

บทที่ 1

บทนำ



1.1 ความเป็นมาและความสำคัญของปัญหา

ในการศึกษาเกี่ยวกับการประมาณค่าสัมประสิทธิ์การถดถอยโดยทั่วไปนั้น ค่าประมาณของสัมประสิทธิ์การถดถอยจะได้อาจจากการใช้วิธีการกำลังสองน้อยที่สุดมาประมาณค่า โดยตัวประมาณที่ได้นั้นจะมีคุณสมบัติเป็น Best Linear Unbiased Estimator ในกรณีที่มีสมมติฐานเกี่ยวกับข้อจำกัดของสัมประสิทธิ์การถดถอยเกิดขึ้น ค่าประมาณของสัมประสิทธิ์การถดถอยที่ได้จะมีค่าแตกต่างกันเป็นสองชุด กล่าวคือค่าประมาณชุดหนึ่งจะได้อาจจากการใช้วิธีการกำลังสองน้อยที่สุดประมาณสัมประสิทธิ์การถดถอยตามตัวแบบการถดถอยเมื่อไม่มีข้อจำกัด และค่าประมาณอีกชุดหนึ่งนั้นจะได้อาจจากการใช้วิธีการกำลังสองน้อยที่สุดประมาณสัมประสิทธิ์การถดถอยตามตัวแบบการถดถอยเมื่อมีข้อจำกัด โดยข้อจำกัดของตัวแบบการถดถอยอาจเกิดขึ้นในกรณีที่มีสมมติฐานเกี่ยวกับค่าของสัมประสิทธิ์การถดถอยบางตัวหรือเกี่ยวกับความสัมพันธ์ระหว่างสัมประสิทธิ์การถดถอยบางส่วนในตัวแบบการถดถอย ดังตัวอย่างเช่น

จากตัวแบบการถดถอยเชิงเส้นตรง

$$Y = X\beta + \epsilon$$

เมื่อกำหนดให้

Y เป็น $n \times 1$ เวกเตอร์ของค่าตัวแปรตาม

X เป็นเมทริกซ์ของค่าตัวแปรอิสระที่มีขนาด $n \times k$

β เป็น $k \times 1$ เวกเตอร์ของสัมประสิทธิ์การถดถอย

และ ϵ เป็น $n \times 1$ เวกเตอร์ของความคลาดเคลื่อนที่เกิดขึ้น

โดยที่ n เป็นขนาดตัวอย่างของข้อมูล และ k เป็นจำนวนสัมประสิทธิ์การถดถอย

สมมติฐานที่ต้องการทดสอบเกี่ยวกับสัมประสิทธิ์การถดถอยนั้นสามารถเขียนในรูปทั่วไปได้เป็น $H\beta = h$ เมื่อ H เป็นเมทริกซ์ของค่าคงที่ที่มีขนาดเป็น $m \times k$ และ h เป็น $m \times 1$ เวกเตอร์ของค่าคงที่ใด ๆ โดยที่ m เป็นจำนวนข้อจำกัดของสัมประสิทธิ์การถดถอยที่ต้องการทดสอบสมมติฐาน

ยกตัวอย่างในกรณี $k=3$ H และ h อาจจะมีอยู่ในรูปดังต่อไปนี้คือ

1. เมื่อสัมประสิทธิ์การถดถอยบางตัวมีค่าเท่ากับค่าคงที่ เช่น $\beta_2=3, \beta_3=5$

$$H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad h = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

