บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

การวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis) เป็นเทคนิคที่ศึกษาถึงความสัมพันธ์ของตัวแปร 2 กลุ่ม กลุ่มหนึ่งที่เรียกว่า ตัวแปรตาม (Dependent Variable) ถูกกำหนดขึ้นโดยตัวแปรอีกกลุ่มหนึ่งที่เราเรียกว่า ตัวแปรอิสระ (Independent Variable) ตัวเดียวหรือหลายตัว และความสัมพันธ์ระหว่างตัวแปรทั้งสองกลุ่มมีลักษณะความสัมพันธ์เชิงเส้น แสดงในรูปแบบทั่วไปได้ดังนี้

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \text{----} + \beta_r X_{rt} + U_t$$

ซึ่งในที่นี้

Y_t คือ ตัวแปรตาม (Dependent Variable)

X_{it} คือ ตัวแปรอิสระ (Independent Variables)

β_i คือ พารามิเตอร์ มีชื่อเรียกว่า สัมประสิทธิ์การถดถอย (Regression Coefficient)

U_t คือ ความคลาดเคลื่อนสุ่มหรือค่าผิดพลาดสุ่ม (Ramdom Error)

เมื่อสามารถค้นหาว่าตัวแปรตามที่สนใจถูกกำหนดขึ้นโดยตัวแปรอิสระใด หรือโดยอาศัยทฤษฎีสนับสนุนแล้ว ขั้นตอนต่อไปคือ การประมาณค่าพารามิเตอร์หรือสัมประสิทธิ์การถดถอย β_i แทนด้วย $\hat{\beta}_i$ ซึ่ง วิธีการประมาณค่าพารามิเตอร์ที่นิยมใช้กันอย่างแพร่หลายวิธีหนึ่ง คือ วิธีกำลังสองน้อยที่สุดแบบสามัญ (Ordinary least squares method) เรียกย่อ ๆ ว่าวิธี OLS วิธีนี้จะให้ตัวประมาณค่าที่มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator : BLUE) ตามทฤษฎีเกาส์ - มาร์โคฟ (Gauss - Markov Theorem) โดยมีข้อตกลงเบื้องต้น (Assumptions) ของความคลาดเคลื่อน (U_t) ดังนี้

1. U_t มีค่าคาดหวัง (Expected Value) เป็น 0 $E(U_t) = 0$

2. ความแปรปรวนของความคลาดเคลื่อนคงที่ $\sigma^2_{u_i} = \sigma^2_{u_j}$; $i \neq j$

3. U_i และ U_j ไม่มีสหสัมพันธ์หรือความแปรปรวนร่วม (Covariance) เป็น 0, $E(U_i U_j) = 0$, $i \neq j$

ข้อตกลงเบื้องต้นมีความจำเป็นอย่างมากในการหาตัวประมาณ BLUE $\hat{\beta}_i$ สำหรับในการวิจัยครั้งนี้ ผู้วิจัยมีความสนใจในข้อตกลงข้อที่ 3 ซึ่งก็เป็นข้อตกลงข้อหนึ่งที่สามารถส่งผลให้เกิดความผิดพลาดร้ายแรงได้ในการอนุมาน ถ้าหากตัวแบบการถดถอยขาดคุณสมบัติข้อนี้

ความคลาดเคลื่อนไม่มีสหสัมพันธ์กันหรือเรียกอีกอย่างว่า ความคลาดเคลื่อนไม่มีอัตตสหสัมพันธ์ (Autocorrelation) กัน เกิดขึ้นได้ด้วยสาเหตุหลายประการ อาทิเช่น การละเลยตัวแปรอิสระที่สัมพันธ์กันกับตัวแปรในตัวแบบ หรือข้อมูลที่เราสนใจ มีธรรมชาติของความสัมพันธ์ที่เกี่ยวเนื่องกันเองในระหว่างคาบเวลา เป็นต้น

เมื่อเกิดปัญหาอัตตสหสัมพันธ์ในความคลาดเคลื่อนขึ้น และยังคงใช้วิธี OLS ในการประมาณค่าพารามิเตอร์ β_i จะได้ตัวประมาณที่ไม่เป็น BLUE (มีความแปรปรวนไม่ต่ำสุด) ถึงแม้จะยังคงเป็นตัวประมาณที่ไม่เอนเอียงก็ตาม จากผลนี้สามารถทำให้การอนุมานผิดพลาดร้ายแรงได้ ตัวอย่างเช่น ถ้าหากค่าผิดพลาดมีอัตตสหสัมพันธ์บวก ค่าเบี่ยงเบนมาตรฐานของตัวประมาณจะเอนเอียงในทางค่าต่ำ เพราะฉะนั้นการทดสอบนัยสำคัญของตัวประมาณอาจปรากฏว่าต่างจากศูนย์ ทั้ง ๆ ที่ความจริงแล้วไม่ต่างจากศูนย์

