บทที่ 3

การดำเนินการวิจัย

การวิจัยครั้งนี้มีลักษณะเป็นการวิจัยเชิงทดลอง เพื่อต้องการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อความคลาดเคลื่อนเกิดอัตสหสัมพันธ์และมีค่าผิดปกติ โดยเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ 5 วิธี คือ วิธีกำลังสองน้อยที่สุด วิธีค่าสัมบูรณ์ต่ำสุด วิธีกำลังสองน้อยที่สุดแบบใช้การแปลงของเพรสและวินส์เทน วิธีการหาค่าพยากรณ์ร่วม และวิธีค่าสัมบูรณ์ต่ำสุดแบบใช้การแปลงของเพรสและวินส์เทน โดยจะเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนความคลาดเคลื่อนกำลังสองของค่าพยากรณ์ล่วงหน้า 12 คาบเวลา ภายใต้สถานการณ์ เมื่อขนาดตัวอย่าง 5 ขนาด คือ 20,30,40,50 และ 60 ค่าสหสัมพันธ์มี 5 ระดับ คือ 0.1,0.3,0.5 0.7 และ 0.9 รูปแบบของตัวแปรอิสระมี 2 รูปแบบ คือ รูปแบบปกติ และรูปแบบอัตสหสัมพันธ์อันดับที่หนึ่ง สัดส่วนการปลอมปน 3 ระดับ ซึ่งจะศึกษาความคลาดเคลื่อนที่มีการแจกแจงแบบปกติ และการแจกแจงแบบปลอมปน โดยการแจกแจงที่มาปลอมปนนี้เป็นการแจกแจงที่มีหางยาวกว่าการแจกแจงแบบปกติ

เทคนิคที่ใช้ในการจำลองข้อมูลครั้งนี้อาศัยเทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) สร้างสถานการณ์ต่างๆ ดังนั้นจะขอกล่าวถึงวิธีการจำลอง โดยใช้เทคนิคมอนติคาร์โลก่อน แล้วจึงแสดงถึงขั้นตอนต่างๆ ในการทำวิจัยและโปรแกรมที่ใช้ในการทำวิจัย

วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคที่ใช้สำหรับการแก้ไขปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีการ จำลองโดยใช้เทคนิคมอนติคาร์โล เป็นวิธีหนึ่งที่นิยมนำมาใช้ในการแก้ปัญหากันอย่างแพร่หลายใน ปัจจุบัน ซึ่งหลักการของการจำลองโดยเทคนิคดังกล่าว จะใช้เลขสุ่ม (Random Numbers) มาช่วยใน การหาคำตอบของปัญหาที่ต้องการศึกษา ขั้นตอนของวิธีการจำลองด้วยเทคนิคมอนติ-คาร์โล แบ่งออกเป็น 3 ขั้นตอน ดังนี้ คือ

ขั้นตอนที่ 1 การสร้างเลขสุ่ม การสร้างเลขสุ่มเป็นสิ่งที่สำคัญมากในเทคนิคนี้ ทั้งนี้เพราะว่าหลักการของการจำลองข้อมูลแบบมอนติคาร์โลนั้นจะใช้เลขสุ่มมาช่วยในการหาคำตอบของปัญหา โดยที่เลขสุ่มที่สร้างขึ้นนี้ จะมีการแจกแจงแบบสม่ำเสมอ (Uniform Distribution) ในช่วง (0,1) ตัวเลขสุ่มแต่ละตัวจะเป็นอิสระกันและมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ (มีวัฏจักรยาว)

ขั้นตอนที่ 2 นำเลขสุ่มที่สร้างขึ้นมาได้มาประยุกต์ใช้กับปัญหาที่ต้องการศึกษา บางปัญหาใช้เลขสุ่มโดยตรง ในบางปัญหาอาจจะต้องใช้ขั้นตอนอื่นอีกหลายขั้นตอน โดยที่มีการใช้เลขสุ่มในบางขั้นตอนเท่านั้น

