บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อต้องการศึกษาค่าประมาณพารามิเตอร์ของตัวแบบความถดถอยเชิงลำดับชั้น ในกรณีที่การแจกแจงของความคลาดเคลื่อนสุ่มในระดับที่ 2 ไม่ได้มีการแจกแจงแบบปกติ โดยข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โล (Monte Carlo Simulation Technique) และทำการเขียนโปรแกรม MATLAB 2006b ในการประมวลผลและวิเคราะห์ข้อมูล โดยวิธีประมาณค่าพารามิเตอร์ในตัวแบบ ที่ศึกษามีดังต่อไปนี้

1. วิธี Iterative Generalized Least Square (IGLS)
2. วิธี Restricted Iterative Generalized Least Square (RIGLS)
3. วิธีประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งของฮูเบอร์ไวท์ (ROBUST)

ผู้วิจัยจะทำการศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ ซึ่งแบ่งออกได้เป็น 3 กลุ่มคือ กลุ่มพารามิเตอร์อิทธิพลคงที่ กลุ่มพารามิเตอร์ความแปรปรวนในระดับที่ 1 และกลุ่มพารามิเตอร์ส่วนประกอบความแปรปรวนในระดับที่ 2 โดยพิจารณาจากเกณฑ์พิจารณา 2 เกณฑ์คือ ค่าความเอนเอียงสัมพัทธ์ และค่าอัตราส่วนของค่าสัมบูรณ์ของค่าผลต่างระหว่างค่าเฉลี่ยของค่าประมาณค่าคลาดเคลื่อนมาตรฐานกับค่าคลาดเคลื่อนมาตรฐานแบบมอนติคาร์โลสัมพัทธ์ ภายใต้สถานการณ์ที่กำหนด และเนื่องจากในการวิจัยครั้งนี้ ผู้วิจัยได้อาศัยเทคนิคการจำลองมอนติคาร์โลมาสร้างข้อมูลในสถานการณ์ต่างๆ ดังนั้นในตอนแรกผู้วิจัยจะกล่าวถึงวิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล แล้วจึงแสดงรายละเอียดของขั้นตอนการวิจัยในลำดับถัดไป ส่วนรายละเอียดของโปรแกรมการวิจัยจะแสดงไว้ในภาคผนวก

3.1 วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคที่ใช้ในการแก้ปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โลเป็นวิธีหนึ่งที่นิยมนำมาแก้ปัญหากันอย่างแพร่หลายในปัจจุบัน ซึ่งหลักการของการจำลองโดยใช้เทคนิคดังกล่าวจะใช้เลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา ขั้นตอนของวิธีการจำลองด้วยเทคนิคมอนติคาร์โลแบ่งออกเป็น 3 ขั้นตอนใหญ่ ดังต่อไปนี้

3.1.1 <u>การสร้างเลขสุ่ม</u> การใช้ตัวเลขสุ่มเป็นสิ่งสำคัญมากในเทคนิคนี้ ทั้งนี้เป็นเพราะว่าหลักการจำลองด้วยเทคนิคมอนติคาร์โลนั้นจะใช้เลขสุ่มมาช่วยในการหาคำตอบของปัญหา โดยลักษณะของตัวเลขสุ่มนำมาใช้จะมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) สำหรับวิธีการสร้างเลขสุ่มมีผู้เสนอไว้หลายวิธี แต่วิธีที่ดีนั้นลักษณะของเลขสุ่มที่ถูกสร้างขึ้นมาจะต้องมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) ตัวเลขสุ่มแต่ละตัวเป็นอิสระต่อกันละมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ (มีวัฏจักรยาว)

3.1.2 <u>การนำตัวเลขสุ่มมาประยุกต์ใช้กับปัญหาที่ต้องการศึกษา</u> ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหา บางปัญหาอาจจะไม่ได้ใช้เลขสุ่มโดยตรงแต่ใช้ในการผลิตเลขสุ่มที่มีการแจกแจงแบบอื่นต่อไป

3.2.3 <u>การทดลองกระทำซ้ำ</u> เมื่อนำตัวเลขสุ่มมาประยุกต์ให้เข้ากับปัญหาที่ต้องการศึกษาได้แล้ว ขั้นต่อไปคือการทดลองโดยใช้กระบวนการของการสุ่ม (Random Process) มากระทำในลักษณะซ้ำๆ กัน หลายๆ ครั้ง เพื่อหาคำตอบของปัญหาที่ต้องการ

3.2 แผนการทดลอง

ผู้วิจัยกำหนดสถานการณ์ในการหาค่าเกณฑ์ที่ใช้ในการพิจารณาทั้งสองของวิธีประมาณพารามิเตอร์ดังที่ได้กล่าวมา โดยจะทำการเปรียบเทียบในสถานการณ์ต่างๆ ดังต่อไปนี้

3.2.1 เลือกตัวอย่างอย่างสุ่มของ $\underline{r}_j$ ในระดับที่ 1 จากประชากรที่มีการแจกแจงแบบปกติหลายตัวแปรที่มีค่าเฉลี่ย $\underline{0}$ และเมตริกซ์ความแปรปรวนร่วม $\sigma^2 \cdot I_{(n)}$ และเลือกตัวอย่างอย่างสุ่มของ $\underline{u}_j = (u_{0j}, u_{1j})^T$ ในระดับที่ 2 จากประชากรที่มีการแจกแจงแบบเอกโพเนนเชียลพาวเวอร์แบบไม่สมมาตรสองตัวแปรที่มีค่าเฉลี่ย $\underline{0}$ และเมตริกซ์ความแปรปรวนร่วม Σ พารามิเตอร์ความโด่ง α และพารามิเตอร์ความเบ้ κ

