บทที่ 1

บทนำ

ที่มาและความสำคัญของปัญหา

การพยากรณ์มีบทบาทสำคัญอย่างมากต่องานทางด้านธุรกิจในปัจจุบันไม่ว่าในสายงานทางด้านการจัดการใดๆ เช่น การตลาด การผลิต การวิจัยและพัฒนา เป็นต้น ความสำคัญของการพยากรณ์เป็นที่ยอมรับและต้องการเป็นอย่างมากทั้งนี้เนื่องจากการแข่งขันทางด้านธุรกิจที่นับวันจะมากขึ้นเรื่อยๆ ประกอบกับบรรยากาศของการแข่งขันเปลี่ยนแปลงไปอย่างรวดเร็ว ผู้ประกอบธุรกิจหรือผู้บริหารที่ฉลาดจะต้องไวต่อสภาพการเปลี่ยนแปลงของธุรกิจ การคาดคะเนหรือการพยากรณ์เหตุการณ์ที่จะเกิดขึ้นในอนาคตจึงได้มีบทบาทสำคัญในขบวนการตัดสินใจของผู้บริหารเพื่อให้ธุรกิจสามารถดำเนินต่อไปได้อย่างมีประสิทธิภาพ

การวิเคราะห์การถดถอย เป็นเทคนิคอย่างหนึ่งที่ใช้ในการคาดคะเนหรือพยากรณ์ค่าของตัวแปรตัวหนึ่งที่สนใจศึกษา เรียกกันทั่วไปว่า "ตัวแปรตาม" โดยอาศัยความสัมพันธ์กับตัวแปรอื่นๆที่เกี่ยวข้องกันเรียกตัวแปรหลังนี้ว่า "ตัวแปรอิสระ"

ในการวิจัยครั้งนี้สนใจศึกษาความถดถอยแบบง่ายมาใช้ในการพยากรณ์ ซึ่งมีรูปแบบดังนี้ คือ

$$y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \quad , t = 1, 2, 3, ..., n$$

โดยที่

y_t เป็นตัวแปรตามที่เป็นตัวแปรสุ่ม

X_t เป็นตัวแปรอิสระที่ไม่เป็นตัวแปรสุ่ม

β_0 เป็นพารามิเตอร์ที่ไม่ทราบค่า แทนจุดตัดบนแกน Y

β_1 เป็นพารามิเตอร์ที่ไม่ทราบค่า แทนความชันของเส้นถดถอย

ε_t เป็นความคลาดเคลื่อนสุ่ม

และ n เป็นขนาดตัวอย่าง

โดยปกติการประมาณค่าพารามิเตอร์ผู้วิจัยมักเลือกใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares : OLS) ซึ่งจะทำให้ตัวประมาณที่ได้ มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียง

เชิงเส้นที่ดีที่สุด(Best Linear Unbiased Estimator : BLUB) ทั้งนี้ต้องมีข้อสมมติพื้นฐานเกี่ยวกับความคลาดเคลื่อน ดังนี้คือ

1. มีค่าคาดหวัง (Expected Value) เป็น 0, $E(\varepsilon_t) = 0$, t=1,2,...,n
2. ความแปรปรวนของความคลาดเคลื่อนคงที่ $Var(\varepsilon_t) = \sigma^2$, t=1,2,...,n
3. ε_s และ ε_t ไม่มีสหสัมพันธ์ต่อกันหรือมีความแปรปรวนร่วม (Covariance) เป็น 0, $Cov(\varepsilon_s, \varepsilon_t) = E(\varepsilon_s \varepsilon_t) = 0$, $s \neq t$

