



บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยในครั้งนี้เป็นการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณแต่ละตัวโดยใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบซึ่งทำการศึกษาในกรณีที่ประชากรทั้งสองกลุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยในกลุ่มที่ 1 เป็น ศูนย์ ค่าเฉลี่ยในกลุ่มที่ 2 เป็น $(\Delta, 0, 0, \dots, 0)$ และความแปรปรวนเป็น 1 ทั้งสองกลุ่ม ขนาดตัวอย่างเป็น 10, 20, 25, 50 และ 100 ขนาดของตัวแปรอิสระเป็น 3, 5, 7 และ 9 ลักษณะการแยกจากกันของประชากรจะเป็นตัวกำหนดการแจกแจงของประชากรกลุ่มที่ 2 นั่นคือค่าของ Δ มีค่าเป็น 0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 แต่ละการทดลองกระทำซ้ำ 500 ครั้ง จำนวนครั้งในการสุ่มตัวอย่างโดยวิธี Bootstrap เท่ากับ 100 ครั้ง

5.1 สรุปผลการวิจัย

5.1.1 ขนาดตัวอย่าง จำนวนตัวแปรอิสระ และ ลักษณะการแยกจากกันของประชากรมีผลต่อความคลาดเคลื่อนกำลังสองเฉลี่ย

5.1.2 เมื่อลักษณะการแยกจากกันของประชากรเด่นชัดขึ้น นั่นคือ Δ มีค่ามากขึ้นค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของทุกตัวประมาณจะมีค่าลดลงแสดงให้เห็นว่าลักษณะการแยกจากกันของประชากรทั้งสองกลุ่มมีผลต่อค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ถ้าประชากรแยกออกจากกันอย่างชัดเจนมากขึ้น การระบุกลุ่มให้กับค่าสังเกตก็มีโอกาสผิดพลาดน้อยลงจึงทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าลดลง

5.1.3 เมื่อขนาดตัวอย่างคงที่การเพิ่มจำนวนตัวแปรอิสระมีผลทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยในตัวประมาณวิธี R และ ตัวประมาณ DS มีค่าเพิ่มขึ้นยกเว้นตัวประมาณวิธี U จะให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ลดลง แสดงให้เห็นว่าขณะที่ขนาดของตัวแปรอิสระในคำสั่งเกตเพิ่มขึ้นวิธี U ได้ปรับปรุงตัวเองดีขึ้น
ตัวประมาณตัวอื่น

5.1.4 เมื่อขนาดตัวอย่างเพิ่มมากขึ้นโดยส่วนใหญ่ตัวประมาณทั้ง 4 ตัวจะมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ลดลง แสดงว่าการเพิ่มขนาดตัวอย่างให้ใหญ่ขึ้นจะช่วยเพิ่มความถูกต้องแม่นยำในการประมาณค่าอัตราความผิดพลาด มากขึ้น

5.1.5 ตัวประมาณโดยวิธี R มีประสิทธิภาพในการประมาณค่าอัตราความผิดพลาดน้อยมาก จากการพิจารณาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยจะสูงกว่าตัวประมาณตัวอื่นๆ โดยเฉพาะในกรณีที่ ขนาดตัวอย่างที่มีขนาดเล็กและลักษณะการแยกจากกันของประชากรทั้งสองกลุ่มแยกจากกันไม่ชัดเจน

5.1.6 ตัวประมาณ Bootstrap หรือ วิธี B จะให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดเมื่อลักษณะการแยกจากกันของประชากรทั้งสองกลุ่มแยกจากกันไม่ชัดเจน นั่นคือเมื่อ Δ มีขนาดอยู่ระหว่าง 0.0 - 0.5 และเมื่อขนาดตัวอย่างมีขนาดเล็กมากคือน้อยกว่าหรือเท่ากับ 10 ขนาดของตัวแปรอิสระมีขนาดใกล้เคียงกับขนาดตัวอย่างตัวประมาณโดยวิธี B จะมีประสิทธิภาพดีกว่าตัวประมาณตัวอื่นๆ

5.1.7 เมื่อขนาดของตัวแปรอิสระมีขนาดเล็กหรือ k มีค่าอยู่ระหว่าง 3 - 7 ตัวประมาณ DS จะให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด

5.1.8 ระหว่างตัวประมาณโดยวิธี U กับตัวประมาณโดยวิธี B เมื่อประชากรทั้ง 2 กลุ่มแยกจากกันค่อนข้างชัดเจน นั่นคือ Δ มากกว่า 0.5 ขึ้นไปวิธี U จะมีประสิทธิภาพดีกว่าวิธี B

5.1.9 เมื่อตัวแปรอิสระมีขนาดใหญ่ นั่นคือ k มากกว่า 7 ขึ้นไปและลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่มมีค่ามากขึ้น นั่นคือค่า Δ มากกว่าหรือเท่ากับ 1 ขึ้นไปตัวประมาณโดยวิธี U จะมีประสิทธิภาพดีกว่าตัวประมาณตัวอื่นๆ

5.2 ข้อเสนอแนะ

สำหรับเกณฑ์การเลือก ใช้ตัวประมาณนั้นจะพิจารณาจากขนาดของตัวแปรอิสระและลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่ม จากสรุปผลการวิจัยที่ได้สามารถสรุปเป็นข้อเสนอแนะได้ดังนี้

