บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในสภาพสังคมปัจจุบันที่มีการแข่งขันกัน องค์กรใดมีข้อมูลมากกว่าก็มีโอกาสที่จะตัดสินใจ เพื่อทำให้เกิดประโยชน์มากกว่าในการกำหนดนโยบาย เพื่อหาลู่ทางแก้ปัญหาต่างๆ และยังใช้เพื่อเปรียบเทียบแนวโน้มเรื่องต่างๆ กับคู่แข่งขันได้อีกด้วย โดยการนำกระบวนการทางสถิติเพื่อการพยากรณ์มาช่วยในการตัดสินใจ เนื่องจากสถิติสามารถชี้ให้เห็นแนวโน้ม ซึ่งสามารถคาดการณ์ได้ล่วงหน้า และคาดคะเนเหตุการณ์ในอนาคต โดยนำมาใช้ในการกำหนดนโยบายการวางแผนการบริหาร การตัดสินใจในเรื่องสำคัญๆ และมีความเสี่ยงสูง ควรศึกษาข้อมูลในอดีตและอาจอาศัยเทคนิคการพยากรณ์ เพื่อจะทำให้ผลพยากรณ์แม่นยำหรือใกล้เคียงความจริงมากที่สุด ดังนั้นในปัจจุบันนี้การพยากรณ์ได้รับความสนใจมากและมีการพัฒนารวดเร็วและก้าวหน้าไปไกลมาก และเข้ามามีบทบาทสำคัญในหน่วยงานต่างๆ เกือบทุกหน่วยงานทั้งในภาครัฐบาลและเอกชน ตัวอย่างการนำเทคนิคการพยากรณ์เชิงปริมาณไปใช้ในเรื่องต่างๆ ที่ก่อให้เกิดประโยชน์และมีส่วนช่วยในการตัดสินใจ เช่น พยากรณ์รายได้หรือรายจ่าย เพื่อการวางแผนทางด้านการเงิน การพยากรณ์ยอดขายหรือความต้องการซื้อ เพื่อการวางแผนการตลาด หรือการผลิต หรือการวางแผนการจัดการสินค้าคงคลัง เป็นต้น

การพยากรณ์เชิงปริมาณจำแนกเป็น 2 ประเภทใหญ่ๆ ประเภทที่ 1 เป็นวิธีการที่มีแนวความคิดว่า พฤติกรรมในอดีตของสิ่งที่จะพยากรณ์ควรจะเพียงพอที่จะพยากรณ์พฤติกรรมในอนาคตได้ วิธีการพยากรณ์ประเภทนี้ได้แก่ อนุกรมเวลา Box-Jenkins, เทคนิคการทำให้เรียบ, อนุกรมเวลาคลาสสิค เป็นต้น การพยากรณ์เชิงปริมาณประเภทที่ 2 เป็นวิธีการที่มีแนวความคิดว่าพฤติกรรมของสิ่งที่จะพยากรณ์ถูกกำหนดขึ้นโดยสิ่งอื่นๆ ซึ่งมีความสัมพันธ์บางลักษณะกับสิ่งที่จะพยากรณ์ วิธีการพยากรณ์ประเภทนี้ได้แก่ การพยากรณ์ด้วยการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอยเป็นวิธีพยากรณ์วิธีหนึ่งที่ได้รับความนิยมมาก โดยเฉพาะในการพยากรณ์ระยะยาว วิธีการนี้จะพิจารณาตัวแปรอื่นที่มีผลกระทบ หรือมีอิทธิพลต่อตัวแปรที่จะพยากรณ์ โดยไม่ได้พิจารณาแต่ตัวแปรที่จะพยากรณ์อย่างเดียว ตัวอย่างเช่น รายจ่ายขึ้นอยู่กับ

