บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์ความถดถอยเชิงพหุ เป็นวิธีวิเคราะห์ที่นิยมใช้กันมากทางจิตวิทยา สังคมศาสตร์ เศรษฐศาสตร์ การแพทย์ ฯลฯ เพื่อศึกษาความสัมพันธ์ของตัวแปรต่าง ๆ ที่เกี่ยวข้องกับเรื่องที่สนใจศึกษา สมการถดถอยจะประกอบด้วยกลุ่มของตัวแปรอิสระ เช่น ศึกษาฟังก์ชันการบริโภค ซึ่งประกอบด้วยกลุ่มของตัวแปรอิสระต่าง ๆ เป็นต้นว่า รายได้ของผู้บริโภค ปริมาณการผลิตสินค้า ระดับราคาสินค้า ปริมาณความต้องการสินค้าของผู้บริโภค เมื่อศึกษาในเรื่องเดียวกันนี้กับกลุ่มตัวอย่างอื่นหรืออาจจะกลุ่มตัวอย่างชุดเดิมแต่ต่างช่วงเวลา ความสัมพันธ์ของฟังก์ชันการบริโภคนั้นยังคงประกอบด้วยกลุ่มของตัวแปรอิสระเดียวกัน จึงต้องการทราบว่า ค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้น ที่ได้จากตัวอย่างทั้ง 2 กลุ่มนั้นเท่ากันหรือไม่ นั่นคืออัตราการเปลี่ยนแปลงของตัวแปรตามต่อตัวแปรอิสระเท่ากันหรือไม่ แต่เนื่องจากช่วงเวลาที่ต่างกันหรือต่างกลุ่มตัวอย่างกัน อาจมีผลทำให้ความแปรปรวนของความคลาดเคลื่อนสุ่มไม่เท่ากัน ดังนั้นเพื่อศึกษาการทดสอบความเท่ากันของสัมประสิทธิ์ความถดถอยเชิงเส้นเมื่อความแปรปรวนของความคลาดเคลื่อนสุ่มไม่เท่ากัน

แบบจำลองทั่วไป (General Model) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามแบบเชิงเส้น มีลักษณะดังนี้คือ

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad ; \quad \underset{\sim}{\varepsilon} \sim N(0, \sigma^2 I)$$

โดยที่ $\underset{\sim}{Y}$ เป็นเมตริกซ์ของตัวแปรตามขนาด $nx1$

X เป็นเมตริกซ์ของตัวแปรอิสระขนาด nxk_1

$\underset{\sim}{\beta}$ เป็นเมตริกซ์ของสัมประสิทธิ์ความถดถอยเชิงเส้นขนาด k_1x1

$\underset{\sim}{\varepsilon}$ เป็นเมตริกซ์ของความคลาดเคลื่อนขนาด $nx1$

k_1 เป็นจำนวนตัวแปรอิสระ + 1 และ n เป็นขนาดตัวอย่าง

โดยปกติการประมาณค่าและการทดล่อบล่มมติฐานเกี่ยวกับพารามิเตอร์นั้น ผู้วิจัยมักจะเลือกใช้วิธีกำลังล่องน้อยที่สุด (Least Squares Method) จะได้ตัวประมาณ $\hat{\beta} = (X'X)^{-1} X'Y$ ซึ่งเป็นตัวประมาณที่ไม่เอนเอียง และให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังล่องน้อยที่สุด สำหรับการทดล่อบล่มมติฐานก็ให้ล่ถิติทดล่อบที่มีอำนาจการทดล่อบสูง ทั้งนี้จะต้องมีข้อตกลงเบื้องต้นเกี่ยวกับความคลาดเคลื่อนดังนี้คือ ความคลาดเคลื่อนจะต้องมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น $\underset{\sim}{0}$ $[E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}]$ มีความแปรปรวนเป็น $\sigma^2 I$ $[E(\underset{\sim}{\varepsilon}_i \underset{\sim}{\varepsilon}_i') = \sigma^2 I]$ และ $\underset{\sim}{\varepsilon}_i, \underset{\sim}{\varepsilon}_j$ ไม่มีล่หลัมพันธ์ต่อกัน $[E(\underset{\sim}{\varepsilon}_i \underset{\sim}{\varepsilon}_j') = \underset{\sim}{0}]$

