

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อที่จะศึกษาเปรียบเทียบการประมาณค่าสัมประสิทธิ์การถดถอยจากวิธีการถดถอยวิดิค โดยการหาค่าคงที่  $k$  จาก 3 วิธี คือ 1.วิธีไฮเออร์นและเคนนาร์ต 2.วิธีค้นหาข้อมูลแบบลำดับ และ 3.วิธีเบส โดยจำลองข้อมูลภายใต้สถานการณ์ต่าง ๆ กล่าว คือ

1. ขนาดตัวอย่างเท่ากับ 10,30,50 และ 100
2. ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน มีค่าเท่ากับ 0.05,0.5,1 และ 3
3. จำนวนตัวแปรอิสระที่ทำการศึกษแบ่งตามลักษณะความสัมพันธ์ ดังนี้

ส่วนที่ 1 จำนวนตัวแปรอิสระมีความสัมพันธ์กัน 1 กลุ่ม ศึกษากรณีที่จำนวนตัวแปรอิสระทั้งหมดเท่ากับ 3 และ 5

ส่วนที่ 2 จำนวนตัวแปรอิสระมีความสัมพันธ์กัน 2 กลุ่ม ศึกษากรณีที่จำนวนตัวแปรอิสระทั้งหมดเท่ากับ 5 โดยแบ่งการศึกษาออกเป็น

- กรณีที่ 1 จำนวนตัวแปรอิสระที่มีความสัมพันธ์กันในกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 2
- กรณีที่ 2 จำนวนตัวแปรอิสระที่มีความสัมพันธ์กันในกลุ่มที่ 1 เท่ากับ 3 และในกลุ่มที่ 2 เท่ากับ 2

ในการจำลองข้อมูลครั้งนี้ใช้เทคนิคมอนติคาร์โลกระทำซ้ำ 500 ครั้ง ในแต่ละสถานการณ์

#### 5.1 สรุปผลการวิจัย

ในการพิจารณาว่าวิธีการประมาณค่าสัมประสิทธิ์การถดถอยวิดิคด้วยวิธีการหาค่า  $k$  จากวิธีต่างๆ นั้น วิธีการใดจะให้ผลดีที่สุด จะพิจารณาจากค่าเฉลี่ยความคลาดเคลื่อนยกกำลังสอง (AMSE) ซึ่งหากว่าวิธีใดให้ค่านี้น้อยที่สุด วิธีนั้นจะให้ผลดีที่สุด การสรุปผลการวิจัยจะแบ่งตามลักษณะความสัมพันธ์ของกลุ่มตัวแปรอิสระ ดังต่อไปนี้

ส่วนที่ 1 จำนวนตัวแปรอิสระมีความสัมพันธ์กัน 1 กลุ่ม

ในการวิจัยครั้งนี้เมื่อจำนวนตัวแปรอิสระมีความสัมพันธ์กัน 1 กลุ่ม ในกรณีที่จำนวนตัวแปรอิสระเท่ากับ 3 และ 5 ให้ผลสรุปเหมือนกัน กล่าวคือ

1. เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่า AMSE ของแต่ละวิธีมีแนวโน้มลดลง เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุได้
2. เมื่อระดับความสัมพันธ์ของตัวแปรอิสระและจำนวนตัวแปรอิสระที่มีความสัมพันธ์กันมีเพิ่มขึ้น ค่า AMSE มีค่าเพิ่มขึ้น เพราะว่าค่า AMSE เป็นฟังก์ชันของเมทริกซ์  $(X'X)^{-1}$  ดังที่ได้กล่าวมาแล้วในบทที่ 1 และบทที่ 2 ดังนั้นเมื่อจำนวนตัวแปรอิสระที่มีความสัมพันธ์กันเพิ่มขึ้นหรือตัวแปรอิสระที่มีความสัมพันธ์กันมากขึ้น จะส่งผลให้ค่าเฉพาะของเมทริกซ์  $X'X$  บางค่ามีค่าน้อย ทำให้ค่า AMSE มีค่าเพิ่มขึ้น
3. จำนวนตัวแปรอิสระเพิ่มขึ้น ค่า AMSE มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น เพราะว่าเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น จึงเหมือนกับเรามีตัวแปรมาอธิบายตัวแปรตามเพิ่มมากขึ้น จึงทำให้ค่า AMSE มีค่าลดลง
4. เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน ( $\sigma$ ) เพิ่มขึ้น ค่า AMSE ของแต่ละวิธีมีจะมีค่าเพิ่มขึ้น เนื่องจากค่า  $MSE = \sigma^2 + \text{bias}^2$