รายได้ รายได้ขึ้นอยู่กับราคาสินค้าและต้นทุนสินค้า ราคาสินค้าขึ้นอยู่กับสภาพการณ์ทางตลาด สภาพการณ์ทางตลาดจะกำหนดส่วนแบ่งในตลาด (market share)ของสินค้าต่างๆ ซึ่งขึ้นอยู่กับจำนวน และราคาของสินค้าเหล่านั้น ต้นทุนสินค้าขึ้นอยู่กับราคาของปัจจัยในการผลิตซึ่งได้แก่วัตถุดิบ แรงงาน ราคาวัตถุดิบขึ้นอยู่กับตลาดของวัตถุดิบ ค่าแรงก็ขึ้นอยู่กับค่าครองชีพและนโยบายของรัฐ เป็นต้น

การพยากรณ์ด้วยการวิเคราะห์การถดถอยมีขั้นตอนที่สำคัญ 2 ขั้นตอนคือ ขั้นตอนที่ 1 เป็นการสร้างตัวแบบเพื่อการพยากรณ์ ในขั้นนี้การประมาณค่าพารามิเตอร์ของสมการถดถอยเป็นเป้าหมายที่สำคัญของเทคนิคการวิเคราะห์การถดถอย การเลือกใช้วิธีประมาณค่าที่เหมาะสมต้องพิจารณาปัจจัยหลายประการอาทิเช่น ลักษณะของข้อมูลที่ต้องการวิเคราะห์ และข้อตกลงเบื้องต้นของวิธีแต่ละวิธีที่ใช้ด้วย ขั้นตอนที่ 2 ของการพยากรณ์ด้วยการวิเคราะห์การถดถอยคือ การนำตัวแบบที่ได้จากขั้นตอนแรกไปใช้ในการพยากรณ์

การศึกษาการถดถอยเชิงเส้นอย่างง่าย มีรูปแบบทั่วไปดังนี้

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \ , \ t = 1, 2, \ldots, n$$

โดยที่ X_t เป็นตัวแปรอิสระที่มีค่าคงที่ Y_t เป็นตัวแปรตาม β_0 และ β_1 เป็นพารามิเตอร์ที่ไม่ทราบค่าแทนจุดตัดบนแกน Y และความชันของเส้นถดถอยตามลำดับ ε_t เป็นความคลาดเคลื่อนสุ่ม และ n เป็นเลขจำนวนเต็มบวก

โดยปกติการประมาณค่า และการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ ผู้วิจัยมักเลือกใช้วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares : OLS) ซึ่งเป็นวิธีที่ให้ตัวประมาณที่มีคุณสมบัติ เป็นตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator : BLUE) ทั้งนี้ต้องอยู่ภายใต้ข้อตกลงเบื้องต้นเกี่ยวกับความคลาดเคลื่อน ดังนี้

1. ε_t มีค่าคาดหมาย (expected value) เป็นศูนย์ หรือ $E(\varepsilon_t) = 0$
2. ε_t มีความแปรปรวนคงที่ นั่นคือ $E(\varepsilon_t^2) = \sigma^2$
3. ε_i และ ε_j มีความแปรปรวนร่วมเป็นศูนย์ หรือ $E(\varepsilon_i \varepsilon_j) = 0$ เมื่อ $i \neq j$

อย่างไรก็ตามในทางปฏิบัติ บ่อยครั้งที่พบว่าข้อมูลที่นำมาใช้ในการวิเคราะห์นั้นมีอยู่ไม่น้อยที่ไม่เป็นไปตามข้อตกลงดังกล่าว ในการวิจัยครั้งนี้สนใจศึกษากรณีที่ไม่เป็นไปตามข้อตกลง

เบื้องต้นข้อที่ 3 และหาหนทางแก้ปัญหาเพื่อนำมาใช้ในการพยากรณ์ กรณีศึกษานี้คือความคลาดเคลื่อนเกิดสหสัมพันธ์ต่อกันซึ่งเรียกสถานการณ์เช่นนี้ว่า อัตตสหสัมพันธ์ (autocorrelation) และรูปแบบที่พบโดยทั่วไปในอัตตสหสัมพันธ์ของความคลาดเคลื่อนของข้อมูลอนุกรมเวลา คือ อัตตถดถอยอันดับที่หนึ่ง(first order autoregressive process) เกิดขึ้นได้ด้วยสาเหตุหลายประการ อาทิเช่น การละเลยตัวแปรอิสระที่มีความสัมพันธ์กันกับตัวแปรตามในตัวแบบ และข้อมูลที่สนใจมีธรรมชาติของความสัมพันธ์ที่เกี่ยวเนื่องกันเองในระหว่างคาบเวลา ซึ่งจะพบมากในการวิเคราะห์ข้อมูลอนุกรมเวลา ทางด้านธุรกิจและเศรษฐศาสตร์ ตัวอย่างของการเกิดอัตตถดถอยอันดับที่หนึ่ง เช่น จำนวนสินค้าที่ส่งออกของปีนี้ขึ้นกับจำนวนสินค้าที่ส่งออกของปีที่ผ่านมา ราคาขายข้าวของปีนี้ขึ้นกับราคาขายข้าวของปีที่ผ่านมา และงบประมาณรายจ่ายของปีนี้ขึ้นกับงบประมาณรายจ่ายของปีที่ผ่านมา เป็นต้น

