บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบวิธีพยากรณ์ เมื่อมีการนำเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลมาใช้ในการวิเคราะห์อนุกรมเวลาที่มีข้อมูลสูญหาย วิธีพยากรณ์ที่นำมาศึกษาเปรียบเทียบมี 3 วิธี คือ วิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท วิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของไรท์ และวิธีพยากรณ์ที่มีการประมาณค่าสูญหาย ซึ่งสองวิธีแรกนั้น จะเป็นการปรับแก้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเพื่อให้สามารถทำการพยากรณ์ได้ โดยไม่ต้องประมาณค่าของข้อมูลที่สูญหาย ส่วนวิธีที่มีการประมาณค่าสูญหายนั้นจะประมาณข้อมูลส่วนที่หายไป ด้วยค่าพยากรณ์จากเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล สำหรับการเปรียบเทียบ จะกระทำภายใต้สถานการณ์ต่างๆ ดังนี้

(1) เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล 2 เทคนิค คือ เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวและซ้ำสองครั้ง

(2) ขนาดตัวอย่าง 4 ขนาด คือ 10, 15, 30 และ 50

(3) สัดส่วนของข้อมูลสูญหาย 3 ระดับ คือ 10%, 20% และ 30%

(4) สัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย 4 ระดับ คือ 10%, 20%, 30% และ 40%

นอกจากจะศึกษาเปรียบเทียบวิธีพยากรณ์ทั้ง 3 วิธีดังกล่าวแล้ว ยังศึกษาถึงผลกระทบของจำนวนข้อมูลสูญหาย และจำนวนข้อมูลหลังช่วงข้อมูลสูญหายที่ต่อความคลาดเคลื่อนของการพยากรณ์ด้วย

การวิจัยครั้งนี้ใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) สร้างสถานการณ์ต่างๆ ดังนั้นจะขอกล่าวถึงวิธีการจำลองโดยใช้เทคนิคมอนติคาร์โลก่อน แล้วจึงแสดงรายละเอียดของขั้นตอนการวิจัยและโปรแกรมที่ใช้ในการวิจัยตามลำดับ

3.1 วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคที่ใช้แก้ปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โลเป็นวิธีหนึ่งที่นิยมนำมาใช้แก้ปัญหากันอย่างแพร่หลายในปัจจุบัน ซึ่งหลักการของการจำลองโดยใช้เทคนิคดังกล่าว จะอาศัยตัวเลขสุ่ม (Random numbers) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ขั้นตอนของวิธีการจำลองด้วยเทคนิคมอนติคาร์โลที่ใช้กันในปัจจุบัน แบ่งได้เป็น 3 ขั้นตอน ดังนี้

3.1.1 การสร้างตัวเลขสุ่ม การใช้ตัวเลขสุ่มเป็นสิ่งที่สำคัญมากในเทคนิคนี้ ทั้งนี้เพราะว่าหลักการของการจำลองแบบมอนติคาร์โลนั้น จะใช้ตัวเลขสุ่มมาช่วยในการค้นหาคำตอบของปัญหา โดยลักษณะของตัวเลขสุ่มที่นำมาใช้ จะมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) สำหรับวิธีการสร้างตัวเลขสุ่ม มีผู้เสนอไว้หลายวิธี แต่วิธีที่ดีนั้นลักษณะของตัวเลขสุ่มที่ถูกสร้างขึ้นจะต้องมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) ตัวเลขสุ่มแต่ละตัวเป็นอิสระต่อกัน และมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ (มีวัฏจักรยาว)

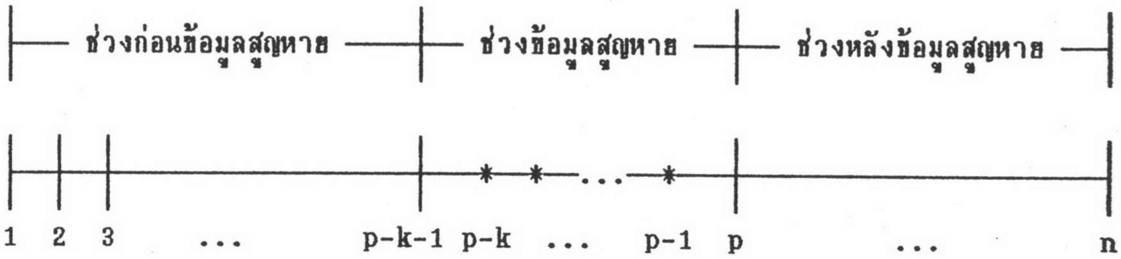
3.1.2 การนำตัวเลขสุ่มมาประยุกต์ใช้กับปัญหาที่ต้องการศึกษา ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจจะมีขั้นตอนอื่นอีกหลายๆ ขั้นตอนที่ต้องการใช้ตัวเลขสุ่ม