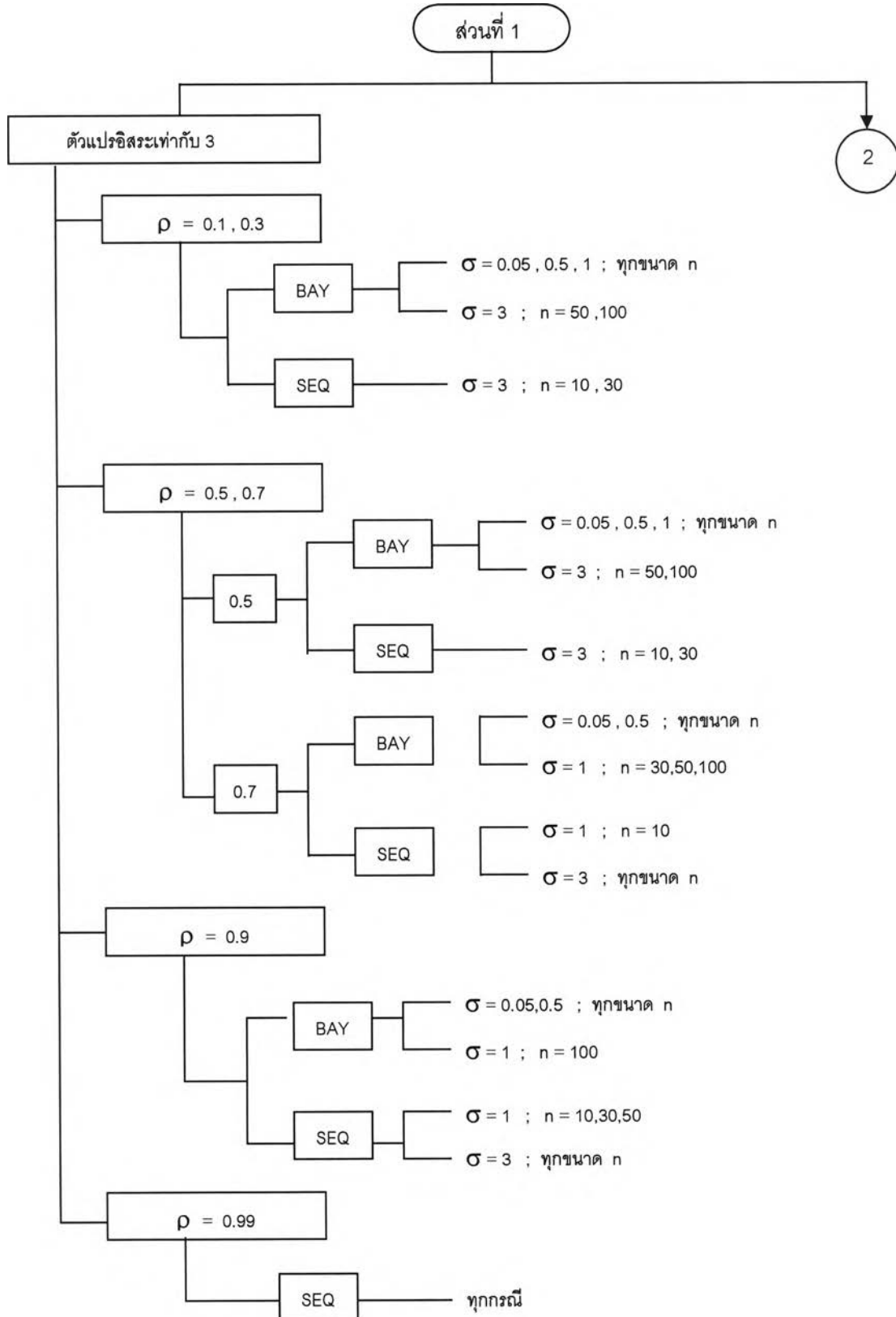
#### ส่วนที่ 2 จำนวนตัวแปรอิสระที่มีความสัมพันธ์กัน 2 กลุ่ม