ในทางปฏิบัติปัญหาที่พบมากปัญหาหนึ่งโดยเฉพาะเมื่อตัวแปรเป็นอนุกรมเวลาคือความคลาดเคลื่อนมีสหสัมพันธ์ต่อกันซึ่งสถานการณ์เช่นนี้เราเรียกว่า อัตตสหสัมพันธ์ (Auto-correlation) ในกรณีเช่นนี้ ถ้าหากผู้วิจัยยังคงใช้การประมาณพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดจะได้ตัวประมาณไม่เป็น BLUE คือมีความแปรปรวนไม่ต่ำสุด ถึงแม้จะเป็นตัวประมาณไม่เอนเอียงก็ตามเช่น ถ้าความคลาดเคลื่อนมีอัตตสหสัมพันธ์ทางบวก (positive autocorrelation) จะมีผลทำให้ค่าความแปรปรวนของตัวประมาณมีค่าน้อยกว่าค่าความแปรปรวนที่แท้จริง ฉะนั้นการทดสอบสมมติฐานของตัวประมาณอาจปรากฏว่าต่างจากศูนย์ ทั้งๆ ที่ความจริงแล้วไม่ต่างจากศูนย์ ซึ่งจะก่อให้เกิดการอนุมานผิดพลาดและส่งผลต่อการพยากรณ์ที่ได้จากสมการถดถอยไม่ถูกต้องแม่นยำ

จากผลกระทบดังกล่าวจึงเป็นที่น่าสนใจอย่างยิ่งในการหาวิธีประมาณค่าพารามิเตอร์ที่เหมาะสมมากกว่ามาแทนการใช้วิธีกำลังสองน้อยที่สุดเพื่อที่จะทำให้การประมาณค่าพารามิเตอร์และการพยากรณ์มีความถูกต้องมากยิ่งขึ้น ซึ่งถ้าเราทราบค่าอัตตสหสัมพันธ์ของข้อมูล (prior information) เราสามารถใช้วิธีกำลังสองต่ำสุดแบบทั่วไป (Generalized Least Square Method) หรือตัวประมาณเบส์ (Bayesian Estimator) ในการประมาณค่าพารามิเตอร์ได้ แต่ในบางครั้งเราไม่ทราบค่าจึงจำเป็นต้องประมาณค่าอัตตสหสัมพันธ์ขึ้นมาเพื่อใช้ในการประมาณค่าพารามิเตอร์ ซึ่งในการศึกษากรณีดังกล่าวนี้นักสถิติหลายท่านได้ทำการศึกษาและคิดค้นตัวสถิติขึ้น เพื่อประมาณค่าพารามิเตอร์ที่ใช้ในการพยากรณ์ สรุปได้ดังนี้

ในปี ค.ศ. 1964 Zellner และ Tiao ได้เสนอตัวประมาณเบส์เพื่อใช้แก้ปัญหาความคลาดเคลื่อนมีอัตตสหสัมพันธ์ในสมการถดถอย ซึ่งต่อมาในปี ค.ศ. 1978 Fomby และ Guilkey ได้ทำการปรับปรุงการแจกแจงโดยหลักเกณฑ์ของตัวประมาณเบส์ พบว่าสามารถแก้ไขและปรับปรุงประสิทธิภาพของตัวประมาณได้ และทำการเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (MSE) ของตัวประมาณค่าพารามิเตอร์ระหว่างตัวประมาณเบส์กับตัวประมาณ OLS ในปี ค.ศ. 1985 Dielman นำตัวประมาณ OLS วิธีการแปลงของคอคแครนและออร์คัต วิธี Iterative ของเพรสและวินส์เทน วิธีของ Beach และ MacKinnon และตัวประมาณเบส์มาเปรียบ

เทียบค่า MSE ของค่าพยากรณ์ สรุปได้ว่าการประมาณค่าอัตตสหสัมพันธ์นั้นมีความสำคัญมาก แม้ว่าผลที่ได้จะมีประสิทธิภาพต่ำลงเมื่อคาบเวลาของการพยากรณ์ยาวขึ้น แต่สำหรับค่าอัตตสหสัมพันธ์มากกว่าหรือเท่ากับ 0.5 และช่วงการพยากรณ์ระยะสั้น พบว่าตัวประมาณเบส์สามารถทำงานได้ดี และในปี ค.ศ. 1993 Latif, A. และ King, M. L. เสนอตัวพยากรณ์ผสมโดยใช้ฟังก์ชันความหนาแน่นเชิงเดี่ยวของค่าอัตตสหสัมพันธ์(marginal probability density function of ρ) ซึ่งวิธีนี้มีข้อดีคือไม่ขึ้นอยู่กับการประมาณค่าอัตตสหสัมพันธ์