เหตุผลที่ผู้วิจัยเลือกศึกษาอัตตสหสัมพันธ์ตำแหน่งที่ 4 เพราะว่าในปัจจุบันการนำเสนอข้อมูลที่เราพบมากแบบหนึ่งทั้งในภาครัฐและเอกชนคือ การนำเสนอข้อมูลในรูปของอนุกรมเวลารายไตรมาส (Quarterly Time Series Data) และข้อมูลอนุกรมเวลารายไตรมาสนี้มีบ่อยครั้งเสมอที่ถูกนำเสนอมาใช้ศึกษาการวิเคราะห์การถดถอย ข้อมูลประเภทนี้เมื่อนำมาใช้ในการวิเคราะห์การถดถอยมักจะพบว่า มีผลทำให้ข้อตกลงเบื้องต้นที่ว่า ความคลาดเคลื่อนสุ่มไม่มีสหสัมพันธ์กันนั้นไม่เป็นจริง หรืออีกนัยหนึ่งก็คือ เกิดอัตตสหสัมพันธ์ตำแหน่งที่ 4 เกิดขึ้น

วิธีการตรวจสอบอัตตสหสัมพันธ์ในโปรแกรมสำเร็จรูปทั่วไป จะนิยมใช้ตัวสถิติทดสอบของเดอร์บิน - วัตสัน (Durbin - Watson test) ซึ่งมีโครงสร้างของฟังก์ชั่นในอัตตสหสัมพันธ์ตำแหน่งที่ 1 (First - Order Autocorrelation) เพราะฉะนั้นถ้านำวิธีของเดอร์บิน - วัตสัน ไปใช้ตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 อาจจะทำให้ได้ข้อสรุปผิดพลาดได้

สำหรับตัวสถิติที่ใช้ทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 มีผู้สนใจศึกษาหลายท่าน ทั้งที่เป็นวิธีพาราเมตริก และวิธีนอนพาราเมตริก ตัวสถิติทดสอบที่นิยมใช้กันมากพอสมควรก็คือ ตัวสถิติทดสอบของวอลลิส (Wallis test) ตัวสถิติทดสอบโทมัสวอลลิส (Thomas - Wallis test) ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาตัวสถิติทดสอบที่มีประสิทธิภาพสูงกว่าตัวสถิติทดสอบของเดอร์บิน - วัตสัน มาใช้ตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 จากผลของวิธี OLS ในการวิจัยครั้งนี้ผู้วิจัยจะศึกษาเปรียบเทียบตัวสถิติทดสอบ 4 ตัว คือ ตัวสถิติทดสอบเดอร์บิน - วัตสัน (Durbin - Watson test) ตัวสถิติทดสอบวอลลิส (Wallis test) ตัวสถิติทดสอบโทมัส - วอลลิส (Thomas - Wallis test) และตัวสถิติทดสอบบ๊อกซ์ - เพี้ยซ (Box - Pierce test) สำหรับตัวสถิติทดสอบบ๊อกซ์ - เพี้ยซ ผู้วิจัยนำมาศึกษาเปรียบเทียบด้วยนั้นเพราะตัวสถิติตัวนี้เป็นที่นิยมใช้กันมากตัวหนึ่งสำหรับการทดสอบอัตตสหสัมพันธ์ทุกตำแหน่งพร้อมกันในตัวแบบอนุกรมเวลา บ๊อกซ์ - เจนกินส์ (Box - Jenkins Model)

การศึกษาเปรียบเทียบนั้นจะศึกษาความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 (Type I Error) และสำหรับตัวสถิติที่สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ จะศึกษาเปรียบเทียบอำนาจการทดสอบ (Power of the test) ของตัวสถิติเหล่านั้นต่อไป เพื่อจะได้เสนอแนะวิธีการทดสอบที่เหมาะสมสำหรับการทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาเปรียบเทียบตัวสถิติที่ใช้ตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 ของความคลาดเคลื่อนในสมการถดถอย

