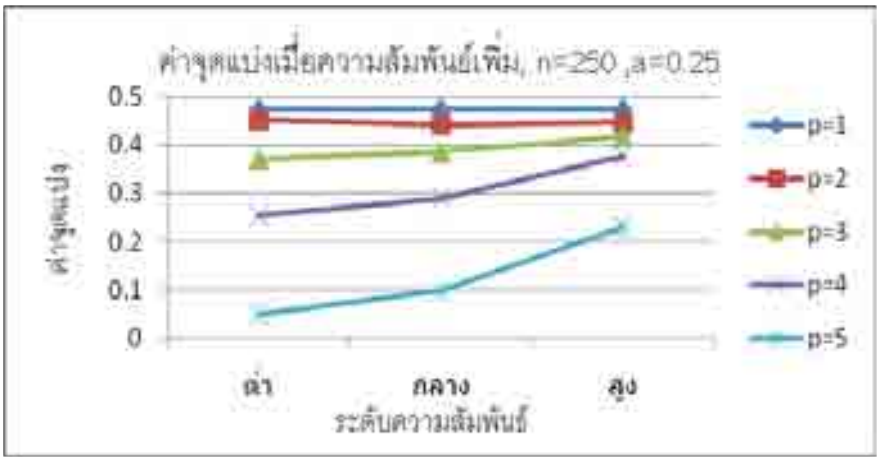
ภาพที่ 4.4 (ต่อ)

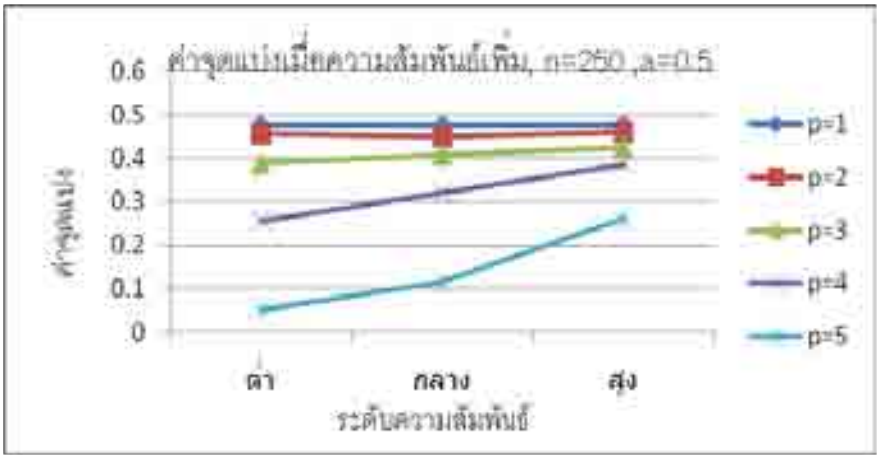

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=200 ,a=0.25
ค่าจุดแบ่ง
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

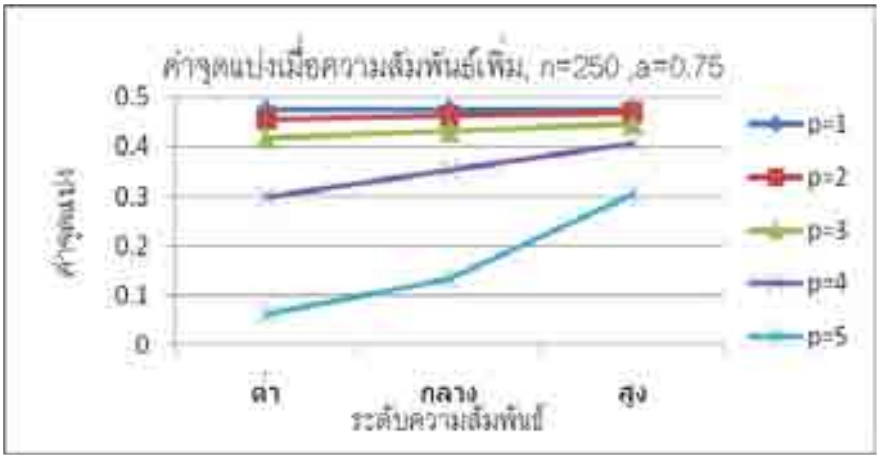

ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=200 ,a=0.5
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=200 ,a=0.75
ค่าจุดแบ่ง
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

ภาพที่ 4.4 (ต่อ)


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=250 ,a=0.25
ค่าจุดแบ่ง
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=250 ,a=0.5
ค่าจุดแบ่ง
0.6
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์


