การกำจัดตัวแปรอิสระแบบถอยหลังในตัวแบบโพรบิท

นางสาวพรทิพย์ คำหล้า

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553

BACKWARD ELIMINATION OF INDEPENDENT VARIABLES IN THE PROBIT MODEL

Miss Pornthip Kumlar

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Statistics

Department of Statistics
Faculty of Commerce and Accountancy
Chulalongkorn University

Academic Year 2010

หัวข้อวิทยานิพนธ์	การกำจัดตัวแปรอิสระแบบถอยหลังในตัวแบบโพรบิท
โดย	นางสาวพรทิพย์ คำหล้า
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.... คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร. อรรณพ ตันละมัย)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธานกรรมการ

(รองศาสตราจารย์ ดร. ธีระพร วีระถาวร)

.... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา)

.... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร. เสกสรร เกียรติสุไพบูลย์)

.... กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ)

พรทิพย์ คำหล้า : การกำจัดตัวแปรอิสระแบบถอยหลังในตัวแบบโพรบิท. (BACKWARD ELIMINATION OF INDEPENDENT VARIABLES IN THE PROBIT MODEL) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร.สุพล ดุรงค์วัฒนา, 105 หน้า.

การศึกษานี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของวิธีการแก้ปัญหาพหุสัมพันธ์ระหว่างตัวแปรอิสระในตัวแบบโพรบิท 2 วิธีคือ วิธีการกำจัดตัวแปรแบบถอยหลังและวิธีการของฮอคกิ้ง โดยเปลี่ยนข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วยการแปลงข้อมูลของนอร์ดเบิร์ก ปัจจัยที่สนใจศึกษาในครั้งนี้คือ จำนวนตัวแปรอิสระ (p) ขนาดตัวอย่าง (n) ระดับความสัมพันธ์ของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) จำลองข้อมูลทั้งหมดที่ใช้ศึกษาด้วยเทคนิคมอนติคาร์โลโดยใช้โปรแกรม R ให้ตัวแปรอิสระมีการแจกแจงแบบปกติ สร้างตัวแปรอิสระ 3,4,5 และ 6 ตัว (p=3,4,5,6) โดยแบ่งตัวแปรอิสระเป็น 2 กลุ่ม คือ กลุ่มที่มีความสัมพันธ์กับตัวแปรตามและกลุ่มที่ไม่มีความสัมพันธ์กับตัวแปรตาม กำหนดค่าขนาดตัวอย่างเป็น 20,60,100 และ 200 (n=20,60,100,200) กำหนดระดับความสัมพันธ์ของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเป็น 0.1,0.5 และ 0.9 (M=0.1,0.5,0.9) ในแต่ละสถานการณ์กระทำซ้ำ 500 รอบ และใช้ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยที่พิจารณาจากจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามที่ถูกคัดเลือกให้อยู่ในตัวแบบสุดท้ายที่ได้ในแต่ละสถานการณ์ของการทดลองเป็นเกณฑ์ ผลการศึกษาพบว่า เมื่อจำนวนตัวแปรอิสระเท่ากับ 3,4,5 วิธีการของฮอคกิ้งมีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยต่ำกว่าวิธีการกำจัดตัวแปรแบบถอยหลัง เมื่อจำนวนตัวแปรอิสระเท่ากับ 6 โดยส่วนใหญ่ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของทั้งสองวิธีมีค่าสูงขึ้น และเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของวิธีการกำจัดตัวแปรแบบถอยหลังมีค่าต่ำกว่าวิธีการของฮอคกิ้ง แต่เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของทั้งสองวิธีมีค่าลดลง

ภาควิชา ....สถิติ.......... ลายมือชื่อนิสิต ....พรทิพย์ คำหล้า....

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

ปีการศึกษา ....2553..........

5181865326 : MAJOR STATISTICS

KEYWORDS : VARIABLE SELECTION /PROBIT MODEL / HOCKING'S PROCEDURE / BACKWARD ELIMINATION / TRANSFORMATION

PORNTHIP KUMLAR : BACKWARD ELIMINATION OF INDEPENDENT VARIABLES IN THE PROBIT MODEL. THESIS ADVISOR : ASSOC.PROF.SUPOL DURONGWATTANA, Ph.D., 105 pp.

The objective of this study is to investigate the efficiency of Backward Elimination Procedure and Hocking's Procedure for remedy of multicollinearity in the probit model. The data are transformed to data in multiple regression using Nordberg's transformation. The interested factors are number of independent variables (p), sample size (n), degree of multicollinearity among independent variables (M), and proportion between amount of independent variables which are associated with dependent variables and non-associated with dependent variables (a). The data are generated using Monte Carlo Technique through R-program. The independent variables are simulated having normal distribution, and the number of independent variables equal to 3, 4, 5, 6 (p = 3, 4, 5, 6). The independent variables are designated to be 2 groups; a group associated with the dependent variables and a group non-associated with the dependent variables. The sample size are set to be 20, 60, 100, 200 (n = 20, 60, 100, 200). The degree of multicollinearity among independent variables is designated to be 0.1, 0.5, 0.9 (M = 0.1, 0.5, 0.9). The simulation is repeated 500 times in each situation. The percentage of mean error is used as the performance measure. The results show that as the number of independent variables equal to 3, 4, 5, the percentage of mean error of Hocking's Procedure is lower than the one from Backward Elimination. When the number of independent variables is 6, the percentage of mean error increase for both methods in almost all of situations and the percentage of mean error of Backward Elimination is lower than the one from Hocking's Procedure. In both methods, when sample size increases, percentage of mean error decreases.

Department : Statistics	Student's Signature Pornthip Kumlar
Field of Study : Statistics	Advisor's Signature
Academic Year : 2010	

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์ช่วยเหลือเป็นอย่างดีจาก รองศาสตราจารย์ ดร. สุพล ดุรงค์วัฒนา อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำปรึกษา แนะนำแนวทางหลักการในการดำเนินงานวิจัย และช่วยตรวจสอบแก้ไขข้อบกพร่องต่างๆ ที่เกิดขึ้น เป็นอย่างดีตลอดระยะเวลาทำงานวิจัยซึ่งมีส่วนสำคัญอย่างมากที่ทำให้วิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความรู้สึกซาบซึ้ง เคารพและสำนึกในพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

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

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ ที่ท่านเสียสละเวลาอันมีค่ามาเป็นกรรมการภายนอกมหาวิทยาลัยและกรุณาตรวจสอบแก้ไขวิทยานิพนธ์ฉบับนี้ให้มีความสมบูรณ์มากยิ่งขึ้น

ท้ายสุดนี้ ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดามารดา ที่ช่วยส่งเสริม สนับสนุน และให้กำลังใจแก่ผู้วิจัยเสมอมาจนสำเร็จการศึกษา ตลอดจนเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่ให้กำลังใจ ให้คำปรึกษา และให้ความช่วยเหลือด้วยดีตลอดมา

สารบัญ

บทที่ หน้า

สารบัญตาราง

ตารางที่		หน้า

สารบัญภาพ

ภาพที่ หน้า

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การคาดคะเนเหตุการณ์ในอนาคตหรือการพยากรณ์เหตุการณ์ที่สนใจ ข้อมูลหรือปัจจัยที่เกี่ยวข้องและส่งผลกระทบต่อการเกิดเหตุการณ์ที่สนใจนั้นเป็นข้อมูลเบื้องต้นที่จะช่วยให้พยากรณ์เหตุการณ์ได้ใกล้เคียงความเป็นจริง หากทราบความสัมพันธ์ระหว่างเหตุการณ์ที่สนใจกับปัจจัยเหล่านั้นก็จะทำให้การพยากรณ์เหตุการณ์มีความถูกต้องมากขึ้นด้วย

ในปัจจุบัน ตัวแบบที่ใช้แสดงความสัมพันธ์ระหว่างเหตุการณ์ที่สนใจหรือตัวแปรตามกับปัจจัยที่ทำให้เกิดเหตุการณ์หรือตัวแปรอิสระมีหลากหลายรูปแบบ ซึ่งการเลือกใช้ตัวแบบในลักษณะต่างๆ ขึ้นอยู่กับชนิดของข้อมูล ตัวแบบความถดถอยเชิงเส้นเป็นตัวแบบที่นิยมใช้ในการวิเคราะห์ข้อมูลเมื่อตัวแปรตามเป็นตัวแปรเชิงปริมาณ และใช้ตัวแบบความถดถอยโลจิทหรือตัวแบบโพรบิทซึ่งเป็นตัวแบบที่นิยมใช้กันอย่างแพร่หลายในการวิเคราะห์ข้อมูลที่ตัวแปรตามเป็นตัวแปรเชิงคุณภาพหรือตัวแปรจำแนกประเภทซึ่งมีได้เพียง 2 ค่า คือ มีค่าเป็น 1 เมื่อเกิดเหตุการณ์ที่สนใจและมีค่าเป็น 0 เมื่อไม่เกิดเหตุการณ์ที่สนใจ

เมื่อได้ตัวแบบที่ใช้พยากรณ์เหตุการณ์ที่เหมาะสมกับชนิดของข้อมูลแล้ว ควรพิจารณาปัจจัยต่างๆ ที่อาจส่งผลกระทบต่อการพยากรณ์ด้วย ซึ่งปัจจัยหนึ่งก็คือจำนวนตัวแปรอิสระในตัวแบบ เนื่องจากหากมีจำนวนตัวแปรอิสระในตัวแบบมากเกินไปอาจเกิดพหุสัมพันธ์ (Multicollinearity) ระหว่างตัวแปรอิสระบางตัว ส่งผลให้ความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์เพิ่มสูงขึ้น ทำให้การพยากรณ์ผิดพลาดได้ อีกปัจจัยหนึ่งคือการเลือกตัวแปรอิสระที่ไม่เกี่ยวข้องเข้ามาอยู่ในตัวแบบ นั่นคือในตัวแบบมีทั้งตัวแปรอิสระที่มีผลกระทบและไม่มีผลกระทบต่อตัวแปรตามรวมอยู่ด้วยกัน ถ้าสามารถแยกตัวแปรอิสระเหล่านั้นออกจากกันได้ โดยมีเพียงตัวแปรอิสระที่มีผลกระทบต่อตัวแปรตามเท่านั้นที่อยู่ในตัวแบบก็จะทำให้การพยากรณ์มีประสิทธิภาพมากขึ้น

จากงานวิจัยของ Nordberg (1981) ได้เสนอวิธีการแก้ปัญหาพหุสัมพันธ์ในตัวแบบโลจิท โดยทำการแปลงข้อมูลที่ใช้ตัวแบบโลจิทซึ่งเป็นปีนตัวแบบที่ใช้เมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพหรือตัวแปรจำแนกประเภทซึ่งมีได้เพียง 2 ค่า คือ 1 กับ 0 แปลงเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้น (Multiple Regression) จากนั้นใช้การคัดเลือกตัวแปรอิสระ 2 วิธี คือ ใช้การเลือกตัวแปรอิสระแบบขั้นบันได (Stepwise Selection) และใช้วิธีการของฮอคกิ้ง (Hocking's Procedure) ซึ่งพบว่าเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระจะส่งผลกระทบต่อ

ประสิทธิภาพของการประมาณค่าพารามิเตอร์ โดยวิธีการของฮอคกิ้ง (Hocking's Procedure) จะมีประสิทธิภาพดีกว่าถ้าตัวแปรอิสระมีพหุสัมพันธ์ในระดับสูง แต่ถ้าตัวแปรอิสระมีพหุสัมพันธ์ในระดับต่ำหรือไม่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระหรือตัวแปรอิสระไม่มีความสัมพันธ์กัน วิธีการคัดเลือกตัวแปรทั้งสองวิธีจะให้ผลในการประมาณค่าพารามิเตอร์ใกล้เคียงกัน

เนื่องจากตัวแบบโพรบิทเป็นอีกตัวแบบหนึ่งที่นิยมใช้ในการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามเมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพเช่นเดียวกับตัวแบบโลจิท และการกำจัดตัวแปรแบบถอยหลังเป็นวิธีการหนึ่งในการแก้ปัญหาพหุสัมพันธ์ ผู้วิจัยจึงสนใจศึกษาวิธีแปลงข้อมูลของนอร์ดเบิร์กในตัวแบบโพรบิท และใช้วิธีการคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure) โดยศึกษาปัจจัยที่อาจส่งผลกระทบต่อวิธีการคัดเลือกตัวแปรดังกล่าวคือจำนวนตัวแปรอิสระ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม และขนาดตัวอย่าง

1.2 วัตถุประสงค์ของการวิจัย

ศึกษาประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิท 2 วิธี คือ

1. การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination)
2. วิธีการของฮอคกิ้ง (Hocking's Procedure)

โดยแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วยวิธีของนอร์ดเบิร์ก เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป

1.3 ขอบเขตของการวิจัย

การวิจัยนี้ศึกษาการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิทโดยมีขอบเขตดังนี้

1.3.1 ตัวแปรอิสระมีการแจกแจงแบบปกติมาตรฐาน

1.3.2 ใช้ตัวแปรอิสระ 3,4,5 และ 6 ตัว

แบ่งเป็น 2 กลุ่ม คือกลุ่มที่มีความสัมพันธ์กับตัวแปรตามและกลุ่มที่ไม่มีความสัมพันธ์กับตัวแปรตาม

1.3.3 ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M)

กำหนดเป็น 0.1, 0.5 และ 0.9 โดยเมทริกซ์สหสัมพันธ์ (Correlation Matrix) มีรูปแบบดังนี้

$$\rho_{pxp} = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \rho_{23} & \cdots & \rho_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \rho_{p3} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \rho & \frac{\rho}{2} & \cdots & \frac{\rho}{p-1} \\ \rho & 1 & \rho & \cdots & \frac{\rho}{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\rho}{p-1} & \frac{\rho}{p-2} & \frac{\rho}{p-3} & \cdots & 1 \end{pmatrix}$$

เมื่อ $\rho_{ij}; i,j=1,2,...,p$ คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

1.3.4 ขนาดตัวอย่างเท่ากับ 20, 60, 100 และ 200

1.3.5 กำหนดสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม

1.3.5.1 เมื่อ p = 3 ให้ a = 1:2 และ 2:1

1.3.5.2 เมื่อ p = 4 ให้ a = 1:3, 2:2 และ 3:1

1.3.5.3 เมื่อ p = 5 ให้ a = 1:4, 2:3, 3:2 และ 4:1

1.3.5.4 เมื่อ p = 6 ให้ a = 1:5, 2:4, 3:3, 4:2 และ 5:1

1.3.6 กำหนดค่าพารามิเตอร์ของข้อมูลจำลอง ดังนี้

1.3.6.1 $\beta_0 = 0.1$

1.3.6.2 $\beta_i = 0.1$ เมื่อ β_i เป็นสัมประสิทธ์ของตัวแปรอิสระตัวที่ i และมีความสัมพันธ์กับตัวแปรตาม

1.3.6.3 $\beta_j = 0$ เมื่อ β_j เป็นสัมประสิทธ์ของตัวแปรอิสระตัวที่ j และไม่มีความสัมพันธ์กับตัวแปรตาม

1.3.7 ค่าความคลาดเคลื่อนของหน่วยตัวอย่าง (ε_i) มีการแจกแจงแบบปกติมาตรฐาน

1.3.8 การวิจัยครั้งนี้ใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation) ในการจำลองข้อมูล และกระทำซ้ำ 500 รอบในแต่ละสถานการณ์

1.4 ข้อตกลงเบื้องต้น

สำหรับการวิจัยนี้ได้ศึกษาตัวอย่างจากการจำลองข้อมูลที่มีข้อตกลงเบื้องต้น ดังนี้

1.4.1 ศึกษาตัวแบบโพรบิท มีรูปแบบคือ

$$\begin{aligned}
\pi_i &= P\,(Y_i = 1) \\
&= P\left(\underset{\sim}{X}_i'\underset{\sim}{\beta} + \varepsilon_i > 0\right) \\
&= P\left(\varepsilon_i > -\underset{\sim}{X}_i'\underset{\sim}{\beta}\right) \\
&= 1 - \Phi\left(-\underset{\sim}{X}_i'\underset{\sim}{\beta}\right) \\
&= \Phi\left(\underset{\sim}{X}_i'\underset{\sim}{\beta}\right) \qquad ;\ i = 1, 2, \ldots, n
\end{aligned}$$

เมื่อ π_i คือค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

Y_i เป็นตัวแปรตามที่เป็นตัวแปรเชิงคุณภาพและมี 2 ค่า คือ มีค่าเป็น 1 เมื่อเกิดเหตุการณ์ที่สนใจ และมีค่าเป็น 0 เมื่อไม่เกิดเหตุการณ์ที่สนใจ

$\underset{\sim}{X}_i = \left(X_{1i}, X_{2i}, \ldots, X_{pi}\right)'$ แทนเวกเตอร์ของตัวแปรอิสระ p ตัว

$\underset{\sim}{\beta} = \left(\beta_0, \beta_1, \beta_2, \ldots, \beta_p\right)'$ แทนเวกเตอร์พารามิเตอร์

ε_i คือค่าความคลาดเคลื่อนโดยที่ $\varepsilon_i \overset{iid}{\sim} N(0,1)$

n คือขนาดตัวอย่าง

และ $\Phi(.)$ คือฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน

1.4.2 ศึกษาตัวแปรอิสระที่มีการแจกแจงแบบปกติ (Normal Distribution) ที่มีฟังก์ชันความน่าจะเป็นของ x อยู่ในรูปของ

$$f(x) = \frac{1}{\sqrt{2\pi}\zeta} e^{-\frac{1}{2\zeta^2}\cdot(x-\mu)^2}$$

1.5 คำจำกัดความที่ใช้ในการวิจัย

1.5.1 ตัวแบบโพรบิท คือตัวแบบที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระเมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่มี 2 ลักษณะ คือ มีค่าเป็น 1 เมื่อเกิดเหตุการณ์ที่สนใจ และมีค่าเป็น 0 เมื่อไม่เกิดเหตุการณ์ที่สนใจ ส่วนตัวแปรอิสระเป็นได้ทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพ หรือมีทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพก็ได้ เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรแล้ว จะนำตัวแบบโพรบิทนี้ไปใช้ในการประมาณค่าตัวแปรตามหรือการพยากรณ์โอกาสที่แต่ละหน่วยจะอยู่ในกลุ่มใดกลุ่มหนึ่ง

1.5.2 การแจกแจงแบบเบอร์นูลลี (Bernoulli distribution)

ตัวแปรสุ่ม Y เรียกว่า ตัวแปรสุ่มเบอร์นูลลี ถ้า

กำหนด $Y=1$ เมื่อเกิดเหตุการณ์ที่สนใจ

และ $Y=0$ เมื่อไม่เกิดเหตุการณ์ที่สนใจ

โดยที่ $P(Y=1)=p \quad ,0<p<1$

และ $P(Y=0)=1-p$

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

$$P(Y=y)=p^{Y}(1-p)^{1-Y} \quad , y=0,1$$

1.5.3 การแปลงข้อมูลของนอร์ดเบิร์ก เป็นการแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้น

1.5.4 พหุสัมพันธ์ (multicollinearity) คือสถานการณ์ที่ตัวแปรอิสระมีความสัมพันธ์กันทำให้ค่าประมาณของพารามิเตอร์ที่ได้จากวิธีภาวะน่าจะเป็นสูงสุดขาดความแม่นยำและมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองสูงขึ้น ซึ่งไม่สามารถนำไปใช้อธิบายอิทธิพลของตัวแปรอิสระที่มีต่อตัวแปรตามเมื่อตัวแปรอิสระอื่นมีค่าคงที่ ทำให้การวิเคราะห์ข้อมูลและการพยากรณ์ผิดพลาดได้

1.5.5 ตัวประมาณริดจ์ เป็นตัวประมาณค่าพารามิเตอร์ในการแก้ปัญหาพหุสัมพันธ์ระหว่างตัวแปรอิสระที่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุด

1.5.6 การกำจัดตัวแปรแบบถอยหลัง เป็นวิธีการคัดเลือกตัวแปรอิสระโดยเริ่มต้นด้วยตัวแบบที่สมบูรณ์ คือประกอบด้วยตัวแปรอิสระทุกตัวที่ใช้พิจารณา แล้วเลือกตัวแปรอิสระออกจากสมการครั้งละ 1 ตัวแปร

1.5.7 การถดถอยแบบขั้นบันได เป็นวิธีการคัดเลือกตัวแปรอิสระโดยเริ่มต้นด้วยการเลือกตัวแปรอิสระเข้าสู่สมการครั้งละหนึ่งตัว แล้วตรวจสอบดูว่ามีตัวแปรอิสระตัวใดที่ควรเลือกออกจากสมการหรือไม่ แล้วจึงเลือกตัวแปรอิสระตัวใหม่เข้ามาในสมการ ทำเช่นนี้ไปเรื่อยๆ และจะหยุดเมื่อไม่มีตัวแปรใดเข้าและออกจากสมการถดถอยแล้ว

1.6 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อทราบประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิท 2 วิธี คือ

1. การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination)
2. วิธีการของฮอคกิ้ง (Hocking's Procedure)

โดยแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วยวิธีของนอร์ดเบิร์ก เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม และระดับความสัมพันธ์ระหว่างตัวแปรอิสระเปลี่ยนแปลงไป และใช้เป็นแนวทางสำหรับผู้วิจัยในการเลือกวิธีการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิทต่อไป

1.7 วิธีดำเนินการวิจัย

1.7.1 ศึกษาค้นคว้าเอกสารและข้อมูลที่เกี่ยวข้องกับการวิจัย

1.7.2 จำลองข้อมูลให้มีลักษณะที่ต้องการศึกษา

1.7.3 แปลงข้อมูลด้วยวิธีของนอร์ดเบิร์ก

1.7.4 ใช้การคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure)

1.7.5 นับจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามแต่ถูกคัดเลือกให้อยู่ในตัวแบบในแต่ละรอบ และทำการทดลองซ้ำจำนวน 500 รอบในแต่ลสถานการณ์

1.7.6 คำนวณเปอร์เซ็นต์ความผิดพลาดของการคัดเลือกตัวแปรอิสระ

1.7.7 สรุปผลการวิจัย

1.8 เกณฑ์ที่ใช้ในการตัดสินใจ

ใช้ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการคัดเลือกตัวแปรอิสระที่พิจาณาจากจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามในตัวแบบสุดท้ายที่ได้ในแต่ละสถานการณ์ของการศึกษาเป็นเกณฑ์ในการตัดสินใจว่าจำนวนตัวแปรอิสระ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ขนาดตัวอย่าง และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม มีผลกระทบต่อประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิทเมื่อแปลงข้อมูลด้วยวิธีการของนอร์ดเบิร์กโดยใช้วิธีการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking’s Procedure) หรือไม่ โดยค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการคัดเลือกตัวแปรอิสระสามารถคำนวณได้จาก

$$\frac{\sum_{i=1}^{m} w_i}{r \cdot m} \times 100\%$$

เมื่อ w_i คือจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามที่ถูกคัดเลือกให้อยู่ในตัวแบบในการวิเคราะห์รอบที่ i

r คือจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามทั้งหมดในตัวแบบเริ่มต้น

และ m คือจำนวนรอบในการทำซ้ำ

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

งานวิจัยนี้ศึกษาประสิทธิภาพของการเลือกตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามในตัวแบบโพรบิท โดยการแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นและใช้วิธีการคัดเลือกตัวแปรอิสระ 2 วิธีคือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure) ซึ่งประกอบด้วย 2 ขั้นตอน คือการประมาณค่าด้วยตัวประมาณริดจ์ และใช้การถดถอยแบบขั้นบันได (Stepwise Regression) ในการคัดเลือกตัวแปรอิสระ ซึ่งมีทฤษฎีและตัวสถิติที่เกี่ยวข้องดังนี้

2.1.1 ตัวแบบโพรบิท

ให้ y_i เป็นตัวแปรตามซึ่งเป็นตัวแปรเชิงคุณภาพที่เป็นได้ 2 ค่า คือ มีค่าเป็น 1 เมื่อเกิดเหตุการณ์ที่สนใจ และมีค่าเป็น 0 เมื่อไม่เกิดเหตุการณ์ที่สนใจ และ $x_{i1},x_{i2},...,x_{ip}$ เป็นตัวแปรอิสระของค่าสังเกตที่ได้จากหน่วยตัวอย่างที่ i โดยที่ $i=1,2,...,n$