จากผลการวิจัยดังกล่าวนี้ ผู้วิจัยจึงสนใจทำการศึกษาเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการพยากรณ์ (Mean Square Forecast Errors : MSFE) ของวิธีการประมาณค่าพารามิเตอร์ 4 วิธีคือ

1. วิธีกำลังสองน้อยที่สุด (Ordinary Least Square : OLS)
2. วิธีการแปลงของเพรสและวินส์เทน (Prais-Winsten Transformation)
3. ตัวประมาณเบส์ (Bayesian Estimator)
4. ตัวพยากรณ์ผสม (Mixture Prediotor)

โดยทำการเปรียบเทียบในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อตัวแปรอิสระมีแนวโน้มรูปแบบต่าง ๆ และการแจกแจงของความคลาดเคลื่อนมีการแจกแจงปกติ ซึ่งรูปแบบความสัมพันธ์เป็นอัตตถดถอยอันดับหนึ่ง

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ เมื่อทราบและไม่ทราบข้อมูลเบื้องต้นเกี่ยวกับพารามิเตอร์ ในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อความคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับหนึ่ง ด้วยวิธีการประมาณค่าพารามิเตอร์ดังต่อไปนี้

1. วิธีกำลังสองน้อยที่สุด (Ordinary Least Square : OLS)
2. วิธีการแปลงของเพรสและวินส์เทน (Prais-Winsten Transformation)
3. ตัวประมาณเบส์ (Bayesian Estimator)
4. ตัวพยากรณ์ผสม (Mixture Prediotor)

ข้อตกลงเบื้องต้น

1. ในการวิจัยครั้งนี้ ได้ศึกษารูปแบบของสมการถดถอยเชิงเส้นอย่างง่าย (Simple

Linear Regression Equation) ซึ่งมีรูปแบบดังนี้

$$y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \quad , t = 1, 2, 3, ..., n$$

โดยที่

y_t เป็นตัวแปรตาม

X_t เป็นตัวแปรอิสระ

β_i เป็นพารามิเตอร์ที่ไม่ทราบค่า , i = 0, 1

ε_t เป็นค่าความคลาดเคลื่อน

n เป็นขนาดตัวอย่าง

2. ค่าความคลาดเคลื่อนมีสหสัมพันธ์กัน โดยกำหนดรูปแบบความสัมพันธ์เป็นอัตดถดถอยอันดับหนึ่ง (First Order Autoregressive model : AR(1)) คือ

$$\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t \quad , t = 1, 2, ..., n$$

เมื่อ ρ คือ ค่าสหสัมพันธ์ระหว่าง ε_t กับ ε_{t-1} และ $|\rho| < 1$
และมีข้อตกลงเบื้องต้นของ ν_t คือ

$$E(\nu_t) = 0$$

$$Var(\nu_t) = \sigma^2$$

$$E(\nu_s\nu_t) = 0 \; ; s \neq t$$

3. ความคลาดเคลื่อนเป็นตัวแปรสุ่มที่มีสหสัมพันธ์และความแปรปรวนคงที่คือ

$$V(\varepsilon_t) = \frac{\sigma_\nu^2}{1-\rho^2} \quad ; t = 1, 2, ..., n$$

สมมติฐานของการวิจัย

ภายใต้ลักษณะความคลาดเคลื่อนที่มีอัตดสหสัมพันธ์อันดับหนึ่ง ณ ระดับค่าต่างๆ ตัวพยากรณ์ผสมจะให้ค่าพยากรณ์ที่มีค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการพยากรณ์น้อยกว่าวิธีกำลังสองน้อยที่สุด วิธีการแปลงของเพรสและวินส์เทน และตัวประมาณเบส์

ขอบเขตของการวิจัย

1. ศึกษาภายใต้ลักษณะการแจกแจงของความคลาดเคลื่อน(ε_t) มีการแจกแจงปกติ (Normal Distribution)

1.1 v_t มีค่าเฉลี่ยเป็น 0 และความแปรปรวนคงที่เป็น σ_v^2 ในการวิจัยครั้งนี้กำหนดให้เท่ากับ 1