ในสถานการณ์ที่ความคลาดเคลื่อนมีสหสัมพันธ์ต่อกัน ถ้าผู้วิจัยยังคงใช้การประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดจะได้ตัวประมาณไม่เป็น BLUE คือ มีความแปรปรวนไม่ต่ำสุด ถึงแม้จะยังคงเป็นตัวประมาณที่ไม่เอนเอียงก็ตาม เช่นถ้าค่าความคลาดเคลื่อนมีอัตตสหสัมพันธ์ทางบวก (positive autocorrelation) จะมีผลทำให้ค่าความแปรปรวนของตัวประมาณจะมีค่าน้อยกว่าค่าความแปรปรวนที่แท้จริง ฉะนั้นการทดสอบสมมติฐานของตัวประมาณอาจปรากฏว่าต่างจากศูนย์ ทั้งๆที่ความจริงแล้วไม่ต่างจากศูนย์ ซึ่งจะก่อให้เกิดการอนุมานผิดพลาด และส่งผลต่อการพยากรณ์ที่ได้จากสมการถดถอยไม่ถูกต้องแม่นยำ

จากผลกระทบดังกล่าวจึงเป็นที่น่าสนใจ ในการหาวิธีที่เหมาะสมกว่าการใช้วิธีกำลังสองน้อยที่สุด ในการประมาณค่าพารามิเตอร์ของสมการถดถอยเชิงเส้นอย่างง่ายเพื่อการพยากรณ์ ผู้วิจัยได้ทำการศึกษาวิธีการหลายวิธีในการประมาณค่าพารามิเตอร์เมื่อความคลาดเคลื่อนสุ่มมีอัตตสหสัมพันธ์ และสนใจที่จะศึกษาวิจัยวิธีการประมาณสามวิธีต่อไปนี้ โดยเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองจากการพยากรณ์ เมื่อความคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับที่หนึ่ง

1. วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares Method)
2. วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น (Nonlinear Least Squares Method)
3. วิธีการแปลงข้อมูลโดยใช้ผลต่างอันดับที่หนึ่ง (First Difference Transformation Method)

1.2 วัตถุประสงค์ของการวิจัย

เพื่อหาข้อสรุปที่เหมาะสม ในการเลือกวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยเพื่อการพยากรณ์ เมื่อความคลาดเคลื่อนสุ่มในสมการถดถอยมีอัตตสหสัมพันธ์อันดับที่หนึ่ง โดยจะศึกษาเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองจากการพยากรณ์ของวิธีการประมาณค่าพารามิเตอร์ 3 วิธีต่อไปนี้

1.2.1 วิธีกำลังสองน้อยที่สุด (Ordinary Least Squares Method)

1.2.2 วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น (Nonlinear Least Squares Method)

1.2.3 วิธีการแปลงข้อมูลโดยใช้ผลต่างอันดับที่หนึ่ง (First Difference Transformation Method)

1.3 สมมติฐานของการวิจัย

ภายใต้ลักษณะความคลาดเคลื่อนที่มีอัตตสหสัมพันธ์อันดับที่หนึ่ง วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น จะให้ค่าพยากรณ์ที่มีค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองน้อยกว่าวิธีกำลังสองน้อยที่สุดและวิธีการแปลงข้อมูลโดยใช้ผลต่างอันดับที่หนึ่ง