ค่าจุดแบ่งเมื่อความสัมพันธ์เพิ่ม, n=250 ,a=0.75
ค่าจุดแบ่ง
0.5
0.4
0.3
0.2
0.1
0
p=1
p=2
p=3
p=4
p=5
ต่ำ
กลาง
สูง
ระดับความสัมพันธ์

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้ต้องการหาจุดแบ่งที่เหมาะสมที่สุดสำหรับการพยากรณ์การจำแนกข้อมูลไม่จัดกลุ่มในตัวแบบการถดถอยโลจิสติกแบบ 2 ประเภทสำหรับแต่ละสถานการณ์ที่ต้องการศึกษาโดยจุดแบ่งที่เหมาะสมที่สุดจะให้อัตราความผิดพลาดในการจำแนกกลุ่มมีค่าต่ำสุด โดยจะทำการเปรียบเทียบที่ขนาดของตัวอย่างมีค่าเท่ากับ 50,100,150,200 และ 250 จำนวนของตัวแปรอิสระเป็น 1,2,3,4 และ 5 อัตราส่วนของความล้มเหลวของลักษณะที่สนใจ คือ 0.25,0.5 และ 0.75 และแบ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น 3 ระดับ คือ ความสัมพันธ์กันในระดับ ต่ำ $(0 < Max\{|r_{ij}|\} < 0.3)$, กลาง $0.3 < Max\{|r_{ij}|\} < 0.6$ และ สูง $(0.6 < Max\{|r_{ij}|\} < 0.9)$ โดยทำการจำลองข้อมูลและกระทำซ้ำ 500 รอบ โดยเทคนิคมอนติคาร์โล และเขียนด้วยโปรแกรม R

5.1 สรุปผลการวิจัย

ผลการวิจัยเรื่องการหาจุดแบ่งสำหรับการพยากรณ์การจำแนกข้อมูลสำหรับตัวแบบถดถอยโลจิสติก 2 ประเภท โดยใช้ฟังก์ชันคอมพลีเมนทารีล็อก ล็อก เป็นฟังกชันเชื่อมโยง ให้ผลของค่าเฉลี่ยของจุดแบ่งในแต่ละสถานการณ์ดังต่อไปนี้

5.1.1 กรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเพิ่มขึ้น

5.1.2 กรณีที่ขนาดของตัวอย่างเพิ่มขึ้น

5.1.3 กรณีที่จำนวนของตัวแปรอิสระเพิ่มขึ้น

5.1.4 กรณีที่ระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนแปลง

ซึ่งทั้ง 4 ปัจจัยนี้ให้ค่าเฉลี่ยของจุดแบ่งมีค่าไม่เกิน 0.5 ซึ่งแต่ละปัจจัยมีผลต่อค่าเฉลี่ยของจุดแบ่งแตกต่างกันไปดังนี้

5.1.1 กรณีที่อัตราส่วนของความล้มเหลวของลักษณะที่สนใจศึกษาเพิ่มขึ้น

ในการวิจัยนี้กำหนดอัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษา(a)คือ 0.25 ,0.5 และ 0.75โดยที่ขนาดตัวอย่าง(n)จำนวนตัวแปรอิสระ(p)และระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)มีค่าคงที่

ผลการวิจัยพบว่า ที่ขนาดตัวอย่างเท่ากับ 50 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0 < Max\{|r_{ij}|\} < 0.3$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 1,3 และ 4 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 2 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ขนาดตัวอย่างเท่าเดิม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.3 < Max\{|r_{ij}|\} < 0.6$ และ $0.6 < Max\{|r_{ij}|\} < 0.9$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3 และ 4 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ที่จำนวนตัวแปรอิสระเท่ากับ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา

ที่ขนาดตัวอย่างเท่ากับ 100 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0 < Max\{|r_{ij}|\} < 0.3$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 1และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 2,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ขนาดตัวอย่างเท่าเดิม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.3 < Max\{|r_{ij}|\} < 0.6$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.6 < Max\{|r_{ij}|\} < 0.9$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3 และ 4 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ที่จำนวนตัวแปรอิสระเท่ากับ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา

ที่ขนาดตัวอย่างเท่ากับ 150 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0 < Max\{|r_{ij}|\} < 0.3$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 1และ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ขนาดตัวอย่างเท่าเดิม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.3 < Max\{|r_{ij}|\} < 0.6$ และ $0.6 < Max\{|r_{ij}|\} < 0.9$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

ที่ขนาดตัวอย่างเท่ากับ 200 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0 < Max\{|r_{ij}|\} < 0.3$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ขนาดตัวอย่างเท่าเดิม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ

$0.3 < Max\{|r_{ij}|\} < 0.6$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.6 < Max\{|r_{ij}|\} < 0.9$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