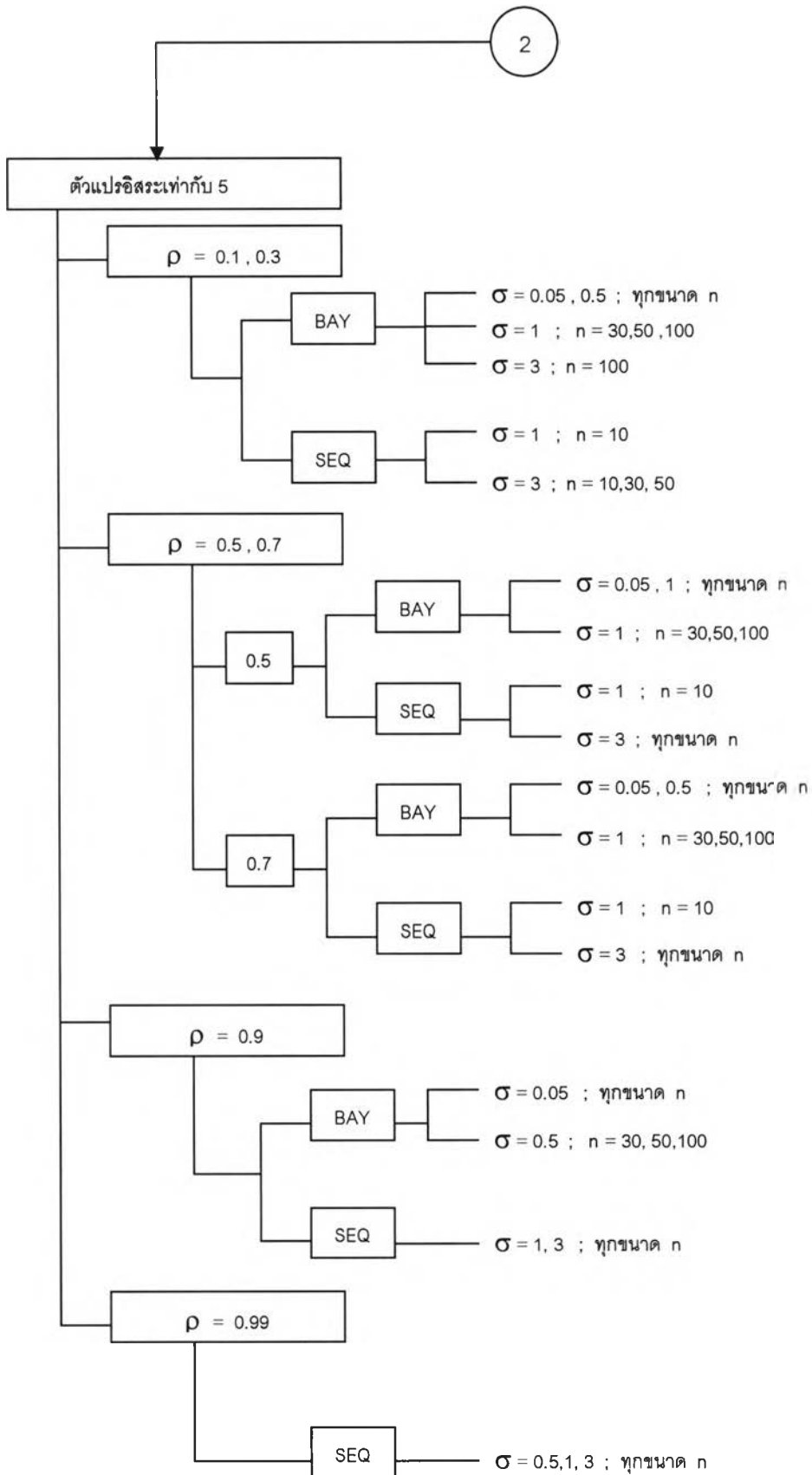
ผู้วิจัยทำการศึกษาในกรณีที่ขนาดตัวอย่างเท่ากับ 5 จากการศึกษาทั้ง 2 กรณี คือ ในกรณีที่จำนวนตัวแปรอิสระที่มีความสัมพันธ์กันในกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากับ 2 และในกรณีที่จำนวนตัวแปรอิสระที่มีความสัมพันธ์กันในกลุ่มที่ 1 เท่ากับ 3 และกลุ่มที่ 2 เท่ากับ 2 ให้ผลสรุปโดยส่วนใหญ่เหมือนกัน กล่าวคือ

1. เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่า AMSE ของแต่ละวิธีมีแนวโน้มลดลง เหตุผลเหมือนกับข้อ 1. ในส่วนที่ 1
2. เมื่อระดับความสัมพันธ์ในแต่ละกลุ่มเพิ่มขึ้น ค่า AMSE ของแต่ละวิธีมีแนวโน้มเพิ่มขึ้น เหตุผลเหมือนกับข้อ 2. ในส่วนที่ 1
3. เมื่อทำการเปรียบเทียบระหว่างในกรณีที่ 1 และกรณีที่ 2 (เมื่อจำนวนตัวแปรอิสระที่มีความสัมพันธ์กันเพิ่มขึ้น) ค่า AMSE ของแต่ละวิธีมีแนวโน้มเพิ่มขึ้น เหตุผลเหมือนกับข้อที่ 2. ในส่วนที่ 1
4. เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่าเพิ่มขึ้น ค่า AMSE ของแต่ละวิธีมีแนวโน้มเพิ่มขึ้น เหตุผลเหมือนกับข้อ 4. ในส่วนที่ 1

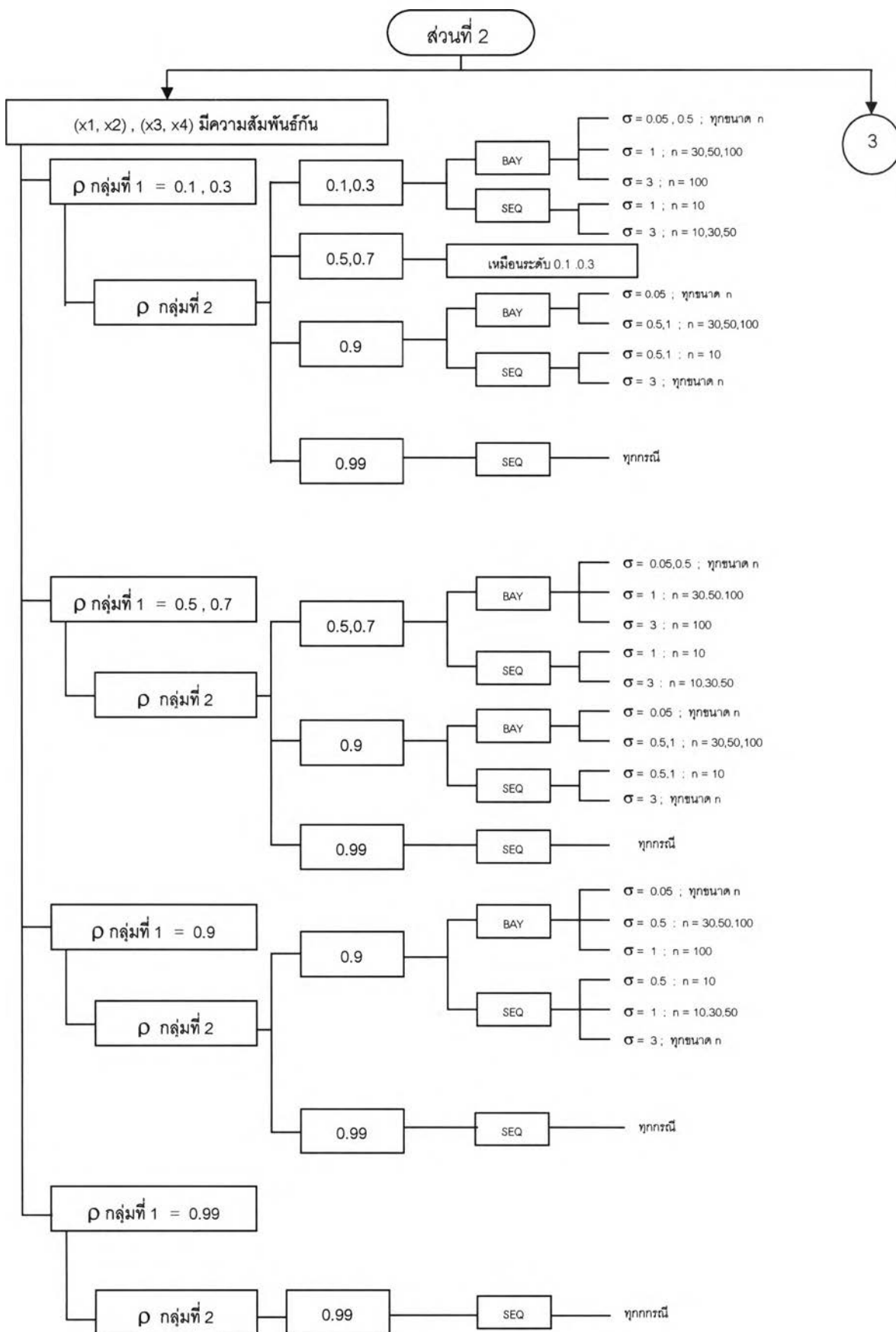
สำหรับผลสรุปในแต่ละส่วน สามารถเขียนได้ดังแผนผังในหน้าต่อไป

