บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis) เป็นการศึกษาถึงความสัมพันธ์ของตัวแปร 2 กลุ่ม กลุ่มหนึ่งที่เรียกว่าตัวแปรตาม (Dependent Variable) ถูกกำหนดขึ้นโดยตัวแปรอีกกลุ่มหนึ่ง ที่เราเรียกว่าตัวแปรอิสระ (Independent Variable) ตัวเดียวหรือหลายตัว และความสัมพันธ์ระหว่างตัวแปรทั้งสองกลุ่มมีลักษณะความสัมพันธ์เชิงเส้น แสดงในรูปแบบทั่วไปได้ดังนี้

$$Y_t = \beta_o + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_p x_{pt} + u_t, \; t = 1,2,\ldots, n$$

ซึ่งในที่นี้

Y_t คือ ตัวแปรตาม (Dependent Variable)

x_{it} คือ ตัวแปรอิสระ (Independent Variables), $i = 1,2,\ldots, p$

β_i คือ พารามิเตอร์ มีชื่อเรียกว่าสัมประสิทธิ์การถดถอย (Regression Coefficients)

u_t คือ ความคลาดเคลื่อนสุ่ม หรือค่าผิดพลาดสุ่ม (Random Errors)

n คือ ขนาดตัวอย่าง

เมื่อสามารถค้นหาตัวแบบสมการถดถอยได้แล้ว ขั้นตอนที่สำคัญต่อไปในการวิเคราะห์ความถดถอยคือ การประมาณค่าพารามิเตอร์ หรือสัมประสิทธิ์การถดถอย β_i แทนด้วย $\hat{\beta}_i$ โดยวิธีใดวิธีหนึ่งที่เหมาะสม และการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์นั้น ซึ่งในการเลือกใช้วิธีที่เหมาะสมนั้นจำเป็นต้องคำนึงถึงข้อตกลงเบื้องต้นของวิธีที่ใช้ด้วย วิธีการประมาณค่าพารามิเตอร์ที่นิยมใช้กันอย่างแพร่หลายวิธีหนึ่งคือ วิธีกำลังสองน้อยที่สุดแบบสามัญ

(Ordinary Least Square method) เรียกย่อ ๆ ว่า วิธี OLS ซึ่งเป็นวิธีที่ให้ตัวประมาณที่มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator : BLUE)* ทั้งนี้จะต้องมีข้อตกลงเบื้องต้น (Assumptions) ของความคลาดเคลื่อน (u_t) ดังนี้

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

2. ความแปรปรวนของความคลาดเคลื่อนคงที่ $\sigma^2_{u_t} = \sigma^2$

3. u_i และ u_j ไม่มีสหสัมพันธ์หรือความแปรปรวนร่วม (Covariance) = 0 ; $E(u_i u_j) = 0, i \neq j$

ข้อตกลงเบื้องต้นดังกล่าวมีความจำเป็นอย่างยิ่งในการหาตัวประมาณ BLUE $\hat{\beta}_i$ สำหรับในการวิจัยครั้งนี้ ผู้วิจัยมีความสนใจในข้อตกลงเบื้องต้น ข้อที่ 3 ซึ่งก็เป็นข้อตกลงข้อหนึ่งที่สามารถส่งผลให้เกิดความผิดพลาดร้ายแรงได้ในการอนุมาน ถ้าหากตัวแบบการถดถอยขาดคุณสมบัติข้อนี้และได้ใช้วิธี OLS ในการประมาณค่าพารามิเตอร์

ความคลาดเคลื่อนมีสหสัมพันธ์กันหรือเรียกอีกอย่างว่าความคลาดเคลื่อนมีอัตตสหสัมพันธ์กัน (Autocorrelation) เกิดขึ้นได้ด้วยสาเหตุหลายประการ อาทิเช่น การละเลยตัวแปรอิสระที่มีความสัมพันธ์กันกับตัวแปรตามในตัวแบบหรือข้อมูลที่สนใจมีธรรมชาติของความสัมพันธ์ที่เกี่ยวเนื่องกันเองในระหว่างคาบเวลา เป็นต้น

