

บทที่ 5

การวินิจฉัยและวิธีแก้ไขปัญหาเกี่ยวกับการวิเคราะห์ความถดถอยพหุคูณ

DRAFT

วัตถุประสงค์

- 1) สามารถอธิบายเหตุและผลของการเกิดปัญหาเชิงเทคนิคของการวิเคราะห์การถดถอย
- 2) สามารถอธิบายและสามารถตรวจสอบข้อมูลที่ผิดปกติ
- 3) สามารถอธิบายและตรวจสอบปัญหาการแจกแจงของค่าความคลาดเคลื่อน
- 4) สามารถอธิบายและตรวจสอบปัญหาความสัมพันธ์ของตัวแปรอิสระ
- 5) สามารถอธิบายและตรวจสอบปัญหาค่าความแปรปรวนของค่าความคลาดเคลื่อนไม่คงที่
- 6) สามารถอธิบายและทดสอบความไม่เป็นเส้นตรงของสมการถดถอย
- 7) สามารถอธิบายและตรวจสอบปัญหาการกำหนดโมเดล
- 8) สามารถอธิบายและตรวจสอบความเป็นอิสระของค่าความคลาดเคลื่อน

5.1 เหตุและผลของการเกิดปัญหาเชิงเทคนิคของการวิเคราะห์การถดถอย

การวิเคราะห์ความถดถอย (Regression Analysis) นั้นเป็นเครื่องมือหนึ่งที่นิยมในการประยุกต์ใช้สำหรับอธิบายความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระทั้งที่เป็นเชิงเส้นตรง (Linear) และไม่ใช่เส้นตรง (Nonlinear) ในสาขาวิชาการต่างๆ โดยความมุ่งหมายของการวิเคราะห์ก็เพื่อที่จะได้ค่าสัมประสิทธิ์ของตัวแปรอิสระต่างๆ (Estimators) และทิศทางความสัมพันธ์สำหรับการอธิบายแบบจำลอง (Model) ของการวิจัย

อย่างไรก็ตาม การวิเคราะห์ความถดถอย นั้นผู้วิจัยต้องคำนึงถึงข้อสมมติที่อยู่เบื้องหลังของการวิเคราะห์ด้วย ได้แก่ (1) ปัญหาการกำหนดโมเดลผิด (Model Specification error) หรือปัญหาที่มีการละทิ้งตัวแปรสำคัญของแบบจำลอง (2) ปัญหาความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) (3) ปัญหาความแปรปรวนไม่คงที่ (Heteroscedasticity) (4) ปัญหาค่าความคลาดเคลื่อน (Error Term หรือ Residual) แจกแจงไม่ปกติ (5) ปัญหาความสัมพันธ์ที่ไม่เป็นเส้นตรง (Nonlinearity) และ (6) รวมไปถึงลักษณะของข้อมูลที่ผิดปกติ เช่น ปัญหา Outlier หากผู้วิจัยได้ละเลยปัญหาที่ได้อธิบายข้างต้นจะทำให้แบบจำลองนั้นอธิบายปรากฏการณ์ที่ผิดพลาด ชี้นำไปในทางที่ผิด และทำให้แบบจำลองนั้นมีความน่าเชื่อถือลดน้อยลงไป

ดังนั้นก่อนที่จะนำแบบจำลองไปใช้อธิบายปรากฏการณ์ต่างๆ ต้องมั่นใจว่าแบบจำลองนั้นได้รับการเยียวยาปัญหาต่างๆ ที่ได้กล่าวข้างต้นเรียบร้อยแล้ว

1.1 ปัญหาเชิงเทคนิคของการวิเคราะห์การถดถอยด้วยวิธีกำลังสองน้อยที่สุด

ปัญหาเกี่ยวกับการใช้วิธีกำลังสองน้อยที่สุด หรือ Ordinary Least Squares (OLS) นั้นมีความสำคัญอย่างมาก สมการถดถอยที่ประมาณขึ้นมาต้องมีการตรวจสอบถึงปัญหาเหล่านี้และหากไม่เป็นไปตามข้อสมมติหรือเงื่อนไขก็ต้องหาวิธีแก้ไข มิฉะนั้นแล้วจะทำให้สมการที่ประมาณนั้นไม่น่าเชื่อถือ ซึ่งในบทนี้ได้กล่าวเพียงสาเหตุของการเกิดปัญหาและผลที่จะเกิดขึ้นตามมาเท่านั้น

1) การกำหนดโมเดลไม่ถูกต้อง (Specification error)

ปัญหานี้มีความสำคัญมาก เพราะจะทำให้ผลการวิเคราะห์เชื่อถือไม่ได้ ซึ่งมีสองกรณีคือ กรณีแรกเป็นการกำหนดรูปแบบฟังก์ชันถดถอยที่ไม่ถูกต้องสองคือ การละทิ้งตัวแปรที่สำคัญ ปัญหาดังกล่าวจะส่งผลให้ค่าสัมประสิทธิ์มีความลำเอียง (Bias)

2) ปัญหาจากความสัมพันธ์ที่เกิดขึ้นพร้อมกัน (Identification problem)

การวิเคราะห์การถดถอยนั้นเป็นการวิเคราะห์แบบสมการเดียว ซึ่งสามารถอธิบายความสัมพันธ์ทั้งหมดได้ ปัญหาหนึ่งในการประมาณสมการอุปสงค์ก็คือ ระดับราคาหนึ่งนั้นอาจเป็นผลมาจากความสัมพันธ์ของสมการอุปสงค์และสมการอุปทานที่เกิดขึ้นพร้อมๆ กันก็ได้ ปัญหาที่เกิดขึ้นนั้นมักเกิดกับข้อมูลอนุกรมเวลา ในความเป็นจริงฟังก์ชันอุปสงค์สามารถเปลี่ยนแปลงได้โดยเฉพาะช่วงเวลาที่ยาวนาน เพราะอิทธิพลของตัวแปรอื่นๆ เช่น รายได้หรือรสนิยมของผู้บริโภคเปลี่ยนไป ในกรณีนี้การขจัดปัญหาอาจทำได้โดยการรวมเอาตัวแปรที่มีส่วนทำให้อุปสงค์เลื่อนระดับได้เข้าไปในสมการถดถอยด้วย มิฉะนั้นแล้วหากประมาณสมการด้วยวิธี OLS ก็จะทำให้สัมประสิทธิ์มีความลำเอียง

3) ปัญหาตัวแปรอิสระมีความสัมพันธ์กัน (Multicollinearity)

การประมาณสมการถดถอยด้วยวิธีกำลังสองน้อยที่สุดที่มีตัวแปรอิสระมากกว่า 1 ตัวแปรนั้น ในทางทฤษฎีนั้นเราได้สมมติว่าตัวแปรอิสระเหล่านั้นต้องไม่มีความสัมพันธ์กัน การเกิด Multicollinearity นั้นอาจจะเกิดได้หลายสาเหตุ ดังนี้