นิยาม ตัวแปรแฝง (Latent Variable) y_i^* ของหน่วยตัวอย่างที่ i เป็นฟังก์ชันเชิงเส้นของตัวแปรอิสระ $x_{i1},x_{i2},...,x_{ip}$ นั่นคือ

$$y_i^* = \sum_{j=0}^{p} \beta_j x_{ij} + \varepsilon_i$$

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

และให้ $\underset{\sim}{\beta} = \left(\beta_0,\beta_1,\beta_2,...,\beta_p\right)'$ แทนเวกเตอร์พารามิเตอร์ จะได้

$$y_i = \begin{cases} 1 & \text{ถ้า } y_i^* > 0 \\ 0 & \text{ถ้า } y_i^* \leq 0 \end{cases}$$

พิจารณาความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ

$$\begin{aligned}\pi_i &= P(y_i = 1)\\ &= P\left(\underset{\sim}{x}_i'\underset{\sim}{\beta} + \varepsilon_i > 0\right)\\ &= P\left(\varepsilon_i > -\underset{\sim}{x}_i'\underset{\sim}{\beta}\right)\\ &= 1 - \Phi\left(-\underset{\sim}{x}_i'\underset{\sim}{\beta}\right)\\ &= \Phi\left(\underset{\sim}{x}_i'\underset{\sim}{\beta}\right) \qquad ; i = 1, 2, ..., n\end{aligned}$$

เมื่อ π_i คือค่าความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจในหน่วยที่ i

$\underset{\sim}{x}_i = \left(x_{1i}, x_{2i}, ..., x_{pi}\right)'$ แทนเวกเตอร์ของตัวแปรอิสระ p ตัว

n คือขนาดตัวอย่าง

และ $\Phi(.)$ คือฟังก์ชันการแจกแจงสะสมปกติมาตรฐาน

2.1.2 **ฟังก์ชันภาวะน่าจะเป็น** (Likelihood function) **ของตัวแบบโพรบิท**

ให้ $g(y_1, y_2, ..., y_n)$ แทนการแจกแจงความน่าจะเป็นร่วมของ $y_1, y_2, ..., y_n$ จะได้ว่า

$$\begin{aligned}g(y_1, y_2, ..., y_n) &= \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i}\\ &= \prod_{i=1}^{n}\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)^{y_i}\left\{1-\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)\right\}^{1-y_i}\end{aligned}$$

ดังนั้น ฟังก์ชันภาวะน่าจะเป็นของ $y_1, y_2, ..., y_n$ คือ

$$L(\beta; y) = \prod_{i=1}^{n}\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)^{y_i}\left\{1-\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)\right\}^{1-y_i}$$

2.1.3 **ตัวประมาณภาวะน่าจะเป็นสูงสุด** (Maximum Likelihood Estimator)

การประมาณค่าพารามิเตอร์ $\underset{\sim}{\beta}$ ในตัวแบบโดยใช้หลักการคือหา $\hat{\underset{\sim}{\beta}}$ (เวกเตอร์ค่าประมาณของพารามิเตอร์) ที่ทำให้ $L\left(\underset{\sim}{\beta}; \underset{\sim}{y}\right)$ มีภาวะน่าจะเป็นสูงสุด นั่นคือ $\hat{\underset{\sim}{\beta}}$ สอดคล้องกับ

$$L\left(\hat{\underset{\sim}{\beta}}; \underset{\sim}{y}\right) = \max_{\underset{\sim}{\beta}} L\left(\underset{\sim}{\beta}; \underset{\sim}{y}\right)$$

จะเรียก $\hat{\underset{\sim}{\beta}}$ ว่าเป็นตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator)

เนื่องจากฟังก์ชันลอการิทึมเป็นฟังก์ชันเพิ่ม จะได้ว่า $\hat{\underset{\sim}{\beta}}$ จะทำให้ $L(\underset{\sim}{\beta};\underset{\sim}{y})$ มีค่ามากที่สุดก็ต่อเมื่อทำให้ฟังก์ชัน $\ln L(\underset{\sim}{\beta};\underset{\sim}{y})$ มีค่ามากที่สุด ซึ่งการหาอนุพันธ์จากฟังก์ชันลอการิทึมทำได้สะดวกกว่าฟังก์ชันภาวะน่าจะเป็นในกรณีที่ตัวแปรสุ่มมีการแจกแจงเป็นวงศ์ชี้กำลัง

สำหรับตัวแบบโพรบิท จะได้

$$\ln L(\underset{\sim}{\beta};\underset{\sim}{y}) = \sum_{i=1}^{n}\left[y_i \cdot \left\{ \ln\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)\right\} + (1-y_i)\cdot \ln\left\{1-\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)\right\}\right]$$

$$= \sum_{y_i=0} \ln\left\{1-\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)\right\} + \sum_{y_i=1}\ln\Phi\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)$$

อนุพันธ์อันดับหนึ่งของ $\ln L(\underset{\sim}{\beta};\underset{\sim}{y})$ เทียบกับ $\underset{\sim}{\beta}$ คือ

$$\frac{\partial \ln L(\underset{\sim}{\beta};\underset{\sim}{y})}{\partial \underset{\sim}{\beta}} = \sum_{i=0}^{n}\left\{\frac{y_i\phi_i}{\Phi_i} + (1-y_i)\cdot\frac{1-\phi_i}{1-\Phi_i}\right\}$$

$$= \sum_{y_i=0}\frac{-\phi_i\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)}{1-\Phi_i\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)}\cdot x_i + \sum_{y_i=1}\frac{\phi_i\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)}{\Phi_i\left(\sum_{j=0}^{p}\beta_{ij}x_{ij}\right)}\cdot x_i$$

$$= 0$$

เมื่อ ϕ_i แทนฟังก์ชันความหนาแน่นของค่าคลาดเคลื่อนที่ได้จากตัวอย่างที่ i ซึ่งมีการ แจกแจกแบบปกติมาตรฐาน

และ Φ_i แทนฟังก์ชันการแจกแจงสะสมของค่าคลาดเคลื่อนที่ได้จากตัวอย่างที่ i ซึ่งมีการแจกแจกแบบปกติมาตรฐาน

2.1.4 **สถิติ** Kaiser-Meyer-Olkin (KMO)

Kaiser (1970) ได้เสนอสถิติ KMO เป็นสถิติที่ใช้วัดระดับพหุสัมพันธ์ระหว่างตัวแปรอิสระ (multicollinearity) ซึ่งคำนวณจากหน่วยตัวอย่างโดยใช้สูตร ดังนี้

$$0 \le \mathrm{KMO} = \frac{\sum_{i=1}^{p}\sum_{i<j=1}^{p} r_{ij}^2}{\sum_{i=1}^{p}\sum_{i<j=1}^{p} r_{ij}^2 + \sum_{i=1}^{p}\sum_{i<j=1}^{p} r_{ij\,|\,1,2,\ldots,(i-1),(i+1),\ldots,(j-1),(j+1),\ldots,p}^2} \le 1$$

โดยที่ r_{ij} คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j และ $r_{ij \mid 1,2,\ldots,(i-1),(i+1),\ldots,(j-1),(j+1),\ldots p}$ คือสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j สำหรับ $i < j = 1, 2, \ldots, p$

2.1.5 **ตัวประมาณริดจ์** (Ridge Estimator)

ให้ $\underset{\sim}{z}$ เป็นเวกเตอร์ค่าสังเกตขนาด n x 1

U เป็น design matrix ขนาด n x (p+1)

$\underset{\sim}{\beta} = \left(\beta_0, \beta_1, \ldots, \beta_p\right)'$ เป็นเวกเตอร์พารามิเตอร์ที่ไม่ทราบค่า

และ $\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ค่าคลาดเคลื่อนขนาด n x 1 ซึ่งมีค่าเฉลี่ยเท่ากับศูนย์และความแปรปรวนเท่ากับ $\zeta^2 I$

พิจารณาตัวแบบความถดถอยเชิงเส้น (Linear Regression Model)

$$\underset{\sim}{z} = U\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

สำหรับค่า k ใดๆ ที่ $k \geq 0$ ตัวประมาณริดจ์ที่ k ของ $\underset{\sim}{\beta}$ คือ

$$\hat{\underset{\sim}{\beta}}_R^{(k)} = \left(U'U + kI\right)^{-1} U'\underset{\sim}{z}$$

ตัวประมาณริดจ์ที่เหมาะสมคือตัวประมาณที่ทำให้ค่าความคลาดเคลื่อนกำลังสอง (mean square error;MSE) มีค่าน้อยที่สุดสำหรับทุกค่า $k \geq 0$ นั่นคือ

$$\mathrm{MSE}\left(\hat{\underset{\sim}{\beta}}_R^{(k_{opt})}\right) \leq \mathrm{MSE}\left(\hat{\underset{\sim}{\beta}}_R^{(k)}\right)$$

2.1.6 **การแปลงข้อมูลด้วยวิธีของนอร์ดเบิร์ก**

การแปลงข้อมูลด้วยวิธีของนอร์ดเบิร์กเป็นการแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้น มีขั้นตอนดังนี้

1) ประมาณค่าพารามิเตอร์ด้วยตัวประมาณภาวะน่าจะเป็นสูงสุด
2) แปลงข้อมูลตัวแปรอิสระ x_{ij} เป็น u_{ij} และตัวแปรตาม y_i เป็น z_i ดังนี้

$$\hat{u}_{ij} = x_{ij}\sqrt{p_i\left(\hat{\underset{\sim}{\beta}}\right)\left(1 - p_i\left(\hat{\underset{\sim}{\beta}}\right)\right)} \quad , i = 1,2,\ldots,n \; , j = 1,2,\ldots,p$$

$$\hat{z}_i = \left(\frac{y_i - p_i\left(\hat{\underset{\sim}{\beta}}\right)}{\sqrt{p_i\left(\hat{\underset{\sim}{\beta}}\right)\left(1 - p_i\left(\hat{\underset{\sim}{\beta}}\right)\right)}}\right) + \sum_{j=1}^{p}\hat{\beta}_j\hat{u}_{ij} \quad , i = 1,2,\ldots,n$$

จะได้ตัวแบบความถดถอยเชิงเส้น (Linear Regression Model)

$$\hat{z}_i = \sum_{j=1}^{p}\hat{\beta}_j\hat{u}_{ij} + \varepsilon_i \quad , i = 1,2,\ldots,n$$

2.1.7 **วิธีการของฮอคกิ้ง** (Hocking's Procedure)

มีขั้นตอนดังนี้

1) เลือกค่า k ที่ทำให้ $MSE\left(\hat{\underset{\sim}{\beta}}_R^{(k)}\right)$ มีค่าต่ำที่สุด โดยที่ $k \geq 0$ และ $MSE\left(\hat{\underset{\sim}{\beta}}_R^{(k)}\right)$ คำนวณได้จาก

$$MSE\left(\hat{\underset{\sim}{\beta}}_R^{(k)}\right) = \zeta^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \underset{\sim}{\beta}' (U'U + kI)^{-2} \underset{\sim}{\beta}$$

โดยประมาณค่า ζ^2 ด้วย $\zeta^2 = \dfrac{\left|z - U\hat{\underset{\sim}{\beta}}_R^{(0)}\right|}{n-p-1}$ และประมาณค่า $\underset{\sim}{\beta}$ ด้วย $\hat{\underset{\sim}{\beta}}_R^{(0)}$

เมื่อ $\underset{\sim}{\lambda} = \left(\lambda_1, \lambda_2, \ldots, \lambda_p\right)'$ เป็นเวกเตอร์ของ eigenvalue ของ $U'U$

2) ขยายตัวแบบความถดถอยเชิงเส้น (Linear Regression Model)

$$\underset{\sim}{z} = U\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

ด้วยตัวแปรหุ่น (dummy variable) จะได้ตัวแบบ

$$\begin{pmatrix} \underset{\sim}{z} \\ 0 \end{pmatrix} = \begin{pmatrix} U \\ \sqrt{k}I \end{pmatrix} \underset{\sim}{\beta} + \begin{pmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\delta} \end{pmatrix}$$

เมื่อ I แทนเมทริกซ์เอกลักษณ์ขนาด (p+1) x (p+1)

p คือจำนวนตัวแปรอิสระ

และ $\underset{\sim}{\delta}$ แทนเวกเตอร์ค่าคลาดเคลื่อนเช่นเดียวกับ $\underset{\sim}{\varepsilon}$

3) ใช้วิธีการเลือกตัวแปรอิสระแบบขั้นบันได โดยชุดของตัวแปรอิสระสุดท้ายที่ได้จะอยู่ในรูปของตัวแปรอิสระ u_j สำหรับบางค่าของ j ซึ่งตัวแปรอิสระ x_j ที่ใช้แปลงข้อมูลจะเป็นตัวแปรอิสระที่ต้องการ

2.1.8 **การกำจัดตัวแปรแบบถอยหลัง** (Backward Elimination)

การกำจัดตัวแปรแบบถอยหลังเป็นวิธีการคัดเลือกตัวแปรอิสระโดยเริ่มต้นด้วยตัวแบบที่สมบูรณ์คือประกอบด้วยตัวแปรอิสระทุกตัวที่ใช้พิจารณา แล้วจะเลือกตัวแปรอิสระออกจากสมการครั้งละ 1 ตัวแปร โดยจะเลือกตัวแปรอิสระออกเมื่อตัวแปรอิสระนั้นความสัมพันธ์กับตัวแปรตามน้อยที่สุด (เมื่อตัวแปรอื่นๆ คงที่) และค่าสัมประสิทธิ์ความถดถอยของตัวแปรอิสระนั้นไม่มีนัยสำคัญ จากนั้นคำนวณหาสมการความถดถอยสำหรับตัวแปรอิสระที่เหลือและเลือกตัวแปรอิสระที่มีนัยสำคัญน้อยที่สุดออก ทำเช่นนี้ไปเรื่อย ๆ จนกระทั่งไม่มีตัวแปรอิสระใดถูกเลือกออก โดยมีขั้นตอนดังนี้

1) สร้างสมการถดถอยโดยเริ่มต้นจากตัวแบบที่สมบูรณ์ (Complete Model) คือมีตัวแปรอิสระทุกตัวอยู่ในสมการ
2) คำนวณค่าสถิติเอฟบางส่วนสำหรับตัวแปรอิสระทุกตัวที่อยู่ในสมการ
3) ค่าสถิติเอฟบางส่วนที่น้อยที่สุด (F_L) จะถูกเปรียบเทียบกับค่าเอฟ ณ ระดับนัยสำคัญที่กำหนด (F_0)

ถ้า $F_L < F_0$ จะตัดตัวแปรอิสระที่มีค่าสถิตเอฟบางส่วนน้อยที่สุดออกจากสมการถดถอย และคำนวณค่าสัมประสิทธิ์ของสมการถดถอยเมื่อตัดตัวแปรอิสระนั้นแล้ว จากนั้นกลับไปทำในข้อ 2)

ถ้า $F_L > F_0$ จะหยุดกระบวนการคัดเลือกตัวแปรอิสระ จะได้สมการถดถอยที่เหมาะสมโดยตัวแปรอิสระที่เหลือจะอยู่ในสมการถดถอยที่ได้

2.1.9 **การถดถอยแบบขั้นบันได** (Stepwise Regression)

การถดถอยแบบขั้นบันไดเป็นวิธีการคัดเลือกตัวแปรอิสระโดยมีหลักการคือ เริ่มต้นด้วยการเลือกตัวแปรอิสระเข้าสู่สมการครั้งละหนึ่งตัว จากนั้นจะตรวจสอบดูว่ามีตัวแปรอิสระตัวใดที่ควรเลือกออกจากสมการหรือไม่ แล้วจึงเลือกตัวแปรอิสระตัวใหม่เข้ามาในสมการ ทำเช่นนี้ไปเรื่อยๆ และจะหยุดเมื่อไม่มีตัวแปรใดเข้าและออกจากสมการถดถอยแล้ว โดยมีขั้นตอนดังนี้

1) สร้างสมการถดถอยโดยเริ่มจากไม่มีตัวแปรอิสระอยู่ในสมการ
2) เลือกตัวแปรอิสระที่มีสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) กับตัวแปรตามสูงสุดเข้าสู่สมการเป็นตัวแปรแรก และทดสอบว่าตัวแปรอิสระดังกล่าวมีนัยสำคัญหรือไม่ หากตัวแปรอิสระดังกล่าวไม่มีนัยสำคัญจะได้สมการถดถอยที่เหมาะสมคือ $y = \bar{y}$ แต่ถ้าตัวแปรอิสระดังกล่าวมีนัยสำคัญจะทำในขั้นตอนต่อไป
3) คำนวณค่าสถิติเอฟบางส่วนสำหรับทุกตัวแปรอิสระที่ไม่อยู่ในสมการ
4) ค่าสถิติเอฟบางส่วนที่มากที่สุด (F_U) จะถูกเปรียบเทียบกับค่าเอฟ ณ ระดับนัยสำคัญที่กำหนด (F_0)

ถ้า $F_U < F_0$ จะไม่นำตัวแปรอิสระที่มีค่าสถิตเอฟบางส่วนมากที่สุดเข้าสู่สมการถดถอย

ถ้า $F_U > F_0$ จะไม่นำตัวแปรอิสระที่มีค่าสถิตเอฟบางส่วนมากที่สุดเข้าสู่สมการถดถอย และคำนวณค่าสัมประสิทธิ์ของสมการถดถอยเมื่อนำตัวแปรอิสระนั้นเข้าสู่สมการ

5) คำนวณค่าสถิติเอฟบางส่วนของทุกตัวแปรอิสระที่อยู่ในสมการถดถอย
6) ค่าสถิติเอฟบางส่วนที่น้อยที่สุด (F_L) จะถูกเปรียบเทียบกับค่าเอฟ ณ ระดับนัยสำคัญที่กำหนด (F_0)

ถ้า $F_L < F_0$ จะตัดตัวแปรอิสระที่มีค่าสถิติเอฟบางส่วนน้อยที่สุดออกจากสมการถดถอย และคำนวณค่าสัมประสิทธิ์ของสมการถดถอยเมื่อตัดตัวแปรอิสระนั้นแล้ว

ถ้า $F_L > F_0$ จะไม่ตัดตัวแปรอิสระที่มีค่าสถิตเอฟบางส่วนน้อยที่สุดออกจากสมการถดถอย

7) ถ้าไม่มีตัวแปรอิสระใดเข้าและออกจากสมการถดถอยแล้ว จะได้สมการถดถอยที่เหมาะสม แต่ถ้ายังมีตัวแปรอิสระใดที่เป็นไปตามเงื่อนไขของการเลือกตัวแปรอิสระเข้าหรือออกจากสมการ ให้กลับไปทำในข้อ 3)

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

จากงานวิจัยของ Nordberg (1981) ได้เสนอวิธีการแก้ปัญหาพหุสัมพันธ์ในตัวแบบโลจิท โดยทำการแปลงข้อมูลที่ใช้ตัวแบบโลจิทซึ่งเป็นปืนตัวแบบที่ใช้เมื่อตัวแปรตามเป็นตัวแปรเชิงคุณภาพหรือตัวแปรจำแนกประเภทซึ่งมีได้เพียง 2 ค่า คือ 1 กับ 0 แปลงเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้น (Multiple Regression) จากนั้นทำการคัดเลือกตัวแปรอิสระ 2 วิธี คือใช้การเลือกตัวแปรอิสระแบบขั้นบันได (Stepwise Selection) และใช้วิธีการของฮอคกิ้ง (Hocking's Procedure) ซึ่งพบว่าเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระจะส่งผลกระทบต่อประสิทธิภาพของการประมาณค่าพารามิเตอร์โดยวิธีการของฮอคกิ้ง (Hocking's Procedure) จะมีประสิทธิภาพดีกว่าถ้าตัวแปรอิสระมีพหุสัมพันธ์ในระดับสูง แต่ถ้าตัวแปรอิสระมีพหุสัมพันธ์ในระดับต่ำหรือไม่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระหรือตัวแปรอิสระไม่มีความสัมพันธ์กันวิธีการคัดเลือกตัวแปรทั้งสองวิธีจะให้ผลในการประมาณค่าพารามิเตอร์ใกล้เคียงกัน

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ได้ศึกษาประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure) โดยจะนำข้อมูลที่ใช้ตัวแบบโพรบิทแปลงเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วยวิธีของนอร์ดเบิร์ก ในการจำลองข้อมูลในแต่ละสถานการณ์ได้ใช้เทคนิคมอนติคาร์โลและใช้โปรแกรม R บนเครื่องคอมพิวเตอร์ในการประมวณผล ดังนั้น ในบทนี้จะกล่าวถึงรายละเอียดของเทคนิคมอนติคาร์โล รวมถึงแผนการดำเนินการวิจัย ขั้นตอนในการดำเนินการวิจัย การจำลองข้อมูลในการวิจัย ตลอดจนขั้นตอนของโปรแกรมที่ใช้ในการวิจัย

3.1 เทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นการจำลองระบบที่ไม่เปลี่ยนแปลงตามเวลา ตัวแบบของการจำลองจะมีลักษณะเป็นตัวแบบทางคณิตศาสตร์ โดยใช้การนำตัวเลขสุ่มมาประยุกต์ใช้ในการแก้ปัญหาหรือหาคำตอบให้กับระบบที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น เนื่องจากเลขสุ่มมาจากแนวคิดเกี่ยวกับการคำนวณค่าความน่าจะเป็น ทำให้การเลือกตัวอย่างไม่มีความเอนเอียง ซึ่งมีขั้นตอนที่สำคัญ 3 ขั้นตอน ดังนี้

1) การสร้างเลขสุ่ม (Generate Random Number) การสร้างเลขสุ่มจะกำหนดให้มีการแจกแจงแบบสม่ำเสมอ (Uniform Distribution) ในช่วง [0,1] และเป็นอิสระซึ่งกันและกัน จากนั้นนำเลขสุ่มที่ได้ไปสร้างตัวแปรตามให้มีลักษณะการแจกแจงที่ต้องการศึกษาเพื่อเป็นข้อมูลสำหรับปัญหานั้น ๆ
2) การประยุกต์ใช้เลขสุ่มในการแก้ปัญหา ขั้นตอนนี้เป็นการนำตัวแปรที่ได้จากขั้นตอนที่ 1) มาใช้ในปัญหาที่ต้องการศึกษา
3) การทดลอง ขั้นตอนนี้เป็นการทำวิธีนั้นซ้ำ ๆ กัน (Replication) จำนวนหลาย ๆ ครั้ง โดยถือว่าการทำซ้ำ ๆ กันนั้นเป็นวิธีการเก็บรวบรวมข้อมูลให้มีจำนวนมากเพื่อลดค่าความไม่แน่นอนของคำตอบในการวิเคราะห์ปัญหาต่าง ๆ

จากหลักการของเทคนิคมอนติคาร์โล จะเห็นได้ว่าการใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหาเป็นวิธีการที่จะนำไปสู่แนวคิดทางทฤษฎีที่เกี่ยวกับกับการคำนวณ โดยเฉพาะทฤษฎีความน่าจะเป็นที่นำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริงเพราะ

ไม่มีผลกระทบจากปัจจัยอื่น ๆ เข้ามาเกี่ยวข้องในการทดลอง เมื่อกระทำซ้ำ ๆ กันเป็นจำนวนมาก ค่าความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่าง ๆ ในแต่ละครั้งจะหมดไป

3.2 แผนการดำเนินการวิจัย

การวิจัยในครั้งนี้ ได้จำลองข้อมูลขึ้นโดยกำหนดสถานการณ์จำลองต่าง ๆ ดังนี้

3.2.1 กำหนดให้ตัวแปรอิสระ (X) มีการแจกแจงแบบปกติมาตรฐาน

3.2.2 กำหนดจำนวนตัวแปรอิสระ 3,4,5 และ 6 ตัว แบ่งเป็น 2 กลุ่ม คือ กลุ่มที่มีความสัมพันธ์กับตัวแปรตามและกลุ่มที่ไม่มีความสัมพันธ์กับตัวแปรตาม

3.2.3 ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) กำหนดเป็น 0.1, 0.5 และ 0.9

3.2.4 กำหนดขนาดตัวอย่างเท่ากับ 20, 60, 100 และ 200

3.2.5 กำหนดตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มี 2 ค่า คือ 1 และ 0 โดยให้ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจเท่ากับความน่าจะเป็นของการไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 0.5

3.2.6 กำหนดสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม

3.2.6.1 เมื่อ p = 3 ให้ a = 1:2 และ 2:1

3.2.6.2 เมื่อ p = 4 ให้ a = 1:3, 2:2 และ 3:1

3.2.6.3 เมื่อ p = 5 ให้ a = 1:4, 2:3, 3:2 และ 4:1

3.2.6.4 เมื่อ p = 6 ให้ a = 1:5, 2:4, 3:3, 4:2 และ 5:1

3.2.7 กำหนดค่าพารามิเตอร์ของข้อมูลจำลอง ดังนี้

3.2.7.1 $\beta_0 = 0.1$

3.2.7.2 $\beta_i = 0.1$ เมื่อ β_i เป็นสัมประสิทธ์ของตัวแปรอิสระตัวที่ i และมีความสัมพันธ์กับตัวแปรตาม