3.1.3 การทดลองกระทำ เมื่อนำตัวเลขสุ่มมาประยุกต์ให้เข้ากับปัญหาที่ต้องการศึกษาได้แล้ว ขั้นต่อไปก็คือ การทดลองโดยใช้กระบวนการของการสุ่ม (Random Process) มากระทำในลักษณะที่ซ้ำๆ กันหลายๆ ครั้ง เพื่อหาคำตอบที่ต้องการ

3.2 แผนการทดลอง

การวิจัยครั้งนี้ต้องการเปรียบเทียบความสามารถในการพยากรณ์ของวิธีพยากรณ์ทั้ง 3 วิธี เมื่อมีการนำเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลมาใช้ในการวิเคราะห์อนุกรมเวลาที่มีข้อมูลสูญหายแบบ 1 ช่วง (single gap) ดังนั้นก่อนที่จะนำข้อมูลมาทำการวิเคราะห์ จะแบ่งข้อมูลออกเป็น 3 ช่วง คือ ช่วงก่อนข้อมูลสูญหาย ช่วงที่มีข้อมูลสูญหาย และช่วงหลังข้อมูลสูญหาย ดังแสดงในรูปที่ 3.1

รูปที่ 3.1 แสดงการแบ่งช่วงของข้อมูลที่นำมาวิเคราะห์



ช่วง 1,..., p-k-1 เป็นช่วงก่อนข้อมูลสูญหาย

ช่วง p-k,..., p-1 เป็นช่วงที่มีข้อมูลสูญหาย

และช่วง p,...,n เป็นช่วงหลังข้อมูลสูญหาย

เนื่องจากข้อมูลในช่วง 1,...,p-k-1 และ p,...,n เป็นช่วงที่ทราบค่าของข้อมูล จึงนำข้อมูลทั้งสองช่วงนี้มาใช้ในการวิเคราะห์ และใช้สำหรับค้นหา (search) ค่าสัมประสิทธิ์ส่วนลด (a) เพราะว่าจำนวนข้อมูลในแต่ละช่วง อาจส่งผลต่อความสามารถในการพยากรณ์ ดังนั้นจะกำหนดจำนวนข้อมูลของแต่ละช่วงไม่คงที่ โดยจำนวนข้อมูลในช่วง p-k ถึง p-1 ซึ่งเป็นช่วงที่มีข้อมูลสูญหายจะมีสัดส่วนเป็น 10%, 20% และ 30% ของจำนวนข้อมูลทั้งหมด จำนวนข้อมูลในช่วง p ถึง n (ช่วงหลังข้อมูลสูญหาย) มีสัดส่วนเป็น 10%, 20%, 30% และ 40% ของจำนวนข้อมูลทั้งหมด ส่วนข้อมูลในช่วง 1 ถึง p-k-1 (ช่วงก่อนข้อมูลสูญหาย) จะมีจำนวนเท่ากับจำนวนข้อมูลทั้งหมดลบด้วยจำนวนข้อมูลที่สูญหาย และจำนวนข้อมูลหลังช่วงที่มีข้อมูลสูญหาย