5.2.1 เมื่อขนาดของตัวแปรอิสระมีขนาดเล็กถึงขนาดกลาง นั่นคือ k มีค่าอยู่ในช่วง 2 - 7 และลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่มค่อนข้างชัดเจน นั่นคือค่าของ Δ มากกว่าหรือเท่ากับ 1 ควรเลือกใช้ตัวประมาณโดยวิธี DS

5.2.2 เมื่อขนาดตัวอย่างมีขนาดเล็กและขนาดของตัวแปรอิสระมีขนาดใหญ่ใกล้เคียงกับขนาดตัวอย่างควรเลือกใช้วิธี B ในการประมาณค่าอัตราความผิดพลาด

5.2.3 เมื่อขนาดของตัวแปรอิสระมีขนาดใหญ่และลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่มค่อนข้างชัดเจน นั่นคือค่าของ Δ มากกว่าหรือเท่ากับ 1 ควรเลือกตัวประมาณ Leave - one - out หรือ วิธี U

5.2.4 เมื่อลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่มไม่ชัดเจน นั่นคือค่าของ Δ อยู่ระหว่าง 0.0 - 0.5 ควรเลือกตัวประมาณ Bootstrap หรือ วิธี B

5.2.5 ตัวประมาณ Resubstitution หรือวิธี R ควรใช้ในกรณีที่ขนาดตัวอย่างมีขนาดใหญ่มาก เช่น ขนาดตัวอย่าง 50 ขึ้นไปและประชากรทั้งสองกลุ่มแยกจากกันค่อนข้างชัดเจนนั่นคือ Δ มากกว่า 1 ขึ้นไป

5.2.6 ควรหลีกเลี่ยงการใช้ตัวประมาณ Resubstitution หรือวิธี R ในการประมาณค่าอัตราความผิดพลาดเพราะความเอนเอียงที่เกิดขึ้นนั้นเนื่องจากการนำเอาค่าของตัวอย่างมาคำนวณหาดีสคริมิแนนท์และใช้ค่าของตัวอย่างเดิมมาประมาณค่าอัตราความผิดพลาดจากการใช้ดีสคริมิแนนท์ที่สร้างขึ้นจากตัวอย่างเดียวกัน

5.2.7 ในทางปฏิบัติหากสงสัยว่าข้อมูลที่นำมาวิเคราะห์ไม่ใช่การแจกแจงแบบปกติก็ให้ทำการทดสอบดูเมื่อพบว่าไม่เป็นการแจกแจงแบบปกติจริงก็ควรทำการแก้ไขก่อนโดยปรับข้อมูลให้มีการแจกแจงแบบปกติก่อนแล้วจึงใช้ตัวประมาณด้วยวิธี DS ซึ่งเป็นวิธีที่นิยมใช้กันมากอีกทั้งยังทำให้ประหยัดเวลาและต้นทุนค่าใช้จ่าย ควรหลีกเลี่ยงตัวประมาณที่ใช้กันมากอีกวิธีหนึ่งคือตัวประมาณ Resubstitution หรือ วิธี R เนื่องจากการประมาณค่าโดยส่วนใหญ่จะประมาณค่าได้ต่ำกว่าที่เป็นจริง แต่ถ้าลักษณะการแยกจากกันของประชากรทั้ง 2 กลุ่มไม่ชัดเจนและขนาดตัวอย่างมีขนาดจำกัดและเรื่องที่สนใจศึกษาเป็นเรื่องที่สำคัญมากต่อการประมาณค่าอัตราความผิดพลาด เช่น เรื่องที่เกี่ยวกับทางการแพทย์ หรือทางด้านเทคโนโลยีที่สำคัญ ควรเลือกใช้ตัวประมาณ Bootstrap หรือตัวประมาณโดยวิธี B เพราะจะให้ประสิทธิภาพในการประมาณดีที่สุด

5.2.8 การศึกษาครั้งนี้เป็นแนวทางในการนำเอาตัวประมาณโดยวิธี B นี้ไปประมาณค่าพารามิเตอร์อื่นๆ หรือเมื่อขนาดตัวอย่างมีจำนวนจำกัด เพราะว่าวิธีนี้จะทำการสร้างตัวอย่างจากข้อมูลที่มีอยู่แล้วจึงนำไปใช้ประมาณค่าพารามิเตอร์ได้ ซึ่งอาจจะเป็นพารามิเตอร์ที่ไม่มีวิธีอื่นหาได้หรือหาได้โดยวิธีการที่ยุงยากมากกว่า

5.2.9 เกณฑ์การพิจารณาโดยใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยนั้นยังอาจไม่เพียงพอต่อการพิจารณาประสิทธิภาพของตัวประมาณ ดังนั้นควรเพิ่มการพิจารณาค่าความคลาดเคลื่อนมาตรฐาน (Standard Deviation ,SD) และ เปอร์เซนต์ของอัตราความผิดพลาดควบคู่กันไปด้วย

5.2.10 แนวทางที่ควรศึกษาต่อไปคือประสิทธิภาพของตัวประมาณทั้ง 3 ตัวคือวิธี R วิธี U และ วิธี B เมื่อการแจกแจงไม่ใช่การแจกแจงแบบปกติ แต่เป็นการแจกแจงในรูปแบบอื่นโดยพิจารณาประสิทธิภาพของตัวประมาณจะเป็นอย่างไรใช้ได้ไหมในกรณีใดบ้างตัวประมาณตัวใดเด่นที่สุดจึงควรนำจะมีการพิจารณานำไปศึกษาต่อ

รูปที่ 5.1 แสดงแผนผังการทำงานของ การสรุปแนวทางการเลือกใช้ตัวประมาณ