3.2.7.3 $\beta_j = 0$ เมื่อ β_j เป็นสัมประสิทธ์ของตัวแปรอิสระตัวที่ j และไม่มีความสัมพันธ์กับตัวแปรตาม

3.2.8 กำนหดให้ความคลาดเคลื่อน (ε) มีการแจกแจงแบบปกติมาตรฐาน

3.2.9 กำหนดจำนวนการกระทำซ้ำในแต่ละสถานการณ์เท่ากับ 500 รอบ

3.3 ขั้นตอนในการดำเนินการวิจัย

3.3.1 สร้างข้อมูลที่ใช้ในการวิจัย

3.3.2 ประมาณค่าพารามิเตอร์ด้วยตัวประมาณภาวะน่าจะเป็นสูงสุด

3.3.3 แปลงข้อมูลด้วยวิธีของนอร์ดเบิร์ก

3.3.4 ใช้การคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure)

3.3.5 นับจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามแต่ถูกคัดเลือกให้อยู่ในตัวแบบในแต่ละรอบ และทำการทดลองซ้ำจำนวน 500 รอบในแต่ลสถานการณ์

3.3.6 คำนวณเปอร์เซ็นต์ความผิดพลาดของการคัดเลือกตัวแปรอิสระ

3.3.7 สรุปผลการวิจัย

3.4 การจำลองข้อมูลในการวิจัย

3.4.1 สร้างข้อมูลตัวแปรอิสระให้มีการแจกแจงแบบปกติมาตรฐาน โดยแบ่งเป็น 2 กลุ่ม คือกลุ่มที่มีความสัมพันธ์กับตัวแปรตามและกลุ่มที่ไม่มีความสัมพันธ์กับตัวแปรตาม

3.4.2 ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) กำหนดเป็น 0.1, 0.5 และ 0.9 โดยเมทริกซ์สหสัมพันธ์ (Correlation Matrix) มีรูปแบบ ดังนี้

$$\rho_{pxp} = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \rho_{23} & \cdots & \rho_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \rho_{p3} & \cdots & \rho_{pp} \end{pmatrix} = \begin{pmatrix} 1 & \rho & \frac{\rho}{2} & \cdots & \frac{\rho}{p-1} \\ \rho & 1 & \rho & \cdots & \frac{\rho}{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\rho}{p-1} & \frac{\rho}{p-2} & \frac{\rho}{p-3} & \cdots & 1 \end{pmatrix}$$

เมื่อ $\rho_{ij};i,j=1,2,...,p$ คือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระตัวที่ i และตัวแปรอิสระตัวที่ j

3.4.3 สร้างตัวแปรตาม $\left(\underset{\sim}{Y}^*\right)$ ให้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระและความคลาดเคลื่อน โดยมีรูปแบบดังนี้

$$\underset{\sim}{Y}^* = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{Y}^*$ เป็นเวกเตอร์ของตัวแปรตาม

X เป็นเมทริกซ์ของตัวแปรอิสระ

$\underset{\sim}{\beta} = \left(\beta_0, \beta_1, \beta_2, \ldots, \beta_p\right)'$ แทนเวกเตอร์พารามิเตอร์เมื่อมีตัวแปรอิสระ p ตัว

และ $\underset{\sim}{\varepsilon}$ แทนเวกเตอร์ความคลาดเคลื่อน

3.4.4 สร้างตัวแปรตาม (Y) ที่มีค่าเป็น 1 หรือ 0 จากค่าของ Y^* ที่สร้างได้

3.4.5 ประมาณค่าพารามิเตอร์โดยใช้ตัวแบบโพรบิทด้วยตัวประมาณภาวะน่าจะเป็นสูงสุด

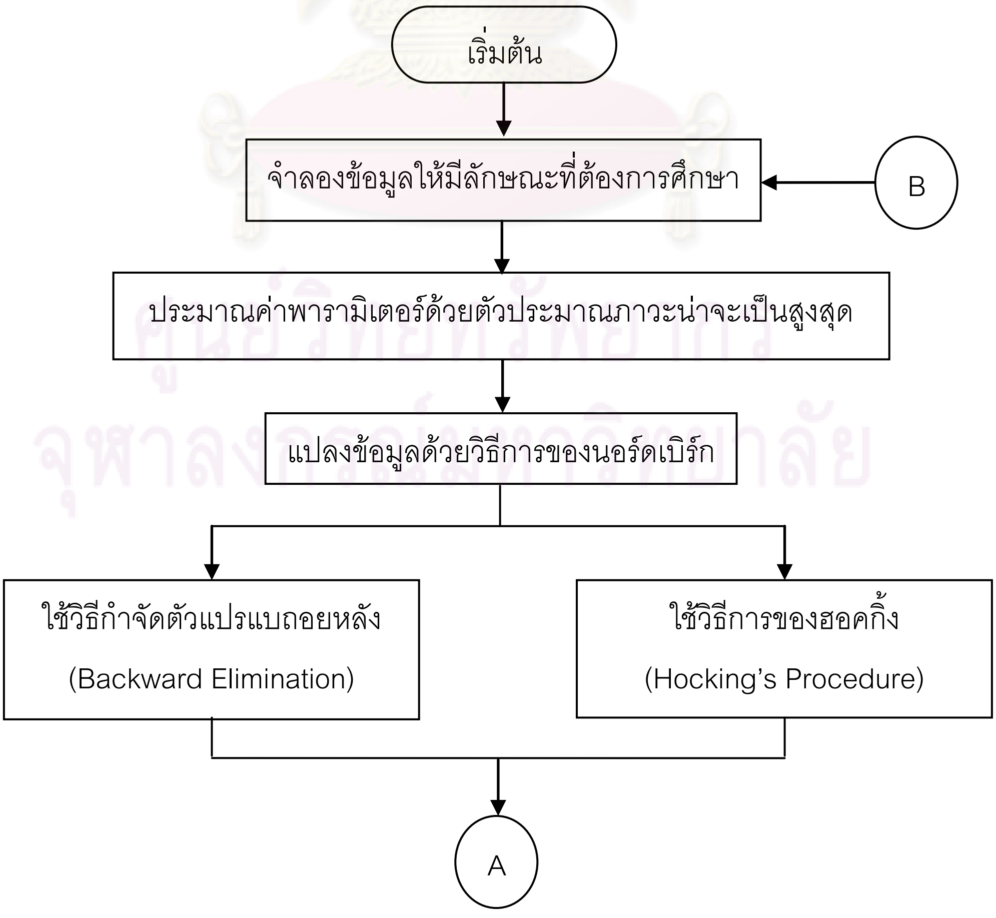
3.4.6 แปลงข้อมูลด้วยวิธีของนอร์ดเบิร์ก

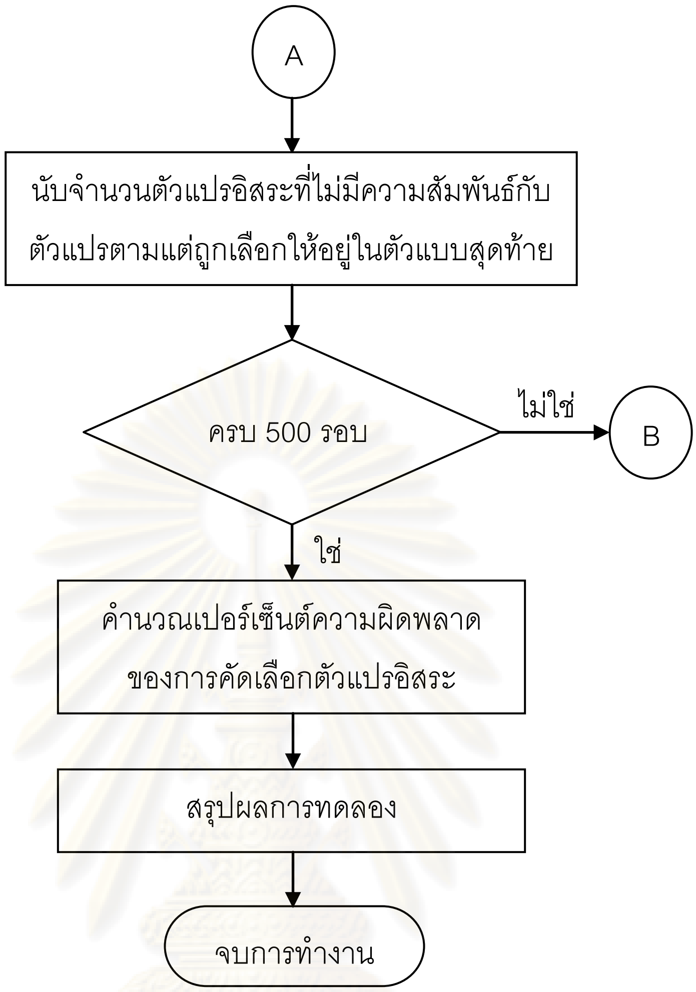
3.4.7 ใช้การคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure)

3.4.8 นับจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามแต่ถูกคัดเลือกให้อยู่ในตัวแบบในแต่ละรอบ และทำการทดลองซ้ำจำนวน 500 รอบในแต่ลสถานการณ์

3.4.9 คำนวณเปอร์เซ็นต์ความผิดพลาดของการคัดเลือกตัวแปรอิสระ

3.5 ขั้นตอนการทำงานของโปรแกรม





รูปที่ 3.1 แสดงขั้นตอนการทำงานของโปรแกรม

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระ 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure) ในตัวแบบโพรบิทโดยแปลงข้ออมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วยวิธีของนอร์ดเบิร์ก เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป โดยนำปัจจัยต่าง ๆ ที่สนใจศึกษานี้มาพิจารณาร่วมกันซึ่งแปรเปลี่ยนไปพร้อมกันโดยศึกษาภายใต้สถานการณ์ดังต่อไปนี้

1) ข้อมูลตัวแปรอิสระมีการแจกแจงแบบปกติ กำหนดให้สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ ขนาดตัวอย่าง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่
2) ข้อมูลตัวแปรอิสระมีการแจกแจงแบบปกติ กำหนดให้ขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 แต่สัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามจำนวนตัวแปรอิสระและระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่
3) ข้อมูลตัวแปรอิสระมีการแจกแจงแบบปกติ กำหนดให้ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 แต่สัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม จำนวนตัวแปรอิสระ และขนาดตัวอย่างคงที่

ในการตัดสินใจว่าปัจจัยต่างๆ ที่สนใจศึกษา มีผลกระทบต่อประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระทั้งสองวิธีการหรือไม่นั้น ใช้ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่พิจารณาจากจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามในตัวแบบสุดท้ายที่ได้ในแต่ละสถานการณ์ของการศึกษาเป็นเกณฑ์

ผู้วิจัยขอใช้สัญลักษณ์แทนความหมายต่าง ๆ เพื่อความสะดวกในการอธิบายผลการวิเคราะห์ ดังนี้

p แทนจำนวนตัวแปรอิสระ

a แทนสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม

n แทนขนาดตัวอย่าง

M แทนระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม

BE แทนวิธีการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination)

HP แทนวิธีของฮอคกิ้ง (Hocking's Procedure)

ผลการวิเคราะห์ในแต่ละสถานการณ์จะนำเสนอในรูปแบบตาราง แสดงในตารางที่ 4.1 – 4.38 โดยมีรายละเอียดดังต่อไปนี้

4.1 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ ขนาดตัวอย่าง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่ นำเสนอในตารางที่ 4.1 – 4.12

4.2 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 แต่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม จำนวนตัวแปรอิสระ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่ นำเสนอในตารางที่ 4.13 – 4.24

4.3 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 แต่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและจำนวนตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม จำนวนตัวแปรอิสระ และขนาดตัวอย่างคงที่ นำเสนอในตารางที่ 4.25 – 4.38

ผลการวิเคราะห์

4.1 กรณีที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่างและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ผลการวิเคราะห์นำเสนอในตารางที่ 4.1 – 4.12 ดังนี้

ตารางที่ 4.1 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.1	20	3	1:2	34.70%	18.80%
			2:1	34.60%	16.00%
		4	1:3	35.00%	7.73%
			2:2	38.30%	23.30%
			3:1	38.80%	14.40%
		5	1:4	37.40%	7.70%
			2:3	34.20%	0.27%
			3:2	37.40%	16.00%
			4:1	38.20%	3.40%
		6	1:5	29.36%	25.60%
			2:4	31.45%	19.40%
			3:3	33.20%	2.80%
			4:2	34.30%	6.60%
			5:1	32.80%	2.80%

จากตารางที่ 4.1 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.2 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.1	60	3	1:2	30.10%	3.80%
			2:1	29.40%	14.00%
		4	1:3	31.60%	4.93%
			2:2	29.80%	11.00%
			3:1	31.00%	10.40%
		5	1:4	31.85%	15.00%
			2:3	30.53%	17.60%
			3:2	34.90%	22.10%
			4:1	35.60%	19.00%
		6	1:5	31.56%	48.60%
			2:4	34.95%	21.80%
			3:3	33.07%	21.00%
			4:2	31.80%	15.30%
			5:1	34.60%	5.50%

จากตารางที่ 4.2 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป เมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.3 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 100, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.1	100	3	1:2	30.50%	15.20%
			2:1	27.40%	13.60%
		4	1:3	29.53%	4.13%
			2:2	29.80%	0.80%
			3:1	30.80%	8.00%
		5	1:4	27.45%	19.25%
			2:3	29.87%	18.13%
			3:2	31.70%	21.00%
			4:1	32.20%	25.40%
		6	1:5	29.88%	37.10%
			2:4	32.45%	22.80%
			3:3	32.00%	29.70%
			4:2	29.80%	21.30%
			5:1	35.40%	14.20%

จากตารางที่ 4.3 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, n = 100, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป เมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.4 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 200, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.1	200	3	1:2	31.40%	13.90%
			2:1	26.80%	19.60%
		4	1:3	27.60%	10.93%
			2:2	28.50%	14.30%
			3:1	30.80%	19.20%
		5	1:4	27.40%	16.45%
			2:3	28.33%	19.73%
			3:2	29.90%	23.00%
			4:1	33.20%	14.60%
		6	1:5	28.12%	49.20%
			2:4	27.35%	41.30%
			3:3	28.40%	33.20%
			4:2	27.30%	23.80%
			5:1	30.60%	12.00%

จากตารางที่ 4.4 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, n = 200, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5, 2:4, 3:3 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.5 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.5	20	3	1:2	36.30%	20.60%
			2:1	33.20%	17.40%
		4	1:3	38.87%	8.13%
			2:2	35.70%	15.80%
			3:1	40.00%	10.00%
		5	1:4	36.50%	6.05%
			2:3	37.07%	4.33%
			3:2	35.40%	11.60%
			4:1	37.20%	4.00%
		6	1:5	31.16%	12.00%
			2:4	32.00%	8.20%
			3:3	33.67%	11.00%
			4:2	31.20%	4.50%
			5:1	34.40%	1.40%

จากตารางที่ 4.5 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.6 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.5	60	3	1:2	30.60%	11.70%
			2:1	30.60%	24.20%
		4	1:3	32.33%	16.53%
			2:2	32.00%	26.00%
			3:1	30.00%	24.20%
		5	1:4	32.80%	11.30%
			2:3	32.80%	17.00%
			3:2	34.60%	1.20%
			4:1	34.40%	24.40%
		6	1:5	33.60%	30.30%
			2:4	35.05%	35.70%
			3:3	36.93%	34.10%
			4:2	34.60%	15.00%
			5:1	39.20%	12.10%

จากตารางที่ 4.6 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป พบว่าเมื่อ p= 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 2:4 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.7 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 100, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.5	100	3	1:2	30.50%	22.60%
			2:1	28.40%	21.20%
		4	1:3	28.07%	15.13%
			2:2	29.40%	7.10%
			3:1	28.80%	15.20%
		5	1:4	28.55%	2.45%
			2:3	31.00%	17.00%
			3:2	30.00%	19.60%
			4:1	32.00%	16.60%
		6	1:5	29.00%	42.30%
			2:4	31.50%	16.40%
			3:3	30.80%	30.70%
			4:2	30.60%	18.00%
			5:1	32.60%	6.50%

จากตารางที่ 4.7 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, n = 100, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ สูงกว่าวิธี BE

ตารางที่ 4.8 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 200, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.5	200	3	1:2	25.10%	15.80%
			2:1	24.80%	15.20%
		4	1:3	26.93%	20.67%
			2:2	30.80%	19.00%
			3:1	24.40%	17.60%
		5	1:4	27.70%	23.15%
			2:3	28.07%	12.93%
			3:2	29.50%	23.70%
			4:1	32.20%	23.60%
		6	1:5	28.96%	35.30%
			2:4	29.65%	45.30%
			3:3	28.47%	26.80%
			4:2	29.80%	23.80%
			5:1	30.60%	11.10%

จากตารางที่ 4.8 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, n = 200, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5, 2:4 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.9 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.9	20	3	1:2	37.00%	4.80%
			2:1	36.60%	9.60%
		4	1:3	38.60%	16.60%
			2:2	36.00%	18.60%
			3:1	34.20%	11.20%
		5	1:4	36.55%	4.55%
			2:3	37.33%	2.67%
			3:2	33.70%	10.80%
			4:1	38.20%	13.20%
		6	1:5	30.72%	1.80%
			2:4	32.50%	13.60%
			3:3	30.47%	7.60%
			4:2	34.50%	0.70%
			5:1	32.00%	0.10%

จากตารางที่ 4.9 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, n = 20, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.10 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.9	60	3	1:2	32.60%	23.00%
			2:1	31.00%	24.00%
		4	1:3	31.87%	20.20%
			2:2	30.00%	23.90%
			3:1	33.40%	3.60%
		5	1:4	32.05%	15.45%
			2:3	32.73%	4.47%
			3:2	35.00%	23.30%
			4:1	34.80%	20.40%
		6	1:5	33.72%	40.70%
			2:4	33.25%	24.90%
			3:3	35.67%	32.60%
			4:2	35.00%	10.70%
			5:1	32.40%	7.90%

จากตารางที่ 4.10 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, n = 60, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

ตารางที่ 4.11 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 100, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.9	100	3	1:2	30.20%	23.10%
			2:1	30.40%	7.60%
		4	1:3	28.60%	16.33%
			2:2	26.50%	19.60%
			3:1	30.20%	18.20%
		5	1:4	29.90%	10.65%
			2:3	29.20%	17.07%
			3:2	32.30%	22.00%
			4:1	31.80%	23.80%
		6	1:5	30.72%	39.20%
			2:4	31.40%	28.30%
			3:3	32.07%	28.60%
			4:2	28.40%	22.30%
			5:1	32.80%	11.10%

จากตารางที่ 4.11 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, n = 100, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

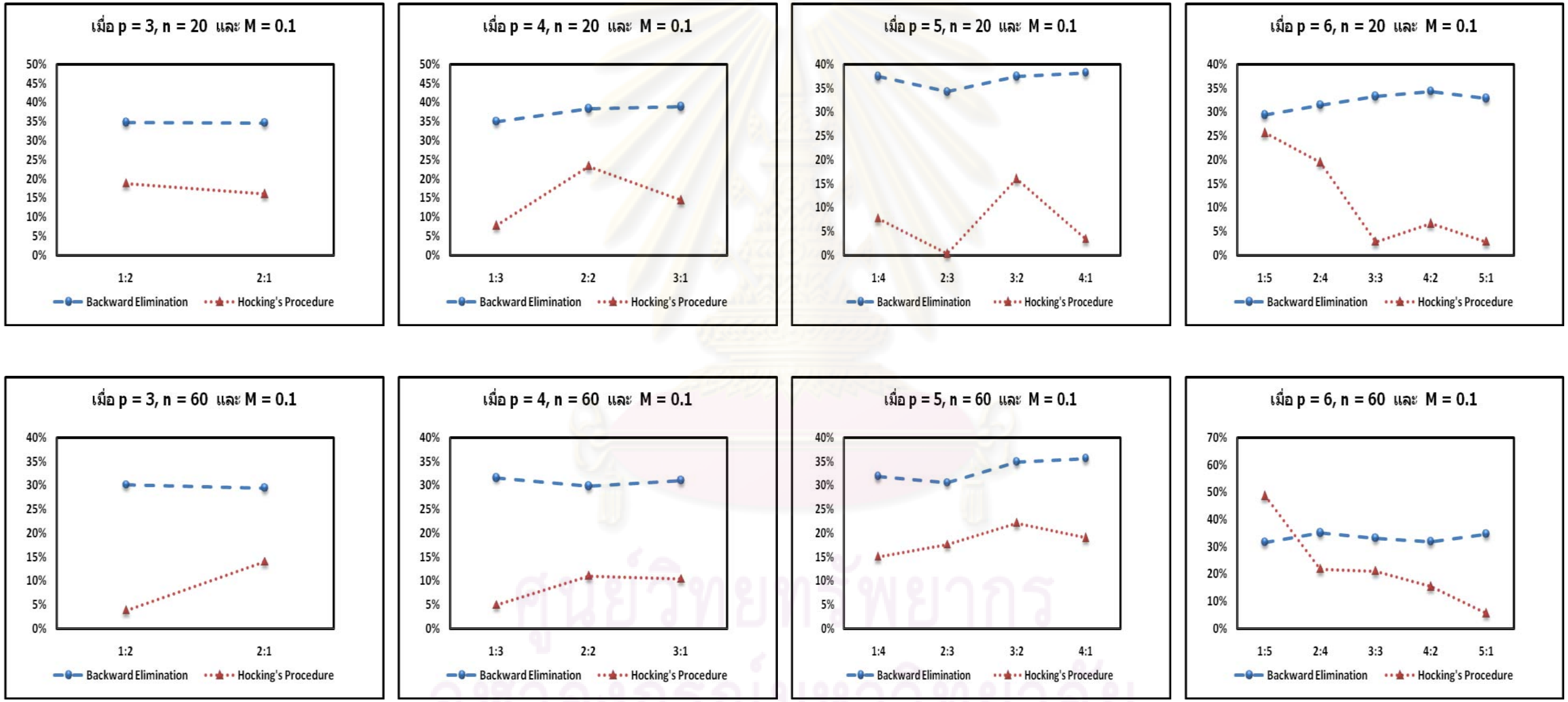
ตารางที่ 4.12 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 200, p = 3, 4, 5, 6 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป

M	n	p	a	Percentage of mean error	
				BE	HP
0.9	200	3	1:2	24.10%	18.40%
			2:1	27.60%	14.60%
		4	1:3	26.87%	19.67%
			2:2	28.00%	17.20%
			3:1	25.60%	11.00%
		5	1:4	26.70%	23.65%
			2:3	30.40%	16.60%
			3:2	30.20%	15.00%
			4:1	29.80%	16.80%
		6	1:5	27.40%	33.80%
			2:4	28.60%	43.20%
			3:3	27.93%	34.50%
			4:2	31.40%	19.90%
			5:1	30.20%	7.90%

จากตารางที่ 4.12 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, n = 200, p = 3, 4, 5, 6 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม เปลี่ยนแปลงไป พบว่าเมื่อ p = 3, 4, 5 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ p = 6 และ a = 1:5, 2:4, 3:3 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระสูงกว่าวิธี BE