เมื่อเกิดปัญหาอัตตสหสัมพันธ์ในความคลาดเคลื่อนขึ้น และยังคงใช้วิธี OLS ในการประมาณค่าพารามิเตอร์ β_i จะได้ตัวประมาณที่ไม่เป็น BLUE (มีความแปรปรวนไม่ต่ำสุด) ถึงแม้จะยังคงเป็นตัวประมาณที่ไม่เอนเอียงก็ตาม จากผลนี้สามารถทำให้การอนุมานผิดพลาดร้ายแรงได้ ตัวอย่างเช่น ถ้าค่าความคลาดเคลื่อนมีอัตตสหสัมพันธ์ทางบวก (Positive Autocorrelation) จะมีผลทำให้ค่าเบี่ยงเบนมาตรฐานของตัวประมาณจะเอนเอียงในทางค่าต่ำ เพราะฉะนั้นการทดสอบนัยสำคัญของตัวประมาณอาจปรากฏว่าต่างจาก

---

* ทฤษฎี Gauss-Markov: ในกลุ่มของ Linear Unbiased Estimators ของ Regression parameters นั้น OLS-Estimators จะเป็นตัวประมาณที่ดีที่สุด (Chow 1983:40-41, Wonnacott and Wonnacott 1970:27-28)

ศูนย์ ทั้ง ๆ ที่ความจริงแล้วไม่ต่างจากศูนย์

จากผลกระทบที่เกิดจากปัญหาอัตตสหสัมพันธ์ จึงมีความจำเป็นที่จะต้องทำการตรวจสอบค่าความคลาดเคลื่อนว่าเป็นไปตามข้อตกลงเบื้องต้นข้อที่ 3 หรือไม่

สำหรับการศึกษาครั้งนี้ ผู้วิจัยสนใจศึกษาตัวแบบการถดถอยที่เรียกว่าตัวแบบ Distributed lag ซึ่งเป็นตัวแบบที่แสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ โดยที่ค่าในอดีตของตัวแปรอิสระส่งผลกระทบต่อความผันแปรของค่าปัจจุบันของตัวแปรตามด้วย ซึ่งพบได้ในทางปฏิบัติเมื่อค่าของตัวแปรตามได้รับอิทธิพลมาจากตัวแปรอิสระที่สนใจต่อเนื่องกันในหลาย ๆ คาบเวลา เช่น พบว่ายอดขายสินค้า ณ คาบเวลา t มีอิทธิพลมาจากการโฆษณาทั้งในคาบเวลาที่ t และ t-1, t-2,... เนื่องจากค่าในอดีตของตัวแปรอิสระมีผลต่อค่าของตัวแปรตามอย่างไม่มีที่สิ้นสุด ในทางทฤษฎีกล่าวคือไม่ว่าเวลาจะล่วงเลยมานานเพียงใดก็ตาม อิทธิพลของตัวแปรอิสระในอดีตจะมีผลต่อค่าตัวแปรตาม ดังนั้นเรียกตัวแบบนี้ว่า Infinite Distributed lag Model เมื่อมีตัวแปรอิสระ x_t หนึ่งตัวจะมีรูปแบบดังนี้

$$Y_t = \sum_{i=0}^{\infty} \beta_i^* x_{t-i} + v_t$$

จากรูปแบบข้างต้นพบว่า การวิเคราะห์ยุ่งยากมากในทางปฏิบัติเนื่องจากต้องประมาณค่าพารามิเตอร์ β_i^* จำนวนมากมายจึงต้องมีการกำหนดรูปแบบของ β_i^* โดยทั่วไปแล้วพบว่าเมื่อเวลาผ่านไปเรื่อย ๆ ตัวแปรอิสระที่อยู่ ณ คาบเวลาใกล้ ๆ กับปัจจุบันจะมีผลต่อตัวแปรตามมากกว่าตัวแปรอิสระ ณ คาบเวลาที่ห่างไกล ดังนั้นการศึกษาครั้งนี้จะกำหนดรูปแบบ β_i^* เป็นแบบ Geometric Lag นั่นคือค่า β_i^* จะมีค่าลดลงในรูปอนุกรมเรขาคณิต ซึ่งมีรูปแบบดังนี้

$$\beta_i^* = \alpha \lambda^i \quad ; \quad 0 < \lambda < 1 \quad ; \quad i = 0,1,2,\ldots$$

$$\alpha = \text{ค่าคงที่}$$

จาก $Y_t \sum_{i=0}^{\infty} \beta_i^* x_{t-i} + v_t$

แทนค่า $\beta_i^* = \alpha\lambda^i$

$$Y_t = \sum_{i=0}^{\infty} (\alpha\lambda^i)\, x_{t-i} + v_t$$

$$= \alpha \sum_{i=0}^{\infty} \lambda^i x_{t-i} + v_t$$

$$= \alpha\,(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \ldots) + v_t$$