ขั้นตอนที่ 3 การทดลองกระทำ เมื่อประยุกต์ปัญหาที่สนใจให้ใช้เลขสุ่มได้แล้ว ขั้นต่อไป คือ การทดลองโดยใช้กระบวนการสุ่ม (Random Process) มากระทำในลักษณะซ้ำๆ กัน (Replication) เพื่อหาคำตอบปัญหาที่ต้องการศึกษา

การวางแผนการทดลอง

ในการวิจัยครั้งนี้ กำหนดสถานการณ์ต่างๆ สำหรับศึกษาเปรียบเทียบของวิธีการประมาณค่าพารามิเตอร์เพื่อการพยากรณ์ โดยการใช้สมการถดถอยเชิงเส้นอย่างง่ายลักษณะความคลาดเคลื่อนเกิดอัตสหสัมพันธ์และมีค่าผิดปกติ

เริ่มต้นจากการสร้างค่าความคลาดเคลื่อนที่มีสหสัมพันธ์กันและมีค่าผิดปกติ และสร้างตัวแปรอิสระตามรูปแบบที่กำหนด โดยมีระดับสหสัมพันธ์และขนาดตัวอย่างต่างๆ กัน ซึ่งลักษณะของความคลาดเคลื่อนที่ต้องการศึกษามีการแจกแจงแบบปกติ และมีการแจกแบบปลอมปนซึ่งการแจกแจงที่มาที่มาปลอมปนนี้มีการแจกแจงหางยาวกว่าการแจกแจงแบบปกติ ต่อจากนั้นนำค่าความคลาดเคลื่อนและค่าตัวแปรอิสระที่ได้ไปสร้างตัวแปรตาม แล้วจึงทำการประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณทั้ง 5 วิธี นำค่าประมาณพารามิเตอร์จากแต่ละวิธีไปเข้าสมการพยากรณ์ ทำการพยากรณ์ล่วงหน้า 12 คาบเวลา หาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อกำลังสองจากการพยากรณ์ในแต่ละคาบเวลา แล้วทำการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อกำลังสองจากการพยากรณ์เฉลี่ย 12 คาบเวลา

วิธีการทดลอง

เขียนโปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรน (Fortran) โดยใช้กับเครื่อง AMDHAL 5850 เพื่อสร้างข้อมูลให้เป็นไปตามการทดลอง ซึ่งวิธีการทดลองแบ่งออกเป็น 4 ขั้นตอนดังต่อไปนี้

1. การสร้างโปรแกรมย่อย (Subroutines) สำหรับตัวแปรอิสระและความคลาดเคลื่อน
2. การสร้างความคลาดเคลื่อน (ε_t)
3. การสร้างข้อมูล (x,y) ที่มีความสัมพันธ์เชิงเส้นตรง
4. การประมาณค่าพารามิเตอร์จากแต่ละวิธีการ
5. การหาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจากการพยากรณ์ของแต่ละวิธี

ซึ่งรายละเอียดแต่ละขั้นตอนเป็นดังนี้

1. การสร้างโปรแกรมย่อย (Subroutines) สำหรับตัวแปรอิสระและความคลาดเคลื่อน

1.1 การแจกแจงแบบปกติ โดยการสร้างโปรแกรมย่อย NORMAL(RMEAN,VAR,EX) รายละเอียดแสดงไว้ในภาคผนวก ข.

1.2 การแจกแจงแบบลาปลาซ โดยการสร้างโปรแกรมย่อย LAP(EX) รายละเอียดแสดงไว้ในภาคผนวก ข.

2. การสร้างความคลาดเคลื่อน (ε_t)

จากรูปแบบของความคลาดเคลื่อนซึ่งมีรูปแบบอัตสหสัมพันธ์อันดับที่หนึ่ง เป็นดังนี้

$$\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t$$

ระดับสหสัมพันธ์ (ρ) มี 5 ระดับ คือ 0.1,0.3,0.5,0.7 และ 0.9 ขนาดตัวอย่าง (n) ที่ใช้มี 5 ระดับ คือ 20,30,40,50 และ 60 ซึ่ง ν_t เป็นความคลาดเคลื่อนสุ่ม ซึ่งการแจกแจงของความคลาดเคลื่อนเคลื่อน ν_t ที่ต้องการศึกษามีดังนี้