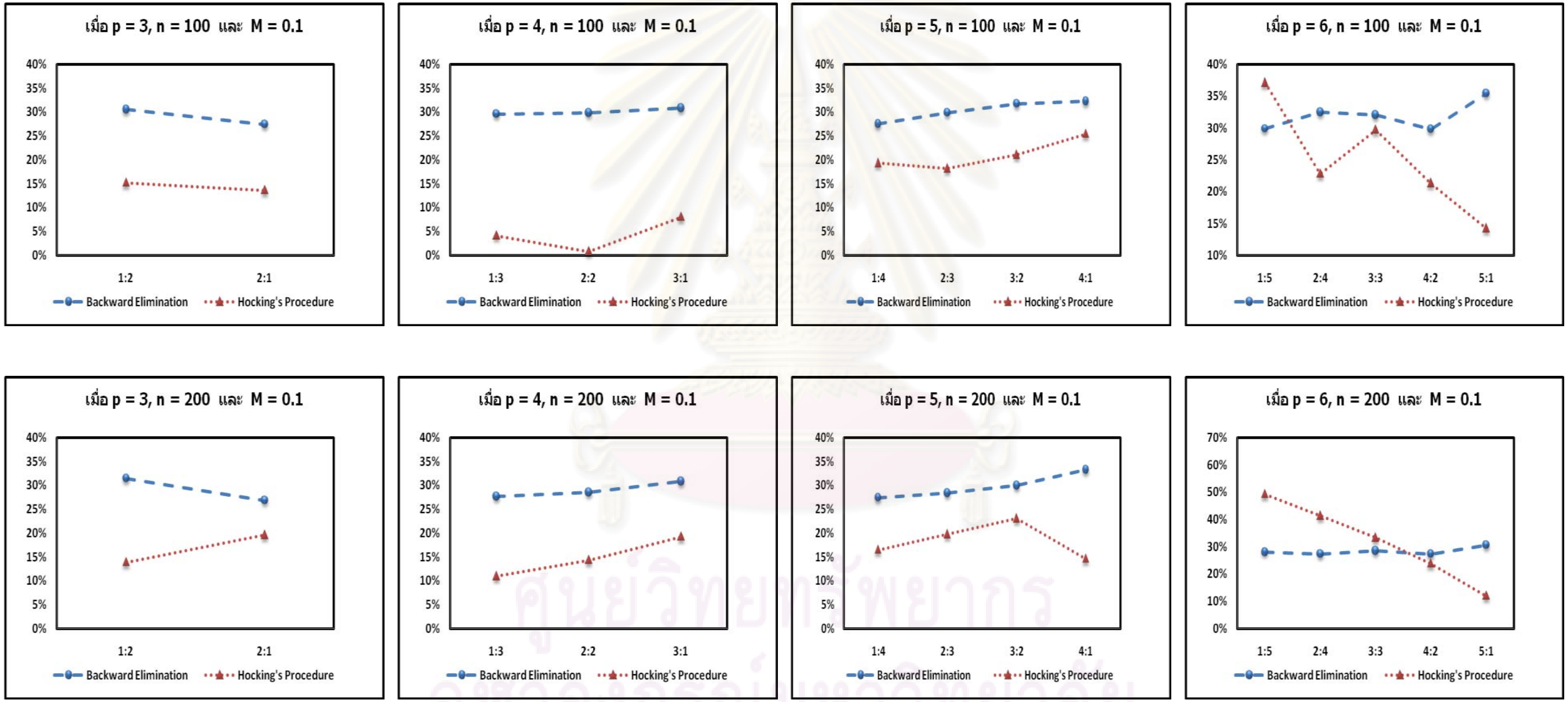
จากตารางที่ 4.1 – 4.12 สรุปได้ว่า ถ้าระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอยู่ในระดับต่ำ เมื่อตัวแปรอิสระมีจำนวนเพิ่มมากขึ้น เนื่องจากปัจจัยต่างๆ มีความสัมพันธ์กันน้อย ทำให้ค่าความแปรปรวนลดลงและประมาณค่าพารามิเตอร์ได้ดีขึ้น ส่งผลให้การคัดเลือกตัวแปรอิสระมีความผิดพลาดลดลง แต่ในกรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าสูงหรือเกิดปัญหาพหุสัมพันธ์มากนั้น เมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นก็จะทำให้ความแปรปรวนเพิ่มสูงขึ้นด้วย ทำให้การประมาณค่าพารามิเตอร์ผิดพลาดมากขึ้น ส่งผลให้การคัดเลือกตัวแปรอิสระมีความผิดพลาดเพิ่มขึ้น สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป แต่ปัจจัยอื่นคงที่ นำเสนอในรูปที่ 4.1 – 4.6 ได้ดังนี้

รูปที่ 4.1 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 20, 60 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6



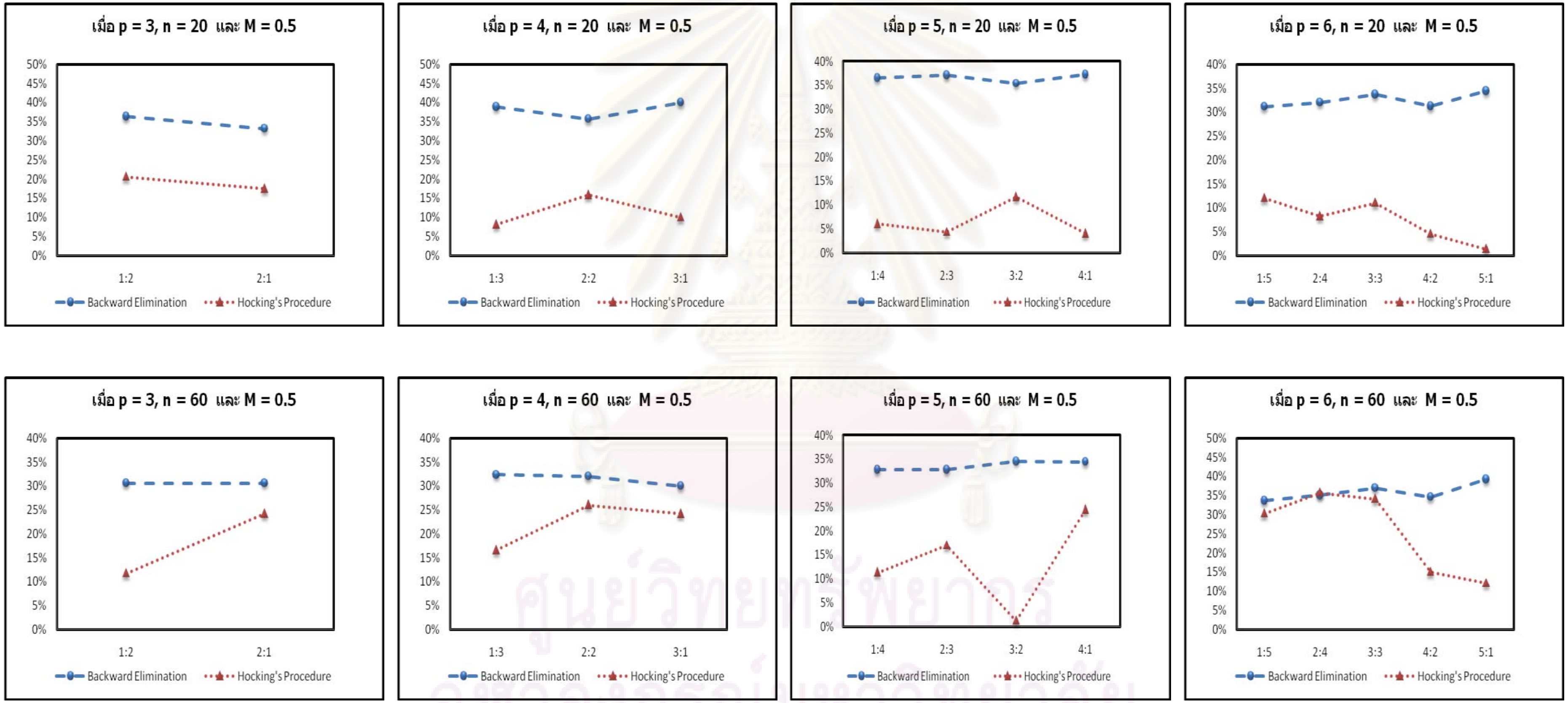
35

รูปที่ 4.2 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, n = 100, 200 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6

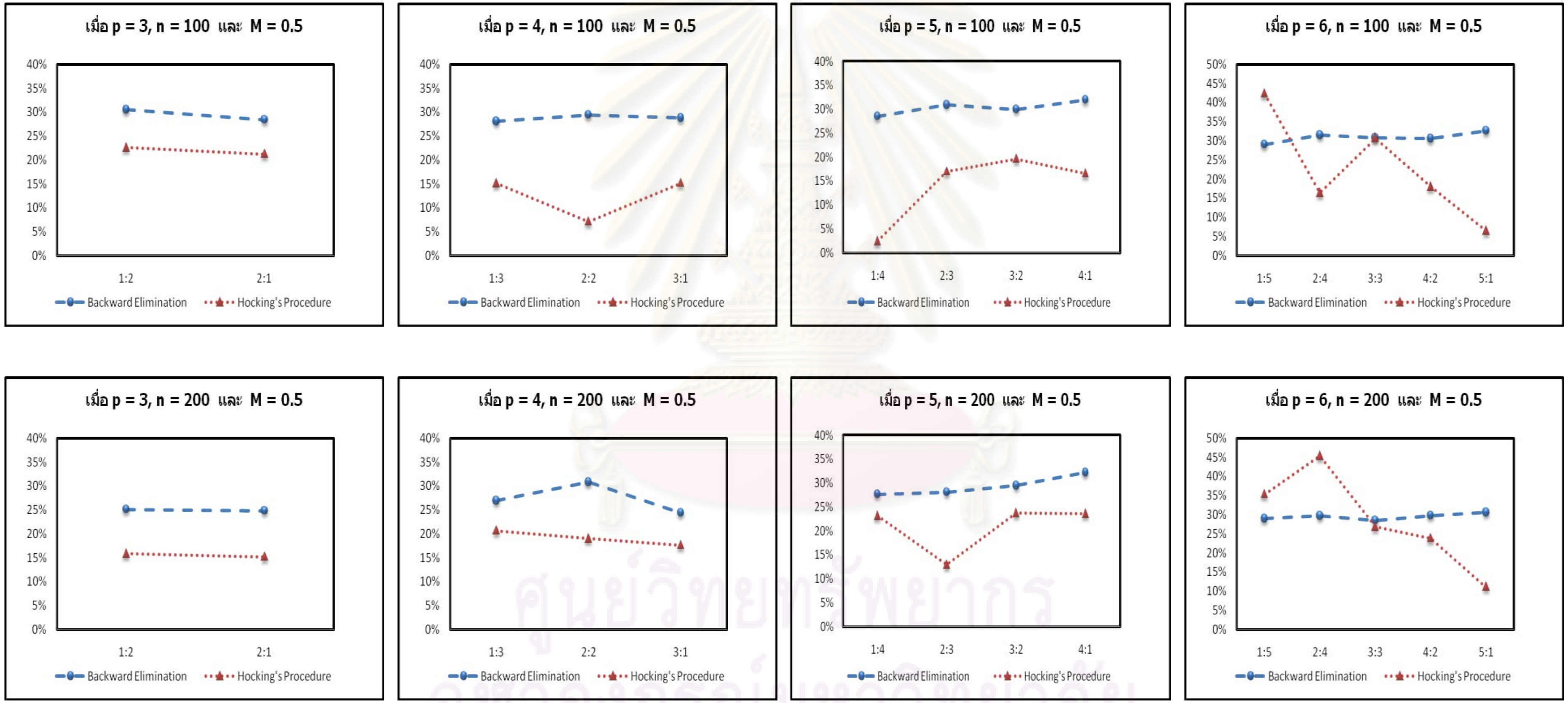


36

รูปที่ 4.3 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 20, 60 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6

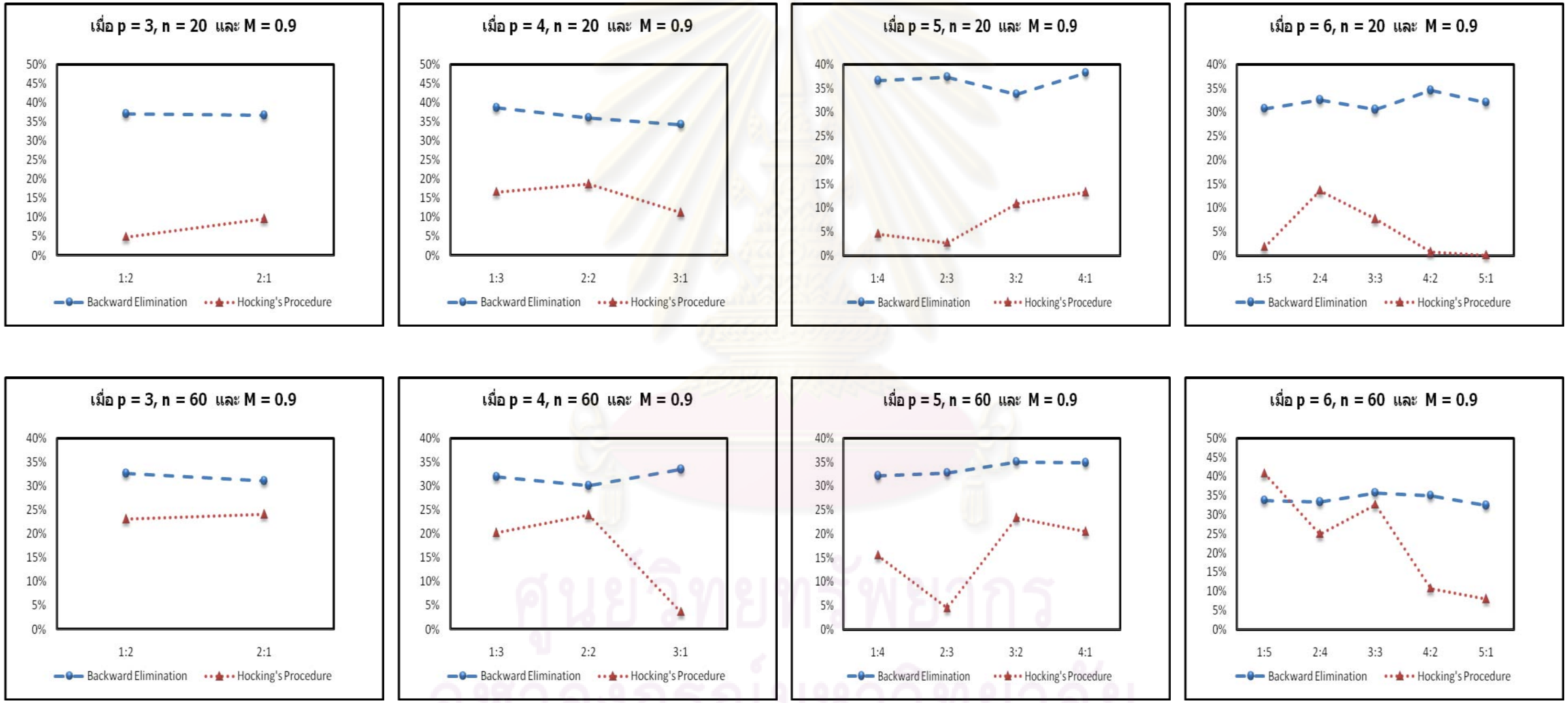


รูปที่ 4.4 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, n = 100, 200 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6



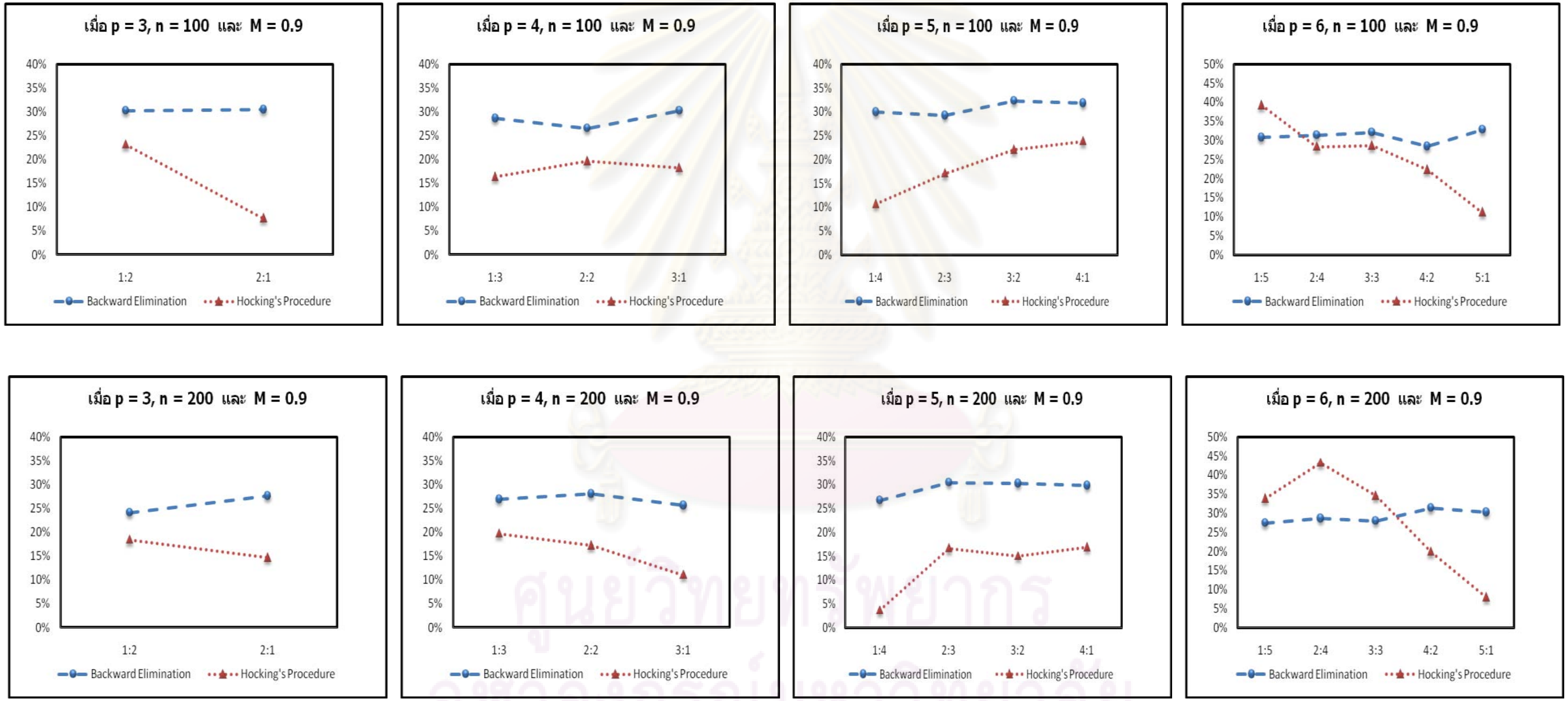
38

รูปที่ 4.5 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 20, 60 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6



39

รูปที่ 4.6 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, n = 100, 200 และสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป โดยแยกตาม p = 3, 4, 5 และ 6



4.2 กรณีที่ขนาดตัวอย่างเปลี่ยนแปลงไป เมื่อสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความ สัมพันธ์กับตัวแปรตาม จำนวนตัวแปรอิสระและระดับความสัมพันธ์ระหว่างตัวแปรอิสระคงที่

ผลการวิเคราะห์นำเสนอในตารางที่ 4.13 – 4.24 ดังนี้

ตารางที่ 4.13 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 3, a = 1:2, 2:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.1	3	1:2	20	34.70%	18.80%
			60	30.10%	3.80%
			100	30.50%	15.20%
			200	31.40%	13.90%
		2:1	20	34.60%	16.00%
			60	29.40%	14.00%
			100	27.40%	13.60%
			200	26.80%	19.60%

จากตารางที่ 4.13 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, p = 3, a = 1:2, 2:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.14 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.1	4	1:3	20	35.00%	7.73%
			60	31.60%	4.93%
			100	29.53%	4.13%
			200	27.60%	10.93%
		2:2	20	38.30%	23.30%
			60	29.80%	11.00%
			100	29.80%	0.80%
			200	28.50%	14.30%
		3:1	20	38.80%	14.40%
			60	31.00%	10.40%
			100	30.80%	8.00%
			200	30.80%	19.20%

จากตารางที่ 4.14 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.15 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.1	5	1:4	20	37.40%	7.70%
			60	31.85%	15.00%
			100	27.45%	19.25%
			200	27.40%	16.45%
		2:3	20	34.20%	0.27%
			60	30.53%	17.60%
			100	29.87%	18.13%
			200	28.33%	19.73%
		3:2	20	37.40%	16.00%
			60	34.90%	22.10%
			100	31.70%	21.00%
			200	29.90%	23.00%
		4:1	20	38.20%	3.40%
			60	35.60%	19.00%
			100	32.20%	25.40%
			200	33.20%	14.60%

จากตารางที่ 4.15 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.16 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.1	6	1:5	20	29.36%	25.60%
			60	31.56%	48.60%
			100	29.88%	37.10%
			200	28.12%	49.20%
		2:4	20	31.45%	19.40%
			60	34.95%	21.80%
			100	32.45%	22.80%
			200	27.35%	41.30%
		3:3	20	33.20%	2.80%
			60	33.07%	21.00%
			100	32.00%	29.70%
			200	28.40%	33.20%
		4:2	20	34.30%	6.60%
			60	31.80%	15.30%
			100	29.80%	21.30%
			200	27.30%	23.80%
		5:1	20	32.80%	2.80%
			60	34.60%	5.50%
			100	35.40%	14.20%
			200	30.60%	12.00%

จากตารางที่ 4.16 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.1, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าเมื่อ n = 60, 100, a = 1:5 และ n= 200, a = 1:5, 2:4, 3:3

วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.17 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 3, a = 1:2, 2:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.5	3	1:2	20	36.30%	20.60%
			60	30.60%	11.70%
			100	30.50%	22.60%
			200	25.10%	15.80%
		2:1	20	33.20%	17.40%
			60	30.60%	24.20%
			100	28.40%	21.20%
			200	24.80%	15.20%

จากตารางที่ 4.17 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, p = 3, a = 1:2, 2:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีค่าลดลง

ตารางที่ 4.18 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.5	4	1:3	20	38.87%	8.13%
			60	32.33%	16.53%
			100	28.07%	15.13%
			200	26.93%	20.67%
		2:2	20	35.70%	15.80%
			60	32.00%	26.00%
			100	29.40%	7.10%
			200	30.80%	19.00%
		3:1	20	40.00%	10.00%
			60	30.00%	24.20%
			100	28.80%	15.20%
			200	24.40%	17.60%

จากตารางที่ 4.18 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.19 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.5	5	1:4	20	36.50%	6.05%
			60	32.80%	11.30%
			100	28.55%	2.45%
			200	27.70%	23.15%
		2:3	20	37.07%	4.33%
			60	32.80%	17.00%
			100	31.00%	17.00%
			200	28.07%	12.93%
		3:2	20	35.40%	11.60%
			60	34.60%	1.20%
			100	30.00%	19.60%
			200	29.50%	23.70%
		4:1	20	37.20%	4.00%
			60	34.40%	24.40%
			100	32.00%	16.60%
			200	32.20%	23.60%

จากตารางที่ 4.19 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.20 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.5	6	1:5	20	31.16%	12.00%
			60	33.60%	30.30%
			100	29.00%	42.30%
			200	28.96%	35.30%
		2:4	20	32.00%	8.20%
			60	35.05%	35.70%
			100	31.50%	16.40%
			200	29.65%	45.30%
		3:3	20	33.67%	11.00%
			60	36.93%	34.10%
			100	30.80%	30.70%
			200	28.47%	26.80%
		4:2	20	31.20%	4.50%
			60	34.60%	15.00%
			100	30.60%	18.00%
			200	29.80%	23.80%
		5:1	20	34.40%	1.40%
			60	39.20%	12.10%
			100	32.60%	6.50%
			200	30.60%	11.10%

จากตารางที่ 4.20 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.5, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่างเปลี่ยนแปลงไป

เป็น 20, 60, 100 และ 200 พบว่าเมื่อ n = 60, a = 2:4 n=100, a= 1:5 และ n= 200, a = 1:5, 2:4 วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP

ตารางที่ 4.21 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 3, a = 1:2, 2:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.9	3	1:2	20	37.00%	4.80%
			60	32.60%	23.00%
			100	30.20%	23.10%
			200	24.10%	18.40%
		2:1	20	36.60%	9.60%
			60	31.00%	24.00%
			100	30.40%	7.60%
			200	27.60%	14.60%

จากตารางที่ 4.21 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, p = 3, a = 1:2, 2:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีค่าลดลง

ตารางที่ 4.22 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.9	4	1:3	20	38.60%	16.60%
			60	31.87%	20.20%
			100	28.60%	16.33%
			200	26.87%	19.67%
		2:2	20	36.00%	18.60%
			60	30.00%	23.90%
			100	26.50%	19.60%
			200	28.00%	17.20%
		3:1	20	34.20%	11.20%
			60	33.40%	3.60%
			100	30.20%	18.20%
			200	25.60%	11.00%

จากตารางที่ 4.22 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, p = 4, a = 1:3, 2:2, 3:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

ตารางที่ 4.23 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.9	5	1:4	20	36.55%	4.55%
			60	32.05%	15.45%
			100	29.90%	10.65%
			200	26.70%	23.65%
		2:3	20	37.33%	2.67%
			60	32.73%	4.47%
			100	29.20%	17.07%
			200	30.40%	16.60%
		3:2	20	33.70%	10.80%
			60	35.00%	23.30%
			100	32.30%	22.00%
			200	30.20%	15.00%
		4:1	20	38.20%	13.20%
			60	34.80%	20.40%
			100	31.80%	23.80%
			200	29.80%	16.80%

จากตารางที่ 4.23 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, p = 5, a = 1:4, 2:3, 3:2, 4:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มลดลง