นิยาม Lag Operation (L) ; $(L^i)\, x_t = x_{t-i}$ และฟังก์ชันของ L ให้เป็น

$$W(L) \quad ; W(L) = 1 + \lambda L + \lambda^2 L^2 + \ldots$$

$$= (1 - \lambda L)^{-1}$$

เพราะฉะนั้น $Y_t = \alpha\,(1 + \lambda L + \lambda^2 L^2 + \ldots)\, x_t + v_t$

$$= \alpha\, W(L)\, x_t + v_t$$

$$= \alpha\,(1 - \lambda L)^{-1} x_t + v_t$$

$$(1 - \lambda L) Y_t = \alpha\, x_t + v_t\,(1 - \lambda L)$$

$$Y_t - \lambda Y_{t-1} = \alpha\, x_t + u_t\ ,\ u_t = v_t\,(1 - \lambda L)$$

$Y_t = \lambda\, Y_{t-1} + \alpha x_t + u_t$ ซึ่งเรียกว่าสมการ

ถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระ

สำหรับการวิเคราะห์สมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระนั้น ก็กระทำเช่นเดียวกับการวิเคราะห์สมการถดถอยโดยทั่วไป นั่นคือสามารถใช้วิธี OLS ในการประมาณค่าพารามิเตอร์ได้ ถ้าตรวจสอบดูแล้วไม่มีอัตตสหสัมพันธ์ในค่าความคลาดเคลื่อน และในกรณีที่พบว่ามีอัตตสหสัมพันธ์ก็ควรจะเลือกใช้วิธีอื่นที่เหมาะสมกว่าในการประมาณค่าพารามิเตอร์ ซึ่งมีอยู่หลายวิธีเช่น Search and Iteration, Wallis Three-Step เป็นต้น

การตรวจสอบอัตตสหสัมพันธ์ในความคลาดเคลื่อนนั้น โดยทั่วไปพบว่าความคลาดเคลื่อนมักจะมีรูปแบบเป็น AR(1) [First-Order Autoregressive Model] ดังจะเห็นได้จากโปรแกรมสำเร็จรูปทั่วไปจะมีการคำนวณค่าตัวสถิติเดอร์บิน-วัตสัน (Durbin-Watson statistic) สำหรับทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 [AR(1)] ของความคลาดเคลื่อน ดังนั้นผู้วิจัยจึงสนใจศึกษาปัญหาอัตตสหสัมพันธ์ในความคลาดเคลื่อนที่มีรูปแบบเป็น AR(1)

สำหรับตัวสถิติที่ใช้ทดสอบอัตตสหสัมพันธ์ มีผู้สนใจศึกษาหลายท่านและตัวสถิติทดสอบที่นิยมใช้กัน สำหรับสมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระ คือตัวสถิติทดสอบ h ของเดอร์บิน (Durbin's h statistic) แต่มีข้อจำกัดในการใช้บางกรณีกับตัวอย่างขนาดเล็ก ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาตัวสถิติทดสอบอื่น ๆ ซึ่งเหมาะสมกว่าตัวสถิติทดสอบ h ของเดอร์บินในการวิจัยครั้งนี้ ผู้วิจัยจะศึกษาเปรียบเทียบตัวสถิติทดสอบ 5 ตัวคือ ตัวสถิติทดสอบเดอร์บิน-วัตสัน (Durbin-Watson statistic) ซึ่งเป็นตัวสถิติทดสอบที่นิยมใช้กันมากสำหรับสมการถดถอยทั่ว ๆ ไป ตัวสถิติทดสอบ h ของเดอร์บิน ซึ่งต่อไปจะเรียกสั้น ๆ ว่าตัวสถิติทดสอบ H ตัวสถิติทดสอบ h ของเดอร์บินที่ได้ปรับปรุงใหม่ (Modified Durbin's h) ซึ่งต่อไปจะเรียกสั้น ๆ ว่าตัวสถิติทดสอบ H-M ตัวสถิติทดสอบ m (m statistic) และตัวสถิติทดสอบบอกซ์-เพียซ (Box-Pierce statistic) สำหรับตัวสถิติทดสอบบอกซ์-เพียซ ผู้วิจัยนำมาศึกษาเปรียบเทียบด้วยนั้นเพราะตัวสถิตินี้เป็นที่นิยมใช้กันมากตัวหนึ่งสำหรับการทดสอบอัตตสหสัมพันธ์ทุกตำแหน่งพร้อมกันในตัวแบบอนุกรมเวลาบอกซ์-เจนกินส์ (Box-Jenkins Medel)

การศึกษาเปรียบเทียบนั้นจะศึกษาความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 (Type-I Error) และสำหรับตัวสถิติทดสอบที่สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ จะศึกษาเปรียบเทียบอำนาจการทดสอบ (Power of the test) ของตัวสถิติเหล่านั้นต่อไป เพื่อจะได้เสนอแนะสถิติทดสอบที่เหมาะสมสำหรับการทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในสมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระ

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยต้องการศึกษาเปรียบเทียบตัวสถิติที่ใช้ตรวจสอบปัญหาอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในการวิเคราะห์สมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระ 5 วิธีดังนี้

1.2.1 การทดสอบเดอร์บิน-วัตสัน (Durbin-Watson test)

1.2.2 การทดสอบ H (Durbin's h test)

1.2.3 การทดสอบ H-M (Modified Durbin's h test)

1.2.4 การทดสอบบอกซ์-เพียซ (Box-Pierce test)

1.2.5 การทดสอบ m (m-test)

เพื่อหาข้อสรุปเกี่ยวกับความเหมาะสมของการทดสอบดังกล่าว โดยพิจารณาจากความน่าจะเป็นของความผิดพลาดประเภทที่ 1 (Type I error) และอำนาจของการทดสอบ (Power of the test)

1.3 สมมติฐานของการวิจัย

ในตัวอย่างขนาดเล็ก สถิติทดสอบทั้ง 5 จะให้อำนาจการทดสอบที่แตกต่างกัน โดยสถิติทดสอบ H-M จะให้อำนาจการทดสอบสูงสุด

1.4 ข้อตกลงเบื้องต้น

สมการถดถอยที่ใช้ในการศึกษาเป็นสมการถดถอยเชิงเส้นที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระกับตัวแปรอิสระ x_t หนึ่งตัว มีรูปแบบดังนี้

$$Y_t = \beta_o + \beta_1 Y_{t-1} + \beta_2 x_t + u_t , \quad t = 1,2,\ldots, n$$

Y_t คือ ตัวแปรตาม (Dependent Variable)

Y_{t-1} คือ ตัวแปรตามย้อนเวลา (Lag Dependent Variable)

x_t คือ ตัวแปรอิสระ (Independent Variable)

$\beta_o, \beta_1, \beta_2$ คือ พารามิเตอร์ (Parameters) สำหรับ β_1 เป็นพารามิเตอร์ซึ่งแสดงอัตตสหสัมพันธ์ของ Y_t

u_t คือ ความคลาดเคลื่อนสุ่ม (Random Error)

โดยกำหนดรูปแบบความสัมพันธ์ของ u_t และ x_t ตามลำดับดังนี้

$$u_t = \rho u_{t-1} + e_t \quad \text{ตัวแบบ AR(1)}$$

เมื่อ ρ คือ สัมประสิทธิ์อัตตสัมพันธ์ตำแหน่งที่ 1, $-1 \leq \rho \leq 1$

e_t คือ ความคลาดเคลื่อนสุ่ม

$$x_t = \Theta x_{t-1} + \eta_t$$

เมื่อ Θ คือ พารามิเตอร์

η_t คือ ความคลาดเคลื่อนสุ่ม

1.5 <u>ขอบเขตการวิจัย</u>

1. e_t เป็นอิสระกันและมีการแจกแจงปกติ ค่าเฉลี่ย 0 และความแปรปรวนคงที่ $= \sigma_e^2$ ซึ่งจะกำหนดค่าที่ทำให้ค่า R^2 ของสมการถดถอยที่ศึกษามีค่าสูงต่ำในระดับต่าง ๆ คือ 0.9, 0.7, และ 0.5

2. η_t เป็นอิสระกันและมีการแจกแจงปกติ ค่าเฉลี่ยเป็น 0 และความแปรปรวนคงที่ $= \sigma_\eta^2 = 10^*$

3. ρ = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9 แบ่งออกเป็น 3 ระดับคือ

