บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis) เป็นการศึกษาถึงความสัมพันธ์ของตัวแปร 2 กลุ่ม กลุ่มหนึ่งเรียกว่า ตัวแปรตาม (Dependent variable) ถูกกำหนดโดยตัวแปรอีกกลุ่มหนึ่ง ซึ่งเรียกว่า ตัวแปรอิสระ (Independent variable) ตัวเดียวหรือหลายตัว และความสัมพันธ์ระหว่างตัวแปรทั้งสองกลุ่ม มีลักษณะความสัมพันธ์เชิงเส้น ซึ่งมีรูปแบบทั่วไปเป็นดังนี้ คือ

$$y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_p X_{pt} + u_t \; ;$$

$$t = 1, 2, \ldots, n$$

โดยที่ y_t คือ ตัวแปรตาม

x_{it} คือ ตัวแปรอิสระ $i = 1, 2\ldots, p$

β_i คือ พารามิเตอร์ มีชื่อเรียกว่า สัมประสิทธิ์การถดถอย (Regression Coefficients)

u_t คือ ความคลาดเคลื่อนสุ่ม หรือค่าผิดพลาดสุ่ม (Random errors)

n คือ ขนาดตัวอย่าง

สำหรับงานวิจัยโดยทั่วไปที่ต้องอาศัยเทคนิคการวิเคราะห์การถดถอยนั้น เป็นที่ทราบกันดีอยู่แล้วว่า เมื่อสามารถค้นหาตัวแบบสมการถดถอยได้แล้ว ขั้นตอนที่สำคัญต่อไปในการวิเคราะห์การถดถอย คือ การประมาณค่าพารามิเตอร์ หรือสัมประสิทธิ์การถดถอย β_i แทนด้วย $\hat{\beta}_i$ โดยวิธีใดวิธีหนึ่งที่เหมาะสม และการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์นั้น ในการเลือกใช้วิธีที่เหมาะสมจำเป็นต้องคำนึงถึงข้อตกลงเบื้องต้นของวิธีที่ใช้ด้วย วิธีการประมาณค่าพารามิเตอร์ที่นิยมใช้กันแพร่หลายวิธีหนึ่ง คือ วิธีกำลังสองน้อยที่สุดแบบสามัญ (Ordinary Least Squares Method) เรียกย่อว่า วิธี OLS ซึ่งเป็นวิธีที่ให้ตัวประมาณที่มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear

Unbiased Estimator : BLUE) ทั้งนี้จะต้องมีข้อตกลงเบื้องต้น 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) $= 0$; $E(u_i u_j) = 0$; $i \neq j$

ข้อตกลงเบื้องต้นดังกล่าว มีความจำเป็นอย่างยิ่งในการหาตัวประมาณ BLUE $\hat{\beta}_i$ สำหรับการวิจัยครั้งนี้ ผู้วิจัยมีความสนใจในข้อตกลงเบื้องต้น ข้อที่ 3 ซึ่งเป็นข้อตกลงหนึ่งที่สามารถส่งผลให้เกิดความผิดร้ายแรงได้ในการอนุมาน ถ้าหากตัวแบบการถดถอยขาดคุณสมบัติข้อนี้และใช้วิธี OLS ในการประมาณค่าพารามิเตอร์

กรณีที่ความคลาดเคลื่อนมีสหสัมพันธ์กัน หรือเรียกอีกอย่างหนึ่งว่า ความคลาดเคลื่อนมีอัตตสหสัมพันธ์ (Autocorrelation) เกิดขึ้นได้ด้วยสาเหตุหลายประการ เช่น การละเลยตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามในตัวแบบหรือข้อมูลที่สนใจมีธรรมชาติของความสัมพันธ์ที่เกี่ยวเนื่องกันเอง ในระหว่างคาบเวลา

ดังนั้นเมื่อเกิดปัญหาอัตตสหสัมพันธ์ในความคลาดเคลื่อนขึ้น แล้วยังคงใช้วิธี OLS ในการประมาณค่าพารามิเตอร์ β_i จะได้ตัวประมาณที่ไม่เป็น BLUE (มีความแปรปรวนไม่ต่ำสุด) ถึงแม้จะเป็นตัวประมาณที่ไม่เอนเอียงก็ตาม จากผลนี้สามารถทำให้การอนุมานผิดพลาดร้ายแรงได้ เช่น ถ้าค่าความคลาดเคลื่อนมีอัตตสหสัมพันธ์ทางบวก (Positive Autocorrelation) จะมีผลทำให้ค่าเบี่ยงเบนมาตรฐานของตัวประมาณจะเอนเอียงในทางค่าต่ำ เพราะฉะนั้นการทดสอบนัยสำคัญของตัวประมาณอาจปรากฏว่าต่างจากศูนย์ ทั้ง ๆ ที่ความจริงแล้วไม่ต่างจากศูนย์