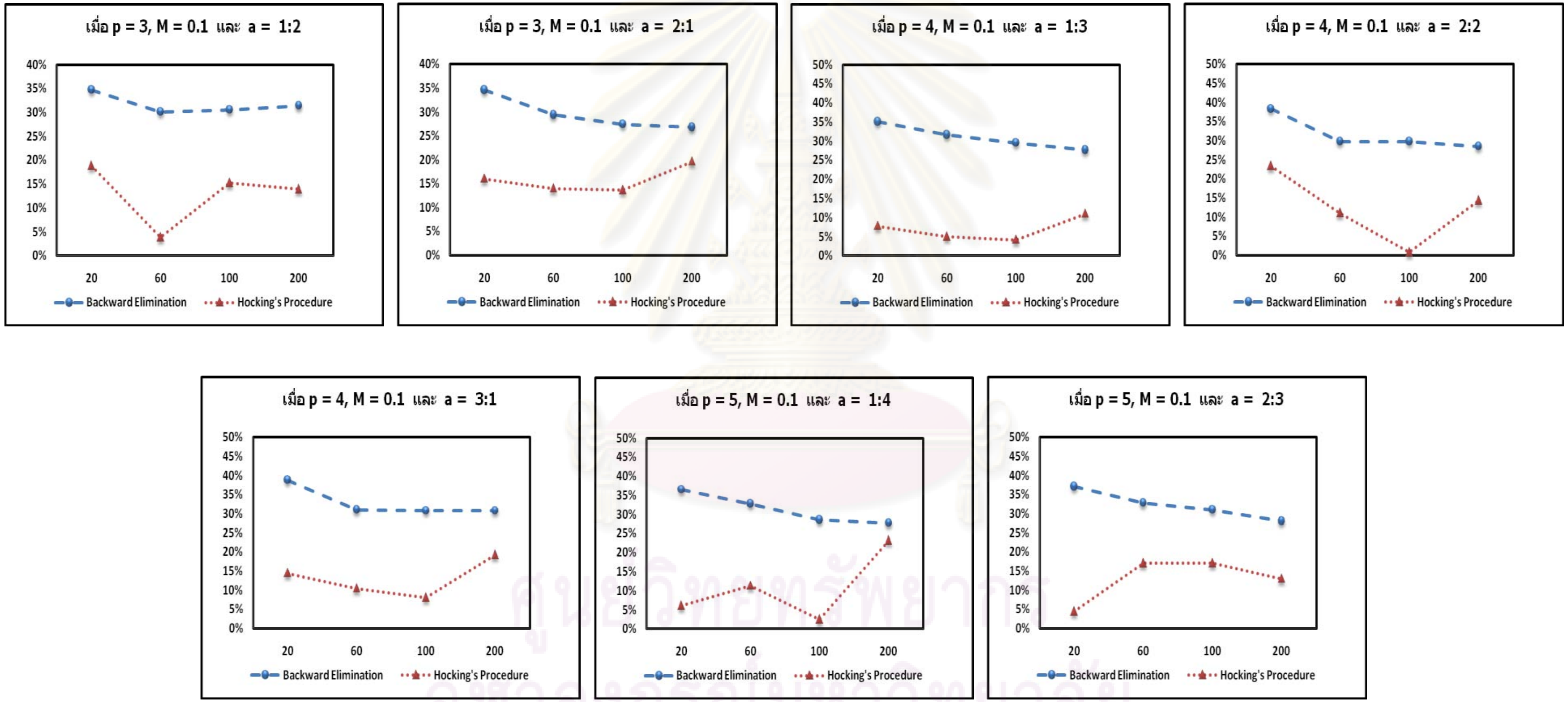
ตารางที่ 4.24 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200

M	p	a	n	Percentage of mean error	
				BE	HP
0.9	6	1:5	20	30.72%	1.80%
			60	33.72%	40.70%
			100	30.72%	39.20%
			200	27.40%	33.80%
		2:4	20	32.50%	13.60%
			60	33.25%	24.90%
			100	31.40%	28.30%
			200	28.60%	43.20%
		3:3	20	30.47%	7.60%
			60	35.67%	32.60%
			100	32.07%	28.60%
			200	27.93%	34.50%
		4:2	20	34.50%	0.70%
			60	35.00%	10.70%
			100	28.40%	22.30%
			200	31.40%	19.90%
		5:1	20	32.00%	0.10%
			60	32.40%	7.90%
			100	32.80%	11.10%
			200	30.20%	7.90%

จากตารางที่ 4.24 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ M = 0.9, p = 6, a = 1:5, 2:4, 3:3, 4:2, 5:1 และขนาดตัวอย่างเปลี่ยนแปลงไปเป็น 20, 60, 100 และ 200 พบว่าเมื่อ n = 60, 100 , a = 1:5 และ n= 200, a = 1:5, 2:4, 3:3 วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP

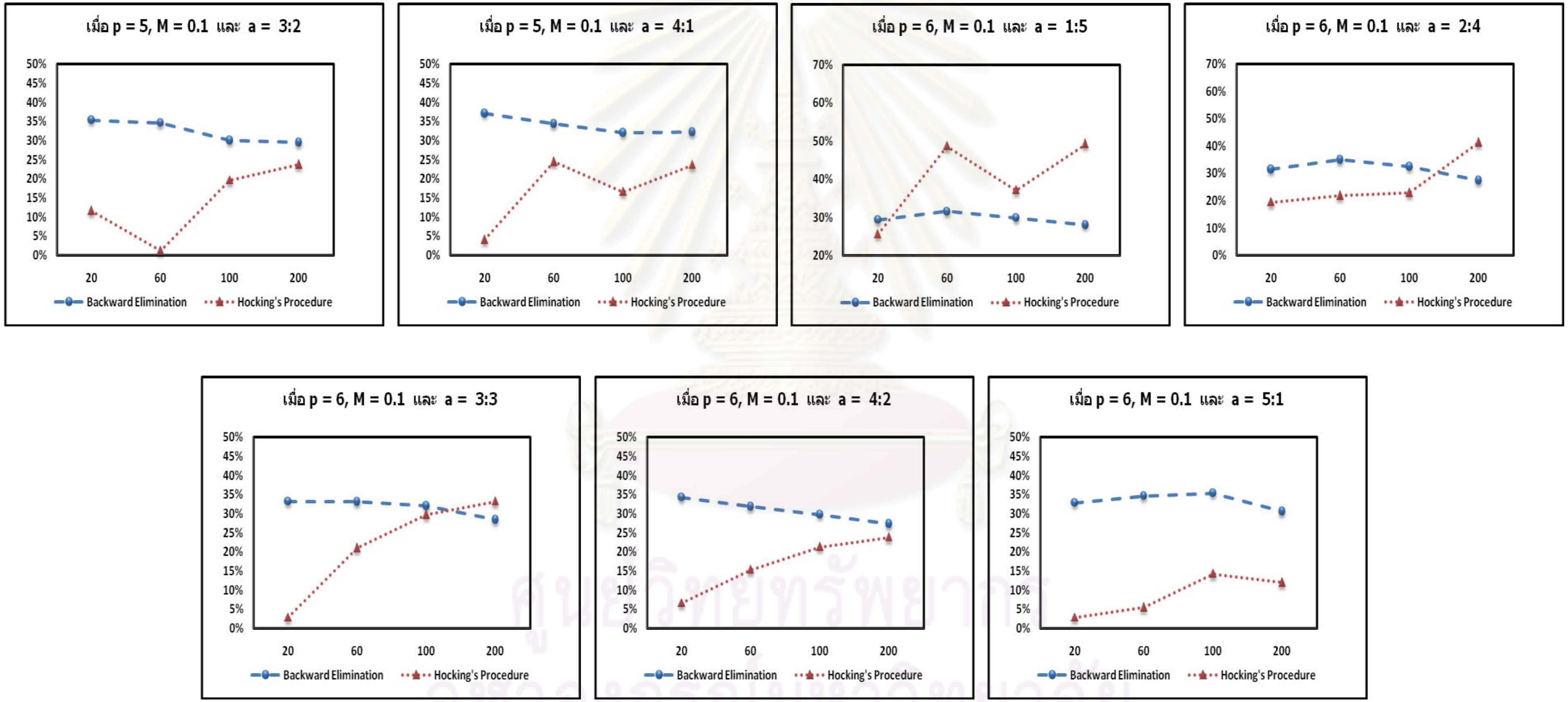
จากตารางที่ 4.13 – 4.24 สรุปได้ว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากมีข้อมูลเพิ่มมากขึ้นจะสามารถสร้างตัวแบบที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้ดีขึ้นทำให้การประมาณค่าพารามิเตอร์มีความถูกต้องมากขึ้น ส่งผลให้การคัดเลือกตัวแปรอิสระมีความผิดพลาดลดลง สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระเมื่อขนาดตัวอย่างเปลี่ยนแปลงไป แต่ปัจจัยอื่นคงที่ นำเสนอในรูปที่ 4.7 – 4.9 ได้ดังนี้

รูปที่ 4.7 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)



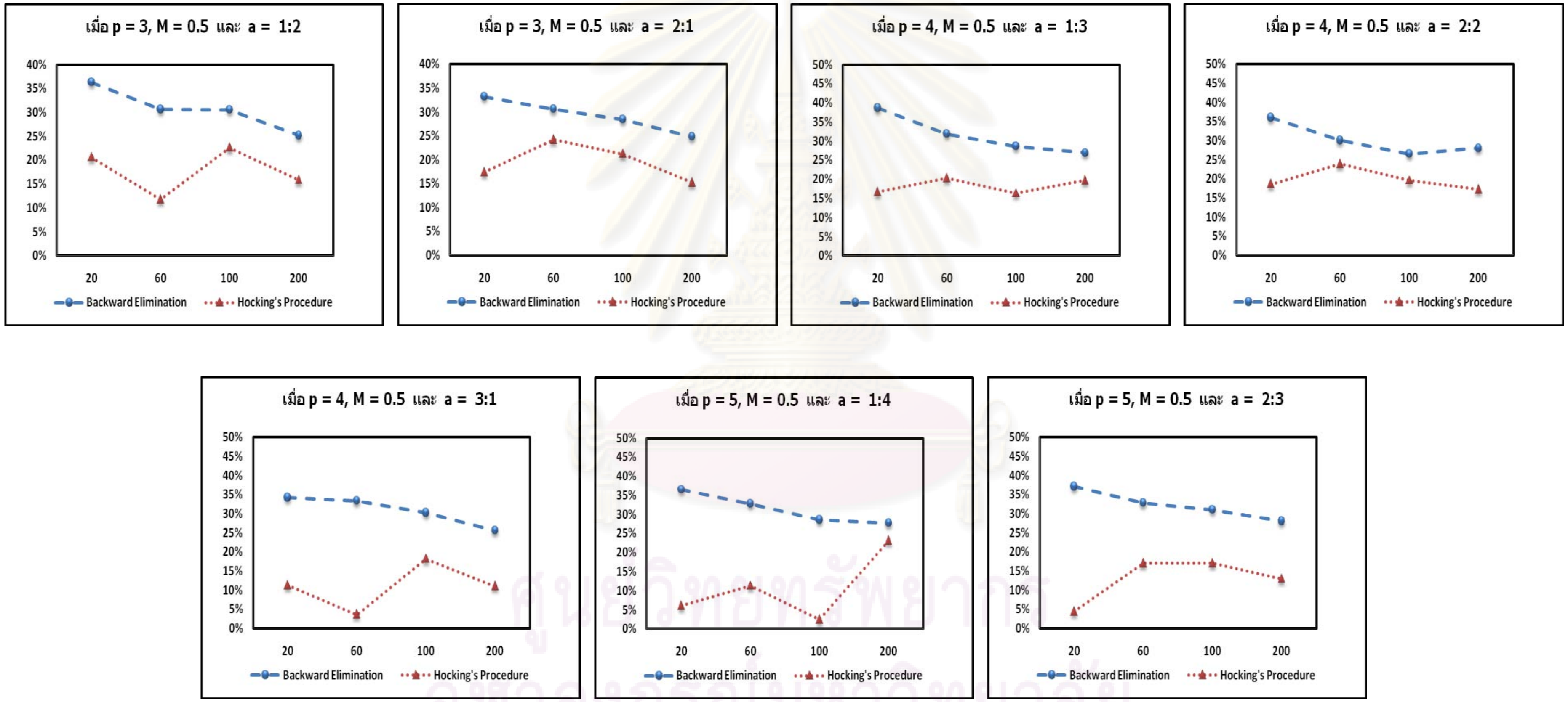
54

รูปที่ 4.7(ต่อ) แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.1, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)



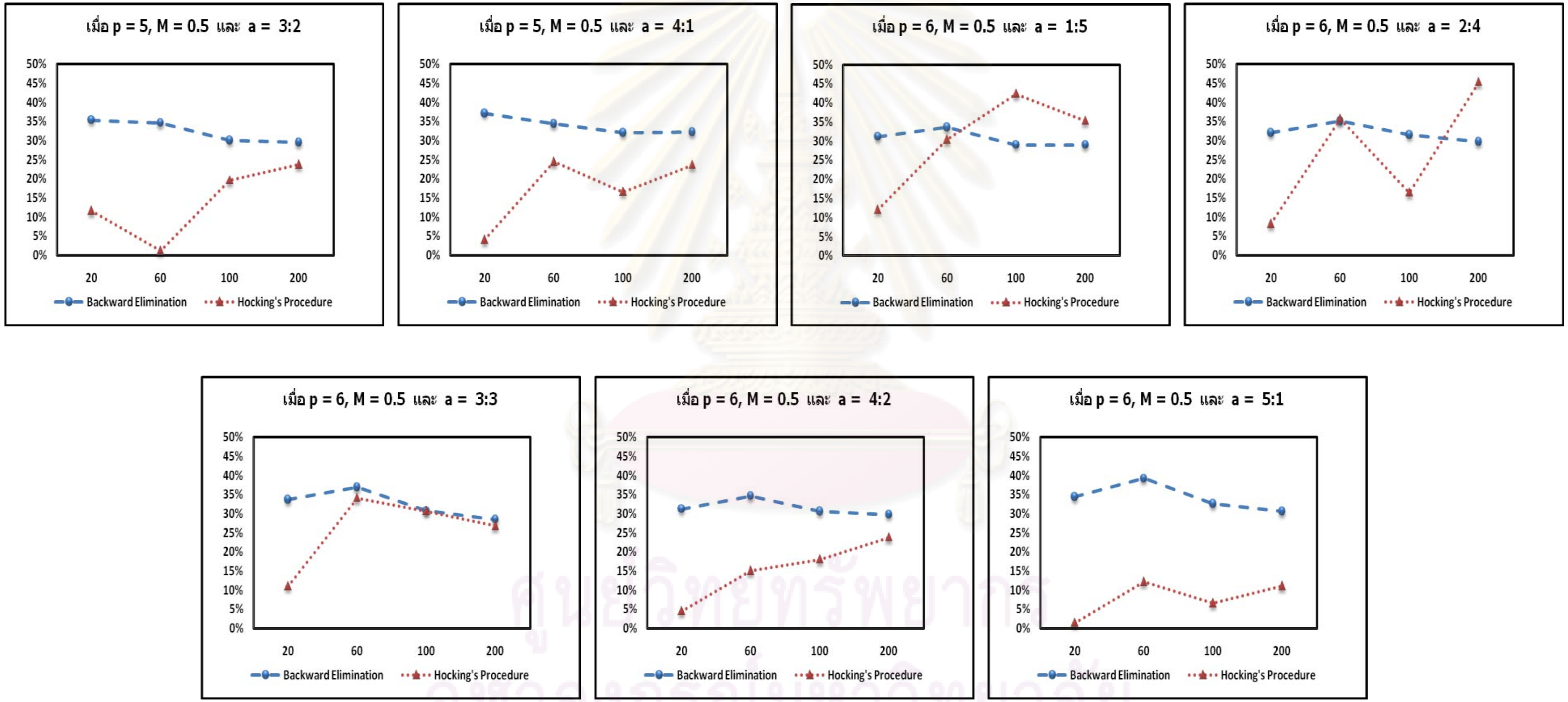
55

รูปที่ 4.8 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)

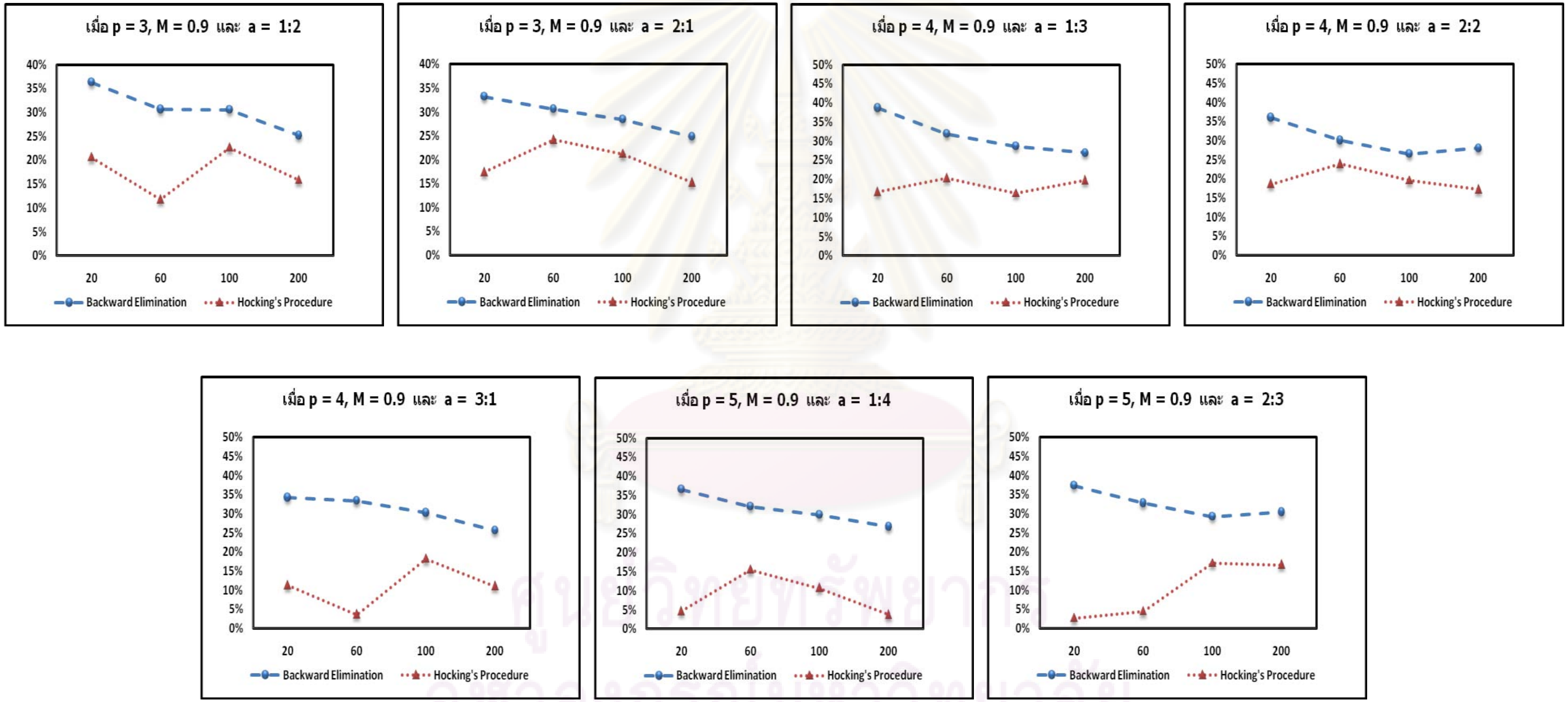


56

รูปที่ 4.8(ต่อ) แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.5, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)

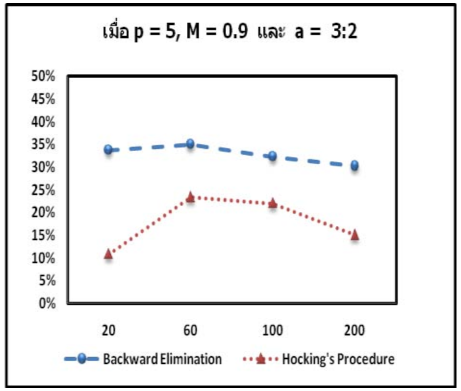


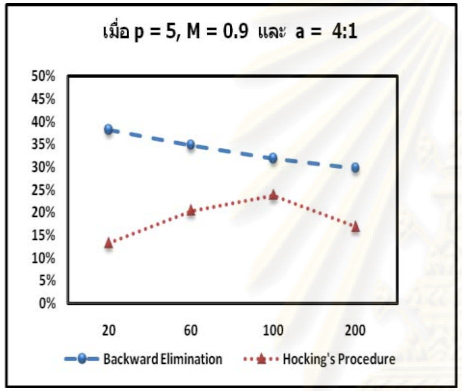
รูปที่ 4.9 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)

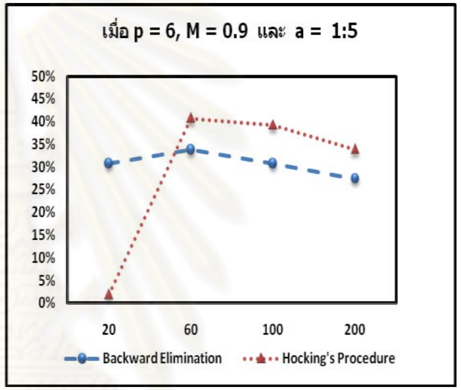


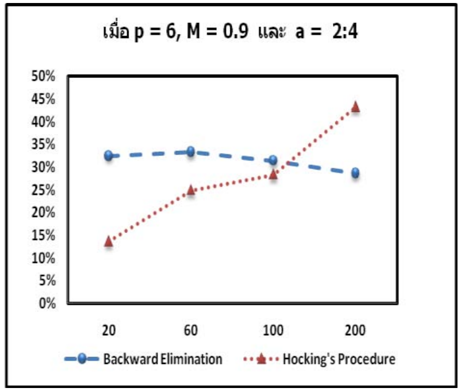
58

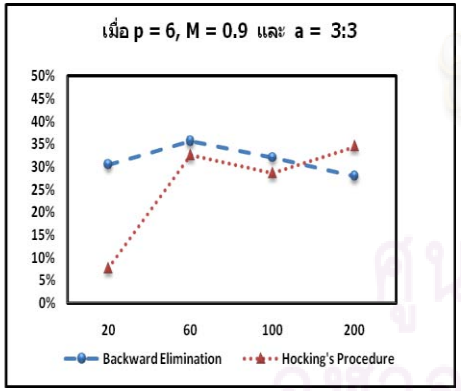
รูปที่ 4.9(ต่อ) แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ M = 0.9, p = 3, 4, 5 และ 6 และขนาดตัวอย่าง (n) เปลี่ยนแปลงไป โดยแยกตามสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a)

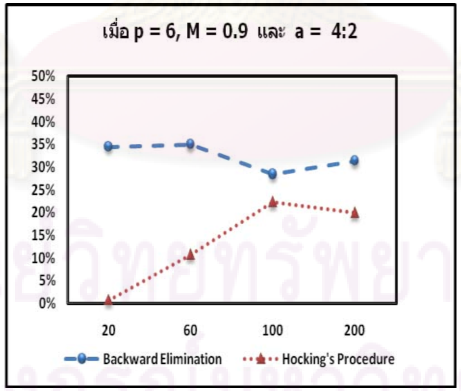


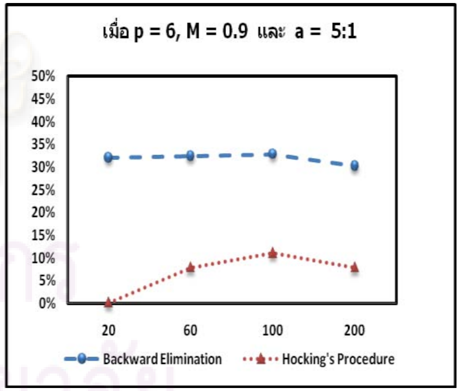












4.3 กรณีระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป เมื่อสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม จำนวนตัวแปรอิสระ และขนาดตัวอย่างคงที่

ผลการวิเคราะห์นำเสนอในตารางที่ 4.25 – 4.38 ดังนี้

ตารางที่ 4.25 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 3, a = 1:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์ กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
3	1:2	20	0.1	34.70%	18.80%
			0.5	36.30%	20.60%
			0.9	37.00%	4.80%
		60	0.1	30.10%	3.80%
			0.5	30.60%	11.70%
			0.9	32.60%	23.00%
		100	0.1	30.50%	15.20%
			0.5	30.50%	22.60%
			0.9	30.20%	23.10%
		200	0.1	31.40%	13.90%
			0.5	25.10%	15.80%
			0.9	24.10%	18.40%

จากตารางที่ 4.25 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 3, a = 1:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE เมื่อ n = 20, 60, 100 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระเพิ่มขึ้นด้วย เมื่อ n = 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระลดลง แต่วิธี HP มีค่าเพิ่มสูงขึ้น