ในการทดล่อบล่มมติฐานเกี่ยวกับความเท่ากันของล้มประล่ิทธิ์ความถดถอยเชิงเล้นได้มีผู้ศึกษาไว้หลายท่าน เช่น การทดล่อบเซ่า (1960 : 591-605) โทโยดา (1974 : 601-608) ล่มิดท์และซิกเคิลล์ (1977: 1293-1298) จายาทิล่ล่า (1977 : 1291-1292) เพื่อพิจารณาความเหมาะล่มในการทดล่อบล่มมติฐานเกี่ยวกับพารามิเตอร์นั้น ล่งที่ควรพิจารณาคือ อำนาจการทดล่อบ (Power of the test) และความแกร่ง (Robustness) โดยจะพิจารณาในลักษณะที่ว่า ล่ถิติทดล่อบนั้นจะต้องมีความไว (Sensitive) ต่อการเปลี่ยนแปลงของปัจจัยที่ต้องการทดล่อบ และจะต้องไม่มีความไว (insensitive) ต่อการเปลี่ยนแปลงของล่งอื่นที่ไม่ใช่ปัจจัยที่ต้องการทดล่อบ ทั้งนี้จะพิจารณาตามคำกล่าวของเนย์แมน (Neyman 1950 : 265 อ้างโดย Direk Srisukho 1974 : 38) ซึ่งกล่าวว่า เมื่อต้องการจะเลือกใช้ล่ถิติทดล่อบ ควรพิจารณาถึงความล่ามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ก่อนแล้วจึงพิจารณาถึงความล่ามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 2 (Type II Error) ให้ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ไม่เกินค่า α ที่กำหนดไว้ และเมื่อเป็นไปตามเงื่อนไขดังกล่าวแล้ว ล่งที่ใช้ในการพิจารณาเลือกล่ถิติทดล่อบอีกก็คือ เลือกล่ถิติทดล่อบที่มีความน่าจะเป็นมากที่สุดที่จะปฏิเล่ธล่มมติฐาน Ho เมื่อล่มมติฐาน Ho ไม่เป็นจริง นั่น หมายความถึงให้อำนาจการทดล่อบสูงสุด

จึงเป็นที่น่าล่นใจว่า ในการทดล่อบล่มมติฐานเกี่ยวกับความเท่ากันของล้มประล่ิทธิ์ความถดถอยเชิงเล้นเมื่อความแปรปรวนสุ่มไม่เท่ากัน โดยใช้ล่ถิติทดล่อบ 3 วิธีคือ การทดล่อบเซ่า การทดล่อบโทโยดา การทดล่อบเซลเนอร์-ธีล-กุปตา วิธีใดจะเหมาะล่มกับการทดล่อบล่มมติฐานเกี่ยวกับพารามิเตอร์ดังกล่าว

ดังนั้น ในการวิจัยครั้งนี้ จึงสนใจที่จะพิจารณาความเหมาะสมของสถิติทดสอบ 3 วิธี สำหรับการทดสอบสมมติฐานเกี่ยวกับความเท่ากันของสัมประสิทธิ์ความถดถอยเชิงเส้นเมื่อความแปรปรวนกลุ่มไม่เท่ากัน โดยจะศึกษาเปรียบเทียบสภาพการณ์ต่าง ๆ โดยใช้เทคนิคมอนติคาร์โลซิมูเลชั่น (Monte carlo Simulation) ซึ่งเป็นวิธีการที่สามารถจำลองสภาพการณ์ต่าง ๆ ที่สนใจศึกษาได้

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อศึกษาความแกร่งของการทดสอบทั้ง 3 วิธี ของสัมประสิทธิ์ความถดถอยเชิงเส้นเมื่อความแปรปรวนกลุ่มมีค่าไม่เท่ากัน

1.2.2 เพื่อศึกษาอำนาจการทดสอบทั้ง 3 วิธี ของสัมประสิทธิ์ความถดถอยเชิงเส้นเมื่อความแปรปรวนกลุ่มมีค่าไม่เท่ากัน

1.3 สมมติฐานของการวิจัย

ภายใต้ลักษณะการกระจายของข้อมูลและขนาดตัวอย่างเดียวกันมีผลทำให้สถิติทดสอบ 3 วิธี แตกต่างกัน

1.4 ข้อตกลงเบื้องต้น

1.4.1 ตัวแปรอิสระแต่ละตัวเป็นค่าคงที่ ที่มีการแจกแจงแบบเดียวกันและเป็นอิสระซึ่งกันและกัน

1.4.2 ความคลาดเคลื่อนเป็นตัวแปรสุ่มแบบต่อเนื่องที่มีการแจกแจงแบบปกติและเป็นอิสระซึ่งกันและกัน

1.4.3 ใช้วิธีประมาณค่ากำลังสองน้อยที่สุด ในการประมาณค่าสัมประสิทธิ์ความถดถอยเชิงเส้น

1.4.4 ใช้ค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1 เป็นเกณฑ์ในการพิจารณาความแกร่งของการทดสอบ

1.4.5 ใช้อำนาจการทดสอบเป็นเกณฑ์ในการพิจารณาความไวของการทดสอบ

1.5 ขอบเขตของการวิจัย

1.5.1 ศึกษาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธี เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ โดยมีอัตราส่วนความแปรปรวนของความคลาดเคลื่อนของประชากรที่ 1 ต่อความคลาดเคลื่อนของประชากรที่ 2 เป็น 1:1 1:2 1:3 1:5 1:10 1:20 2:1 3:1 5:1 10:1 และ 20:1