เมื่อกำหนดให้ n แทน จำนวนข้อมูลทั้งหมด

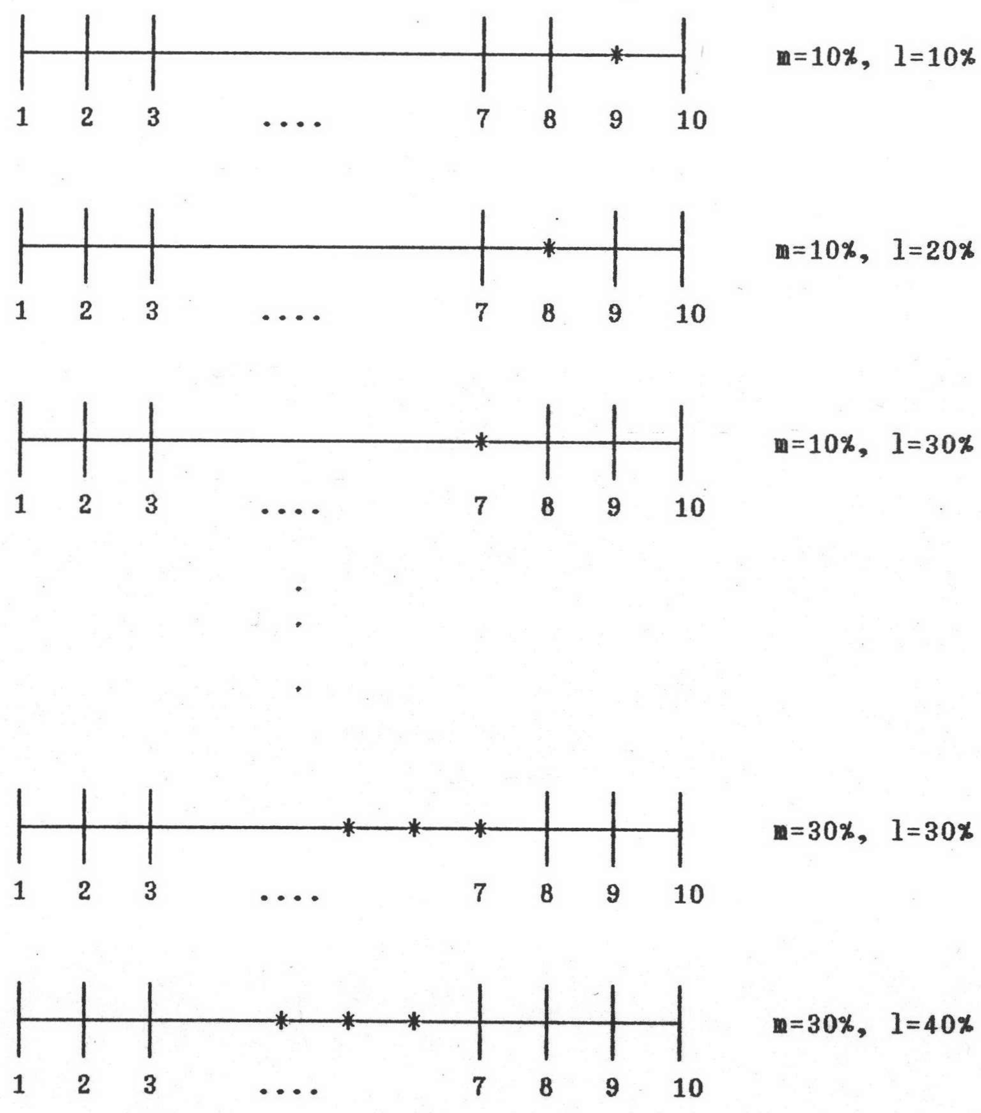
m แทน สัดส่วนของข้อมูลสูญหาย

l แทน สัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย

จะพบว่าการเปลี่ยนขนาดของ m และ l เปรียบเสมือนการเลื่อนตำแหน่งของข้อมูลที่สูญหายไป ณ ตำแหน่งต่างๆ ของข้อมูลอนุกรมเวลา นั่นคือ ถ้า m และ l มีค่าต่ำ ตำแหน่งของข้อมูลที่สูญหายจะเกิดในตอนท้ายของข้อมูลชุดนั้น แต่ถ้า m และ l มีค่าสูง ตำแหน่งของข้อมูลที่สูญหายจะเกิดในตอนต้นของข้อมูล ดังจะแสดงให้เห็นในตัวอย่างที่ 3.1

ตัวอย่างที่ 3.1 การแสดงตำแหน่งที่มีข้อมูลสูญหาย เมื่อขนาดของอนุกรมเวลา (n) = 10 สัดส่วนของข้อมูลสูญหาย (m) และสัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย (l) มีค่าต่างๆ (* แทนตำแหน่งที่มีข้อมูลสูญหาย)

รูปที่ 3.2 แสดงตำแหน่งที่มีข้อมูลสูญหาย



เมื่อแบ่งข้อมูลออกเป็น 3 ช่วงตามที่กล่าวมาแล้วนั้น จะทำการเปรียบเทียบความคลาดเคลื่อนจากการพยากรณ์ของวิธีพยากรณ์ทั้ง 3 วิธี ณ สัดส่วนของข้อมูลสูญหาย และสัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหายจากน้อยไปมาก 3 และ 4 ระดับตามลำดับ ขนาดตัวอย่าง 4 ขนาด และเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล 2 เทคนิค (เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวและซ้ำสองครั้ง) รวมทั้งสิ้น 96 สถานการณ์ การเปรียบเทียบจะพิจารณาจากค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ของวิธีพยากรณ์ทั้ง 3 วิธี เพื่อหาวิธีที่ดีที่สุดในแต่ละสถานการณ์ต่อไป

3.3 ขั้นตอนในการวิจัย

ขั้นตอนในการวิจัย แบ่งเป็น 6 ขั้นตอนหลัก คือ

1. จำลองค่าความคลาดเคลื่อน (ε_t) ของข้อมูลอนุกรมเวลาจากการแจกแจงแบบปกติ $N(\mu,\sigma^2)$
2. จำลองข้อมูลอนุกรมเวลา (y_t) ที่มีลักษณะตามรูปแบบต่อไปนี้

   2.1 ลักษณะของข้อมูลอนุกรมเวลาที่มีระดับค่อนข้างคงที่

   $$y_t = \mu + \varepsilon_t \quad , \; t = 1,2,\ldots,n,n+1,\ldots,n+12$$

   2.2 ลักษณะของข้อมูลอนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรง

   $$y_t = \mu + \beta t + \varepsilon_t$$

3. กำหนดค่าเริ่มต้น (S_o) ของการทำให้เรียบแบบเอกซโพเนนเชียล
4. ค้นหาค่าสัมประสิทธิ์ส่วนลด (a)
5. คำนวณหาค่าพยากรณ์จากวิธีพยากรณ์ทั้ง 3 วิธี
6. หาค่าความคลาดเคลื่อนจากการพยากรณ์ แล้วทำการเปรียบเทียบ

แต่ละขั้นตอนมีรายละเอียดดังนี้

3.3.1 การจำลองค่าความคลาดเคลื่อนของข้อมูลอนุกรมเวลา

จำลองค่าความคลาดเคลื่อน ε_t จากการแจกแจงแบบปกติ มีค่าเฉลี่ย = 0 ความแปรปรวนคงที่ = σ_ε^2 = 10 ด้วยวิธี Box-Muller (รายละเอียดแสดงในภาคผนวก ก)

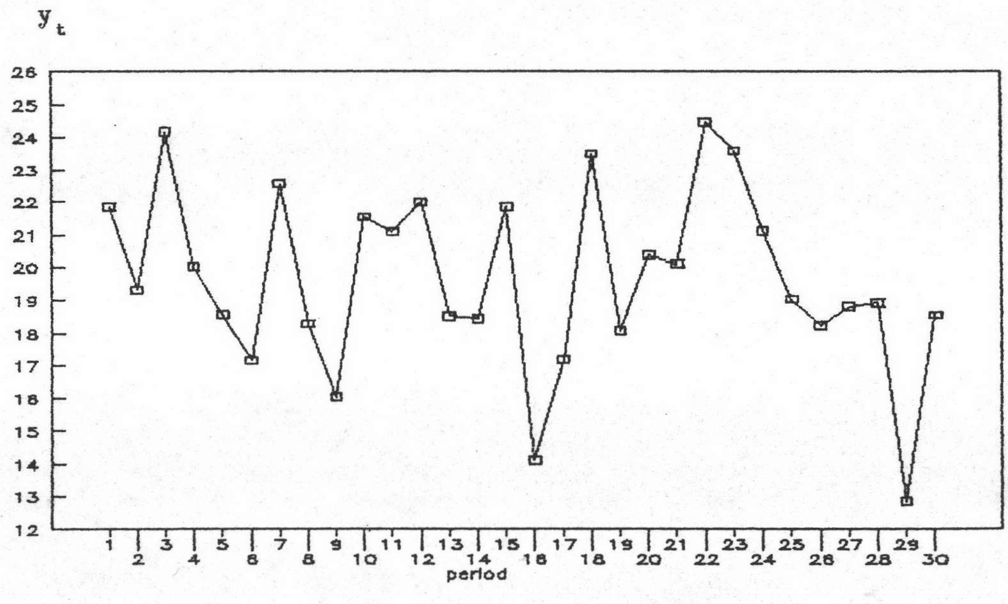
3.3.2 การจำลองข้อมูลอนุกรมเวลา

3.3.2.1 ข้อมูลอนุกรมเวลาที่มีระดับค่อนข้างคงที่

จำลอง y_t , $t = 1,2,\ldots,n,n+1,\ldots,n+12$ ให้มีรูปแบบความสัมพันธ์เป็น $y_t = \mu + \varepsilon_t$ โดยที่ μ เป็นพารามิเตอร์ที่กำหนดขึ้นมาแทนระดับของข้อมูล (ในการวิจัยครั้งนี้ กำหนด $\mu = 20$) จะได้ y_t ที่เวลาต่างๆ ดังนี้

$$
\begin{aligned}
y_1 &= \mu + \varepsilon_1 \\
y_2 &= \mu + \varepsilon_2 \\
&\vdots \\
y_{n+12} &= \mu + \varepsilon_{n+12}
\end{aligned}
$$

สำหรับลักษณะของข้อมูล ($y_t = \mu + \varepsilon_t$) ที่สร้างขึ้นจากการทดลอง จะแสดงในรูปของกราฟได้ดังนี้



รูปที่ 3.3 แสดงลักษณะของข้อมูล, $y_t = \mu + \varepsilon_t$ ที่สร้างขึ้นจากการทดลอง

3.3.2.2 ข้อมูลอนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรง

จำลอง y_t ให้มีรูปแบบความสัมพันธ์เป็น $y_t = \mu + \beta t + \varepsilon_t$ โดยที่ μ และ β เป็นพารามิเตอร์ที่กำหนดขึ้นมาแทนระดับ และความชันของข้อมูลตามลำดับ (ในการวิจัยครั้งนี้กำหนด $\mu = 20$ และ $\beta = 5$) จะได้ y_t ที่เวลาต่างๆ ดังนี้