2. เมื่อสัมประสิทธิ์การถดถอยตัวหนึ่งมีค่ามากกว่าเป็นจำนวนเท่าของสัมประสิทธิ์

อีกตัวหนึ่ง เช่น $\beta_2=3\beta_3$

$$H = \begin{bmatrix} 0 & 1 & -3 \end{bmatrix}; \quad h=0$$

3. เมื่อสัมประสิทธิ์การถดถอยบางส่วนมีความสัมพันธ์กันเชิงเส้นตรง เช่น

$\beta_2+\beta_3=4$

$$H = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}; \quad h=4$$

วิธีการโดยทั่วไปในการที่จะเลือกค่าประมาณชุดใดเป็นค่าประมาณที่แท้จริงของสัมประสิทธิ์การถดถอยนั้นจะใช้ข้อมูลทั้งหมดที่มีในการทดสอบสมมติฐาน ผลจากการทดสอบถ้ายอมรับสมมติฐานตัวประมาณที่ไปก็จะเป็นตัวประมาณสัมประสิทธิ์การถดถอยตามตัวแบบการถดถอยเมื่อมีข้อจำกัด แต่ถ้าผลที่ได้จากการทดสอบปฏิเสธสมมติฐานที่สนใจก็จะใช้ตัวประมาณสัมประสิทธิ์การถดถอยตามตัวแบบการถดถอยเมื่อไม่มีข้อจำกัด แล้วใช้ข้อมูลชุดเดียวกันนั้นคำนวณค่าประมาณสัมประสิทธิ์การถดถอยตามผลที่ได้จากการทดสอบสมมติฐาน ปัญหาในที่นี้คือวิธีการโดยทั่วไปนั้นใช้ข้อมูลทั้งหมดในการทดสอบสมมติฐานและประมาณค่าซึ่งจะทำให้ตัวสถิติที่ใช้ในการทดสอบสมมติฐานและตัวประมาณค่ามีความสัมพันธ์กันอันอาจทำให้ตัวประมาณที่ไม่มีคุณภาพไม่ดีเท่าที่ควรในบางกรณี แต่ถ้าแบ่งขนาดข้อมูลทั้งหมดออกเป็นสองส่วน แล้วใช้ข้อมูลส่วนหนึ่งในการทดสอบสมมติฐานและอีกส่วนหนึ่งในการประมาณค่า แม้ขนาดของข้อมูลจะเล็กลงแต่ก็อาจได้ผลดีกว่าวิธีการโดยทั่วไปได้ จึงควรศึกษาเปรียบเทียบคุณภาพของวิธีการประมาณค่าทั้งสองวิธีว่าวิธีใดจะดีกว่าในกรณีใด และขึ้นอยู่กับตัวพารามิเตอร์ใดบ้าง

1.2 วัตถุประสงค์

วิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเมื่อแบ่ง

ขนาดข้อมูลทั้งหมดออกเป็นสองส่วน โดยใช้ข้อมูลส่วนหนึ่งในการทดสอบสัมมติฐานและข้อมูลอีกส่วนหนึ่งในการประมาณค่า เมื่อทราบผลการทดสอบแล้ว เพื่อเปรียบเทียบวิธีการนี้กับวิธีการที่ใช้กันอยู่โดยทั่วไป ทั้งนี้โดยมีข้อจำกัดว่าข้อมูลทั้งสองส่วนมีขนาดเท่ากัน เป็นอิสระต่อกัน และมีค่าของตัวแปรอิสระชุดเดียวกัน

1.3 ขอบเขตของการศึกษาค้นคว้า

1.3.1 ในการแบ่งขนาดข้อมูลทั้งหมดออกเป็นสองส่วนเพื่อใช้ทดสอบสัมมติฐานส่วนหนึ่งและประมาณค่าสัมประสิทธิ์การถดถอยส่วนหนึ่งนั้น ในที่นี้จะศึกษาในกรณีที่มีข้อมูลสองส่วนนั้นมีขนาดเท่ากัน เป็นอิสระต่อกัน และมีค่าของตัวแปรอิสระชุดเดียวกัน

1.3.2 การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยโดยทั่วไปกับวิธีการที่ใช้ในการศึกษาครั้งนี้จะทำการเปรียบเทียบโดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง เป็นเกณฑ์

1.4 วิธีดำเนินการศึกษาค้นคว้า

ในการศึกษาค้นคว้าครั้งนี้ได้กระทำตามลำดับขั้นตอนดังต่อไปนี้คือ

1.4.1 ศึกษารายละเอียดเกี่ยวกับการประมาณค่า การทดสอบสัมมติฐานเกี่ยวกับค่าหรือความสัมพันธ์ของสัมประสิทธิ์การถดถอย และการคำนวณหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่เกิดขึ้นจากการประมาณค่าสัมประสิทธิ์การถดถอยในกรณีที่ใช้ข้อมูลทั้งหมดที่มีทดสอบสัมมติฐานและประมาณค่า

1.4.2 ประมาณค่าสัมประสิทธิ์การถดถอยเมื่อแบ่งขนาดข้อมูลทั้งหมดออกเป็นสองส่วนที่มีขนาดเท่ากัน โดยข้อมูลทั้งสองส่วนนั้นเป็นอิสระต่อกัน และมีค่าของตัวแปรอิสระชุดเดียวกัน แล้วใช้ข้อมูลส่วนหนึ่งในการทดสอบสัมมติฐานและส่วนหนึ่งในการประมาณค่า

1.4.3 คำนวณหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณในข้อ 1.4.2

1.4.4 เปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณสัมประสิทธิ์การถดถอยที่เกิดขึ้นจากการใช้วิธีการโดยทั่วไปกับวิธีการที่ใช้ในการศึกษาครั้งนี้ประมาณค่า

1.4.5 พิสูจน์ทฤษฎีหรือข้อความที่จำเป็น

1.4.6 ตีความหมายจากผลที่ได้และสรุปผล

1.5 ประโยชน์ที่จะได้รับจากการศึกษาค้นคว้าครั้งนี้

1.5.1 เป็นการศึกษาข้อเท็จจริงเพื่อหาผลสรุปเกี่ยวกับวิธีการใหม่ในการประมาณค่าสัมประสิทธิ์การถดถอยในกรณีที่มีการทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอย

1.5.2 เพื่อเปรียบเทียบวิธีการนี้กับวิธีการที่ใช้อยู่โดยทั่วไปเพื่อที่จะสามารถเลือกใช้ตัวประมาณที่เหมาะสมในการประมาณค่าสัมประสิทธิ์การถดถอยเมื่อต้องการทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอยด้วย



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย