

ผลสรุปการวิเคราะห์และการอภิปรายผล

การวิจัยนี้ได้เสนอวิธีการในการจำแนกกลุ่ม 2 กลุ่ม โดยเปรียบเทียบระหว่างการวิเคราะห์จำแนกประเภทและการวิเคราะห์การถดถอยทวิ เพื่อแก้ไขปัญหาที่พบตามที่กล่าวมาแล้วในบทที่ 1 การศึกษาจะพิจารณาในกรณีที่ประชากรของตัวแปรอิสระจะมีการแจกแจงแบบเบ้ แต่การแจกแจงตามข้อสมมติเบื้องต้นของการจำแนกกลุ่มตัวแปรอิสระมีการแจกแจงแบบปกติ สำหรับในการวิจัยครั้งนี้ได้พิจารณาในกรณีของการแจกแจงแบบปกติและการวิเคราะห์การถดถอยโดยวิธีกำลังสองน้อยที่สุดแบบธรรมดา (OLS) ประกอบการศึกษาด้วยเหตุผลที่ว่า

1. กรณีตัวแปรอิสระมีการแจกแจงแบบปกติ ถ้าเราวิเคราะห์ด้วยตัวประมาณตามวิธี OLS กับวิธีใหม่ที่น่าเสนอ จะทำให้ทราบว่าเราควรจะใช้วิธีการใดเมื่อเปรียบเทียบกับวิธี OLS ที่เคยใช้อยู่

2. ถ้าตัวแปรอิสระมีการแจกแจงที่ไม่ใช่การแจกแจงแบบปกติ (ในที่นี้จะศึกษากรณีการแจกแจงแบบเบ้) จะให้ผลแตกต่างกันหรือไม่ระหว่างวิธี OLS กับวิธีใหม่ที่น่าเสนอ เนื่องจากการเปลี่ยนแปลงในลักษณะที่เพิ่มขึ้นของสัดส่วนตัวแปรตามซึ่งมีผลทำให้ความแตกต่างระหว่างวิธีการมีความชัดเจนมากยิ่งขึ้น ไม่ว่าจะพิจารณาจากค่าเฉลี่ยของร้อยละที่พยากรณ์ถูกต้องจากทั้งหมดหรือค่าเฉลี่ยของร้อยละที่พยากรณ์ถูกต้องบางส่วนจากส่วนที่เหลือ (เมื่อหักออกด้วยจำนวนที่ถูกต้องเหมือนกัน) ทั้ง 2 กรณี จะให้ผลลัพธ์สอดคล้องกัน แต่กรณีที่ 2 จะให้ความชัดเจนมากกว่าและสามารถบอกรายละเอียดอื่น ๆ ได้อีก เช่น ในบางวิธีการ ถึงแม้ว่าค่าเฉลี่ยของร้อยละที่พยากรณ์ถูกต้องจากทั้งหมดเท่ากัน แต่เมื่อพิจารณาในรายละเอียดอาจพบว่ามีส่วนที่พยากรณ์ได้ไม่ตรงกัน ดังนั้น การสรุปผลการวิเคราะห์จะพิจารณาจากกรณีค่าเฉลี่ยของร้อยละที่พยากรณ์ถูกต้องบางส่วน

การวิจัยนี้ศึกษาจากข้อมูลในแต่ละประชากร ที่มีค่าสัมประสิทธิ์ความแปรผันโดยเฉลี่ย 100% และที่สำคัญคือ กำหนดให้มีค่าเฉลี่ยและความแปรปรวนใกล้เคียงกันมากที่สุด สำหรับการนำผลการวิจัยนี้ไปใช้นักวิจัยสามารถอ้างอิงได้กับประชากรที่มีค่าสัมประสิทธิ์ความแปรผันต่ำกว่านี้ และยังให้ผลสอดคล้องกับค่าสัมประสิทธิ์ความแปรผันที่ใช้ในการวิจัยครั้งนี้

ในการจำลองค่าข้อมูลที่ใช้ในการทดลองจำนวนตัวแปรอิสระนับได้ว่าเป็นปัจจัยหนึ่งที่ควรสนใจศึกษาซึ่งอาจจะมีอิทธิพลต่อการเปลี่ยนแปลงของแต่ละวิธีการพยากรณ์ ดังนั้นการวิจัยจึงได้ศึกษาที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวแปรและ 2 ตัวแปร เพื่อหาข้อสรุปเกี่ยวกับปัจจัยดังกล่าว ซึ่งการศึกษานี้พบว่าจำนวนตัวแปรอิสระไม่มีอิทธิพลต่อวิธีการพยากรณ์ถึงแม้ว่าผลที่ได้ในบางกรณีดูเหมือนว่าจำนวนตัวแปรอิสระที่มากกว่า 1 ตัวแปรจะให้ผลดีกว่าแต่จากการทดสอบความแตกต่างทางสถิติจะพบว่าความแตกต่างนั้นไม่มีนัยสำคัญ แต่ในทางปฏิบัติผู้วิจัยควรพิจารณาถึงลักษณะของตัวแปรอิสระด้วย