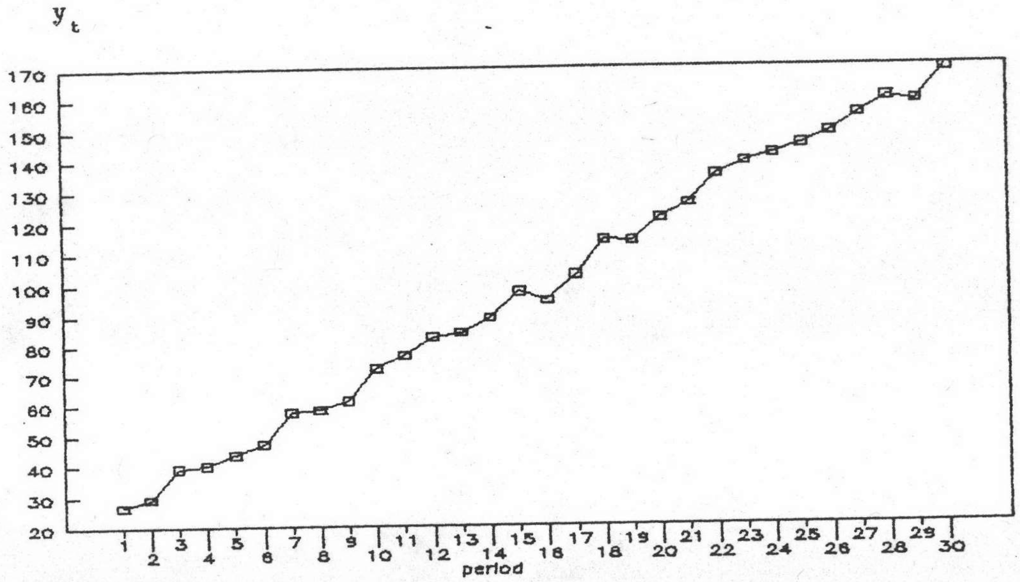
$$y_1 = \mu + \beta + \varepsilon_1$$

$$y_2 = \mu + 2\beta + \varepsilon_2$$

$$\vdots$$

$$y_{n+12} = \mu + (n+12)\beta + \varepsilon_{n+12}$$

สำหรับลักษณะของข้อมูล ($y_t = \mu + \beta t + \varepsilon_t$) ที่สร้างขึ้นจากการทดลองจะแสดงในรูปของกราฟได้ดังนี้



รูปที่ 3.4 แสดงลักษณะของข้อมูล, $y_t = \mu + \beta t + \varepsilon_t$ ที่สร้างขึ้นจากการทดลอง

3.3.3 การกำหนดค่าเริ่มต้นของการทำให้เรียบแบบเอกซโพเนนเชียล

ณ เวลาที่ t=1 การคำนวณหาค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่ r ; $S_t^{(r)}$, r = 1,2,...,i จะต้องทราบค่าเริ่มต้น ($S_o^{(r)}$) ด้วยเหตุนี้จึงต้องกำหนดค่าเริ่มต้นขึ้นมาก่อน และการกำหนดค่าเริ่มต้นว่าควรจะมีค่าเป็นเท่าไรนั้น ได้มีผู้เสนอแนวความคิดและวิธีการในการกำหนดค่าเริ่มต้นหลายวิธีด้วยกัน สำหรับการวิจัยครั้งนี้ จะใช้ข้อมูลตัวแรกประมาณค่าของ $S_o^{(r)}$ ดังนี้

กรณีที่ใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว จะกำหนดค่าเริ่มต้น $S_o^{(1)} = y_1$ ส่วนกรณีที่ใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง จะกำหนด $S_o^{(1)} = S_o^{(2)} = y_1$ การเลือกใช้วิธีดังกล่าวในการกำหนดค่าเริ่มต้น เพราะว่าจุดประสงค์หลักของการวิจัยนี้ต้องการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ ซึ่งในแต่ละวิธีพยากรณ์ที่นำมาเปรียบเทียบ จะกำหนดค่าเริ่มต้นด้วยวิธีเดียวกัน และภายใต้สถานการณ์ต่างๆ เดียวกัน ดังนั้นการกำหนดค่าเริ่มต้นไม่ว่าจะใช้วิธีใดก็ตาม ผลสรุปของการเปรียบเทียบจะไม่เปลี่ยนแปลง

3.3.4 การคำนวณหาค่าสัมประสิทธิ์ส่วนลด (a)

เนื่องจากปัญหาในการพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล คือ ควรกำหนดค่า a เป็นเท่าไรจึงจะเหมาะสมนั้น มักจะกำหนดโดยการทดลองหาค่าความคลาดเคลื่อนของการพยากรณ์ล่วงหน้า 1 คาบเวลา (one-step-ahead forecast errors), $e_{t-1}(1)$, ณ ค่า a ต่างๆ

$$e_{t-1}(1) = y_t - \hat{y}_{t-1}(1) \quad , t=1,2,\ldots,n$$

โดยที่ $\hat{y}_{t-1}(1) = S_{t-1} = (1-a)y_{t-1} + aS_{t-2}$ สำหรับการพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว ส่วนการพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง จะคำนวณหาค่า $\hat{y}_{t-1}(1)$ จากสูตร

$$\hat{y}_{t-1}(1) = [2+(1-a)/a]S_{t-1}^{(1)} - [1+(1-a)/a]S_{t-1}^{(2)} \quad , t=1,2,\ldots,n$$