และสุดท้ายที่ขนาดตัวอย่างเท่ากับ 250 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0 < Max\{|r_{ij}|\} < 0.3$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 1 และ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น ที่ขนาดตัวอย่างเท่าเดิม เมื่อระดับความสัมพันธ์ของตัวแปรอิสระคือ $0.3 < Max\{|r_{ij}|\} < 0.6$ และ $0.6 < Max\{|r_{ij}|\} < 0.9$ อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเปลี่ยนไป ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

5.1.2 กรณีขนาดตัวอย่างเพิ่มขึ้น

ในการวิจัยครั้งนี้กำหนดให้ขนาดตัวอย่างมีค่าเท่ากับ 50,100,150,200 และ 250 เมื่ออัตราส่วนความล้มเหลวของลักษณะสนใจศึกษา(a) ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ) และจำนวนตัวแปรอิสระ(p) มีค่าคงที่ ผลการวิจัยพบว่า ค่าเฉลี่ยของจุดแบ่งในทุกสถานการณ์มีค่าไม่เกินค่า 0.5

เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 1,2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ที่อัตราส่วนความล้มเหลวเดียวกัน ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 2,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ที่อัตราส่วนความล้มเหลวเดียวกัน ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.3 <$

$Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 1,2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ที่อัตราส่วนความล้มเหลวเดียวกัน ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.3 < Max\{|r_{ij}|\} < 0.6$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 2 และ 4 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 3 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อขนาดตัวอย่างเพิ่มขึ้น จำนวนตัวแปรอิสระเท่ากับ 2 และ 3 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา จำนวนตัวแปรอิสระเท่ากับ 4 และ 5 ตัว ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

5.1.3 กรณีจำนวนของตัวแปรอิสระเพิ่มขึ้น

ในการวิจัยครั้งนี้กำหนดให้จำนวนตัวแปรอิสระ (p) มีค่าเท่ากับ 1,2,3,4 และ 5 ตัว เมื่อสัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา (a) ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ)และขนาดตัวอย่าง (n) มีค่าคงที่ผลของการวิจัยพบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าไม่เกิน 0.5 ในทุกสถานการณ์

เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$, $0.3 < Max\{|r_{ij}|\} < 0.6$ และ$0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ที่ขนาดตัวอย่างเท่ากับ 50,100,150,200 และ 250 ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น เมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$, $0.3 < Max\{|r_{ij}|\} < 0.6$ และ $0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ที่ขนาดตัวอย่างเท่ากับ 50,100,150,200 และ 250 ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น และสุดท้ายเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 ระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ $0 < Max\{|r_{ij}|\} < 0.3$, $0.3 <$

$Max\{|r_{ij}|\} < 0.6$ และ$0.6 < Max\{|r_{ij}|\} < 0.9$ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ที่ขนาดตัวอย่างเท่ากับ 50,100,150,200 และ 250 ให้ค่าเฉลี่ยของจุดแบ่งเพิ่มสูงขึ้น

5.1.4 กรณีระดับความสัมพันธ์ของตัวแปรอิสระเพิ่มขึ้น

ในการวิจัยครั้งนี้กำหนดให้ ระดับความสัมพันธ์ของตัวแปรอิสระ(ρ) มี 3 ระดับคือ ระดับต่ำคือ $(0 < Max\{|r_{ij}|\} < 0.3)$ ระดับกลางคือ $(0.3 < Max\{|r_{ij}|\} < 0.6)$และระดับสูงคือ $(0.6 < Max\{|r_{ij}|\} < 0.9)$ ขนาดตัวอย่าง(n) สัดส่วนของความล้มเหลวของลักษณะที่สนใจศึกษา(a) จำนวนตัวแปรอิสระ(p) มีค่าคงที่ ผลการวิจัยพบว่า ค่าเฉลี่ยของจุดแบ่งมีค่าไม่เกิน 0.5 ดังนี้

เมื่อขนาดตัวอย่างเท่ากับ 50 อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4846 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4848 ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4883 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

เมื่อขนาดตัวอย่างเท่ากับ 100 อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4846 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4807 ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4833 จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

เมื่อขนาดตัวอย่างเท่ากับ 150 อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4782 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4744 ที่จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4831 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

เมื่อขนาดตัวอย่างเท่ากับ 200 อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4726 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4766 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4718 จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