สำหรับผลสรุปในแต่ละส่วน สามารถเขียนแผนผังได้ดังต่อไปนี้  
ส่วนที่ 1 เมื่อจำนวนตัวแปรอิสระมีความสัมพันธ์กัน 1 กลุ่ม

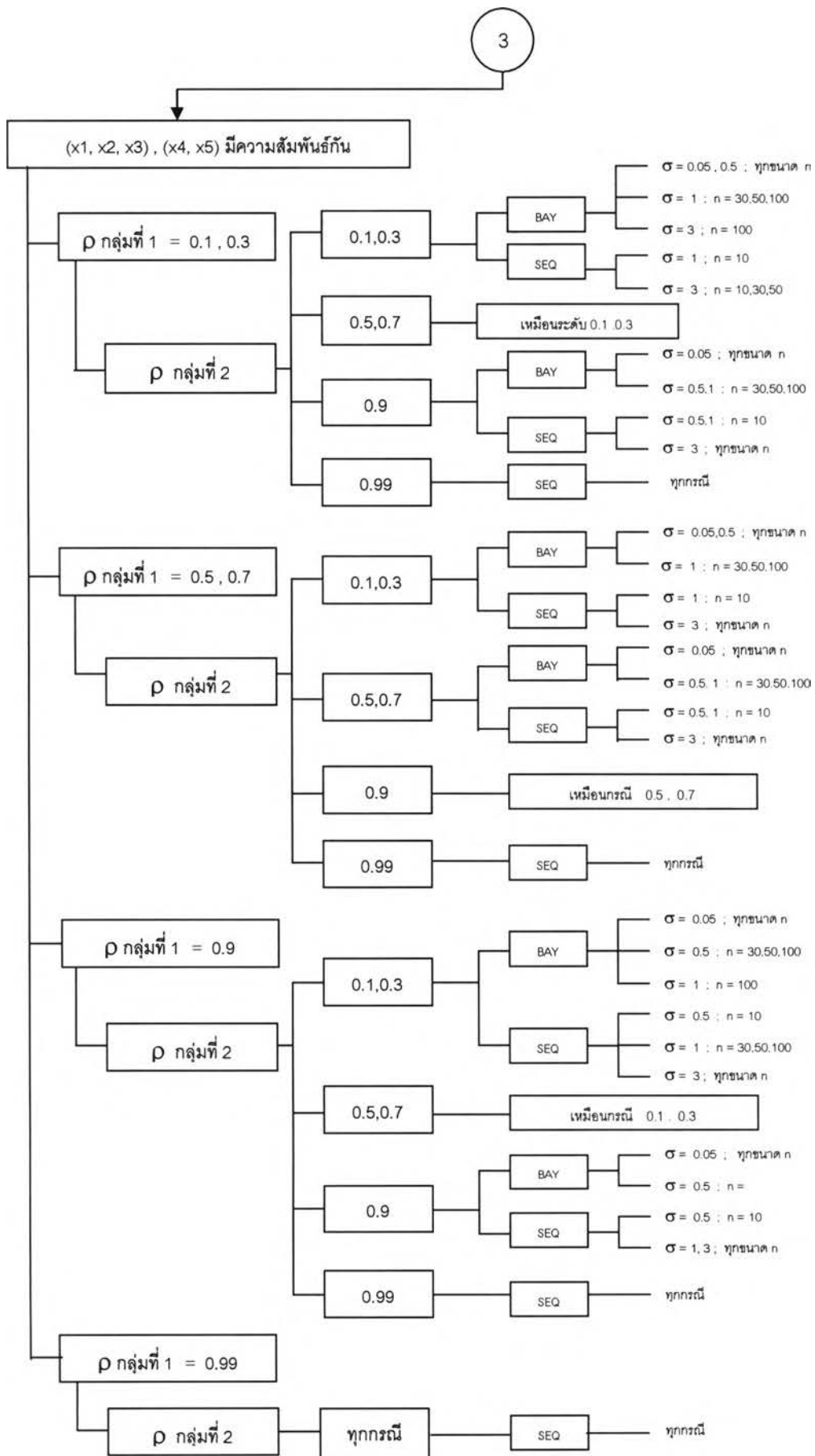




ส่วนที่ 2 เมื่อจำนวนตัวแปรอิสระสัมพันธ์กัน 2 กลุ่ม



หมายเหตุ : ρ (กลุ่มที่ 1, กลุ่มที่ 2) = ρ (กลุ่มที่ 2, กลุ่มที่ 1)



จากแผนผังข้างต้นบางครั้งอาจจะยุ่งยากในการตัดสินใจ ผู้วิจัยจึงสรุปเป็นภาพรวมโดยแบ่งเป็นกลุ่มระดับความสัมพันธ์ดังต่อไปนี้

- ระดับต่ำ  $\rho = 0.1$  และ  $0.3$
- ระดับปานกลาง  $\rho = 0.5$  และ  $0.7$
- ระดับสูง  $\rho = 0.9$
- ระดับสูงมาก  $\rho = 0.99$

ตารางที่ 5.1 สรุปภาพรวมส่วนที่ 1

จำนวนตัวแปรอิสระ	ระดับความสัมพันธ์	$\sigma$			
		0.05	0.5	1	3
3	ต่ำ	BAY	BAY	BAY	BAY**
	ปานกลาง	BAY	BAY	BAY	BAY**
	สูง	BAY	BAY	BAY	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ
5	ต่ำ	BAY	BAY	BAY*	SEQ***
	ปานกลาง	BAY	BAY	BAY*	SEQ
	สูง	BAY	BAY	SEQ***	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ

หมายเหตุ \* หมายถึง ยกเว้นขนาดตัวอย่างเท่ากับ 10 วิธี SEQ ให้ผลดี

\*\* หมายถึง ยกเว้นขนาดตัวอย่างเท่ากับ 10 และ 30 วิธี SEQ ให้ผลดี

\*\*\* หมายถึง ยกเว้นขนาดตัวอย่างเท่ากับ 100 วิธี BAY ให้ผลดี

จากตารางที่ 5.1 จะเห็นได้ว่า วิธี BAY จะให้ผลดีโดยส่วนใหญ่ เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ ปานกลาง และสูง ในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เท่ากับ 0.05, 0.5 และ 1 และกรณีที่จำนวนตัวแปรอิสระเท่ากับ 5 ระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในระดับต่ำ ปานกลาง และสูง ในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เท่ากับ 0.05 และ 0.5 แต่เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เท่ากับ 1 วิธี BAY จะให้ผลดี ในกรณีที่ขนาดตัวอย่างเท่ากับ 30, 50 และ 100 ส่วนวิธี SEQ จะให้ผลดีโดยส่วนใหญ่ เมื่อระดับความสัมพันธ์ของตัวแปรอิสระมีค่าสูงมากหรือในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เท่ากับ 3 (ทั้งจำนวนตัวแปรอิสระเท่า 3 และ 5)