จากนั้นคำนวณหาค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองที่ค่าต่างๆ ของ a จากสูตร

$$MSE(a) = [\sum_{t=1}^{n} e^2_{t-1}(1)] / n$$

ค่า a ที่ให้ค่า MSE ต่ำสุด จะเป็นค่าสัมประสิทธิ์ส่วนลดที่ถูกนำไปใช้เพื่อการพยากรณ์ต่อไป เนื่องจาก a มีค่าระหว่าง 0 ถึง 1 ($0<a<1$) ดังนั้นในการวิจัยครั้งนี้ จะคำนวณหาค่า a โดยเริ่มด้วยค่า a = 0.01 จากนั้นจะเพิ่มค่า a ครั้งละ 0.01 และจะหยุดเมื่อ a = 0.99 จากนี้จะเปรียบเทียบค่า MSE(a) สำหรับทุกค่า a และจะเลือกค่า a ที่ให้ MSE(a) ต่ำสุด

โดยปกติถ้าไม่มีข้อมูลสูญหาย ผู้วิเคราะห์จะนำข้อมูลทั้งหมดมาใช้ในการค้นหาค่า a แต่การวิจัยครั้งนี้เป็นการวิเคราะห์ในกรณีที่มีข้อมูลสูญหายไปแบบ 1 ช่วง ดังนั้นเมื่อจำลองข้อมูล (y_t) และกำหนดสัดส่วนของช่วงต่างๆ ของข้อมูลได้แล้ว จะนำข้อมูลที่เราทราบค่า (ข้อมูลในช่วงก่อนและหลังสูญหาย) มาใช้ค้นหาค่า a ที่ทำให้ MSE จากการพยากรณ์ของข้อมูลทั้งสองช่วงมีค่าต่ำสุด

ขั้นตอนของการค้นหาค่า a โดยใช้ข้อมูลช่วงก่อนและหลังสูญหายมีรายละเอียดดังนี้

1. หาผลรวมของความคลาดเคลื่อนจากการพยากรณ์ล่วงหน้า 1 คาบเวลาของข้อมูลในช่วงก่อนข้อมูลสูญหาย

$$SSE(a)_{before} = \sum_{t=1}^{P-k-1} e^2_{t-1}(1)$$

2. หาผลรวมของความคลาดเคลื่อนจากการพยากรณ์ล่วงหน้า 1 คาบเวลาของข้อมูลในช่วงหลังข้อมูลสูญหาย (ในกรณีที่มีจำนวนข้อมูลในช่วงหลังข้อมูลสูญหายเพียง 1 ค่า ซึ่งไม่สามารถคำนวณหาค่าความคลาดเคลื่อนจากการพยากรณ์ได้ จะไม่นำข้อมูลในส่วนนี้มาใช้ในการค้นหาค่า a)

$$SSE(a)_{after} = \sum_{t=P+1}^{n} e^2_{t-1}(1)$$

3. หาค่าเฉลี่ยของความคลาดเคลื่อนจากการพยากรณ์ของข้อมูลทั้งสองช่วง

$$MSE(a) = \frac{SSE(a)_{before}}{p-k-1} + \frac{SSE(a)_{after}}{n-p}$$

4. เปรียบเทียบค่า MSE(a) จากค่า a ทุกค่า แล้วเลือกค่า a ที่ให้ค่า MSE ต่ำสุดเป็นค่าสัมประสิทธิ์ส่วนลดเพื่อใช้ในการพยากรณ์ต่อไป

3.3.5 การคำนวณหาค่าพยากรณ์จากวิธีพยากรณ์ทั้ง 3 วิธี

เนื่องจากเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลที่ใช้ในการวิจัยครั้งนี้มี 2 เทคนิค คือ เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวและซ้ำสองครั้ง ดังนั้นการคำนวณหาค่าพยากรณ์จากวิธีพยากรณ์ทั้งสามจะพิจารณาทีละเทคนิคดังนี้

3.3.5.1 เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว

เมื่อจำลองข้อมูลอนุกรมเวลา y_t $(t=1,2,...,n,n+1,...,n+12)$ ที่มีระดับ (level) ของข้อมูลคงที่ ซึ่งเป็นลักษณะของข้อมูลที่เหมาะกับการพยากรณ์โดยใช้เทคนิคดังกล่าวแล้ว จะนำข้อมูล n ค่าแรกมากำหนดสัดส่วนของข้อมูลสูญหาย (m) และสัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย (l) จากนั้นจะนำข้อมูลในช่วงก่อนและหลังสูญหาย $(y_1,y_2,...,y_{P-k-1}$ และ $y_P,y_{P+1},...,y_n)$ มาค้นหาค่า a ที่เหมาะสมก่อน ในขั้นต่อไปจึงคำนวณหาค่าพยากรณ์ล่วงหน้า 12 คาบเวลา จากวิธีพยากรณ์ทั้งสาม โดยมีรายละเอียดดังต่อไปนี้

1. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท</u> มีขั้นตอนการคำนวณดังนี้

1.1 คำนวณหาค่า S_t, $t = 1,2,...,p-k-1$ จากสมการที่ (2.8) ในบทที่ 2 (หัวข้อ 2.1.1)

1.2 คำนวณหาค่า S_t, $t = p,p+1,p+2,...,n$

1.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า S_p จากสมการที่ (2.9)

1.2.2 ค่า S_t ที่เวลา $t = p+1,p+2,...,n$ จะคำนวณจากสมการที่ (2.8)

1.3 คำนวณหาค่าพยากรณ์จากสมการที่ (2.11)

2. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการปรับแก้ด้วยวิธีของไรท์</u> มีขั้นตอนการคำนวณดังนี้

2.1 คำนวณหาค่า S_t, $t = 1,2,...,p-k-1$ จากสมการที่ (2.8)

2.2 คำนวณหาค่า S_t, $t = p,p+1,p+2,...,n$

2.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า S_p จากสมการที่ (2.12)

2.2.2 ที่เวลา $t = p+1,p+2,...,n$ จะคำนวณหาค่า S_t จากสมการที่ (2.13)

2.3 คำนวณหาค่าพยากรณ์ด้วยสมการที่ (2.14)

3. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการประมาณค่าสูญหาย</u> มีขั้นตอนการคำนวณดังนี้

3.1 คำนวณหาค่า S_t, $t = 1,2,...,p-k-1$ จากสมการที่ (2.8)

3.2 ใช้ค่า S_{p-k-1} ซึ่งคำนวณได้จากสมการที่ (2.15) เป็นค่าประมาณของข้อมูลสูญหายทุกคาบเวลาที่มีข้อมูลสูญหาย ($t = p-k,p-k+1,...,p-1$)

3.3 คำนวณหาค่า S_t, $t = p-k,p-k+1,...,p-1,p,p+1,...,n$ จากสมการที่ (2.8)

3.4 คำนวณหาค่าพยากรณ์จากสมการที่ (2.11)

3.3.5.2 เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง

เมื่อจำลองข้อมูลอนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรง (linear trend) ซึ่งเป็นลักษณะของข้อมูลที่เหมาะกับการพยากรณ์โดยใช้เทคนิคนี้ จะกำหนดค่า m และ l เพื่อแบ่งช่วงของข้อมูล จากนั้นจะนำข้อมูลในช่วงก่อนและหลังสูญหายมาค้นหาค่า a ที่เหมาะสม แล้วจึงหาค่าพยากรณ์ล่วงหน้า 12 คาบเวลาจากวิธีพยากรณ์ทั้งสามในขั้นต่อไป

1. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท</u> มีขั้นตอนดังนี้

1.1 ประมาณค่า μ_t และ β_t, $t=1,2,...,p-k-1$ จากสมการที่ (2.22) และ (2.23) ในบทที่ 2 (หัวข้อ 2.1.2)

1.2 ประมาณค่า μ_t และ β_t, $t = p,p+1,p+2,...,n$

1.2.1 ที่เวลา $t = p$ จะประมาณค่า μ_p และ β_p จากสมการที่ (2.24) และ (2.25) ตามลำดับ

1.2.2 ที่เวลา $t = p+1,p+2,...,n$ จะประมาณค่า μ_t และ β_t จากสมการที่ (2.22) และ (2.23)

1.3 คำนวณหาค่าพยากรณ์จากสมการที่ (2.28)

2. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการปรับแก้ด้วยวิธีของไรท์</u> มีขั้นตอนดังนี้

2.1 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t, $t = 1,2,...,p-k-1$ จากสมการที่ (2.29), (2.30), (2.31) และ (2.32) ตามลำดับ

2.2 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t, $t = p,p+1,p+2,...,n$

2.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า $S_p^{(1)}$ และ $S_p^{(2)}$ จากสมการที่ (2.33) และ (2.34) ตามลำดับ สำหรับ μ_p และ β_p จะประมาณจากสมการที่ (2.35) และ (2.36)

2.2.2 ที่เวลา $t = p+1,p+2,\ldots,n$ จะคำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ จากสมการที่ (2.37) และ (2.38) และจะประมาณค่า μ_t และ β_t จากสมการที่ (2.39) และ (2.40) ตามลำดับ

2.3 คำนวณหาค่าพยากรณ์จากสมการที่ (2.28)

3. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการประมาณค่าสูญหาย</u> มีขั้นตอนดังนี้

3.1 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t, $t = 1,2,\ldots,p-k-1$ จากสมการที่ (2.29), (2.30), (2.31) และ (2.32) ตามลำดับ

3.2 หาสมการพยากรณ์ที่เวลา $t = p-k-1$ จากสูตร

$$\hat{y}_{p-k-1}(\tau) = \hat{\mu}_{p-k-1} + \tau\,\hat{\beta}_{p-k-1} \quad , \tau = 1,2,\ldots,k$$