ระดับต่ำ	ρ = 0.1, 0.3
ระดับปานกลาง	ρ = 0.5
ระดับสูง	ρ = 0.7, 0.9

ตามที่ทราบกันทั่วไปแล้วว่าสัมประสิทธิ์อัตตสหสัมพันธ์ (ρ) มีค่าอยู่ในช่วง $-1 \leqslant \rho \leqslant 1$ แต่เนื่องจากในทางปฏิบัติโดยเฉพาะข้อมูลทางธุรกิจและเศรษฐศาสตร์ ส่วนมากจะพบว่ามีค่าเป็นบวก และเนื่องจากความสมมาตรของค่า ρ ผู้วิจัยจึงศึกษาเฉพาะ $\rho \geqslant 0$

4. θ = 0.3, 0.9

ซึ่งการกำหนดดังกล่าวมีวัตถุประสงค์ให้ตัวแปรอิสระมีอนุกรมเวลาอยู่ในภาวะสมดุลย์ (Stationary time Series); $|\theta| < 1$ และแสดงระดับอัตตสหสัมพันธ์กันเป็น 2 ระดับ คือ มีความสัมพันธ์กันอย่างไม่ชัดเจน (θ = 0.3) และมีความสัมพันธ์กันอย่างชัดเจน (θ = 0.9)

5. $\beta_o = 0, \beta_2 = 1.0^*$

6. β_1 = 0.1, 0.3, 0.5, 0.7

---

* การวิจัยครั้งนี้กำหนดค่า σ_e^2 และค่าพารามิเตอร์ (β_o, β_2) เป็นค่าดังกล่าว เนื่องจากได้ทดลองกระทำ ณ ค่าต่าง ๆ และจากการทดลองของ Tse, Y.K. Some Modified Versons of Durbin's h Statistic (The Review of Economics and Statistics, 1985) แล้วปรากฏว่าไม่ว่าจะเป็นค่าใดก็ตาม ผลสรุปไม่เปลี่ยนแปลง

ค่า β_1 คือค่า λ ในการศึกษาที่มาของสมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระ (กล่าวไว้ในตอน 1.1 ความสำคัญและที่มาของปัญหา) ซึ่งมีข้อกำหนดว่า $0 < \lambda < 1$ และผู้วิจัยสนใจศึกษาค่า β_1 ใน 3 ระดับคือ ค่าน้อย ($\beta_1 = 0.1, 0.3$) ค่าปานกลาง (0.5) และค่ามาก (0.7)

7. ขนาดตัวอย่างแบ่งเป็น 3 ขนาดคือ ขนาดเล็ก ($n = 30$) ขนาดกลาง ($n = 60$) และขนาดใหญ่ ($n = 100$)

8. ระดับนัยสำคัญ = 0.05 การวิจัยครั้งนี้กำหนดระดับนัยสำคัญเพียงค่าเดียว เนื่องจากการศึกษาความผิดพลาดประเภทที่ 1 เป็นการทดสอบแบบ 2 หาง แต่ตัวสถิติทดสอบเดอร์บิน-วัตสันมีเพียงตารางทดสอบที่มีระดับนัยสำคัญ 0.01, 0.025, 0.05 เท่านั้น

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

1.6 <u>คำจำกัดความ</u>

1.6.1 ความผิดพลาดประเภทที่ 1 (Type I error) เป็นความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐาน H_o เมื่อ H_o ถูกต้อง

1.6.2 ความผิดพลาดประเภทที่ 2 (Type II error) เป็นความผิดพลาดที่เกิดจากการยอมรับสมมติฐาน H_o เมื่อ H_o ไม่ถูกต้อง

1.6.3 อำนาจการทดสอบ (Power of the test) เป็นความน่าจะเป็นที่ปฏิเสธสมมติฐาน H_o เมื่อ H_o ไม่ถูกต้อง

1.6.4 ระดับความรุนแรงของปัญหาอัตตสหสัมพันธ์ตำแหน่งที่ 1 แบ่งออกเป็น 3 ระดับดังนี้

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับต่ำคือ $\rho = 0.1, 0.3$

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับปานกลาง คือ $\rho = 0.5$

ปัญหาอัตตสหสัมพันธ์มีความรุนแรงระดับสูง คือ $\rho = 0.7, 0.9$

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เพื่อเป็นแนวทางในการตัดสินใจ เลือกตัวสถิติทดสอบในการตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในการวิเคราะห์สมการถดถอยที่มีตัวแปรตามย้อนเวลาร่วมเป็นตัวแปรอิสระที่ให้อำนาจการทดสอบสูงสุด

1.7.2 เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบตัวสถิติอื่น ๆ ต่อไป