1. ตัวสถิติทดสอบเดอร์บิน – วัตสัน (Durbin - Watson test)
2. ตัวสถิติทดสอบวอลลิส (Wallis test)
3. ตัวสถิติทดสอบโทมัส – วอลลิส (Thomas - Wallis test)
4. ตัวสถิติทดสอบบ็อกซ์ – เพียซ (Box - Pierce test)

โดยศึกษาเปรียบเทียบค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 (Type I Error) และอำนาจการทดสอบ (Power of the test) ของตัวสถิติทดสอบทั้ง 4 ตัว

1.3 สมมติฐานการวิจัย

1.3.1 ในตัวอย่างขนาดเล็ก ตัวสถิติทดสอบวอลลิสจะให้อำนาจการทดสอบสูงสุด

1.3.2 ในการทดสอบอัตสหสัมพันธ์ตำแหน่งที่ 4 ตัวสถิติทดสอบเดอร์บิน – วัตสัน จะให้อำนาจการทดสอบต่ำที่สุดในทุกกรณีที่ศึกษา

1.4 ข้อตกลงเบื้องต้น

สมการถดถอยที่ใช้ในการศึกษาเป็นสมการถอถอยแบบพหุเชิงเส้น (Multiple Linear Regression Equation) มีรูปแบบดังนี้

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + U_t$$

Y_t คือ ตัวแปรตาม (Dependent Variable), $t = 1,2, \ldots, n$

X_{it} คือ ตัวแปรอิสระ (Independent Variable), $i = 1,2$

$\beta_0, \beta_1, \beta_2$ คือ พารามิเตอร์ (Parameter)

U_t คือ ความคลาดเคลื่อนสุ่ม (Ramdom Error)

โดยกำหนดรูปแบบความสัมพันธ์ของ U_t และ X_{it} ดังนี้

$$U_t = \rho U_{t-4} + e_t$$

เมื่อ ρ คือ สัมประสิทธิ์อัตสหสัมพันธ์ตำแหน่งที่ 4, $-1 \leq \rho \leq 1$

$$X_{it} = \theta_{i1} X_{it-1} + \theta_{i4} X_{it-4} + V_t \; ; \; i = 1,2$$

เมื่อ θ_{i1}, θ_{i4} คือ พารามิเตอร์

V_t คือ ความคลาดเคลื่อนสุ่ม

ในการกำหนดรูปแบบการเปลี่ยนแปลงฤดูกาลของ U_t X_{1t} และ X_{2t} ข้างต้นนั้น เป็นวิธีการหนึ่งสำหรับการสร้างข้อมูล Y_t ให้มีความโน้มเอียงที่เป็นข้อมูลอนุกรมเวลารายไตรมาส[1/] ตามที่ต้องการ

1.5 ขอบเขตการวิจัย

1. e_t เป็นอิสระกันและมีการแจกแจงปรกติ ค่าเฉลี่ย 0 และความแปรปรวนคงที่ $= \sigma_e^2 = 1$

2. V_t เป็นอิสระกันและมีการแจกแจงปรกติ ค่าเฉลี่ยเป็น 0 และความแปรปรวนคงที่ $= \sigma_v^2 = 1$

3. $\rho = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9$ แบ่งออกเป็น 3 ระดับ คือ

ระดับต่ำ $\rho = 0.1, 0.3$

ระดับปานกลาง $\rho = 0.5$

ระดับสูง $\rho = 0.7, 0.9$

ตามที่ทราบกันทั่วไปแล้วว่าอัตตสหสัมพันธ์ (ρ) มีค่าอยู่ในช่วง $-1 \leq \rho \leq 1$ แต่เนื่องจากในทางปฏิบัติโดยเฉพาะข้อมูลทางธุรกิจและเศรษฐศาสตร์ ส่วนมากจะพบค่าอัตตสหสัมพันธ์เป็นบวก และเนื่องจากความสมมาตรของค่า ρ ผู้วิจัยจึงขอเลือกกำหนดว่า $\rho \geq 0$

4. $\theta_{11} = 0.1, 0.5$, $\theta_{14} = 0.1, 0.5$

$\theta_{21} = 0.1, 0.5$, $\theta_{24} = 0.1, 0.5$

1/ King, M.L. and D.E.A Giles, A Comparison of Some Tests For Fourth - Order Autocorrelation (Australian Economic Papers, December, 1978)