2.1 การแจกแจงแบบปกติ

สร้าง ν_t ให้มีการแจกแจงแบบปกติ ด้วยค่าเฉลี่ย 0 ความแปรปรวน 5

การสร้างความคลาดเคลื่อน ε_t ; t=1,2,...,n ในแต่ละคาบเวลาทำได้ดังรูปแบบนี้

$$\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t$$

$$v_0 \sim N(0, \frac{5}{1-\rho^2})$$

2.2 การแจกแจงแบบปกติปลอมปน (Comtaminated Normal) จากฟังก์ชันการแจกแจง

$$f(x) = (1-p)N(0,5) + pN(0,5c^2)$$

สร้าง v_t ให้มีการแจกแจงแบบปกติ ด้วยค่าเฉลี่ย 0 ความแปรปรวน 5 โดยการเรียกโปรแกรมย่อย NORMAL(RMEAN,VARN,EX) ด้วยความน่าจะเป็นเท่ากับ (1-p) และจะสร้าง v_t ให้มีการแจกแจงแบบปกติ มีค่าเฉลี่ย 0 ความแปรปรวน $5c^2$ ด้วยความน่าจะเป็นเท่ากับ p โดยการเรียกโปรแกรมย่อย NORMAL(RMEAN,VARC,EX)

การสร้างความคลาดเคลื่อน ε_t ; t=1,2,...,n ในแต่ละคาบเวลาทำได้ดังรูปแบบนี้

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$
$$v_0 \sim N(0, \frac{5}{1-\rho^2})$$

ในที่นี้กำหนดให้ c = 5,10 และ p = 0.05,0.08,0.10

2.3 ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$f(x) = (1-p)N(0,5) + pL(0,\beta)$$

สร้าง v_t ให้มีการแจกแจงแบบปกติ ด้วยค่าเฉลี่ย 0 ความแปรปรวน 5 โดยการเรียกโปรแกรมย่อย NORMAL(RMEAN,VARN,EX) ด้วยความน่าจะเป็นเท่ากับ (1-p) และจะสร้าง v_t ให้มีการแจกแจงแบบลาปลาซ มีค่าเฉลี่ย 0 ความแปรปรวน $2\beta^2$ ด้วยความน่าจะเป็นเท่ากับ p โดยการเรียกโปรแกรมย่อย LAP(EX)

การสร้างความคลาดเคลื่อน ε_t ; t=1,2,...,n ในแต่ละคาบเวลาทำได้ดังรูปแบบนี้

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$
$$v_0 \sim N(0, \frac{5}{1-\rho^2})$$

ในที่นี้กำหนดให้ β = 8 ,15 และ p = 0.05,0.08,0.10

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3. การสร้างข้อมูล (x,y) ที่มีความสัมพันธ์เชิงเส้นตรง

3.1 การสร้าง x ซึ่งเป็นตัวแปรที่กำหนดรูปแบบได้ ในการวิจัยครั้งนี้จะศึกษารูปแบบของตัวแปรอิสระ 2 รูปแบบ ดังนี้

$$x_t = u_t \qquad ; t = 1,2,.....,n$$

เมื่อ $u_t \sim N(10,5)$

$$x_t = 0.6x_{t-1} + u_t \qquad ; t = 1,2,.....,n$$

เมื่อ $u_t \sim N(0,5)$

3.2 การสร้างตัวแปรตาม y ตามรูปแบบของสมการถดถอยเชิงเส้นอย่างง่าย ซึ่งมีรูปแบบดังนี้

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

โดยการกำหนดให้ค่า $\beta_0 = 10, \beta_1 = 1$ ซึ่ง ε_t เป็นค่าความเคลื่อนสุ่มที่สร้างขึ้นมาตามรูปแบบข้างต้น