สำหรับการอภิปรายผลจะแบ่งตามสัดส่วนของตัวแปรตามโดยพิจารณาจากกรณีที่จำนวนตัวแปรอิสระเท่ากับ 1 ตัวแปร

ในการสรุปผล พิจารณาที่ปัจจัย ต่อไปนี้

- 1.) วิธีการพยากรณ์หรือวิธีการจำแนกกลุ่ม
- 2.) สัดส่วนตัวแปรตาม
- 3.) การแจกแจงของตัวแปรอิสระ
- 4.) ขนาดตัวอย่าง
- 5.) จำนวนตัวแปรอิสระ

ซึ่งสามารถสรุปผลการวิเคราะห์เพื่อสนับสนุนสมมติฐานที่กำหนดไว้ดังนี้

5.1 วิธีการพยากรณ์และสัดส่วนของตัวแปรตาม

5.1.1 กรณีสัดส่วนตัวแปรตามมีค่ามากกว่า 0.50 สำหรับทุกการแจกแจงและทุกขนาดตัวอย่างวิธีการที่ใช้พยากรณ์ต่าง ๆ ให้ประสิทธิภาพเรียงตามลำดับจากมากไปน้อยดังนี้

1. การวิเคราะห์การจำแนกประเภทเมื่อพิจารณาค่า prior probability ประกอบการจำแนกกลุ่ม
2. การวิเคราะห์การถดถอยทวิ เมื่อใช้การปรับค่าพยากรณ์โดยเส้นโค้งปกติ (NOR)
3. การวิเคราะห์การถดถอยทวิ เมื่อใช้การปรับค่าพยากรณ์โดยเส้นโค้งโลจิสติก (LOG)
4. การวิเคราะห์การถดถอย โดยวิธีกำลังสองน้อยที่สุดแบบธรรมดา (OLS) ให้ผลลัพธ์ไม่แตกต่างกับการวิเคราะห์การจำแนกประเภทกรณีที่ไม่ได้พิจารณาค่า prior probability ในการจำแนกกลุ่ม (DSC)

5. การวิเคราะห์การถดถอยทวิ เมื่อใช้การปรับค่าพยากรณ์โดยเส้นโค้งคอมเพิร์ทซ์ (COM)

ถ้าพิจารณาเฉพาะในกลุ่มของวิธีการพยากรณ์จะพบว่า

1. การวิเคราะห์จำแนกประเภท กรณีพิจารณาค่า Prior Probability ประกอบการจำแนกกลุ่ม จะให้ประสิทธิภาพดีกว่าการจำแนกกลุ่มโดยไม่พิจารณาค่า prior probability
2. สำหรับการวิเคราะห์การถดถอย จะให้ผลเรียงตามลำดับจากมากไปน้อย ดังนี้
 - 2.1 Normit model (NOR)
 - 2.2 Logit model (LOG)
 - 2.3 Ordinary Least Square (OLS)
 - 2.4 Gompit model (COM)

5.1.2 กรณีสัดส่วนตัวแปรตามมีค่าเท่ากับ 0.5 ทุกวิธีการจะให้ประสิทธิภาพในการจำแนกกลุ่มใกล้เคียงกันมากที่สุด เมื่อเปรียบเทียบกับสัดส่วนอื่น ๆ

ดังนั้น ไม่ว่าจะพยากรณ์หรือจำแนกกลุ่มโดยวิธีใดก็ตาม ถ้าสัดส่วนของตัวแปรตามเท่ากับ 0.50 จะให้ความเสี่ยงต่ำสุด แต่ถ้าจะให้เรียงลำดับความสำคัญจากมากไปน้อยของแต่ละวิธีการ สามารถสรุปได้ดังตารางที่ 5.1

5.1.3 อันดับความสำคัญของวิธีการพยากรณ์

จากปัจจัยต่าง ๆ ที่ศึกษาคือ สัดส่วนของตัวแปรตาม จำนวนตัวแปรอิสระ ความน่าจะเป็นโดยหลักเกณฑ์ ขนาดตัวอย่างและการแจกแจงของตัวแปรอิสระ เมื่อเรากำหนดปัจจัยเหล่านี้ให้กับวิธีการพยากรณ์ทั้ง 6 วิธี จะพบว่าความสำคัญของอันดับทั้ง 6 วิธีจะไม่เปลี่ยนแปลงทุกกรณีที่ศึกษาส่วนอิทธิพลที่มีต่อร้อยละที่พยากรณ์ถูกต้องนั้น จะต้องศึกษาเป็นกรณีสำหรับแต่ละวิธีการพยากรณ์ที่ต้องการ .