ซึ่งสามารถเขียนค่าพารามิเตอร์ของตัวแปรอิสระ X_{1t}, X_{2t} เป็น $(\theta_{11},\theta_{14},\theta_{21},\theta_{24})$ ออกเป็น 16 รูปแบบ ซึ่งการกำหนดรูปแบบดังกล่าวมีวัตถุประสงค์ให้ตัวแปรอิสระมีอนุกรมเวลาอยู่ในภาวะสมดุลย์* (Stationary time series)* และไม่อยู่ในภาวะสมดุลย์ (nonstationary time series) เพื่อศึกษาผลกระทบต่อประสิทธิภาพของตัวสถิติ

5. $\beta_0 = 1, \beta_1 = 2, \beta_2 = 3$

เนื่องจากได้มีผู้ทำการศึกษาในเรื่องอัตตสหสัมพันธ์ในความคลาดเคลื่อน รวมทั้งผู้วิจัยได้ทดลองเช่นกัน พบว่า β_i เมื่อเปลี่ยนค่าไป จะไม่มีผลกระทบต่อผลสรุปของการตรวจสอบอัตตสหสัมพันธ์ในความคลาดเคลื่อนแต่ละวิธี[2/] ดังนั้นผู้วิจัยจึงขอกำหนดค่า β_0, β_1, β_2 = 1, 2, 3 เป็นค่าคงที่ เพื่อลดสถานการณ์ในการเปลี่ยนแปลงค่าพารามิเตอร์ต่าง ๆ

6. ขนาดตัวอย่างแบ่งเป็น 3 ขนาด คือ ขนาดเล็ก (n = 20) ขนาดกลาง (n = 40, 60) และขนาดใหญ่ (n = 100)

7. ระดับนัยสำคัญ = 0.05 การวิจัยครั้งนี้กำหนดระดับนัยสำคัญเพียงค่าเดียว เนื่องจากตัวสถิติทดสอบวอลลิสมีเพียงตารางทดสอบที่มีระดับนัยสำคัญเพียง0.05 เท่านั้น

8. การวิจัยครั้งนี้ จำลองข้อมูลขึ้นตามสถานการณ์ที่ต้องการศึกษา โดยใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) จากเครื่องคอมพิวเตอร์ AMDAHL 5860 โดยใช้ภาษาฟอร์แทรน (Fortran) ทำการจำลองแบบซ้ำ ๆ กัน 1,000 ครั้ง ในแต่ละสถานการณ์ของการวิจัย

---

*<u>สภาวะสมดุลย์เชิงสถิติ</u> ของรูปแบบของตัวแปรอิสระ X_{1t}, X_{2t} มีข้อกำหนดดังต่อไปนี้ทุกข้อ

1. $\theta_{i1} + \theta_{i4} < 1$, 2. $\theta_{i4} - \theta_{i1} < 1$, 3. $|\theta_{i4}| < 1$

2/ Wallis, K.F. Testing for Fourth - Order Autocorrelation in Quarterly Regression Equations. (Econometrica. Vol.40, No.4 July, 1972)

1.6 คำจำกัดความ

คำจำกัดความที่ใช้ในการวิจัยครั้งนี้มีดังนี้

1.6.1 ความผิดพลาดประเภทที่ 1 (Type I Error) เป็นความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐาน Ho: เมื่อ Ho: ถูกต้อง

1.6.2 ความผิดพลาดประเภทที่ 2 (Type II Error) เป็นความผิดพลาดที่เกิดจากการยอมรับสมมติฐาน Ho: เมื่อ Ho: ไม่ถูกต้อง

1.6.3 อำนาจการทดสอบ (Power of the test) เป็นความน่าจะเป็นที่ปฏิเสธสมมติฐาน Ho: เมื่อ Ho: ไม่ถูกต้อง

1.6.4 ระดับความรุนแรงของปัญหาอัตตสหสัมพันธ์ตำแหน่งที่ 4 แบ่งออกเป็น 3 ระดับดังนี้

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับต่ำ คือ $\rho = 0.1, 0.3$

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับปานกลาง คือ $\rho = 0.5$

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับสูง คือ $\rho = 0.7, 0.9$

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เป็นแนวทางในการตัดสินใจ เลือกตัวสถิติในการตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 ที่ให้อำนาจการทดสอบสูงสุด

1.7.2 เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบตัวสถิติอื่นต่อไป