ตารางที่ 4.26 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 3, a = 2:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
3	2:1	20	0.1	34.60%	16.00%
			0.5	33.20%	17.40%
			0.9	36.60%	9.60%
		60	0.1	29.40%	14.00%
			0.5	30.60%	24.20%
			0.9	31.00%	24.00%
		100	0.1	27.40%	13.60%
			0.5	28.40%	21.20%
			0.9	30.40%	7.60%
		200	0.1	26.80%	19.60%
			0.5	24.80%	15.20%
			0.9	27.60%	14.60%

จากตารางที่ 4.26 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 3, a = 2: 1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มเพิ่มขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น

ตารางที่ 4.27 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 4, a = 1:3, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
4	1:3	20	0.1	35.00%	7.73%
			0.5	38.87%	8.13%
			0.9	38.60%	16.60%
		60	0.1	31.60%	4.93%
			0.5	32.33%	16.53%
			0.9	31.87%	20.20%
		100	0.1	29.53%	4.13%
			0.5	28.07%	15.13%
			0.9	28.60%	16.33%
		200	0.1	27.60%	10.93%
			0.5	26.93%	20.67%
			0.9	26.87%	19.67%

จากตารางที่ 4.27 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 4, a = 1:3, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี HP มีแนวโน้มเพิ่มขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น

ตารางที่ 4.28 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 4, a = 2:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
4	2:2	20	0.1	38.30%	23.30%
			0.5	35.70%	15.80%
			0.9	36.00%	18.60%
		60	0.1	29.80%	11.00%
			0.5	32.00%	26.00%
			0.9	30.00%	23.90%
		100	0.1	29.80%	0.80%
			0.5	29.40%	7.10%
			0.9	26.50%	19.60%
		200	0.1	28.50%	14.30%
			0.5	30.80%	19.00%
			0.9	28.00%	17.20%

จากตารางที่ 4.28 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 4, a = 2:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.29 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 4, a = 3:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
4	3:1	20	0.1	38.80%	14.40%
			0.5	40.00%	10.00%
			0.9	34.20%	11.20%
		60	0.1	31.00%	10.40%
			0.5	30.00%	24.20%
			0.9	33.40%	3.60%
		100	0.1	30.80%	8.00%
			0.5	28.80%	15.20%
			0.9	30.20%	18.20%
		200	0.1	30.80%	19.20%
			0.5	24.40%	17.60%
			0.9	25.60%	11.00%

จากตารางที่ 4.29 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 4, a = 3:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.30 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5, a = 1:4, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
5	1:4	20	0.1	37.40%	7.70%
			0.5	36.50%	6.05%
			0.9	36.55%	4.55%
		60	0.1	31.85%	15.00%
			0.5	32.80%	11.30%
			0.9	32.05%	15.45%
		100	0.1	27.45%	19.25%
			0.5	28.55%	2.45%
			0.9	29.90%	10.65%
		200	0.1	27.40%	16.45%
			0.5	27.70%	23.15%
			0.9	26.70%	23.65%

จากตารางที่ 4.30 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 5, a = 1:4, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.31 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5, a = 2:3, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
5	2:3	20	0.1	34.20%	0.27%
			0.5	37.07%	4.33%
			0.9	37.33%	2.67%
		60	0.1	30.53%	17.60%
			0.5	32.80%	17.00%
			0.9	32.73%	4.47%
		100	0.1	29.87%	18.13%
			0.5	31.00%	17.00%
			0.9	29.20%	17.07%
		200	0.1	28.33%	19.73%
			0.5	28.07%	12.93%
			0.9	30.40%	16.60%

จากตารางที่ 4.31 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 5, a = 2:3, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และส่วนมากค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี BE มีแนวโน้มเพิ่มขึ้น แต่วิธี HP มีแนวโน้มลดลง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น

ตารางที่ 4.32 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5, a = 3:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
5	3:2	20	0.1	37.40%	16.00%
			0.5	35.40%	11.60%
			0.9	33.70%	10.80%
		60	0.1	34.90%	22.10%
			0.5	34.60%	1.20%
			0.9	35.00%	23.30%
		100	0.1	31.70%	21.00%
			0.5	30.00%	19.60%
			0.9	32.30%	22.00%
		200	0.1	29.90%	23.00%
			0.5	29.50%	23.70%
			0.9	30.20%	15.00%

จากตารางที่ 4.32 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 5, a = 3:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.33 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5, a = 4:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
5	4:1	20	0.1	38.20%	3.40%
			0.5	37.20%	4.00%
			0.9	38.20%	13.20%
		60	0.1	35.60%	19.00%
			0.5	34.40%	24.40%
			0.9	34.80%	20.40%
		100	0.1	32.20%	25.40%
			0.5	32.00%	16.60%
			0.9	31.80%	23.80%
		200	0.1	33.20%	14.60%
			0.5	32.20%	23.60%
			0.9	29.80%	16.80%

จากตารางที่ 4.33 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 5, a = 4:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าวิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.34 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6, a = 1:5, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
6	1:5	20	0.1	29.36%	25.60%
			0.5	31.16%	12.00%
			0.9	30.72%	1.80%
		60	0.1	31.56%	48.60%
			0.5	33.60%	30.30%
			0.9	33.72%	40.70%
		100	0.1	29.88%	37.10%
			0.5	29.00%	42.30%
			0.9	30.72%	39.20%
		200	0.1	28.12%	49.20%
			0.5	28.96%	35.30%
			0.9	27.40%	33.80%

จากตารางที่ 4.34 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 6, a = 1:5,n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าเมื่อ M = 0.1, 0.9 และ n = 20 , M = 0.5 และ n = 20, 60 วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE แต่เมื่อ M = 0.1, 0.9 และ n = 60, 100, 200 , M = 0.5 และ n = 100, 200 วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP

ตารางที่ 4.35 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6, a = 2:4, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
6	2:4	20	0.1	31.45%	19.40%
			0.5	32.00%	8.20%
			0.9	32.50%	13.60%
		60	0.1	34.95%	21.80%
			0.5	35.05%	35.70%
			0.9	33.25%	24.90%
		100	0.1	32.45%	22.80%
			0.5	31.50%	16.40%
			0.9	31.40%	28.30%
		200	0.1	27.35%	41.30%
			0.5	29.65%	45.30%
			0.9	28.60%	43.20%

จากตารางที่ 4.35 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 6, a = 2:4,n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าเมื่อ M = 0.5 และ n = 60 , M = 0.1, 0.5, 0.9 และ n = 200 วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP แต่ในกรณีอื่นๆ วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.36 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6, a = 3:3, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
6	3:3	20	0.1	33.20%	2.80%
			0.5	33.67%	11.00%
			0.9	30.47%	7.60%
		60	0.1	33.07%	21.00%
			0.5	36.93%	34.10%
			0.9	35.67%	32.60%
		100	0.1	32.00%	29.70%
			0.5	30.80%	30.70%
			0.9	32.07%	28.60%
		200	0.1	28.40%	33.20%
			0.5	28.47%	26.80%
			0.9	27.93%	34.50%

จากตารางที่ 4.36 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 6, a = 3:3,n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่าเมื่อ M = 0.1, 0.9 และ n = 200 วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี HP แต่ในกรณีอื่นๆ วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

ตารางที่ 4.37 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6, a = 4:2, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
6	4:2	20	0.1	34.30%	6.60%
			0.5	31.20%	4.50%
			0.9	34.50%	0.70%
		60	0.1	31.80%	15.30%
			0.5	34.60%	15.00%
			0.9	35.00%	10.70%
		100	0.1	29.80%	21.30%
			0.5	30.60%	18.00%
			0.9	28.40%	22.30%
		200	0.1	27.30%	23.80%
			0.5	29.80%	23.80%
			0.9	31.40%	19.90%

จากตารางที่ 4.37 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 6, a = 4:2,n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่า วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE และค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระของวิธี HP ส่วนมากมีแนวโน้มลดลง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น

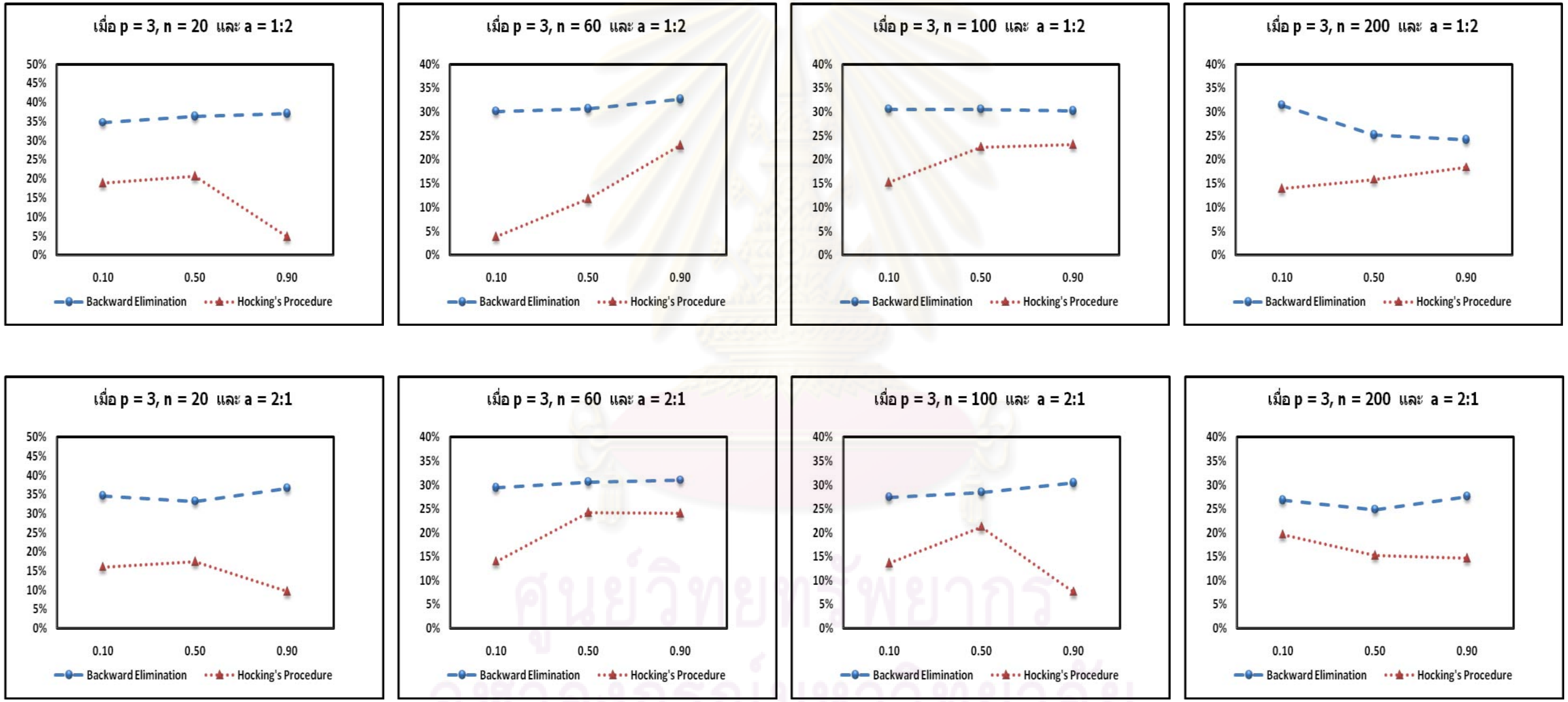
ตารางที่ 4.38 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6, a = 5:1, n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9

p	a	n	M	Percentage of mean error	
				BE	HP
6	5:1	20	0.1	32.80%	2.80%
			0.5	34.40%	1.40%
			0.9	32.00%	0.10%
		60	0.1	34.60%	5.50%
			0.5	39.20%	12.10%
			0.9	32.40%	7.90%
		100	0.1	35.40%	14.20%
			0.5	32.60%	6.50%
			0.9	32.80%	11.10%
		200	0.1	30.60%	12.00%
			0.5	30.60%	11.10%
			0.9	30.20%	7.90%

จากตารางที่ 4.38 เมื่อพิจารณาเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระที่ p = 6, a = 5:1,n = 20, 60, 100, 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปเป็น 0.1, 0.5 และ 0.9 พบว่า วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระต่ำกว่าวิธี BE

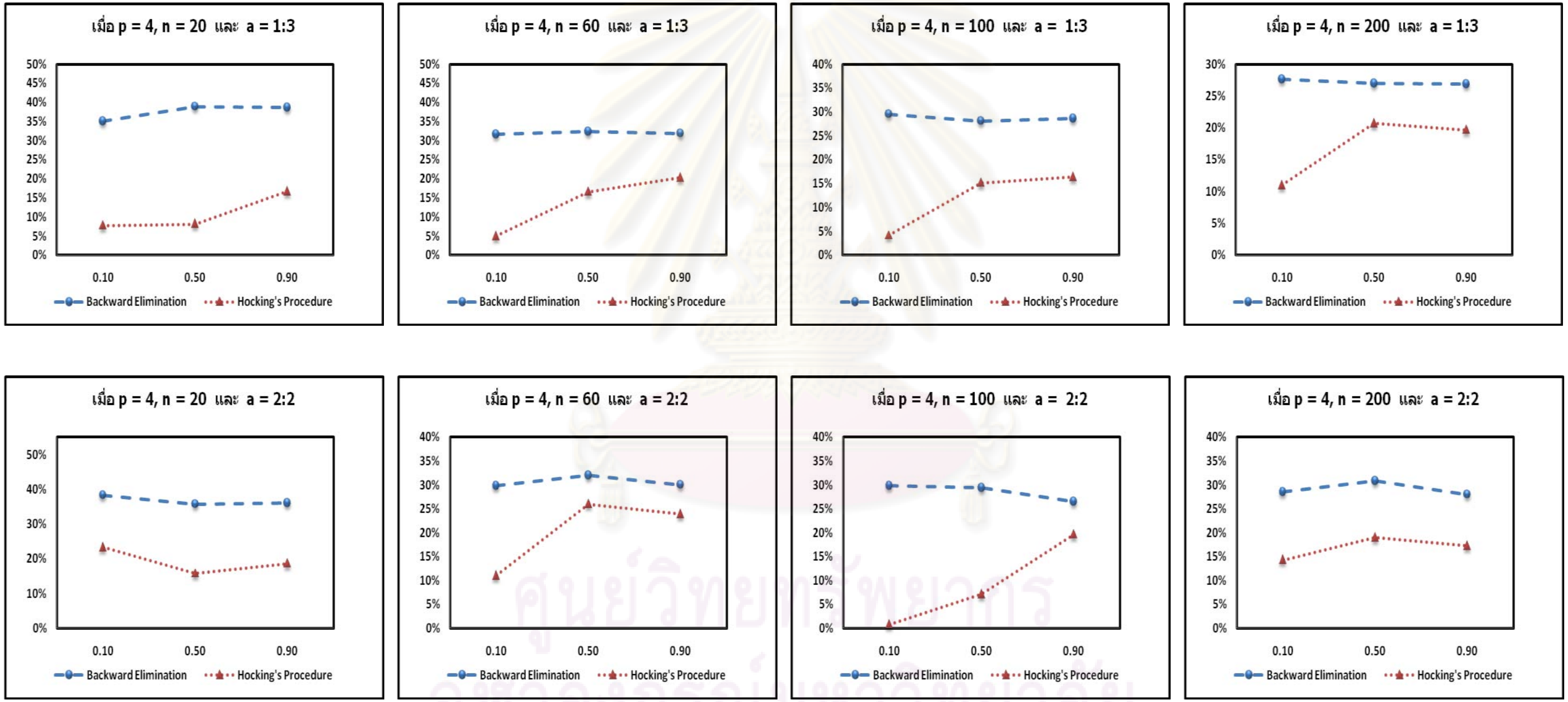
จากตารางที่ 4.25 – 4.38 สรุปได้ว่า เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น นั่นคือเกิดปัญหาพหุสัมพันธ์มาก ซึ่งอาจส่งผลกระทบต่อการประมาณค่าพารามิเตอร์ ทำให้ค่าประมาณที่ได้ผิดพลาดมากขึ้นและมีความแปรปรวนเพิ่มสูงขึ้น ส่งผลให้การคัดเลือกตัวแปรอิสระผิดพลาดมากขึ้นเช่นกัน สามารถแสดงรูปภาพเพื่อดูแนวโน้มค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไปแต่ปัจจัยอื่นคงที่ นำเสนอในรูปที่ 4.10 – 4.16 ได้ดังนี้

รูปที่ 4.10 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 3 (a = 1:2, 2:1), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป

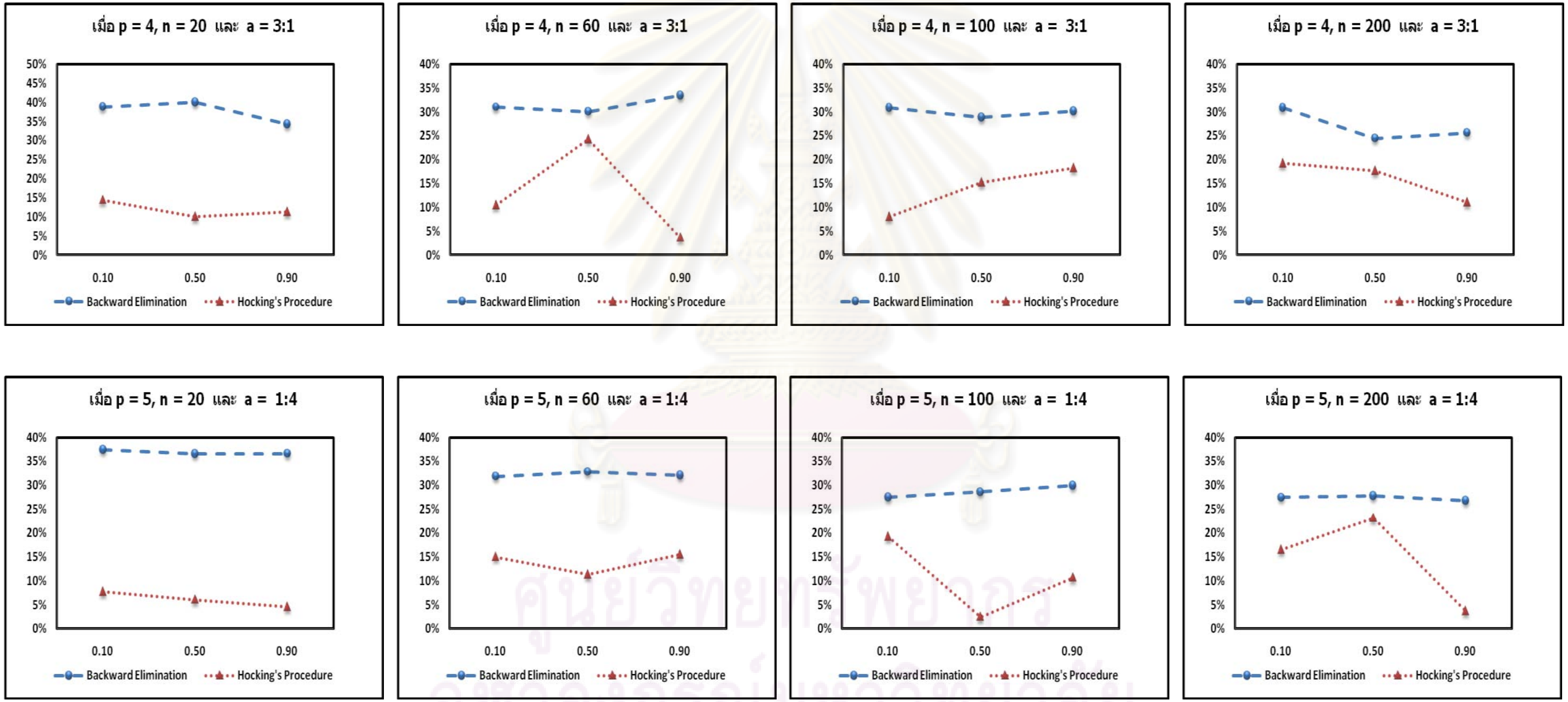


74

รูปที่ 4.11 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 4 (a = 1:3, 2:2), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป

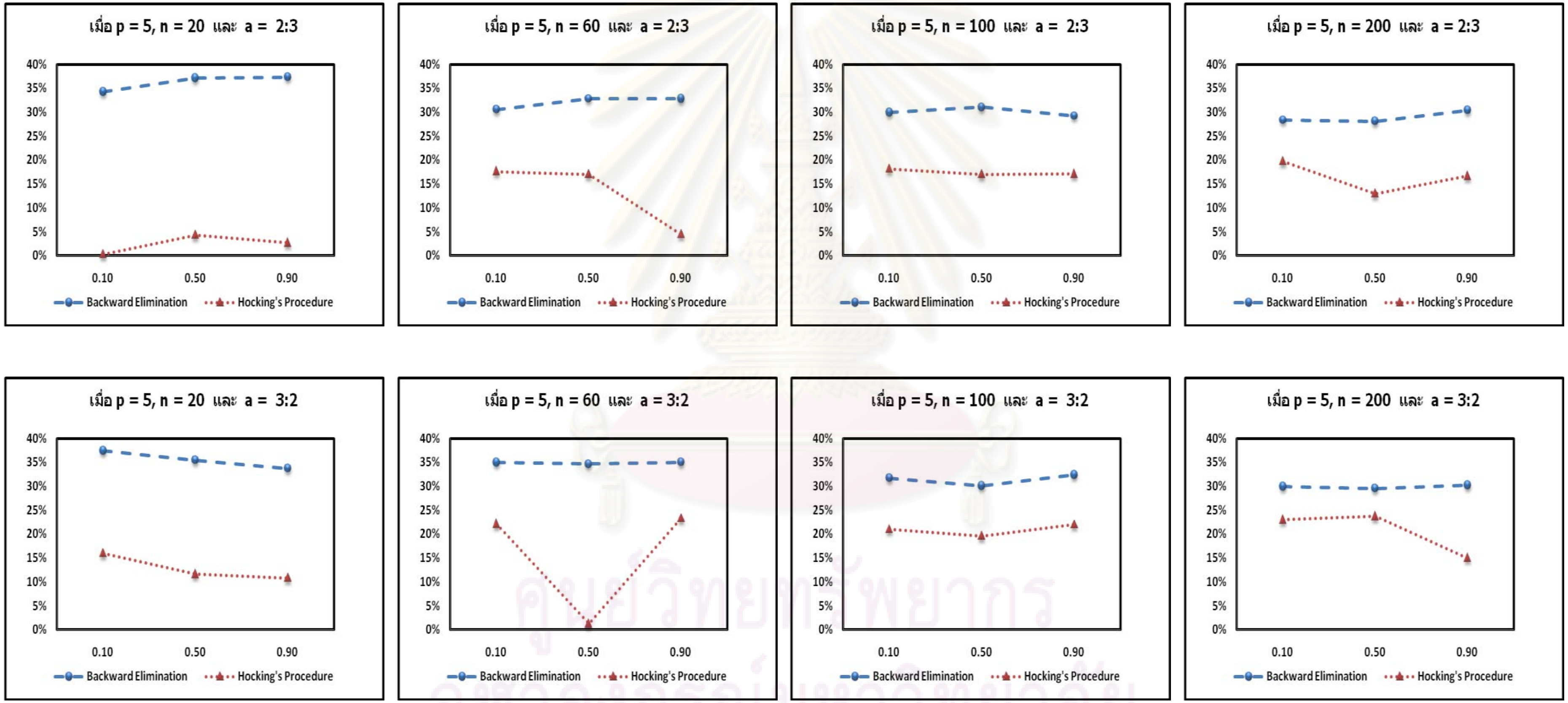


รูปที่ 4.12 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 4 (a = 3:1), p = 5 (a = 1:4), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป

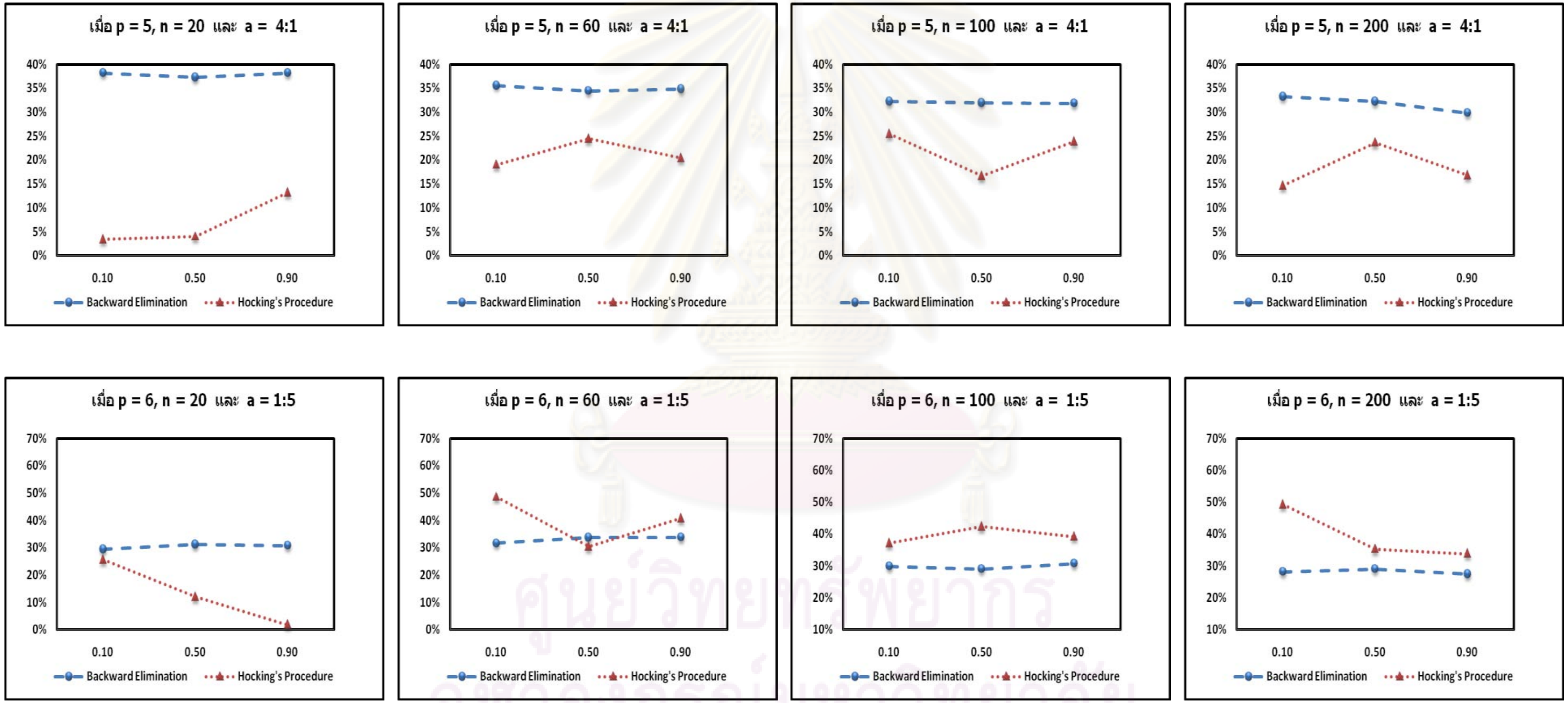


76

รูปที่ 4.13 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5 (a = 2:3, 3:2), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป

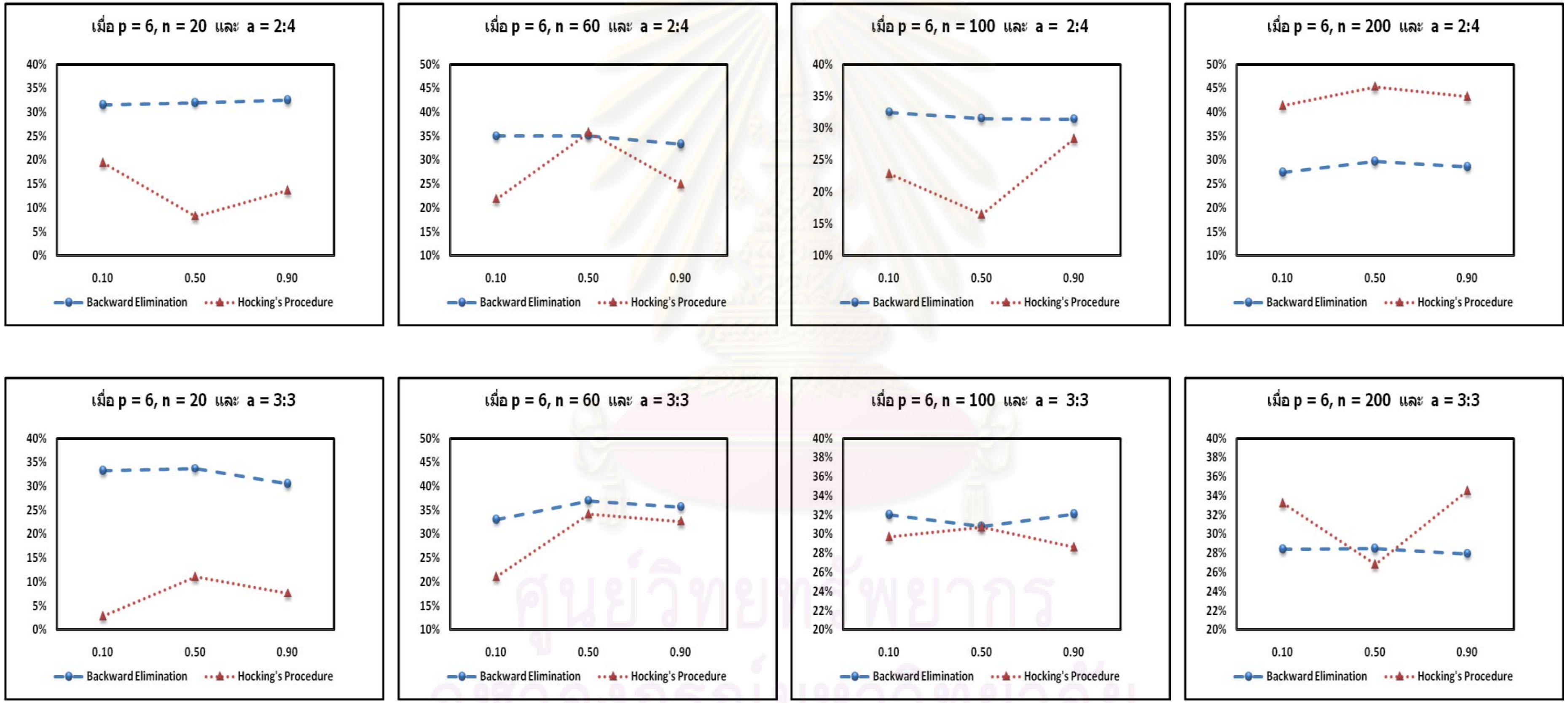


รูปที่ 4.14 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 5 (a = 4:1), p = 6 (a = 1:5), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป



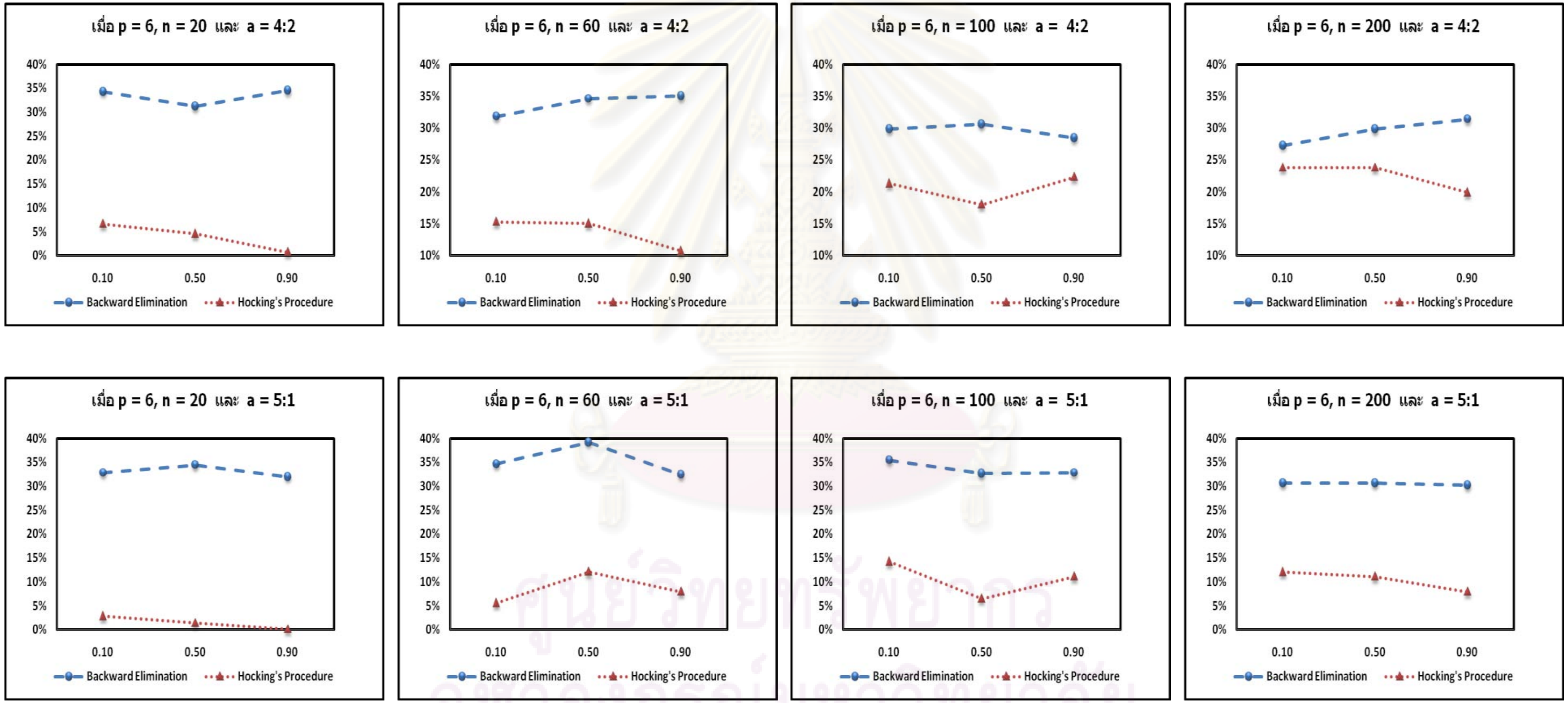
78

รูปที่ 4.15 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6 (a = 2:4, 3:3), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป



79

รูปที่ 4.16 แสดงค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการคัดเลือกตัวแปรอิสระ เมื่อ p = 6 (a =4:2, 5:1), n = 20, 60, 100 และ 200 และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป



80

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของวิธีการคัดเลือกตัวแปรอิสระในตัวแบบโพรบิท 2 วิธี คือ การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking's Procedure) โดยอาศัยวิธีของนอร์ดเบิร์กในการแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้นด้วย เมื่อจำนวนตัวแปรอิสระ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม ขนาดตัวอย่างและสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป ข้อมูลในแต่ละสถานการณ์จำลองขึ้นโดยใช้เทคนิคมอนติคาร์โลด้วยโปรแกรม R ผลการวิจัยมีข้อสรุปแยกตามสถานการณ์ต่างๆ ที่สนใจศึกษา ดังนี้

5.1.1 เมื่อสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป แต่ปัจจัยอื่นๆ คงที่

ตารางที่ 5.1 แสดงผลสรุปของค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระกรณีที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) เปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ (p) ขนาดตัวอย่าง (n) และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) คงที่

สถานการณ์	ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ
M=0.1, n=20,60,100,200, p=3	วิธี HP มีค่าต่ำกว่าวิธี BE
M=0.1, n=20,60, p=4,5	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนใกล้เคียงกัน แต่มีค่าต่ำกว่าวิธี BE
M=0.1, n=100,200, p=4,5	วิธี BE มีแนวโน้มสูงขึ้นและมีค่าสูงกว่าวิธี HP
M=0.1, n=20, p=6	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น แต่มีค่าต่ำกว่าวิธี BE
M=0.1, n=60,100, 200, p=6	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น
M=0.5, n=20,60,100,200, p=3	วิธี BE มีแนวโน้มลดลงแต่มีค่าสูงกว่าวิธี HP

ตารางที่ 5.1 (ต่อ)

สถานการณ์	ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ
M=0.5, n=20,60, p=4,5	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนใกล้เคียงกัน แต่มีค่าต่ำกว่าวิธี BE
M=0.5, n=100,200, p=4	วิธี BE มีค่าสูงเมื่อ a=2:2 และมีค่าสูงกว่าวิธี HP
M=0.5, n=100,200, p=5	วิธี HP มีค่าต่ำกว่าวิธี BE
M=0.5, n=20, p=6	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น แต่มีค่าต่ำกว่าวิธี BE
M=0.5, n=60,100, 200, p=6	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น
M=0.9, n=20,60,100,200, p=3	วิธี HP มีค่าต่ำกว่าวิธี BE
M=0.9, n=20,60,100, p=4	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนใกล้เคียงกัน แต่มีค่าต่ำกว่าวิธี BE
M=0.9, n=200, p=4	วิธี HP มีแนวโน้มลดลงแต่มีค่าต่ำกว่าวิธี BE
M=0.9, n=20, p=6	วิธี HP มีค่าต่ำกว่าวิธี BE
M=0.9, n=60,100, 200, p=6	วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น

จากตารางที่ 5.1 ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ กรณีที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ ขนาดตัวอย่าง และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่ สรุปผลได้ดังนี้

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าเท่ากับ 0.1, 0.5,0.9 ที่จำนวนตัวแปรอิสระ 3,4,5 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี HP มีค่าต่ำกว่าวิธี BE ในทุกสถานการณ์ เมื่อขนาดตัวอย่างมีขนาดเล็ก วิธี HP มีค่าสูงขึ้นเมื่อตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนใกล้เคียงกัน

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม มีค่าเท่ากับ 0.1, 0.5,0.9 ที่จำนวนตัวแปรอิสระ 6 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี HP มีค่าสูงขึ้น เมื่อตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีจำนวนมากขึ้น ที่ขนาดตัวอย่างเท่ากับ 60,100,200

5.1.2 เมื่อขนาดตัวอย่าง (n) เปลี่ยนแปลงไป แต่ปัจจัยอื่นๆ คงที่

ตารางที่ 5.2 แสดงผลสรุปของค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระกรณีที่ขนาดตัวอย่าง (n) เปลี่ยนแปลงไป แต่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) จำนวนตัวแปรอิสระ (p) และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) คงที่

<table>
<tr><th>สถานการณ์</th><th>ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ</th></tr>
<tr><td>M=0.1, p=3, a=1:2,2:1</td><td>วิธี BE มีค่าสูงเมื่อขนาดตัวอย่างมีค่าน้อยและมีค่าสูงกว่าวิธี HP</td></tr>
<tr><td>M=0.1, p=4, a=1:3,2:2,3:1</td><td rowspan="2">วิธี BE มีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP</td></tr>
<tr><td>M=0.1, p=5, a=1:4,2:3,3:2,4:1</td></tr>
<tr><td>M=0.1, p=6, a=1:5,2:4,3:3,4:2,5:1</td><td>วิธี HP มีค่าสูงเมื่อขนาดตัวอย่างมีค่ามาก</td></tr>
<tr><td>M=0.5, p=3 ,a=1:2,2:1</td><td rowspan="3">วิธี BE มีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP</td></tr>
<tr><td>M=0.5, ,p=4, a=1:3,2:2,3:1</td></tr>
<tr><td>M=0.5, p=5, a=1:4,2:3,3:2,4:1</td></tr>
<tr><td>M=0.5, p=6, a=1:5,2:4,3:3,4:2,5:1</td><td>วิธี BE มีค่าเพิ่มขึ้นเมื่อ n=60 หลังจากนั้นมีค่าลดลง</td></tr>
<tr><td>M=0.9, p=3, a=1:2,2:1</td><td rowspan="3">วิธี BE มีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP</td></tr>
<tr><td>M=0.9, p=4, a=1:3,2:2,3:1</td></tr>
<tr><td>M=0.9, p=5, a=1:4,2:3,3:2,4:1</td></tr>
<tr><td>M=0.9, p=6, a=1:5,2:4,3:3</td><td>วิธี BE มีค่าเพิ่มขึ้นเมื่อ n=60 หลังจากนั้นมีค่าลดลง
วิธี HP มีค่าสูงเมื่อขนาดตัวอย่างมีค่ามาก</td></tr>
<tr><td>M=0.9, p=6, a=4:2,5:1</td><td>วิธี HP มีค่าสูงเมื่อขนาดตัวอย่างมีค่ามาก</td></tr>
</table>

จากตารางที่ 5.2 ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระในกรณีที่ขนาดตัวอย่างเปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามและระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่ สรุปผลได้ดังนี้

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าเท่ากับ 0.1, 0.5, 0.9 ที่จำนวนตัวแปรอิสระ 3, 4, 5 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP ในทุกสถานการณ์

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าเท่ากับ 0.1 ที่จำนวนตัวแปรอิสระ 6 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี HP มีค่าสูง เมื่อขนาดตัวอย่างมาก

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าเท่ากับ 0.5 ที่จำนวนตัวแปรอิสระ 6 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่างเท่ากับ 60 หลังจากนั้นมีค่าลดลง

เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมีค่าเท่ากับ 0.9 ที่จำนวนตัวแปรอิสระ 6 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี HP มีค่าสูงเมื่อขนาดตัวอย่างมีค่ามาก แต่วิธี BE มีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่างเท่ากับ 60 หลังจากนั้นมีค่าลดลง ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเป็น 1:5, 2:4 และ 3:3

5.1.3 เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป แต่ปัจจัยอื่นๆ คงที่

ตารางที่ 5.3 แสดงผลสรุปของค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระกรณีที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป แต่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) จำนวนตัวแปรอิสระ (p) และขนาดตัวอย่าง (n) คงที่

สถานการณ์	ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ
p=3, a=1:2, n=20,60	วิธี BE มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น และมีค่าสูงกว่าวิธี HP
p=3, a=1:2, n=100,200	วิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP
p=3, a=2:1, n=20	วิธี BE มีค่าสูงกว่าวิธี HP
p=3, a=2:1, n=60,100	วิธี BE มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น และมีค่าสูงกว่าวิธี HP

ตารางที่ 5.3 (ต่อ)

สถานการณ์	ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ
p=3, a=2:1, n=200	วิธี HP มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น และมีค่าต่ำกว่าวิธี BE
p=4, a=1:3, n=20,60,100	วิธี HP มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าต่ำกว่าวิธี BE
p=4, a=1:3, n=200	วิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP
p=4, a=2:2,3:1, n=20,60, 200	วิธี BE มีค่าสูงกว่าวิธี HP
p=4, a=2:2,3:1, n=100	วิธี HP มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าต่ำกว่าวิธี BE
p=5, a=1:4, n=20,60,100,200	วิธี BE มีค่าสูงกว่าวิธี HP
p=5, a=2:3, n=20,60,100,200	
p=5, a=3:2, n=20	ทั้งวิธี BE และ HP มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น โดยวิธี HP มีค่าต่ำกว่าวิธี BE
p=5, a=3:2, n=60,100,200	วิธี BE มีค่าสูงกว่าวิธี HP
p=5, a=4:1, n=20	วิธี HP มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าต่ำกว่าวิธี BE
p=5, a=4:1, n=60	วิธี BE มีค่าสูงกว่าวิธี HP
p=5, a=4:1, n=100,200	วิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP
p=6, a=1:5, n=20	วิธี HP มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น และมีค่าต่ำกว่าวิธี BE
p=6, a=1:5, n=60,100,200	วิธี BE มีค่าต่ำกว่าวิธี HP
p=6, a=2:4, n=20	วิธี BE มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น และมีค่าสูงกว่าวิธี HP
p=6, a=2:4, n=60	วิธี BE มีค่าสูง เมื่อ M = 0.5 แต่มีค่าต่ำกว่าวิธี HP
p=6, a=2:4, n=200	วิธี BE มีค่าต่ำกว่าวิธี HP

ตารางที่ 5.3 (ต่อ)

สถานการณ์	ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระ
p=6, a=2:4, n=100	วิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP
p=6, a=3:3, n=20,60	ทั้งวิธี BE และวิธี HP มีค่าสูง เมื่อ M = 0.5
p=6, a=3:3, n=100	วิธี BE มีค่าสูง เมื่อ M = 0.9 วิธี HP มีค่าสูง เมื่อ M = 0.5 และวิธี BE มีค่าสูงกว่าวิธี HP
p=6, a=3:3, n=200	วิธี HP มีค่าสูง เมื่อ M = 0.9 และมีค่าสูงกว่าวิธี BE เมื่อ M = 0.1,0.9
p=6, a=4:2, n=20,60,100,200	วิธี BE มีค่าสูงกว่าวิธี HP
p=6, a=5:1, n=20,60,100,200	

จากตารางที่ 5.3 ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระในกรณีที่ขนาดตัวอย่างเปลี่ยนแปลงไป แต่จำนวนตัวแปรอิสระ สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามและระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามคงที่ สรุปผลได้ดังนี้

เมื่อจำนวนตัวแปรอิสระ 3 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าสูงกว่าวิธี HP ในทุกสถานการณ์ และที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 1:2 ขนาดตัวอย่างเท่ากับ 20, 60 และ 2:1 ขนาดตัวอย่างเท่ากับ 60, 100 วิธี BE มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าลดลงที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 1:2 และขนาดตัวอย่างเท่ากับ 100, 200 ส่วนวิธี HP มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้นที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 2:1 และขนาดตัวอย่างเท่ากับ 200

เมื่อจำนวนตัวแปรอิสระ 4 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าสูงกว่าวิธี HP ในทุกสถานการณ์ และที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 1:3 ขนาดตัวอย่างเท่ากับ 20, 60, 100 และ 2:2, 3:1 ขนาดตัวอย่างเท่ากับ 100 วิธี HP

มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่วิธี BE มีค่าลดลง ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความ สัมพันธ์กับตัวแปรตามเท่ากับ 1:3 ขนาดตัวอย่างเท่ากับ 200

เมื่อจำนวนตัวแปรอิสระ 5 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าสูงกว่าวิธี HP ในทุกสถานการณ์ และที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 3:2 ขนาดตัวอย่างเท่ากับ 20 ทั้งวิธี BE และ HP มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่วิธี HP มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 4:1 ขนาดตัวอย่างเท่ากับ 20 และวิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 4:1 ขนาดตัวอย่างเท่ากับ 100, 200

เมื่อจำนวนตัวแปรอิสระ 6 ตัว ที่ขนาดตัวอย่างเท่ากับ 20 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 1:5 พบว่า วิธี HP มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระลดลง เมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้นและมีค่าต่ำกว่าวิธี BE

เมื่อจำนวนตัวแปรอิสระ 6 ตัว พบว่า วิธี BE มีค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการเลือกตัวแปรอิสระต่ำกว่าวิธี HP ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีค่าเป็น

1) 1:5 ขนาดตัวอย่างเท่ากับ 60, 100, 200
2) 2:4 ขนาดตัวอย่างเท่ากับ 60, 200
3) 3:3 ขนาดตัวอย่างเท่ากับ 200

เมื่อจำนวนตัวแปรอิสระ 6 ตัว ที่ขนาดตัวอย่างเท่ากับ 20 และสัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 2:4 พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าสูงขึ้นเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้นและมีค่าสูงกว่าวิธี HP

เมื่อจำนวนตัวแปรอิสระ 6 ตัว ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามเท่ากับ 2:4 ขนาดตัวอย่างเท่ากับ 100 พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี BE มีค่าลดลงเมื่อระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเพิ่มขึ้น แต่มีค่าสูงกว่าวิธี HP

เมื่อจำนวนตัวแปรอิสระ 6 ตัว พบว่า ค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระของวิธี HP มีค่าต่ำกว่าวิธี BE ที่สัดส่วนของตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามมีค่าเป็น

1) 3:3 ขนาดตัวอย่างเท่ากับ 100
2) 4:2 ขนาดตัวอย่างเท่ากับ 20, 60, 100, 200
3) 5:1 ขนาดตัวอย่างเท่ากับ 20, 60, 100, 200