1.4 ข้อตกลงเบื้องต้น

1.4.1 ในการวิจัยครั้งนี้จะใช้สมการถดถอยเชิงเส้นอย่างง่าย โดยมีรูปแบบสมการดังนี้

$$Y_t = \beta_o + \beta_1 X_t + \varepsilon_t \ , \ t = 1, 2, \ldots, n$$

โดยที่

Y_t เป็นตัวแปรตามและเป็นตัวแปรสุ่ม

X_t เป็นตัวแปรอิสระ

β_i เป็นพารามิเตอร์ที่ไม่ทราบค่า , i = 0,1

ε_t เป็นความคลาดเคลื่อนสุ่ม

n เป็นขนาดตัวอย่าง

1.4.2 ค่าความคลาดเคลื่อนมีสหสัมพันธ์กัน โดยกำหนดรูปแบบความสัมพันธ์ เป็นอัตตถดถอยอันดับที่หนึ่ง (first order autoregressive model)

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

โดยที่ $|\rho| < 1$ และมีข้อตกลงเบื้องต้นของ u_t คือ

$E(u_t) = 0$

$V(u_t) = \sigma^2$

$E(u_i u_j) = 0 \, , \, i \neq j$

1.4.3 ความคลาดเคลื่อนเป็นตัวแปรสุ่มที่มีสหสัมพันธ์ และความแปรปรวนคงที่ คือ

$V(\varepsilon_t) = \sigma^2 / (1 - \rho^2) \, , \, t = 1, 2, ..., n$

1.5 ขอบเขตของการวิจัย

1.5.1 ศึกษาภายใต้ลักษณะการแจกแจงของความคลาดเคลื่อน (ε_t) ที่มีการแจกแจงแบบปกติ (normal distribution) มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนคงที่

1.5.2 ศึกษาเมื่อตัวแปรอิสระ X_t มีลักษณะดังนี้

1.5.2.1 แนวโน้มอย่างง่าย (Simple Time Trend)

$X_t = t \, , \, t = 1, 2, ..., n$

1.5.2.2 แนวโน้วอย่างสุ่ม (Stochastic Trend)

$X_t = t + v_t \, , \, t = 1, 2, ..., n$

$V_t \sim N(0,10)$

1.5.2.3 แนวโน้มอย่างเป็นระยะคาบ (Periodic Trend)

$X_t = t + \cos(2\pi t/12) \, , \, t = 1, 2, ..., n$

1.5.3 ศึกษาวิธีประมาณค่าพารามิเตอร์ในสมการถดถอยเพื่อการพยากรณ์ด้วยวิธีกำลังสองน้อยที่สุด วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น และวิธีการแปลงข้อมูลโดยใช้ผลต่างอันดับที่หนึ่ง โดยจะเปรียบเทียบค่าเฉลี่ยของรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) จากการพยากรณ์ ดังนี้

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1.5.3.1 เปรียบเทียบค่าเฉลี่ยแต่ละคาบเวลาตั้งแต่คาบเวลาที่ 1 ถึง 12 คาบเวลา

1.5.3.2 เปรียบเทียบค่าเฉลี่ยเป็นช่วงๆ ในช่วง 3, 6 และ 12 คาบเวลา

1.5.4 ศึกษาเมื่อค่าสหสัมพันธ์ (ρ) เป็น 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, และ 0.9

1.5.5 ศึกษาเมื่อขนาดตัวอย่าง (n) เท่ากับ 20, 30, และ 50

1.5.6 การประมาณค่าพารามิเตอร์ ในแต่ละครั้งจะทำการประมาณ 500 ค่าต่อพารามิเตอร์ 1 ตัว และนำมาหาค่าเฉลี่ยเพื่อใช้เป็นค่าประมาณพารามิเตอร์ ใช้ในตัวแบบพยากรณ์

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 ให้ข้อสรุปในการเลือกใช้วิธีประมาณค่าพารามิเตอร์เพื่อการพยากรณ์ได้อย่างเหมาะสม เมื่อความคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับที่หนึ่ง

1.6.2 เพื่อเป็นแนวทางในการวิจัยต่อไป