1.2 ε_t มีค่าเฉลี่ยเป็น 0 และความแปรปรวนคงที่เป็น $\frac{\sigma_v^2}{1-\rho^2}$

2. ศึกษาเมื่อตัวแปรอิสระ X_t มีลักษณะดังนี้

2.1 รูปแบบเส้นตรงตามเวลา (Simple Time Trend)

$$X_t = t$$

2.2 รูปแบบแนวโน้มไม่คงที่ (Stochastic Trend)

$$X_t = t + u_t$$
$$u_t \sim N(0, 1)^*$$

2.3 รูปแบบแนวโน้มตามคาบเวลา (Periodic Trend)

$$X_t = t + \cos(2\pi t / 12)$$

2.4 รูปแบบอัตถดถอยสหสัมพันธ์อันดับที่หนึ่ง (First Order Autoregressive)

$$X_t = \lambda X_{t-1} + u_t$$
$$u_t \sim N(0, 1)^*$$
$$X_0 \sim N(0, 1/(1-\lambda^2))$$

3. ค่า β_0 และ β_1 มีค่าเท่ากับ 1**

4. ศึกษาเมื่อค่าสหสัมพันธ์ (ρ) เป็น 0.2, 0.4, 0.6, 0.8, 0.9, และ 0.95

5. ศึกษาเมื่อค่าสหสัมพันธ์ (λ) เป็น 0.2, 0.4, และ 0.8

---

* การวิจัยครั้งนี้ได้ทดลองที่ค่าเฉลี่ยและความแปรปรวนค่าอื่นๆ ได้ผลสรุปไม่แตกต่าง

** การวิจัยครั้งนี้ได้ทดลองค่าสัมประสิทธิ์การถดถอยที่ค่าอื่นๆ ได้ผลสรุปไม่แตกต่าง

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

6. ศึกษาเมื่อขนาดตัวอย่าง (n) เท่ากับ 15, 30, 50, และ 70

เกณฑ์การตัดสินใจ

ในการวิจัยครั้งนี้ ผู้วิจัยศึกษาค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการพยากรณ์ของวิธีกำลังสองน้อยที่สุด วิธีการแปลงของเพรสและวินส์เทน ตัวประมาณเบส์ และตัวพยากรณ์ผสม โดยมีสูตรการคำนวณดังนี้

$$MSFE_i = \sum_{j=1}^{300} \frac{(y_{ij} - \hat{y}_{ij})^2}{300}$$

โดยที่

y_{ij} คือ ค่าข้อมูลจริง ณ คาบเวลาที่ i และรอบที่ j

$\hat{y}_{ij}$ คือ ค่าพยากรณ์ ณ คาบเวลาที่ i และรอบที่ j

j คือ จำนวนรอบของการทำซ้ำ

i คือ คาบเวลาในการพยากรณ์ ; i = 1, 2,... ,12

คำจำกัดความ

1. อัตสหสัมพันธ์ (Autocorrelation) คือ เหตุการณ์ที่ตัวแปรสุ่มมีความสัมพันธ์ในตัวเอง กล่าวคือ $Cov(\varepsilon_s, \varepsilon_t) \neq 0$ เมื่อ $s \neq t$

2. ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการพยากรณ์ (Mean Squared Forecast Errors : MSFEs) คือค่าที่แสดงว่า ค่าจากการพยากรณ์แตกต่างจากค่าข้อมูลจริงเพียงไรโดยวัดในรูปค่าเฉลี่ยของกำลังสองของความแตกต่างระหว่างค่าพยากรณ์และค่าข้อมูลจริง

ประโยชน์ที่คาดว่าจะได้รับ

1. ให้ข้อสรุปในการเลือกใช้วิธีการประมาณค่าพารามิเตอร์ เพื่อการพยากรณ์เมื่อความคลาดเคลื่อนมีอัตสหสัมพันธ์อันดับหนึ่ง

2. สามารถใช้เป็นแนวทางในการศึกษา เพื่อเลือกวิธีการประมาณค่าพารามิเตอร์ที่นำไปใช้ในการพยากรณ์ เมื่อความคลาดเคลื่อนมีอัตสหสัมพันธ์ในรูปแบบอื่นๆ ต่อไป