5.2 ข้อเสนอแนะ

ผลการวิจัยในครั้งนี้มีข้อเสนอแนะ 2 ด้าน คือ

5.2.1 ด้านการศึกษาวิจัย

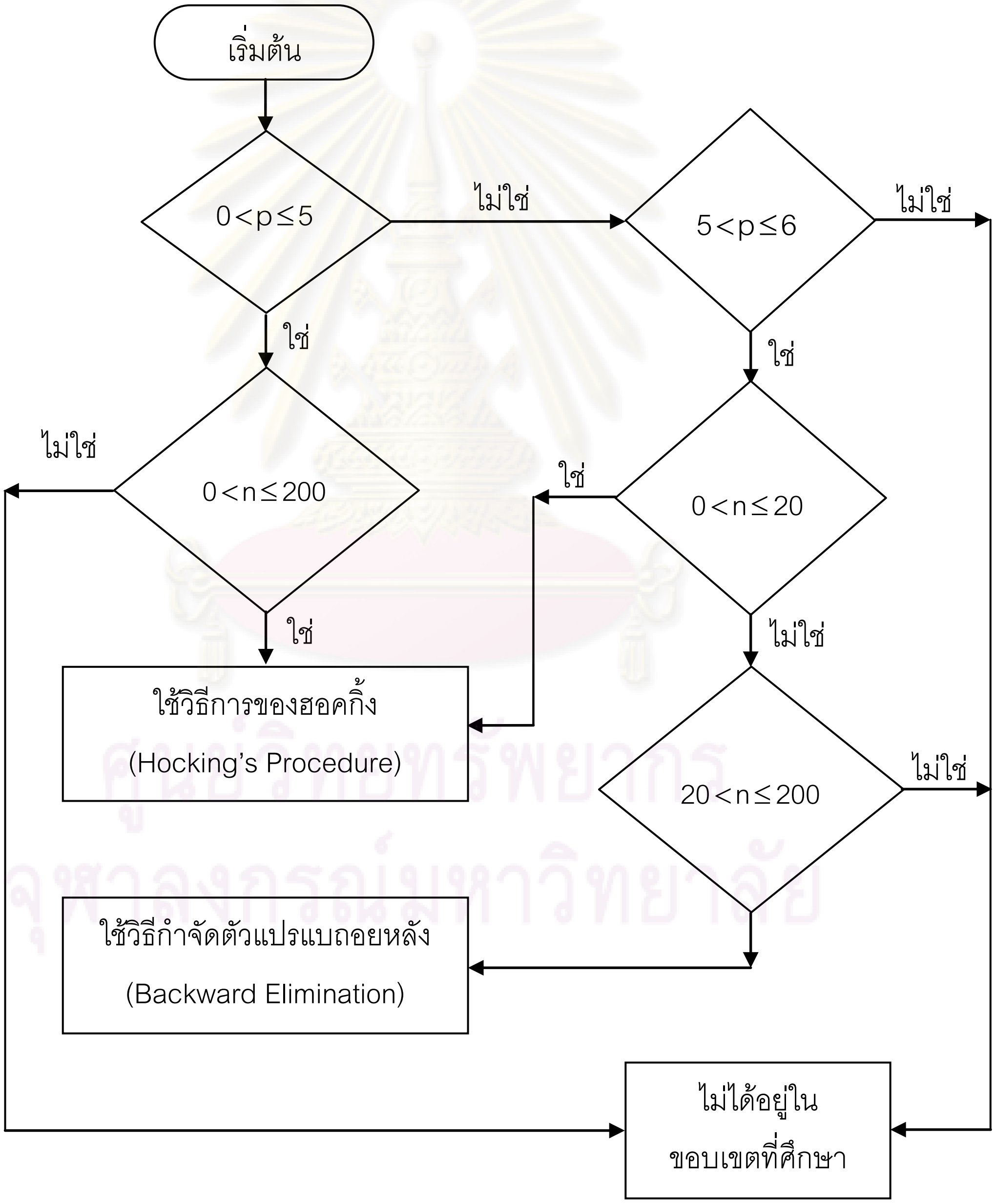
5.2.1.1 ในการวิจัยครั้งนี้ได้ศึกษาค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระในตัวแบบโพรบิทที่จำแนกข้อมูลเป็น 2 ประเภทเท่านั้น คือ กลุ่มที่สนใจและกลุ่มที่ไม่สนใจ ในการวิจัยครั้งต่อไปอาจทำการศึกษาค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระในตัวแบบโพรบิทแบบหลายกลุ่ม (multinomial probit model)

5.2.1.2 ในการวิจัยครั้งนี้ได้ศึกษาค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยในการเลือกตัวแปรอิสระในตัวแบบโพรบิทเมื่อข้อมูลตัวแปรอิสระมีการแจกแจงแบบปกติเท่านั้น ในการวิจัยครั้งต่อไปอาจทำการศึกษาเมื่อข้อมูลตัวแปรอิสระมีการแจกแจงแบบอื่นๆ

5.2.2 ด้านการนำไปใช้ประโยชน์

การวิจัยนี้ได้นำเสนอค่าเปอร์เซ็นต์ความผิดพลาดเฉลี่ยของการเลือกตัวแปรอิสระในตัวแบบโพรบิทด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และวิธีการของฮอคกิ้ง (Hocking’s Procedure) โดยอาศัยวิธีของนอร์ดเบิร์กในการแปลงข้อมูลที่ใช้ตัวแบบโพรบิทเป็นข้อมูลที่ใช้ตัวแบบความถดถอยเชิงเส้น สามารถใช้เป็นส่วนช่วยตัดสินใจเลือกใช้วิธีการเลือกตัวแปรอิสระเพื่อแก้ปัญหาพหุสัมพันธ์ที่เกิดขึ้นในตัวแบบโพรบิท

เนื่องจากเมื่อสัดส่วนของจำนวนตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามและตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตาม (a) และระดับความสัมพันธ์ระหว่างตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (M) เปลี่ยนแปลงไป วิธีการของฮอคกิ้ง (Hocking's Procedure) มีประสิทธิภาพดีกว่าวิธีการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) ในทุกกรณี ดังนั้น ปัจจัยที่ใช้พิจารณาเพื่อตัดสินใจเลือกใช้วิธีการคัดเลือกตัวแปรอิสระให้เกิดความผิดพลาดในการเลือกตัวแปรอิสระน้อยที่สุด คือ จำนวนตัวแปรอิสระ (p) และขนาดตัวอย่าง (n) ซึ่งแสดงเป็นแผนผังในการเลือกใช้วิธีการคัดเลือกตัวแปรอิสระ ได้ดังนี้



รูปที่ 5.1 แผนผังแสดงการลือกใช้วิธีการคัดเลือกตัวแปรอิสระ

รายการอ้างอิง

ภาษาไทย

กัลยา วานิชบัญชา. การวิเคราะห์ข้อมูลหลายตัวแปร. กรุงเทพฯ : ธรรมสาร, 2550.

ชนิศวรา ฉัตรแก้ว. การวิเคราะห์การถดถอยเมื่อตัวแปรตามมีสองลักษณะโดยใช้ตัวแบบความน่าจะเป็นเชิงเส้น ตัวแบบโพรบิท และตัวแบบโลจิท. วิทยานิพนธ์ปริญญามหาบัณฑิต, บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์, 2543.

ธีระพร วีระถาวร. การอนุมานเชิงสถิติขั้นกลาง. กรุงเทพฯ : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2536.

ธีระพร วีระถาวร. ความน่าจะเป็นกับการประยุกต์. กรุงเทพฯ : นำอักษรการพิมพ์, 2539.

นพมาศ อัครจันทโชติ. การเปรียบเทียบวิธีการสร้างตัวแบบในการวิเคราะห์ความถดถอยพหุนามกรณีที่มี 2 ตัวแปรอิสระซึ่งเกิดอันตรกิริยา. วิทยานิพนธ์ปริญญามหาบัณฑิต, บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2539.

เปรมวดี ชูไสว. การเปรียบเทียบวิธีการแก้ไขพหุสัมพันธ์ในการถดถอยเชิงเส้นพหุคูณ. วิทยานิพนธ์ปริญญามหาบัณฑิต, บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2548.

วราภรณ์ บุญยไพศาลเจริญ. วิธีการประมาณค่าพารามิเตอร์ในตัวแบบการถดถอยเชิงเส้นพหุคูณเมื่อเกิดพหุสัมพันธ์. วิทยานิพนธ์ปริญญามหาบัณฑิต, บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2546.

ภาษาอังกฤษ

Aldrich, J.H., and Nelson, F.D. Linear probability, logit and probit models. 9th ed. Newbury Park : SAGE Publications, Inc, 1990.

Finney, D.J. Probit Analysis. 3rd ed. Cambridge : Cambridge University Press, 1971.

Nordberg,L. Stepwise Selection of Explanatory Variables in the Binary Logit Model. Scandinavian Journal Statistics. 1981(8) : 17-26.

Power ,D.A., and Xie,Y. Statistics Methods for Catagorical Data Analysis. San Diego : Academic Press, 2002.

ภาคผนวก

ตัวอย่างการใช้โปรแกรม R ในการดำเนินการวิจัย

```
# Correlation Matrix
cor.matrix <- function(p1, p2, cor.value, dnames=NULL){
  fullmat <- matrix(0.01, p1+p2, p1+p2)
  if(p1==1){
    mat <- matrix(c(1), p1, p1)
  }else if(p1==2){
    mat <- matrix(c(1,2,2,1), p1, p1)
  }else if(p1==3){
    mat <- matrix(c(1,2,3,2,1,4,3,4,1), p1, p1)
  }else if(p1==4){
    mat <- matrix(c(1,2,3,4,2,1,5,6,3,5,1,7,4,6,7,1), p1, p1)
  }else if(p1==5){
    mat <- matrix(c(1,2,3,4,5,2,1,6,7,8,3,6,1,9,10,4,7,9,1,11,5,8,10,11,1), p1, p1)
  }else{
    stop("input error")
  }
  cmat <- 1/mat*cor.value
  fullmat[1:p1, 1:p1] <- cmat
  diag(fullmat) <- 1
  dimnames(fullmat) <- list(dnames, dnames)
  output <- fullmat
  return(output)
}

# Generate Independent Variables
x.generate <- function(n, p1, p2, Mean=0, Var=1, cor.value, error=0.01, nloop=1,
prcorrect=0.7, cnames=NULL){
  cmat <- cor.matrix(p1, p2, cor.value)
  cholmat <- chol(cmat)
```

```
    num <- new("list")
    for(i in 1:nloop){
        loopcount <- 0
        while(loopcount<1){
            num[[i]] <- matrix(rnorm(n*(p1+p2), Mean, Var), n)
            num[[i]] <- num[[i]]%*%cholmat
            if((sum(abs(cor(num[[i]])-cmat)<=error)/length(cholmat))>=prcorrect){
                loopcount <- loopcount+1
            }
            loopcount
        }
        colnames(num[[i]]) <- cnames
    }
    output <- num
    return(output)
}

# Calculate KMO (Kaiser-Meyer-Olkin)
kmo.test <- function(df)
{
  cor.sq <- cor(df)^2
            cor.sumsq <- (sum(cor.sq)-dim(cor.sq)[1])/2
            library(corpcor)
            pcor.sq <- cor2pcor(cor(df))^2
            pcor.sumsq <- (sum(pcor.sq)-dim(pcor.sq)[1])/2
            kmo <- cor.sumsq/(cor.sumsq+pcor.sumsq)
            return(kmo)
}

# Generate error term
e.generate <- function(n, Mean=0, Var=1, nloop=1, cnames=NULL){
```

```
    num <- new("list")
    for(i in 1:nloop){
        num[[i]] <- as.matrix(rnorm(n, Mean, Var))
        colnames(num[[i]]) <- cnames
    }
    output <- num
    return(output)
}

# Generate Dependent Variables
y.generate <- function(datalist, cbeta, cnames=NULL){
    numtmp <- cbeta[1]+apply(datalist[[1]]%*%as.matrix(cbeta[-1]), 1, sum)
    num <- datalist[[2]]+matrix(rep(numtmp, ncol(datalist[[2]])), nrow(datalist[[2]]))
    colnames(num) <- cnames
    output <- num
    return(output)
}

# Transform data to binary (0,1)
y.dummy <- function(y.data, cprob=0.5, cnames=NULL){
    prob.dt <- pnorm(y.data)
    dum.dt <- ifelse(prob.dt>cprob,1,0)
    colnames(dum.dt) <- cnames
    output <- dum.dt
    return(output)
}

# Rerange of two lists
relist2 <- function(list1, list2){
    if(length(list1)==length(list2)){
        data.list <- new("list")
```

```
        for(i in 1:length(list1)){
          data.list[[i]] <- new("list")
          for(j in 1:2){
            if(j==1){
              data.list[[i]][[j]] <- list1[[i]]
            }else{
              data.list[[i]][[j]] <- list2[[i]]
            }
          }
        }
    }else{
        warning("length of two data sets not same")
    }
    output <- data.list
    return(output)
}

# Rerange of three lists
relist3 <- function(list1, list2, list3){
    if(length(unique(length(list1),length(list2),length(list3))==1)){
        data.list <- new("list")
        for(i in 1:length(list1)){
          data.list[[i]] <- new("list")
          for(j in 1:3){
            if(j==1){
              data.list[[i]][[j]] <- list1[[i]]
            }else if(j==2){
              data.list[[i]][[j]] <- list2[[i]]
            }else{
              data.list[[i]][[j]] <- list3[[i]]
            }
```

```
            }
        }
    }else{
        warning("length of three data sets not same")
    }
    output <- data.list
    return(output)
}

# Apply function cbind for use with lapply
cbind.list <- function(list1){
    listDt <- cbind(list1[[1]], list1[[2]])
    output <- listDt
    return(output)
}

# Calcutation y.est, prob.est, u and z
multi.calc <- function(xylist, beta.est, adjvalue=1e-15){
    options(digits=20)
    y.est <- xylist[[3]][1]+xylist[[1]]%*%as.matrix(xylist[[3]][-1])
    prob.y.est <- pnorm(y.est)
    prob.y.est <- ifelse(prob.y.est==0, prob.y.est+adjvalue, prob.y.est)
    prob.y.est <- ifelse(prob.y.est==1, prob.y.est-adjvalue, prob.y.est)
    n <- ncol(xylist[[1]])
    u <- new("list")
    for(i in 1:n){
        u[[i]] <- as.vector(xylist[[1]][,i])*sqrt(prob.y.est*(1-prob.y.est))
    }
    u <- matrix(unlist(u), nrow(xylist[[1]]), n, byrow=F)
    z <- (xylist[[2]]-prob.y.est)/sqrt(prob.y.est*(1-prob.y.est))+y.est
    colnames(z) <- "z"
```

```
   #output section
   output <- list(y.est, prob.y.est, u, z)
   names(output) <- c("y.est","prob.y.est","u","z")
   return(output)
}

#  Calcutation mse at k value
msek.calc <- function(u.tr.u.mtr, beta.u.mtr, sigma.sq, k.const=0){
   nr <- nrow(beta.u.mtr)
   nc <- nrow(beta.u.mtr)
   zero.mtr <- matrix(0, nr, nc)
   #Scalar matrix, diagonal elements are equal to k.initial value
   kscl.mtr.st <- zero.mtr
   diag(kscl.mtr.st)  <- k.const
   #Adjust U transpose U matrix by k initial scalar matrix
   u.tr.u.adj <- u.tr.u.mtr+kscl.mtr.st
   uinv.mmult <- solve(u.tr.u.adj)%*%solve(u.tr.u.adj)
   atemp <- eigen(u.tr.u.mtr)$values/(eigen(u.tr.u.mtr)$values+k.const)^2
   btemp <- t(beta.u.mtr)%*%uinv.mmult
   avalue <- sum(atemp)
   bvalue <- btemp%*%beta.u.mtr
   mse.k  <- sigma.sq*avalue+k.const^2*bvalue
   output <- mse.k
   return(output)
}

# Search k value
search.k <- function(u.tr.u.mtr, beta.u.mtr, sigma.sq, kint=0, d1st=100){
   mse_kint <- 0
   mse_k01  <- 0
   while(mse_kint>=mse_k01){
```

```
      k01 <- kint+d1st
      mse_kint <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, kint)
      mse_k01  <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, k01)
      if(mse_kint>=mse_k01){
         kint <- k01
      }
   }
   output <- kint-d1st
   return(output)
}

# Optimize k
optimize.k <- function(u.tr.u.mtr, beta.u.mtr, sigma.sq, kint=0, d1st=1, d2nd=0.01,
d3rd=0.001){
   kint1 <- search.k(u.tr.u.mtr, beta.u.mtr, sigma.sq, kint, d1st)
   kint2 <- ifelse((kint1-d1st)<0, 0, kint1-d1st)
   #kint2 <- kint1-d1st
   mse_kint <- 0
   mse_k01  <- 0
   while(mse_kint>=mse_k01&kint2<=kint1){
      k01 <- kint2+d2nd
      mse_kint <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, kint2)
      mse_k01  <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, k01)
      if(mse_kint>=mse_k01){
         kint2 <- k01
      }
   }
   kopt <- kint2
   if(kopt!=kint1){
      #Process A
      kstop <- kopt-d2nd
```

```
    k001  <- kopt
    kaf01 <- k001+d3rd
    mse_kaf01 <- 0
    mse_k001  <- 0
    while(kaf01>=kstop&mse_k001>=mse_kaf01){
        mse_k001  <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, k001)
        mse_kaf01 <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, kaf01)
        #Update kaf01 and k001
        k001  <- kaf01
        kaf01 <- k001+d3rd
    }
    #Output Process A
    kfinal <- k001  #Stop!!!
}else{
    #Process B
    kbf01 <- kopt-d2nd
    k001  <- kbf01+d3rd
    mse_kbf01 <- 0
    mse_k001  <- 0
    while(k001>=kopt&mse_kbf01>=mse_k001){
        mse_kbf01 <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, kbf01)
        mse_k001  <- msek.calc(u.tr.u.mtr, beta.u.mtr, sigma.sq, k001)
        #Update kbf01 and k001
        kbf01 <- k001
        k001  <- kbf01+d3rd
    }
    #Output Process B
    kfinal <- kbf01 #Stop!!!
}
options(digits=7)
```

```
   output <- kfinal
   return(output)
}

# Initial Parameters
options(digits=20)
init.beta <- c(0.1 ,0,0)
x1stsec <- c("d1")
x2ndsec <- c("d2","d3")
uznames <- c("u1","u2","u3","z")
myformula <- formula(yd~d1+d2+d3)
uzformula <- formula(z~u1+u2+u3)
cprob.piont <- 0.5
simu.error  <- 0.05
# Correlation for generate X
corr.x  <- 0.1
samsize <- 20
numloop <- 500
numseed <- 1

# Generate independent variables
set.seed(numseed+round(corr.x,2)*100)
xdata <- x.generate(samsize, length(x1stsec), length(x2ndsec), 0, 1, corr.x,
   simu.error, numloop, cnames=c(x1stsec, x2ndsec))

#Calculate KMO
KMO_Calc <- unlist(lapply(xdata, FUN=kmo.test))
KMO <- as.data.frame(cbind(c(1:length(xdata)), KMO_Calc))
names(KMO) <- c("LOOP","KMO")
```

```
#Generate error term with normal distribution
error <- e.generate(samsize, 0, 1, numloop, cnames="e")
xerror <- relist2(xdata, error)

#Generate dependent variable from indepent variables and error term
y_beta <- lapply(xerror, FUN=y.generate, init.beta, "y")

#Transform dependent data from numeric to binary (0,1) scale
y_dum <- lapply(y_beta, FUN=y.dummy, cnames="yd")

#Combine independent data and dependent binary data together
xy <- relist2(xdata, y_dum)
xy <- lapply(xy, FUN=cbind.list)
xy <- lapply(xy, FUN=as.data.frame)

#Fitting Generalized Linear Models
model <- new("list")
for(i in 1:numloop){
   model[[i]] <- glm(myformula, family=binomial(link=probit), data=xy[[i]])
}

beta  <- lapply(model, FUN=function(x){as.vector(coefficients(x))})

#Calculation "y.est", "prob.y.est", "u" and "z"
xy_beta <- relist3(xdata, y_dum, beta)
uz <- lapply(xy_beta, FUN=multi.calc)

# Backward Elimination and Stepwise Regression
data_uz <- new("list")
bem_fitt <- new("list")
bem_coef <- new("list")
```

```
bem_vars <- new("list")
u_trans_u  <- new("list")
beta_est_u <- new("list")
sigma_sq <- new("list")
k   <- new("list")
mtr <- new("list")
data_uz_add <- new("list")
hsw_fitt <- new("list")
hsw_coef <- new("list")
hsw_vars <- new("list")
for(i in 1:numloop){
   data_uz[[i]] <- as.data.frame(cbind(uz[[i]]$u, uz[[i]]$z))
   NCol <- ncol(data_uz[[i]])
   names(data_uz[[i]]) <- uznames
   bem_fitt[[i]] <- lm(uzformula, data=data_uz[[i]])                #Fit regression model
   bem_fitt[[i]] <- step(bem_fitt[[i]], direction="backward")       #Backward elimination
   bem_coef[[i]] <- as.vector(bem_fitt[[i]]$coefficients)           #Coefficients in final model
   bem_vars[[i]] <- names(bem_fitt[[i]]$coefficients)               #Variable in final model

   #Hocking's Procedure
   u_trans_z <- t(uz[[1]]$u)%*%uz[[1]]$z
   u_trans_u[[i]] <- t(uz[[i]]$u)%*%uz[[i]]$u
   beta_est_u[[i]] <- solve(u_trans_u[[i]])%*%u_trans_z
   u_mul_beta <- uz[[1]]$u%*%beta_est_u[[i]]
   z_adjusted <- uz[[1]]$z-u_mul_beta
   sigma_sq[[i]] <- t(z_adjusted)%*%z_adjusted/(nrow(uz[[1]]$u)
      -ncol(uz[[1]]$u)-1)
   k[[i]] <- optimize.k(u_trans_u[[i]], beta_est_u[[i]],
      sigma_sq[[i]], kint=0, d1st=1, d2nd=0.01, d3rd=0.001)
   mtr[[i]] <- matrix(0,NCol-1, NCol-1)
   diag(mtr[[i]]) <- sqrt(k[[i]])
```

```
    mtr[[i]] <- as.data.frame(cbind(mtr[[i]], rep(0, NCol-1)))
    names(mtr[[i]])  <- uznames
    data_uz_add[[i]] <- rbind(data_uz[[i]], mtr[[i]])
    hsw_fitt[[i]] <- lm(uzformula, data=data_uz_add[[i]])          #Fit regression model
    hsw_fitt[[i]] <- step(hsw_fitt[[i]], direction="both")          #Stepwise regression
    hsw_coef[[i]] <- as.vector(hsw_fitt[[i]]$coefficients)          #Coefficients in final model
    hsw_vars[[i]] <- names(hsw_fitt[[i]]$coefficients)              #Variable in final model
}

NLen <- length(uznames)
varT <- c("(Intercept)", uznames[-NLen])
vars_bem <- matrix(0, numloop, NLen)
vars_hsw <- matrix(0, numloop, NLen)
coef_bem <- matrix(0, numloop, NLen)
coef_hsw <- matrix(0, numloop, NLen)
for(i in 1:numloop){
    vars_bem[i, ] <- ifelse(is.na(match(varT, bem_vars[[i]])), 0, 1)
    vars_hsw[i, ] <- ifelse(is.na(match(varT, hsw_vars[[i]])), 0, 1)
    coef_bem[i, ] <- bem_coef[[i]][match(varT, bem_vars[[i]])]
    coef_hsw[i, ] <- hsw_coef[[i]][match(varT, hsw_vars[[i]])]
}
colnames(vars_bem) <- varT
colnames(vars_hsw) <- varT
colnames(coef_bem) <- varT
colnames(coef_hsw) <- varT

#Matrix to Data Frame
vars_bem <- as.data.frame(vars_bem)
vars_hsw <- as.data.frame(vars_hsw)
coef_bem <- as.data.frame(coef_bem)
coef_hsw <- as.data.frame(coef_hsw)
```

```
#Export Output into Microsoft Excel
write.csv(vars_bem,paste("E:/X(",length(x1stsec),",",length(x2ndsec),
   ")_vars_bem_cor_",round(corr.x,2),".csv",sep=""), row.names=F)
write.csv(vars_hsw,paste("E:/X(",length(x1stsec),",",length(x2ndsec),
   ")_vars_hsw_cor_",round(corr.x,2),".csv",sep=""), row.names=F)
write.csv(coef_bem,paste("E:/X(",length(x1stsec),",",length(x2ndsec),
   ")_coef_bem_cor_",round(corr.x,2),".csv",sep=""), row.names=F)
write.csv(coef_hsw,paste("E:/X(",length(x1stsec),",",length(x2ndsec),
   ")_coef_hsw_cor_",round(corr.x,2),".csv",sep=""), row.names=F)
#KMO
write.csv(KMO,paste("E:/X(",length(x1stsec),",",length(x2ndsec),
   ")_kmo_cor_",round(corr.x,2),".csv",sep=""), row.names=F)
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวพรทิพย์ คำหล้า เกิดวันพุธที่ 16 กุมภาพันธ์ พุทธศักราช 2526 ได้รับทุนโครงการพัฒนาและส่งเสริมผู้มีความสามารถพิเศษทางวิทยาศาสตร์และเทคโนโลยี (พสวท.) ในการศึกษาระดับปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาคณิตศาสตร์ จากภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น โดยสำเร็จการศึกษาในปีการศึกษา 2548 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตรมหาบัณฑิต สาขาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2551