ตารางที่ 5.2 สรุปภาพรวมในส่วนที่ 2

ระดับ ความสัมพันธ์ กลุ่มที่ 1	ระดับ ความสัมพันธ์ กลุ่มที่ 2	$\sigma$			
		0.05	0.5	1	3
ต่ำ	ต่ำ	BAY	BAY	BAY*	SEQ***
	ปานกลาง	BAY	BAY	BAY*	SEQ***
	สูง	BAY	BAY	BAY*	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ
ปานกลาง	ต่ำ	BAY	BAY	BAY*	SEQ***
	ปานกลาง	BAY	BAY	BAY*	SEQ***
	สูง	BAY	BAY*	BAY*	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ
สูง	ต่ำ	BAY	BAY*	BAY*	SEQ
	ปานกลาง	BAY	BAY*	BAY*	SEQ
	สูง	BAY	BAY*	SEQ***	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ
สูงมาก	ต่ำ	SEQ	SEQ	SEQ	SEQ
	ปานกลาง	SEQ	SEQ	SEQ	SEQ
	สูง	SEQ	SEQ	SEQ	SEQ
	สูงมาก	SEQ	SEQ	SEQ	SEQ

หมายเหตุ \* หมายถึง ยกเว้นขนาดตัวอย่างเท่ากับ 10 วิธี SEQ ให้ผลดี

\*\*\* หมายถึง ยกเว้นขนาดตัวอย่างเท่ากับ 100 วิธี BAY ให้ผลดี

จากตารางที่ 5.2 วิธี BAY จะให้ผลดีโดยส่วนใหญ่ เมื่อระดับความสัมพันธ์ของตัวแปรอิสระทั้งในกลุ่มที่ 1 และกลุ่มที่ 2 มีค่าอยู่ในระดับต่ำ ปานกลาง ในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่าเท่ากับ 0.05 แต่เมื่อระดับความสัมพันธ์ของตัวแปรอิสระในกลุ่มใดกลุ่มหนึ่งมีค่าสูง (เฉพาะกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่าเท่ากับ 1) หรือในกรณีที่ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 1 วิธี BAY จะให้ผลดี ในกรณีที่ขนาดตัวอย่างเท่ากับ 30, 50 และ 100 ส่วนวิธี SEQ จะให้ผลดีโดยส่วนใหญ่ เมื่อระดับความสัมพันธ์ของตัวแปรอิสระในกลุ่มใดกลุ่มหนึ่งมีค่าสูงมากหรือในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่าเท่ากับ 3



จากการวิจัยในครั้งนี้ ผลสรุปที่ได้คือ วิธี BAY จะเป็นวิธีการหาค่า  $k$  ในการถดถอยแบบบริดจ์ที่ให้ผลดีในการวิจัยครั้งนี้ เพราะว่าสามารถแก้ปัญหาในกรณีที่เกิดพหุสัมพันธ์ระดับต่ำ ปานกลาง และสูงได้ดี โดยเฉพาะในกรณีที่ขนาดตัวอย่างใหญ่ ส่วนวิธี SEQ จะเป็นวิธีที่สามารถแก้ปัญหาในกรณีที่เกิดพหุสัมพันธ์ในระดับสูงมากหรือในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่ามากได้ แต่ในกรณีที่เกิดพหุสัมพันธ์ในระดับสูงมาก (ในกรณีที่  $\sigma = 0.05, 0.5$  และขนาดตัวอย่างขนาดใหญ่) วิธี BAY กับวิธี SEQ ให้ผลใกล้เคียงกัน เพราะเมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่าน้อย แสดงว่าข้อมูลที่เราวิจัยมีความเบี่ยงเบนไม่มากนัก อีกทั้งเมื่อขนาดตัวอย่างขนาดใหญ่จะสามารถลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ ทำให้ประสิทธิภาพของทั้งสองวิธีมีค่าใกล้เคียงกัน