4. การประมาณค่าพารามิเตอร์จากแต่ละวิธี

เมื่อสร้างข้อมูล (x,y) ที่มีความสัมพันธ์เชิงเส้นตรงและรูปแบบความคลาดเคลื่อนตามที่กำหนดแล้ว ขั้นต่อไปคือการนำข้อมูล (x,y) ที่ได้ไปประมาณค่าพารามิเตอร์ $\underset{\sim}{\beta}$ และนำค่าพารามิเตอร์ที่ได้ไปพยากรณ์ค่าล่วงหน้าไป 12 คาบเวลา ซึ่งในการวิจัยครั้งนี้ได้เสนอวิธีการประมาณค่าพารามิเตอร์ไว้ 5 วิธี ดังต่อไปนี้

4.1 วิธีกำลังสองน้อยที่สุด ประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยที่สุด โดยให้ $\underset{\sim}{\hat{\beta}}$ เป็นค่าประมาณของ $\underset{\sim}{\beta}$ ซึ่งมีสูตรการคำนวณดังนวณดังนี้

$$\underset{\sim}{\hat{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\underset{\sim}{\mathbf{y}}$$

4.2 วิธีค่าสัมบูรณ์ต่ำสุด

ขั้นตอนที่ 1 จากข้อมูล (x,y) ที่สร้างขึ้นมา นำมาแปลงให้อยู่ในรูปของสมการเชิงเส้น

ขั้นตอนที่ 2 ประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีซิมเพล็กซ์ โดยให้ $\underset{\sim}{\tilde{\beta}}$ เป็นค่าประมาณของ $\underset{\sim}{\beta}$ จะได้ว่า

$$\tilde{\beta}_0 = \tilde{\beta}_0^+ - \tilde{\beta}_0^-$$
$$\tilde{\beta}_1 = \tilde{\beta}_1^+ - \tilde{\beta}_1^-$$

4.3 วิธีกำลังสองน้อยที่สุดแบบใช้การแปลงของเพรสและวินส์เทน

ขั้นตอนที่ 1 นำค่าความคลาดเคลื่อน e_t มาประมาณค่า ρ โดยให้ $\hat{\rho}$ เป็นค่าประมาณของ ρ ซึ่งมีสูตรการคำนวณดังนี้

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n-1} e_t^2}$$

โดยที่ $e_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t$

$\hat{\beta}_0, \hat{\beta}_1$ คือ ค่าประมาณของ β_0, β_1 ที่ได้จากวิธีกำลังสองน้อยที่สุด

ขั้นตอนที่ 2 นำค่า $\hat{\rho}$ ที่ได้แทนในสมการ

$$y_1\sqrt{1-\hat{\rho}^2} = \beta_0\sqrt{1-\hat{\rho}^2} + \beta_1 x_1\sqrt{1-\hat{\rho}^2}$$

$$y_t - \hat{\rho}y_{t-1} = \beta_0(1-\hat{\rho}) + \beta_1(x_t - \hat{\rho}x_{t-1})$$

ขั้นตอนที่ 3 นำค่าที่ได้จากขั้นตอนที่ 2 มาประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยที่สุด โดยให้ $\underset{\sim}{\hat{\beta}}^*$ เป็นค่าประมาณของ $\underset{\sim}{\beta}$ ซึ่งมีสูตรการประมาณค่าพารามิเตอร์ ดังนี้

$$\underset{\sim}{\hat{\beta}}^* = (\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'}\underset{\sim}{\mathbf{y}}^*$$

โดยที่ $\underset{\sim}{\mathbf{y}}^* = \begin{bmatrix} y_1\sqrt{1-\hat{\rho}^2} \\ y_2 - \hat{\rho}y_1 \\ \vdots \\ \vdots \\ y_n - \hat{\rho}y_{n-1} \end{bmatrix}$ $\mathbf{X}^* = \begin{bmatrix} \sqrt{1-\hat{\rho}^2} & x_1\sqrt{1-\hat{\rho}^2} \\ (1-\hat{\rho}) & x_2 - \hat{\rho}x_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ (1-\hat{\rho}) & x_n - \hat{\rho}x_{n-1} \end{bmatrix}$