จากผลกระทบที่เกิดจากปัญหาอัตตสหสัมพันธ์ จึงมีความจำเป็นที่จะต้องทำการตรวจสอบค่าความคลาดเคลื่อนว่าเป็นไปตามข้อตกลงเบื้องต้นข้อที่ 3 หรือไม่ ก่อนที่จะเลือกวิธีประมาณค่าพารามิเตอร์

อัตตสหสัมพันธ์ในความคลาดเคลื่อนนั้น โดยทั่วไปในทางปฏิบัติโดยเฉพาะข้อมูลด้านเศรษฐศาสตร์และธุรกิจจะพบว่า ความคลาดเคลื่อนมักจะมีรูปแบบเป็น AR(1) [First-Order Autoregressive Model] เพราะฉะนั้นโปรแกรมสำเร็จรูปทั่วไปมักจะคำนวณค่าตัวสถิติทดสอบเดอร์บินวัตสัน (Durbin-Watson statistic) สำหรับทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 [AR(1)] ของความคลาดเคลื่อน นอกจากตัวสถิติเดอร์บินวัตสันแล้วยังมีตัวสถิติอื่น ๆ ที่ได้มีผู้คิดค้นขึ้นสำหรับการทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 AR(1) ด้วยเช่นกัน และผู้วิจัยมีความคิดเห็นว่าในบรรดาตัวสถิติเหล่านี้ น่าจะมีตัวสถิติบางตัวที่มีอำนาจการทดสอบสูงกว่าตัวสถิติเดอร์บินวัตสันในบางเงื่อนไขหรือในบางสถานการณ์ ด้วยเหตุนี้ผู้วิจัยจึงสนใจที่จะศึกษาการทดสอบอัตตสหสัมพันธ์ในความคลาดเคลื่อนที่มีรูปแบบเป็น AR(1) โดยสนใจที่จะศึกษาตัวสถิติอื่น ๆ นอกเหนือจากตัวสถิติเดอร์บินวัตสัน

สำหรับตัวสถิติที่ใช้ทดสอบอัตตสหสัมพันธ์ มีผู้สนใจศึกษาหลายท่าน ผู้วิจัยได้ศึกษาพิจารณาเลือกตัวสถิติทดสอบ 3 ตัว ที่ยังไม่ได้มีการเปรียบเทียบมาก่อน ตามที่ผู้วิจัยได้ศึกษาค้นคว้ามา คือ ตัวสถิติทดสอบเดอร์บิน-วัตสัน (Durbin-Watson statistic) ตัวสถิติทดสอบอัลเตอร์เนทีฟเดอร์บินวัตสัน (Alternative Durbin Watson) และตัวสถิติทดสอบการวิ่ง (Run test)

การศึกษาเปรียบเทียบนั้นจะศึกษาความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 (type I error) และสำหรับตัวสถิติทดสอบที่สามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ จะศึกษาเปรียบเทียบอำนาจการทดสอบ (Power of the test) ของตัวสถิติเหล่านั้นต่อไป ภายใต้สถานการณ์ต่าง ๆ เพื่อจะได้เสนอแนะตัวสถิติที่เหมาะสมสำหรับการทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในสมการถดถอยเชิงเส้นอย่างง่ายในสถานการณ์เหล่านั้น

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยต้องการศึกษาเปรียบเทียบตัวสถิติทดสอบ 3 ตัวต่อไปนี้ ในการทดสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในการวิเคราะห์สมการถดถอยเชิงเส้นอย่างง่าย เพื่อหาข้อสรุปเกี่ยวกับความเหมาะสมของการทดสอบดังกล่าว โดยพิจารณาจากความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบ

1.2.1 ตัวสถิติทดสอบเดอร์บิน-วัตสัน

1.2.2 ตัวสถิติทดสอบอัลเตอร์เนทีฟเดอร์บินวัตสัน

1.2.3 ตัวสถิติทดสอบการวิ่ง

1.3 สมมติฐานของการวิจัย

ในกรณีที่ความรุนแรงของอัตตสหสัมพันธ์ (ρ) มีค่าน้อยกว่า 0.5 ตัวสถิติอัลเตอร์เนทีฟเดอร์บินวัตสัน จะให้อำนาจการทดสอบสูงสุด