3.3 นำสมการที่ได้จากข้อ 3.2 พยากรณ์ไปทุกช่วงเวลาที่มีข้อมูลสูญหาย

3.4 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t, $t = p-k,p-k+1,\ldots,p-1,p,p+1,\ldots,n$ จากสมการที่ (2.29), (2.30), (2.31) และ (2.32) ตามลำดับ

3.5 คำนวณหาค่าพยากรณ์จากสมการที่ (2.28)

3.3.6 <u>การคำนวณหาค่าความคลาดเคลื่อนจากการพยากรณ์และทำการเปรียบเทียบ</u>

การทดลองในสถานการณ์หนึ่งๆ เมื่อได้ค่าพยากรณ์ล่วงหน้า 12 คาบเวลาครบทั้ง 3 วิธีแล้ว จะนำค่าพยากรณ์มาเปรียบเทียบกับค่าจริง $(y_{n+1},y_{n+2},\ldots,y_{n+12})$ เพื่อคำนวณหาค่าความคลาดเคลื่อนกำลังสอง (Square Error : SE) ของแต่ละคาบเวลา และทำซ้ำเช่นเดิมจนครบ 1,000 ครั้ง แล้วจึงคำนวณหาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองในแต่ละคาบเวลา ($RMSE_t$) จากสูตร

$$RMSE_t = \sqrt{\sum_{i=1}^{1000} \frac{(y_{ti} - \hat{y}_{ti})^2}{1000}} \quad , t = n+1, n+2, \ldots, \quad n+12$$

จากนั้นจะนำค่า $RMSE_t$ ของวิธีพยากรณ์ทั้ง 3 วิธีมาทำการเปรียบเทียบเพื่อหาวิธีพยากรณ์ที่ดีที่สุดของแต่ละสถานการณ์ สำหรับวิธีการเปรียบเทียบจะพิจารณาทีละคาบเวลาจนครบทั้ง 12 คาบเวลา แล้วจึงพิจารณาค่าเฉลี่ยของ RMSE ทั้ง 12 คาบเวลา เพื่อเปรียบเทียบโดยรวมอีกครั้งหนึ่ง เมื่อทราบวิธีพยากรณ์ที่ดีที่สุดในแต่ละสถานการณ์แล้ว จะศึกษาถึงผลกระทบของจำนวนข้อมูลสูญหาย และจำนวนข้อมูลหลังช่วงข้อมูลสูญหาย ที่มีต่อค่าความคลาดเคลื่อนของการพยากรณ์ ภายใต้สมมติฐานว่า "ถ้าข้อมูลหลังช่วงข้อมูลสูญหายมีจำนวนเพิ่มขึ้น (ขณะที่จำนวนข้อมูลสูญหายคงที่) ค่าพยากรณ์ จะมีความคลาดเคลื่อนลดลง" ดังนั้นถ้าสมมติฐานเป็นจริง จะแสดงว่า จำนวนข้อมูลสูญหาย และจำนวนข้อมูลหลังช่วงข้อมูลสูญหายมีผลกระทบต่อค่าความคลาดเคลื่อนของการพยากรณ์

การคำนวณหาค่า $RMSE_t$ จากการพยากรณ์ของวิธีพยากรณ์ทั้งสาม ในการทดลองที่ขนาดตัวอย่างหนึ่งๆ จะเปลี่ยนค่าสัดส่วนของข้อมูลสูญหายและสัดส่วนของข้อมูลหลังช่วงที่มีข้อมูลสูญหายทีละค่าจนครบ 3 และ 4 ค่าตามลำดับ โดยแต่ละสถานการณ์จะทดลองซ้ำๆ จำนวน 1,000 ครั้ง จนครบทุกสถานการณ์ ซึ่งขั้นตอนของการทดลองดังกล่าวนี้ สรุปเป็นผังงานได้ดังรูปที่ 3.5

3.4 โปรแกรมที่ใช้ในการวิจัย

โปรแกรมที่ใช้ในการวิจัยทั้งหมด เขียนด้วยภาษาฟอร์แทรน (FORTRAN) โดยใช้กับเครื่อง AMDAHL 5860 ซึ่งในแต่ละสถานการณ์ของการทดลอง ลักษณะการทำงานของโปรแกรมจะเหมือนกัน สำหรับรายละเอียดของโปรแกรมจะแสดงไว้ในภาคผนวก ข และภาคผนวก ค (ภาคผนวก ข และภาคผนวก ค จะเป็นโปรแกรมการทำงานเมื่อนำเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว และซ้ำสองครั้งมาใช้ในการวิเคราะห์อนุกรมเวลาตามลำดับ)

รูปที่ 3.5 แสดงผังงานสำหรับหาค่าความคลาดเคลื่อนจากการพยากรณ์ ด้วยวิธีพยากรณ์ทั้ง 3 วิธี