4.4 วิธีการหาค่าพยากรณ์ร่วม

นำค่าพารามิเตอร์ที่คำนวณได้จากวิธีค่าสัมบูรณ์ต่ำสุดและวิธีกำลังสองน้อยที่สุดแบบใช้การแปลงของเพรสและวินส์เทน มาสร้างสมการพยากรณ์ y_{n+m} ณ คาบเวลา t =n+1,n+2,.....,m เป็นดังนี้

$$\hat{y}_{n+m} = \frac{1}{2}(\tilde{\beta}_0 + \tilde{\beta}_1 x_{n+m}) + \frac{1}{2}(\hat{\beta}_0^* + \hat{\beta}_1^* x_{n+m} + \hat{\rho}^m \hat{s}_n) \quad ; \text{m} = 1,2,.....,12$$

โดยที่ น้ำหนักที่ให้แต่ละวิธีเท่ากับ 0.5

$$\hat{s}_n = y_n - \hat{\beta}_0^* - \hat{\beta}_1^* x_n$$

4.5 วิธีค่าสัมบูรณ์ต่ำสุดแบบใช้การแปลงของเพรสและวินส์เทน

จากรูปแบบความสัมพันธ์ของความคลาดเคลื่อน

$$e_t = \tilde{\rho} e_{t-1} + v_t$$

ขั้นตอนที่ 1 นำค่าความคลาดเคลื่อน e_t มาประมาณค่า ρ ด้วยค่า $\tilde{\rho}$ โดยการแปลงค่าความคลาดเคลื่อน e_t ให้อยู่ในรูปสมการเชิงเส้น

โดยที่ $e_t = y_t - \tilde{\beta}_0 - \tilde{\beta}_1 x_t$

$\tilde{\beta}_0, \tilde{\beta}_1$ คือ ค่าประมาณของ β_0, β_1 ที่ได้จากวิธีค่าสัมบูรณ์ต่ำสุด

คำนวณหาค่า $\tilde{\rho}$ ด้วยวิธีซิมเพล็กซ์ จะได้ว่า

$$\tilde{\rho} = \tilde{\rho}^+ - \tilde{\rho}^-$$

ขั้นตอนที่ 2 นำ $\tilde{\rho}$ ที่ได้แทนในสมการ

$$y_1\sqrt{1-\tilde{\rho}^2} = \beta_0\sqrt{1-\tilde{\rho}^2} + \beta_1 x_1\sqrt{1-\tilde{\rho}^2}$$

$$y_t - \tilde{\rho} y_{t-1} = \beta_0(1-\tilde{\rho}) + \beta_1(x_t - \tilde{\rho} x_{t-1})$$

เขียนใหม่ได้ดังนี้

$$y_t^* = \beta_0 w_t^* + \beta_t x_t^* + v_t$$

เมื่อ t = 1

$$w_1^* = \sqrt{1-\tilde{\rho}^2}$$

$$y_1^* = y_1\sqrt{1-\tilde{\rho}^2}$$

$$x_1^* = x_1\sqrt{1-\tilde{\rho}^2}$$

เมื่อ t = 2,3,.....,n

$$w_t^* = 1-\tilde{\rho}$$

$$y_t^* = y_t - \tilde{\rho} y_{t-1}$$

$$x_t^* = x_t - \tilde{\rho} x_{t-1}$$

ขั้นตอนที่ 3 นำค่าที่ได้จากขั้นตอนที่ 2 นำมาแปลงให้อยู่ในรูปของสมการเชิงเส้น ประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีซิมเพล็กซ์ โดยให้ $\tilde{\beta}^*$ เป็นค่าประมาณของ $\underset{\sim}{\beta}$ จะได้ว่า

$$\tilde{\beta}_0^* = \tilde{\beta}_0^{*+} - \tilde{\beta}_0^{*-}$$

$$\tilde{\beta}_1^* = \tilde{\beta}_1^{*+} - \tilde{\beta}_1^{*-}$$

4. การหาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจากการพยากรณ์ มีขั้นตอนต่างๆ ดังนี้

4.1 เมื่อได้ค่าประมาณค่าพารามิเตอร์จากทั้ง 5 วิธีการแล้ว นำค่าพารามิเตอร์ที่ได้ไปสร้างสมการพยากรณ์ เพื่อหาค่าพยากรณ์ y_t ล่วงหน้า 12 คาบเวลา ซึ่งสมการพยากรณ์ของแต่ละวิธีเป็นดังนี้