1.4 ข้อตกลงเบื้องต้น

สมการถดถอยที่ใช้ในการศึกษาเป็นแบบสมการถดถอยเชิงเส้นอย่างง่าย มีรูปแบบดังนี้

$$y_t = \beta_0 + \beta_1 x_t + u_t \; ; \; t = 1, 2, \ldots, n$$

y_t คือ ตัวแปรตาม

x_t คือ ตัวแปรอิสระ

β_0, β_1 คือ พารามิเตอร์

u_t คือ ความคลาดเคลื่อนสุ่ม

โดยกำหนดรูปแบบของอัตตสหสัมพันธ์ u_t เป็นแบบอัตตสหสัมพันธ์ตำแหน่งที่ 1 [AR(1)]

$$u_t = \rho u_{t-1} + e_t$$

เมื่อ ρ คือ สัมประสิทธิ์อัตตสหสัมพันธ์ ตำแหน่งที่ 1

e_t คือ ความคลาดเคลื่อนสุ่ม

1.5 ขอบเขตการวิจัย

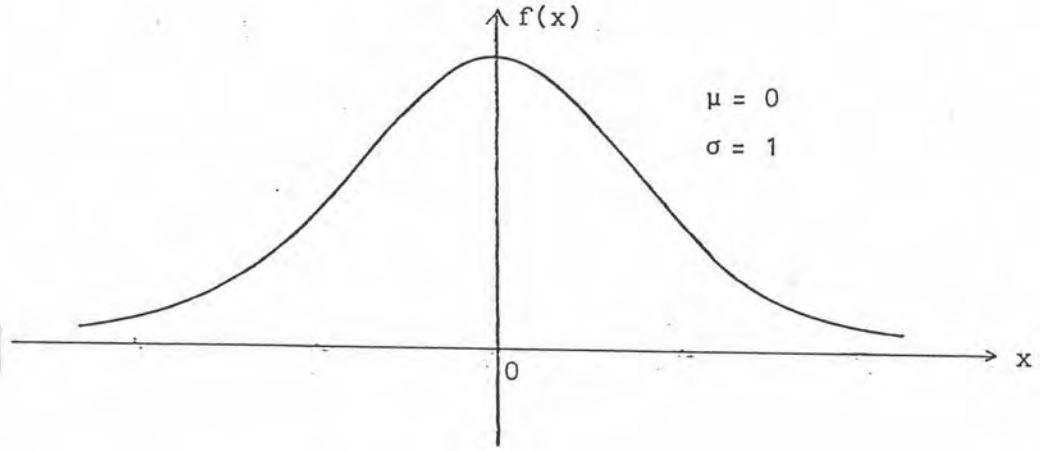
1.5.1 ศึกษาภายใต้ลักษณะการแจกแจงของความคลาดเคลื่อนสุ่ม e_t ดังนี้

1.5.1.1 การแจกแจงปกติ (Normal Distribution)

มีฟังก์ชันความหนาแน่น f และรูปกราฟ 1.5.1.1 ดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(x-\mu)^2 / 2\sigma^2\right] \quad -\infty < x < \infty$$

ผู้วิจัยกำหนดค่าพารามิเตอร์ $\mu = 0$ และ $\sigma^2 = 1$ (ผลจากการทดลองพบว่า ไม่ขึ้นอยู่กับค่าพารามิเตอร์)



รูปที่ 1.5.1.1

1.5.1.2 การแจกแจงดับเบิ้ลเอกซโพเนนเชียล (Double Exponential Distribution)

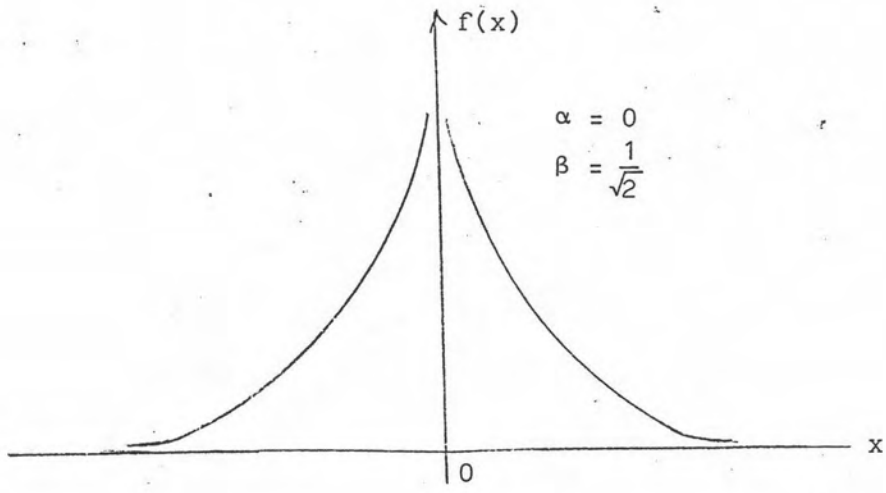
มีฟังก์ชันความหนาแน่น f และรูปกราฟ 1.5.1.2 ดังนี้

$$f(x) = \frac{1}{2\beta} \cdot \exp\left[-|x-\alpha|/\beta\right] \quad -\infty < x < \infty$$

$$E(X) = \alpha$$

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

ผู้วิจัยกำหนดค่าพารามิเตอร์ $\alpha = 0$ และ $\beta = \frac{1}{\sqrt{2}}$ (โดยคำนวณจากการกำหนดค่า $E(X) = 0$ และ $V(X) = 1$)