ตารางที่ 5.1 แสดงอันดับของประสิทธิภาพในการจำแนกกลุ่ม โดยจำแนกตามวิธีการพยากรณ์และขนาดตัวอย่าง เมื่อสัดส่วนตัวแปรตามเท่ากับ 0.50

การแจกแจง	ขนาดตัวอย่าง	วิธีการพยากรณ์กรณีตัวแปรอิสระ 1 ตัวแปร					
		OLS	NOR	LOG	COM	DBC	EMC
แบบปกติ	10	3	1*	3	2	3	3
	30	5	3	4	1*	5	2
	60	5	3	4	2	5	1*
แบบลอกนอร์มอล	10	3	3	1*	4	3	2
	30	4	2	3	5	4	1*
	60	3	2	2	4	3	1*
แบบไวบูลล์	10	3	4	1*	2	3	3
	30	2	3	4	5	2	1*
	60	3	4	5	2	3	1*
แบบแกมมา	10	1*	3	2	4	1*	1*
	30	5	3	4	2	5	1*
	60	3	5	4	3	2	1*

* หมายถึง วิธีการพยากรณ์ที่เหมาะสมกับการแจกแจงใดการแจกแจงหนึ่งจากทั้งหมด 4 การแจกแจง ในแต่ละขนาดตัวอย่าง

5.2 การแจกแจงของตัวแปรอิสระ

ในการวิเคราะห์จำแนกประเภทและการวิเคราะห์การถดถอยทวิ ที่ศึกษาในครั้งนี้ ตัวแปรตามเป็นข้อมูลเชิงคุณภาพ มีค่าเป็น 1 หรือ 0 ให้ข้อสรุปว่า

5.2.1. กรณีสัดส่วนตัวแปรตามมากกว่า 0.50

สำหรับแต่ละวิธีการพยากรณ์ เมื่อนำการแจกแจงต่าง ๆ เปรียบเทียบกัน จะพบว่าลักษณะการแจกแจงของตัวแปรอิสระมีอิทธิพลต่อประสิทธิภาพการจำแนกกลุ่มโดยที่ การแจกแจงแบบปกติจะให้ประสิทธิภาพสูงกว่าการแจกแจงแบบเบ้เมื่อจำนวนตัวอย่างมีขนาดใหญ่ ส่วนการแจกแจงแบบเบ้จะให้ประสิทธิภาพใกล้เคียงกันทุกการแจกแจงและลดลงในลักษณะเบ้ขวา เมื่อสัดส่วนของตัวแปรตามมีค่าเข้าใกล้ 1 ซึ่งผลที่ได้นี้จะเกิดขึ้นในลักษณะเดียวกันทุกกรณี¹⁶ ที่ศึกษา

5.2.2 กรณีสัดส่วนตัวแปรตามเท่ากับ 0.50

เมื่อนำมาพิจารณาในรายละเอียดโดยรวมแล้ว อิทธิพลของการแจกแจงแบบต่าง ๆ ในแต่ละวิธีการพยากรณ์จะให้ผลลัพธ์ไม่แตกต่างกันมากและความแตกต่างมีรูปแบบไม่แน่นอน ในแต่ละวิธี ในทางปฏิบัติควรพิจารณาในละเอียดแต่ละกรณีที่น่าเสนอในภาคผนวก ข. ในที่นี้จะพิจารณาตามผลที่เกิดขึ้นจริงเรียงลำดับจากมากไปหาน้อย ดังตารางที่ 5.2

5.3 ขนาดตัวอย่าง

ในการวิจัยครั้งนี้การเพิ่มขนาดตัวอย่างไม่มีอิทธิพลต่อประสิทธิภาพในการจำแนกกลุ่ม เนื่องจากการสุ่มตัวอย่างจากประชากรในครั้งนี้จะสุ่มคละกันไปในทุก ๆ ลักษณะของประชากร ดังนั้นการเพิ่มขนาดตัวอย่างจึงไม่มีอิทธิพลต่อการจำแนกกลุ่มถ้าประชากรทั้งสองกลุ่มยังมีลักษณะที่ไม่แตกต่างกัน

ถ้าพิจารณาที่การแจกแจงของตัวแปรอิสระจะให้ข้อสรุปว่าที่ขนาดตัวอย่างขนาดเล็ก การแจกแจงแบบเบ้จะให้ประสิทธิภาพดีกว่าแต่ถ้าเพิ่มขนาดตัวอย่างให้มีขนาดใหญ่การแจกแจงแบบปกติจะให้ผลดีกว่าและมีประสิทธิภาพคงเส้นคงวามากกว่า สำหรับการแจกแจงแบบเบ้เมื่อ สัดส่วนตัวแปรตามมีค่าเข้าใกล้ 1 ประสิทธิภาพจะลดลงในลักษณะเบ้ขวา