4.1.1 วิธีกำลังสองน้อยที่สุด สมการพยากรณ์ คือ

$$\hat{y}_{n+m} = \hat{\beta}_0 + \hat{\beta}_1 x_{n+m} \qquad ; m = 1,2,.....,12$$

4.1.2 วิธีค่าสัมบูรณ์ต่ำสุด สมการพยากรณ์ คือ

$$\hat{y}_{n+m} = \tilde{\beta}_0 + \tilde{\beta}_1 x_{n+m} \qquad ; m = 1,2,.....,12$$

4.1.3 วิธีกำลังสองน้อยที่สุดแบบใช้การแปลงของเพรสและวินส์เทน

$$\hat{y}_{n+m} = \hat{\beta}_0^* + \hat{\beta}_1^* x_{n+m} + \hat{\rho}^m \hat{s}_n \qquad ; m = 1,2,.....,12$$

โดยที่ $\hat{s}_n = y_n - \hat{\beta}_0^* - \hat{\beta}_1^* x_n$

4.1.4 วิธีการหาค่าพยากรณ์ร่วม

$$\hat{y}_{n+m} = \frac{1}{2}(\tilde{\beta}_0 + \tilde{\beta}_1 x_{n+m}) + \frac{1}{2}(\hat{\beta}_0^* + \hat{\beta}_1^* x_{n+m} + \hat{\rho}^m \hat{s}_n) \; ; m = 1,2,.....,12$$

โดยที่ $\hat{s}_n = y_n - \hat{\beta}_0^* - \hat{\beta}_1^* x_n$

4.1.5 วิธีค่าวิธีค่าสัมบูรณ์ต่ำสุดแบบใช้การแปลงของเพรสและวินส์เทน

$$\hat{y}_{n+m} = \tilde{\beta}_0^* + \tilde{\beta}_1^* x_{n+m} + \tilde{\rho}^m \tilde{s}_n \qquad ; m = 1,2,.....,12$$

โดยที่ $\tilde{s}_n = y_n - \tilde{\beta}_0^* - \tilde{\beta}_1^* x_n$

4.2 สร้างข้อมูลจริง ตามวิธีการที่กล่าวข้างต้นไปล่วงหน้า 12 คาบเวลาเช่นกัน

4.3 หาค่าผลรวมกำลังสองความคลาดเคลื่อนจากการพยากรณ์ของแต่ละคาบเวลา โดยการนำค่าพยากรณ์ที่ได้ลบออกจากค่าจริงในคาบเวลาเดียวกัน แล้วยกกำลังสองสะสมไว้ในแต่ละครั้ง ทำการทดลองเช่นเดิมจนกระทั่งครบ 700 รอบสำหรับทุกคาบเวลา แล้วจึงคำนวณหาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนจากการพยากรณ์ (RMSFE) เฉลี่ย 12 คาบเวลา ซึ่งมีสูตรการคำนวณดังนี้

$$\text{RMSFE} = \frac{1}{12}\sum_{j=1}^{12}\sqrt{\sum_{i=1}^{700}\frac{(y_{ji} - \hat{y}_{ji})^2}{700}}$$

y_{ji} คือ ค่าของข้อมูลจริง ณ คาบเวลาที่ j รอบที่ i

$\hat{y}_{ji}$ คือ ค่าพยากรณ์ ณ คาบเวลาที่ j รอบที่ i

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

i คือ จำนวนรอบของการทำซ้ำ ; i = 1,2,.....,700

4.4 ทำการเปรียบเทียบค่า RMSFE เฉลี่ย 12 คาบเวลา ของวิธีการประมาณทั้ง 5 วิธี

4.5 ทำการทดลองเช่นนี้โดยการเปลี่ยนรูปแบบตัวแปรอิสระ ระดับสหสัมพันธ์ ขนาดความผิดปกติ สัดส่วนการปลอมปนและขนาดตัวอย่างจนกระทั่งครบรูปแบบที่ต้องการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 3.1 แสดงผังงานสำหรับหาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจากการพยากรณ์ ทั้ง 5 วิธี