เมื่อขนาดตัวอย่างเท่ากับ 250 อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.25 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4730 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น อัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.5 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4766 ตัวแปรอิสระเท่ากับ 2 ตัว ให้ค่าเฉลี่ยของจุดแบ่งสูงต่ำสลับไปมา ที่จำนวนตัวแปรอิสระเท่ากับ 3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น และเมื่ออัตราส่วนความล้มเหลวของลักษณะที่สนใจศึกษาเท่ากับ 0.75 เมื่อระดับความสัมพันธ์ของตัวแปรอิสระเปลี่ยนไป จำนวนตัวแปรอิสระเท่ากับ 1 ตัว ให้ค่าเฉลี่ยของจุดแบ่งมีค่าคงที่เท่ากับ 0.4882 จำนวนตัวแปรอิสระเท่ากับ 2,3,4 และ 5 ตัวให้ค่าเฉลี่ยของจุดแบ่งมีค่าเพิ่มขึ้น

5.2 ข้อเสนอแนะ

1. ในงานวิจัยครั้งต่อไปทดสอบที่จำนวนขนาดตัวอย่างที่สูงขึ้นเพื่อดูว่าค่าเฉลี่ยของจุดแบ่งจะมีแนวโน้มใกล้เคียง 0.5 หรือไม่

2. ในการวิจัยครั้งนี้ เรากำหนดให้แต่ละสถานการณ์คงที่ เมื่อมีตัวแปรหนึ่งเปลี่ยนแปลง ในงานวิจัย ครั้งต่อไปอาจลองใช้ มีตัวแปรที่เปลี่ยนไปพร้อมกันหลายๆ ตัวแปร เพื่อดู Interaction ระหว่างแต่ละปัจจัยไปในตัว

รายการอ้างอิง

ภาษาไทย

กัลยา วานิชย์บัญชา.(2545). การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย. พิมพ์ครั้งที่ 6.โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

กัลยา วานิชย์บัญชา.(2551). การวิเคราะห์ข้อมูลหลายตัวแปร. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: สำนักพิมพ์ ธรรมสาร.

กัลยา วานิชย์บัญชา.(2551). หลักสถิติ. พิมพ์ครั้งที่ 10. กรุงเทพมหานคร: สำนักพิมพ์ ธรรมสาร.

ภาษาอังกฤษ

Amemiya,T.(1985):Advanced Econometrics.Harvard University Press,Cambridge,M.A.

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

Atkinson,A. (1970):A Method for Discriminating Between Models (with Discussion) , Journal of the Royal Statistical Society,B,B32,323-353.

ภาคผนวก

ตัวอย่างการใช้โปรแกรม R ในการวิจัย

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

n<-50;
p<-5;
```

```
a<-0.25
r0<-0;
r<-0.3;

beta<-15;
beta0<-log(-log(a));



cutoff<-c()
set.seed(555)
for(i in 1:500){

errorTerm<-rnorm(n,0,10);
x<-matrix(runif(n*p,-1,1),n);
if(p>1&r>0){
                    cor.mat<-matrix(r,p,p);
                    diag(cor.mat)<-1;
                    chol.mat<-chol(cor.mat);
                    x<-x%*%chol.mat;
                    rho.mat<-cor(x);
                    R<-rho.mat[row(rho.mat)!=col(rho.mat)];
                    while(max(R)>r|max(R)<=r0){
                              x<-matrix(runif(n*p),n);
                              x<-x%*%chol.mat;
                              rho.mat<-cor(x);
                              R<-rho.mat[row(rho.mat)!=col(rho.mat)];
                    }
                    x<-t(t(x)/colSums(chol.mat))
          }
```

```
Ystar<-beta0+beta*rowSums(x)+errorTerm;
Prob<-exp(-exp(Ystar));
Y<-rbinom(n,1,Prob);
data1<-data.frame(x,Y);
fitmodel<-glm(Y~.,data=data1,family=binomial(link="cloglog"));
pihat<-fitmodel$fitted;
MI<-rank(pihat,ties.method="max");
data1<-cbind(data1,pihat,MI);
data1<-data1[order(MI),];
AI<-ai.fun(data1);
data1<-cbind(data1,AI);



cutoff[i]<-max(data1$pihat[data1$AI==max(data1$AI)])
        }

c(mean(cutoff),quantile(cutoff,c(0.025,0.975)))
hist(cutoff)
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวนิภาพรรณ ไพจินดา เกิดเมื่อวันที่ 11 มีนาคม พ.ศ.2527 สำเร็จการศึกษาระดับปริญญาวิทยาศาสตรบัณฑิต ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ปีการศึกษา 2548 อาจารย์ประจำวิชาคณิตศาสตร์และวิทยาศาสตร์ ภาควิชาพื้นฐาน โรงเรียนไทยบริหารธุรกิจและพาณิชยการ พ.ศ. 2550 และ ศึกษาต่อระดับปริญญาสถิติศาสตรมหาบัณฑิต สาขาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2552