ตารางที่ 5.2 แสดงอันดับของประสิทธิภาพในการจำแนกกลุ่ม โดยเรียงอันดับตามการแจกแจงในแต่ละวิธีการพยากรณ์

ขนาดตัวอย่าง	การแจกแจง	วิธีการพยากรณ์กรณีตัวแปรอิสระ 1 ตัวแปร					
		OLB	NOR	LOG	COM	DBC	EMC
10	แบบปกติ	2	1*	4	2	1*	3
	แบบลอกนอร์มอล	2	2	2	3	3	2
	แบบไวบูลล์	2	3	1*	1*	3	3
	แบบแกมมา	1*	2	3	3	2	1*
30	แบบปกติ	3	2	2	1*	3	2
	แบบลอกนอร์มอล	2	1*	1*	3	2	1*
	แบบไวบูลล์	1*	3	3	3	1*	1*
	แบบแกมมา	3	2	2	2	3	1*
60	แบบปกติ	3	1*	2	1*	3	1*
	แบบลอกนอร์มอล	2	1*	1*	2	2	1*
	แบบไวบูลล์	1*	2	4	1*	1*	1*
	แบบแกมมา	1*	3	3	1*	1*	1*

* หมายถึง วิธีการพยากรณ์ที่เหมาะสมกับการแจกแจงใดการแจกแจงหนึ่งจากทั้งหมด 4 การแจกแจง

5.4 ข้อเสนอแนะ

ในการจำแนกกลุ่ม 2 กลุ่ม วิธีการที่ใช้ในการจำแนกกลุ่มที่ควรเลือกใช้ คือ การวิเคราะห์จำแนกประเภท กรณีที่พิจารณาใช้ค่า prior probability ในการจำแนกกลุ่ม แต่สำหรับการวิเคราะห์การถดถอยควรเลือกใช้วิธี Estimated Generalize Least Square (EGLS) ที่มีการปรับค่าพยากรณ์ด้วยเส้นโค้งปกติ (Normal model)

สำหรับการศึกษาในครั้งต่อไปผู้เขียนมีข้อเสนอแนะสำหรับผู้สนใจศึกษาดังนี้

1. การวิจัยครั้งนี้ศึกษากรณี ตัวแปรตามเป็นข้อมูลเชิงคุณภาพที่มีค่าเป็น 0 หรือ 1 คือมี 2 กลุ่มประชากร ผู้สนใจอาจจะศึกษาในกรณีที่ตัวแปรตามมีจำนวนกลุ่มมากกว่า 2 กลุ่มประชากรและตัวแปรอิสระเป็นตัวแปรหุ่น (Dummy Variable)

2. ในการปรับค่าพยากรณ์สำหรับตัวแปรตามอาจจะใช้เส้นโค้งของการแจกแจงที่มีลักษณะเฉพาะเช่น เส้นโค้งของการแจกแจงเบ้ แบบหางยาวกว่าการแจกแจงแบบปกติ

3. นอกจากสัดส่วนของตัวแปรตามที่ได้เห็นได้ชัดเจนแล้วว่ามียุทธศาสตร์ต่อการจำแนกกลุ่ม สิ่งที่ต้องพิจารณาประกอบการจำแนกกลุ่มคือ ลักษณะเฉพาะต่าง ๆ ในแต่ละประชากรที่ต้องคำนึงถึง เพราะว่าถึงแม้เราจะสุ่มตัวอย่างเพิ่มขึ้นจากเดิมก็ตามแต่ผลลัพธ์ที่ได้จะไม่แตกต่างจากเดิม ซึ่งอาจเนื่องมาจากประชากรทั้ง 2 กลุ่ม มีลักษณะคล้ายคลึงกันมาก ถ้ามีเกณฑ์ในการพิจารณาศึกษาจะทำให้ทราบรายละเอียดแต่ละวิธีการพยากรณ์มากขึ้น เกณฑ์ที่ควรพิจารณา เช่น

3.1 ประชากรทั้งสองกลุ่มมีลักษณะคล้ายคลึงกันมาก

3.2 ประชากรทั้งสองกลุ่มมีลักษณะแตกต่างกันมาก

4. เนื่องจากการแจกแจงแบบปกติให้ประสิทธิภาพสูงกว่า ดังนั้น การแปลงข้อมูลจากประชากรที่มีการแจกแจงแบบเบ้เพื่อให้ข้อมูลมีลักษณะเข้าใกล้การแจกแจงแบบปกติจึงเป็นสิ่งที่น่าสนใจศึกษาอีกกรณีหนึ่ง

5.4 ผังงานแสดงการใช้ผลการวิจัย