แต่อย่างไรก็ตามจากการศึกษาในครั้งนี้ได้พบว่า ในกรณีที่ขนาดความสัมพันธ์ระหว่างตัวแปรอิสระมีค่าสูงมาก ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่ามาก ( $\sigma = 3$ ) และขนาดตัวอย่างน้อย ( $N=10$ ) ทั้ง 3 วิธีจะให้ค่า AMSE ที่มีค่ามาก เนื่องจากจำนวนข้อมูลที่เราวิจัยจะไม่เพียงพอในการอธิบายตัวแปรตามได้ถูกต้อง

จากการวิจัยทั้งในส่วนที่ 1 และส่วนที่ 2 เมื่อพิจารณาถึงค่า AMSE ซึ่งเป็นเกณฑ์ที่ใช้เปรียบเทียบตัวประมาณสัมประสิทธิ์การถดถอยบริดจ์ทั้ง 3 วิธี ในครั้งนี้นั้น ปัจจัยที่ผลต่อค่าดังกล่าวได้แก่

1. ขนาดตัวอย่าง
2. ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน
3. จำนวนตัวแปรอิสระ
4. จำนวนตัวแปรอิสระที่มีความสัมพันธ์กัน
5. ขนาดความสัมพันธ์ของตัวแปรอิสระ

โดยที่ปัจจัยที่แปรผกผันกับค่า AMSE คือ ขนาดตัวอย่าง จำนวนตัวแปรอิสระ ส่วนปัจจัยที่แปรผันตามกับค่า AMSE คือ ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน จำนวนตัวแปรอิสระที่มีความสัมพันธ์กัน และ ขนาดความสัมพันธ์ของตัวแปรอิสระ

## 5.2 ข้อเสนอแนะ

### 1. ด้านการศึกษา

- ควรศึกษาเพิ่มเติมในกรณีจำนวนตัวแปรอิสระมากกว่า 3 และ 5
- ควรศึกษาเพิ่มเติมในกรณีที่ขนาดตัวอย่างมากกว่า 100
- ควรศึกษาเพิ่มเติมในกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่ามากกว่า 3
- ควรศึกษาวิธีการสถิติอื่นๆ ที่สามารถแก้ปัญหาในกรณีที่ตัวแปรอิสระมีความสัมพันธ์กัน เช่น วิธีวิเคราะห์องค์ประกอบหลัก(Principle component) เป็นต้น
- ควรทำการศึกษาเพิ่มเติม ในส่วนของการหาค่า  $k$  จากวิธีการอื่นๆ นอกเหนือจากวิธีของไฮเอิร์นเคนนาร์ด วิธีค้นหาข้อมูลแบบลำดับ และวิธีเบส

### 2. ด้านการนำไปใช้ประโยชน์

- ควรตรวจสอบลักษณะของข้อมูล เช่น จำนวนตัวแปรอิสระ ระดับความสัมพันธ์ ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเบื้องต้นก่อนว่ามีลักษณะอย่างไร เพื่อที่จะได้เลือกใช้วิธีประมาณค่าได้อย่างเหมาะสม
- เนื่องจากในบางกรณีวิธีการประมาณค่าทั้ง 3 วิธีในการศึกษาค้างนี้ อาจให้ค่าส่วนเบี่ยงเบนมาตรฐาน (SD) ของค่า AMSE ที่มีค่ามาก โดยเฉพาะกรณีที่ขนาดตัวอย่างน้อย ระดับความสัมพันธ์ระหว่างตัวแปรอิสระมีค่าสูงมากและกรณีที่ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีค่ามาก ดังนั้นในกรณีดังกล่าวควรจะมีการศึกษาถึงผลของการนำไปใช้ด้วย หรืออาจจะต้องมีการหาวิธีการหาค่า  $k$  อื่นๆ นอกเหนือจากนี้เพื่อให้การประมาณค่าถูกต้องมากขึ้น