1.5.2 กำหนดให้ $\beta' = [5\ 1\ 1\ \ldots\ 1]_{1xp}$ ในประชากรที่ศึกษาในการวิจัยครั้งนี้ที่กำหนดค่าพารามิเตอร์ดังกล่าว เนื่องจากการทดลองกระทำที่ขนาดตัวอย่างเท่ากันกับที่ต้องการศึกษาแล้วพบว่า ค่าพารามิเตอร์ไม่มีผลต่อค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบ

1.5.3 จำนวนตัวอย่างแต่ละชุดที่ทำการศึกษาเป็น 20 30 และ 50

1.5.4 จำนวนตัวแปรอิสระที่ศึกษาคือ 2 3 4 และ 5

1.5.5 พิจารณาการทดสอบที่ระดับนัยสำคัญ 0.01 และ 0.05

1.5.6 ศึกษาอำนาจการทดสอบและความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบสมมติฐาน เมื่อค่าคงที่ของพารามิเตอร์เป็น 5 และค่าพารามิเตอร์ของสัมประสิทธิ์ความถดถอยเชิงเส้นมีค่าดังนี้คือ $\beta' = [5\ a\ a\ \ldots\ a]_{1xp}$ เมื่อ a เท่ากับ .6 .7 .8 .85 .9 1.0 1.1 1.2 1.25 1.3 1.4 เมื่ออัตราส่วนของความแปรปรวนของความคลาดเคลื่อนเป็น 1:1 1:2 1:3 1:5 2:1 3:1 และ 5:1

1.5.7 ศึกษาอำนาจการทดสอบและความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบสมมติฐานเมื่อค่าคงที่พารามิเตอร์เป็น 5 และค่าพารามิเตอร์ของสัมประสิทธิ์ถดถอยเชิงเส้นเป็น .5 .6 .7 .8 .9 1.0 1.1 1.2 1.4 1.5 เมื่ออัตราส่วนความแปรปรวนของความคลาดเคลื่อนเป็น 1:10 1;20 10:1 และ 20:1

หมายเหตุ สำหรับการพิจารณาอำนาจการทดสอบ ได้แยกพิจารณาเป็น 2 กลุ่ม ทั้งนี้เพราะค่าพารามิเตอร์อยู่นอกช่วงดังกล่าวใน 1.5.6 และ 1.5.7 จะให้ค่าอำนาจการทดสอบเป็น 1 จึงแยกศึกษาในช่วงดังกล่าว

1.5.8 การวิจัยครั้งนี้จำลองการทดลองโดยใช้เทคนิคมอนติคาร์โลจากเครื่องคอมพิวเตอร์ IBM 370/3031 ซึ่งจะศึกษาเกี่ยวกับความคลาดเคลื่อนของประชากรมีการแจกแจง

แบบปกติ โดยมีฟังก์ชันความน่าจะเป็น ค่าคาดหวัง และความแปรปรวนของการแจกแจง เป็นดังนี้

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

$$E(X) = 0$$

$$V(X) = \sigma^2$$

1.6 คำจำกัดความ

1.6.1 ความคลาดเคลื่อนประเภทที่ 1 เป็นความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐาน H_o เมื่อสมมติฐาน H_o เป็นจริง

1.6.2 ความคลาดเคลื่อนประเภทที่ 2 เป็นความผิดพลาดที่เกิดจากการยอมรับสมมติฐาน H_o เมื่อสมมติฐาน H_o ไม่เป็นจริง

1.6.3 ความแกร่งของการทดสอบ หมายถึงคุณสมบัติของการทดสอบที่ไม่ไวต่อการเปลี่ยนแปลงของปัจจัยอื่นที่ไม่ใช่ปัจจัยที่ต้องการทดสอบ เช่น การฝ่าฝืนข้อตกลงเบื้องต้นของการทดสอบ พิจารณาความแกร่งของการทดสอบได้จากค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

1.6.4 อำนาจการทดสอบ หมายถึงคุณสมบัติของการทดสอบที่ไวต่อการเปลี่ยนแปลงของปัจจัยทดสอบหรือคือความน่าจะเป็นที่จะปฏิเสธสมมติฐาน H_o เมื่อสมมติฐาน H_o ไม่เป็นจริง

1.7 ประโยชน์ของการวิจัย

1.7.1 เพื่อทราบถึงวิธีการทดสอบทั้ง 3 วิธี

1.7.2 เพื่อพิจารณาความเหมาะสมของสถิติทดสอบ 3 วิธี เมื่อความคลาดเคลื่อนของประชากร 2 ชุด ไม่เท่ากัน โดยใช้ความแกร่งของการทดสอบและอำนาจการทดสอบเป็นเกณฑ์ในการพิจารณาเปรียบเทียบ

1.7.3 เพื่อเป็นแนวทางสำหรับผู้วิเคราะห์และผู้วิจัยทั่วไปที่จะนำไปใช้