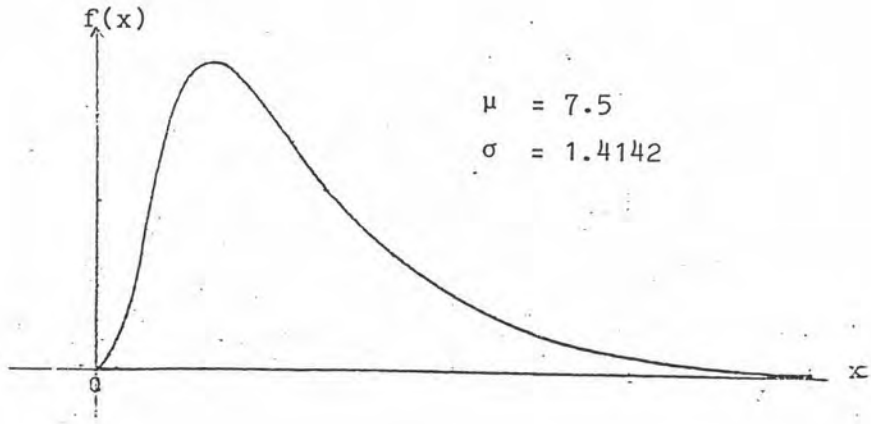
รูปที่ 1.5.1.2

1.5.1.3 การแจกแจงลอกนอร์แมล (Lognormal Distribution) มีฟังก์ชันความหนาแน่นในรูป f และรูปกราฟ 1.5.1.3 ดังนี้

$$f(x) = \frac{1}{x.\sigma\sqrt{2\pi}} \exp\left[-(\ln x-\mu)^2 / 2\sigma^2\right] ; x>0, \sigma>0$$

$$-\infty<\mu<\infty$$

ผู้วิจัยสนใจศึกษา ณ. พารามิเตอร์ $\mu = 7.5$ และ $\sigma = 1.4142$



รูปที่ 1.5.1.3

1.5.2 ตัวแปรอิสระ (x_t) มีรูปแบบเป็นดังนี้

- x_t มีการแจกแจงปกติ มีค่าเฉลี่ยเป็น 0 ความแปรปรวน 1
- $x_t = 0.8x_{t-1} + n_t$ เมื่อ $n_t \sim N(0,1)$
- $x_t = t + n_t$ เมื่อ $n_t \sim N(0,1)$
- $x_t = t + \cos\left(\frac{2\pi t}{12}\right)$
- $x_t = t$

1.5.3 กำหนดให้ $\beta_0 = 1$, $\beta_1 = 1$*

1.5.4 ค่า ρ ที่ศึกษาแบ่งออกเป็น 3 ระดับ

ระดับต่ำ 0.1, 0.15, 0.2, 0.25, 0.3, 0.35

ระดับกลาง 0.4, 0.45, 0.5, 0.55, 0.6

ระดับสูง 0.65, 0.7, 0.75, 0.8, 0.85, 0.9

---

* การวิจัยครั้งนี้ได้ทำการทดลองที่ค่าอื่น ๆ ของ β_0, β_1 นอกเหนือจากค่า $\beta_0 = 1$ และ $\beta_1 = 1$ ปรากฏว่าผลสรุปไม่ต่างกัน จึงขอนำเสนอในกรณี $\beta_0 = 1$ และ $\beta_1 = 1$

1.5.5 ขนาดตัวอย่างแบ่งเป็น 3 ขนาด

ขนาดเล็ก n = 15

ขนาดกลาง n = 30

ขนาดใหญ่ n = 60

1.5.6 ระดับนัยสำคัญ = 0.05

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

1.6 คำจำกัดความ

ความผิดพลาดประเภทที่ 1 (Type I error) เป็นความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐาน Ho เมื่อ Ho ถูกต้อง

ความผิดพลาดประเภทที่ 2 (Type II error) เป็นความผิดพลาดที่เกิดจากการยอมรับสมมติฐาน Ho เมื่อ Ho ไม่ถูกต้อง

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

อัตตสหสัมพันธ์ (Autocorrelation) หมายถึงตัวแปรสุ่ม u_t, $t = \pm1, \pm2, \ldots$, มีความสัมพันธ์ต่อกัน กล่าวคือ $E(u_i u_j) \neq 0$ เมื่อ $i \neq j$

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เพื่อเป็นแนวทางในการตัดสินใจ เลือกตัวสถิติทดสอบในการตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 ในการวิเคราะห์สมการถดถอยเชิงเส้นอย่างง่าย ที่ให้อำนาจการทดสอบสูงสุด

1.7.2 เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบตัวสถิติอื่น ๆ ต่อไป