บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

การพยากรณ์นั้นเป็นกระบวนการที่มีวัตถุประสงค์เพื่อที่จะทำนายหรือคาดการณ์ในอนาคต ซึ่งมีบทบาทสำคัญอย่างมากในปัจจุบัน ไม่ว่าในสายงานทางด้าน การจัดการ การตลาด การผลิต การจัดการสินค้าคงคลัง เป็นต้น การนำเทคนิคการพยากรณ์มาใช้นั้น ผู้ใช้ควรที่จะเลือกวิธีการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีอยู่ ซึ่งจะทำให้ได้ค่าพยากรณ์ที่ถูกต้องและเชื่อถือได้สูงและวิธีการที่ดีที่สุดสำหรับข้อมูลชนิดหนึ่งนั้นไม่จำเป็นว่าจะต้องดีที่สุดเสมอไปสำหรับข้อมูลชุดอื่น ๆ ด้วย ทั้งนี้ขึ้นอยู่กับข้อจำกัดของข้อมูลที่นำมาวิเคราะห์

การพยากรณ์อนุกรมเวลา (Time Series Forecasting) เป็นการพยากรณ์ที่อาศัยข้อมูลในอดีต โดยจะศึกษาถึงลักษณะการเปลี่ยนแปลงของข้อมูลเมื่อเวลาเปลี่ยนไปว่ามีลักษณะเช่นไร และทำการกำหนดรูปแบบของการเคลื่อนไหวที่เหมาะสมที่สุดสำหรับข้อมูลนั้น โดยจะอยู่ในรูปของความสัมพันธ์กับเวลา การพยาการณ์อนุกรมเวลามีอยู่ด้วยกันหลายวิธี เช่น เทคนิคการทำให้เรียบ (Smoothing Techniques) อนุกรมเวลาแบบคลาสิค (Classical Methods) และการพยากรณ์แบบบอกซ์และเจนกินส์ (Box-Jenkins Methods) เป็นต้น ในแต่ละวิธีการของการพยากรณ์อนุกรมเวลานั้นมีความแตกต่างกันไป ดังนั้นในการนำไปใช้งานจะขึ้นอยู่กับข้อจำกัดของข้อมูลและวัตถุประสงค์ เช่น ในกรณีที่ต้องการความถูกต้องของค่าพยากรณ์สูงและต้องการค่าพยากรณ์ในระยะสั้นหรือปานกลาง อาจใช้การพยากรณ์แบบบอกซ์และเจนกินส์ ซึ่งใช้ได้กับข้อมูลที่มีการเคลื่อนไหวทุกรูปแบบ แต่วิธีการนี้อาจมีความยุ่งยากและซับซ้อนในการกำหนดรูปแบบสมการพยากรณ์ ในกรณีที่ต้องการวิธีการที่คำนวณง่ายและรวดเร็ว แต่อาจได้ความถูกต้องของค่าพยากรณ์ไม่สูงมากนัก การพยากรณ์ด้วยเทคนิคการทำให้เรียบจะเป็นวิธีการที่ใช้ได้วิธีหนึ่ง ซึ่งวิธีการนี้เหมาะสมที่จะใช้ในการพยากรณ์ในคาบระยะเวลาสั้น

เทคนิคการพยากรณ์การทำให้เรียบแบบเอกซโพเนนเชียล เป็นวิธีการหนึ่งในเทคนิคการทำให้เรียบ ซึ่งจะเป็นการเฉลี่ยข้อมูลโดยกำหนดน้ำหนักหรือให้ความสำคัญกับข้อมูลที่อยู่ปัจจุบันมากที่สุดและลดลงเรื่อย ๆ สำหรับข้อมูลที่ระยะห่างไกลออกไปแบบเรขาคณิต การ

ทำให้เรียบแบบเอกซโพเนนเชียลที่นิยมใช้ในทางธุรกิจจะแบ่งได้เป็น 3 กลุ่ม คือการทำให้เรียบแบบเอกซโพเนเชียลครั้งเดียว (Single Exponential Smoothing) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (Double Exponential Smoothing) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสามครั้ง (Triple Exponential Smoothing) ซึ่งการเลือกใช้แต่ละวิธีขึ้นอยู่กับลักษณะของข้อมูลว่ามีการเคลื่อนไหวในลักษณะใด จากเทคนิคการพยากรณ์ทำให้เรียบแบบเอกซโพเนนเชียลข้อมูลที่นำมาวิเคราะห์นั้นจะไม่มีความผิดปกติเกิดขึ้น กล่าวคือ ความคลาดเคลื่อนมีการแจกแจงเป็นแบบปกติ แต่ในทางปฏิบัติข้อมูลอาจเกิดมีค่าผิดปกติ (outliers) โดยความคลาดเคลื่อนมีการแจกแจงแบบปลอมปนเกิดขึ้น ดังนั้นในกรณีเช่นนี้การใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลจะไม่เหมาะสมจึงควรที่จะหาวิธีปรับแก้ ทั้งนี้เพื่อที่จะทำให้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลใช้ได้ในกรณีที่มีข้อมูลผิดปกติ และให้การพยากรณ์มีความคลาดเคลื่อนน้อยที่สุด

ดังนั้นในการวิจัยครั้งนี้จึงสนใจที่จะศึกษาวิธีการพยากรณ์เมื่อมีข้อมูลผิดปกติโดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล โดยข้อมูลอนุกรมเวลาที่ศึกษามีรูปแบบลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ยไม่คงที่ตลอดช่วงเวลา (locally constant mean model) และข้อมูลอนุกรมเวลาที่มีแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเวลา (locally constant linear trend model) ภายใต้ความคลาดเคลื่อนที่มีการแจกแแจงปลอมปนแบบต่าง ๆ และวิธีพยากรณ์ที่ใช้ในการศึกษาเปรียบเทียบ คือ

1. เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล
2. เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (Exponential Smoothing Based on Least Absolute Value)
3. เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (Huber Exponential Smoothing)

<u>วัตถุประสงค์ของการวิจัย</u>

1. เพื่อศึกษาวิธีพยากรณ์เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนโดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด และเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์

2. เพื่อเปรียบเทียบความถูกต้องของการพยากรณ์ด้วยเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด และเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ในข้อมูลอนุกรมเวลาเมื่อมีข้อมูลผิดปกติ

<u>สมมติฐานของการวิจัย</u>

เมื่อมีข้อมูลผิดปกติเกิดขึ้น เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุดและวิธีการของฮูเบอร์จะให้ความคลาดเคลื่อนโดยเฉลี่ยต่ำกว่าเมื่อใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล

<u>ขอบเขตของการวิจัย</u>

1. ข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์มี 2 ลักษณะคือ

1.1 ข้อมูลอนุกรมเวลาที่มีการเคลื่อนไหวในลักษณะระดับค่าเฉลี่ยไม่คงที่ตลอดช่วงคงที่ตลอดช่วงเวลา T ที่พิจารณา

$$Y_t = a + \varepsilon_t$$

โดยที่ a คือ ระดับหรือค่าเฉลี่ยของข้อมูล ซึ่งมีค่าแปรเปลี่ยนช้า ๆ ตามเวลา
กำหนดให้ a = 5
ε_t คือ ค่าความคลาดเคลื่อนสุ่มที่เวลา t

1.2 ข้อมูลอนุกรมเวลาที่มีการเคลื่อนไหวแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเวลา T ที่พิจารณา

$$Y_t = a + bt + \varepsilon_t$$

โดยที่ a คือ ระดับค่าเฉลี่ย ซึ่งค่าแปรเปลี่ยนช้า ๆ ตามเวลา

b คือ ความชันหรือองค์ประกอบเชิงเส้น ซึ่งค่าแปรเปลี่ยนช้า ๆ ตามเวลา

ε_t คือ ค่าความคลาดเคลื่อนสุ่มที่เวลา t

กำหนดให้ a = 2 และ b = 2

2. ลักษณะการแจกแจงของความคลาดเคลื่อน (ε_t) ที่ศึกษามีดังนี้

2.1 การแจกแจงปกติ (Normal Distribution)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad ; -\alpha < x < \alpha$$

ในที่นี้กำหนดให้ $\mu = 0$ และ $\sigma = 1$

2.2 ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$$

เมื่อ c คือ สเกลแฟกเตอร์ (scale factor)

ในที่นี้ใช้ c = 3 และ 10

p คือ เปอร์เซ็นต์การปลอมปน (percent of contamination)

ในที่นี้ใช้ p = 5, 10 และ 20

2.3 ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$$

เมื่อ p คือ เปอร์เซ็นต์การปลอมปน ในที่นี้ใช้ p = 5, 10 และ 20

$L(0,\beta)$ คือ การแจกแจงลาปลาซ (Laplace Distribution) ซึ่งมีฟังก์ชันความหนาแน่น

$$f(x) = \frac{1}{2\beta} \exp\left(\frac{-|x-\alpha|}{\beta}\right)$$

$-\infty < x < \infty$, $-\infty < \alpha < \infty$, $\beta > 0$

โดยที่ $E(X) = \alpha$ และ $Var(X) = 2\beta^2$

ในที่นี้ให้ α = 0 และ β = 1 และ 10

2.4 ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$f(x) = (1-p)\ N(0,1) + p\ U(a,b)$$

เมื่อ p คือ เปอร์เซ็นต์การปลอมปน ในที่นี้ใช้ p = 5, 10 และ 20

U(a,b) คือ การแจกแจงยูนิฟอร์ม (Uniform Distribution) ซึ่งมีฟังก์ชัน

ความหนาแน่น $f(x) = \frac{1}{(b-a)}$; $a < x < b$

โดยที่ $E(X) = \frac{(a+b)}{2}$, $V(X) = \frac{(b-a)^2}{12}$

ในที่นี้ให้ (a,b) เท่ากับ (-5,5) และ (-10,10)

3. ขนาดตัวอย่างที่ใช้ในการศึกษามี 4 ระดับคือ 10, 20, 30 และ 50

4. การวิจัยครั้งนี้ได้จำลองข้อมูลให้มีสถานการณ์ตามที่กำหนดข้างต้น โดยใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) จากเครื่องคอมพิวเตอร์ IBM AMDAHL 5860 เขียนโปรแกรมด้วยภาษาฟอร์แทรน (Fortran) ทำการจำลองข้อมูลซ้ำ 500 รอบในแต่ละสถานการณ์

<u>ประโยชน์ที่คาดว่าจะได้รับ</u>

1. ผลจากการศึกษาเปรียบเทียบสามารถแสดงได้ว่า ถ้าข้อมูลอนุกรมเวลาในรูปแบบระดับค่าเฉลี่ยไม่คงที่ตลอดช่วงเวลาหรือมีแนวโน้มแชิงเส้นไม่คงที่ตลอดช่วงเวลา เมื่อมีค่าผิดปกติผู้วิจัยควรใช้วิธีใดภายใต้เงื่อนไขใด จึงจะทำให้ผลการพยากรณ์มีค่าผิดพลาดน้อยที่สุด

2. ผลจากการศึกษาในการพยากรณ์ค่าโดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล เมื่อมีข้อมูลผิดปกติสามารถที่จะเป็นแนวทางในการศึกษาในรูปแบบอื่นของการพยากรณ์ต่อไป เช่น ในกรณีข้อมูลมีฤดูกาล