3.2.2 สนใจศึกษาเมื่อกำหนดพารามิเตอร์ความโด่งและพารามิเตอร์ความเบ้ โดยกำหนดค่าจากสัมประสิทธิ์ความโด่ง และสัมประสิทธิ์ความเบ้ เป็นดังนี้

กรณีที่ 1 กำหนดให้พารามิเตอร์ความเบ้คงที่โดยมีค่าเท่ากับ 1 ($\kappa = 1$) และพารามิเตอร์ความโด่งเป็น $\alpha = 1, 1.25, 1.6, 3, 8$ และ 32

กรณีที่ 2 กำหนดพารามิเตอร์ความโด่งคงที่โดยมีค่าเท่ากับ 2 ($\alpha = 2$) และพารามิเตอร์ความเบ้เป็น $\kappa = 0.3518, 0.6437$ และ 0.1

3.2.3 กำหนดขนาดตัวอย่างในระดับที่ 1 (n) มีขนาด 5, 10 และ15 และกำหนดขนาดตัวอย่างในระดับที่ 2 (J) มีขนาด 15, 30 และ50

3.2.4 ตัวแปรอิสระที่ใช้ในการศึกษาในครั้งนี้มี 1 ตัวแปร โดยมีการแจกแจงแบบปกติ ค่าเฉลี่ย 3 และความแปรปรวน 16

3.2.5 กำหนดให้ค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม (Intra-class Correlation: ICC) มีค่าเท่ากับ 0.1, 0.2 และ 0.3

3.3 ขั้นตอนการวิจัย

การวิจัยได้ดำเนินการตามขั้นตอนดังต่อไปนี้

3.3.1 กำหนดลักษณะการแจกแจงของความคลาดเคลื่อนสุ่มในทั้งสองระดับ พารามิเตอร์ความเบ้และความโด่งในความคลาดเคลื่อนสุ่มในระดับที่ 2 ขนาดตัวอย่างในทั้งสองระดับ และค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม

3.3.2 สร้างข้อมูลของตัวแปรอิสระ (X) ให้มีการแจกแจงตามที่กำหนด เพื่อใช้ในการสร้างตัวแปรตาม

3.3.3 สร้างข้อมูลตัวแปรตาม (y) ที่ใช้สำหรับหาค่าประมาณพารามิเตอร์ จากตัวแปรอิสระ และความคลาดเคลื่อนที่มีลักษณะตามการแจกแจงตามที่ต้องการศึกษา โดยตัวแปรตามมีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระ

3.3.4 การประมาณค่าพารามิเตอร์จะใช้วิธีประมาณ 2 วิธี ได้แก่

ก) วิธี Iterative Generalized Least Square (IGLS)

ข) วิธี Restricted Iterative Generalized Least Square (RIGLS)

และประมาณค่าประมาณความคลาดเคลื่อนมาตรฐานโดยใช้วิธีการประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งของฮูเบอร์ไวท์ (Huber/White Robust Standard Error Estimator)

3.3.5 การหาค่าเกณฑ์การพิจารณาทั้งสองเกณฑ์ ดังนี้

ก) เกณฑ์ความเอนเอียงสัมพัทธ์(RB) เพื่อศึกษาความถูกต้องในการประมาณค่าพารามิเตอร์

ข) เกณฑ์อัตราส่วนของค่าสัมบูรณ์ของค่าผลต่างระหว่างค่าเฉลี่ยของค่าประมาณค่าคลาดเคลื่อนมาตรฐานกับค่าคลาดเคลื่อนมาตรฐานแบบมอนติคาร์โลสัมพัทธ์ (RRADASE) เพื่อศึกษาความถูกต้องในการประมาณค่าคลาดเคลื่อนมาตรฐาน

3.3.6 สรุปผลในรูปของตาราง และกราฟ

3.4 การสร้างตัวเลขสุ่มและการจำลองตัวแปรสุ่ม

3.4.1 การสร้างตัวเลขสุ่มโดยใช้ตัวแบบจำลองสมภาคการคูณ

ในการวิจัยครั้งนี้ ผู้วิจัยได้ทำการสร้างเลขสุ่มโดยใช้ตัวแบบจำลองสมภาคการคูณ (Multiplicative Congruential Simulation) ซึ่งมีรูปแบบดังนี้

(3.1) $$X_i = (aX_i) mod\ m \qquad ;i = 1,2,\dots,n$$

โดย a และ m เป็นค่าคงที่ซึ่งเป็นจำนวนเต็มบวก คำว่า mod หรือ modulus หมายถึงการหารเอาเศษ และ X_i เป็นค่าจำนวนเต็มที่เกิดจากการหารเอาเศษด้วย m ทำให้ X_i มีค่าไม่เกิน m เมื่อคำนวณค่า X_i ได้ก็จำนำมาหาเลขสุ่มเทียม (Pseudo-Random Number: R_i) ซึ่ง R_i จะมีค่าอยู่ในช่วง (0,1) ดังนี้

(3.2) $$R_i = \frac{X_i}{m} m \qquad ;i = 1,2,\dots,$$

ในการสร้างเลขสุ่มนั้นต้องทำการกำหนดค่าให้กับ a และ m และ X_0 เป็นค่าเริ่มต้นเสียก่อน โดนเรียก X_0 ว่า ตัวเลขซีด (Seed Number) มีค่าไม่เกิน m ซึ่งค่า X_0 นี้เป็นเลขจำนวนเต็มบวกใดๆ ที่ต้องกำหนดเองให้มีค่าไม่เกิน m เมื่อทำการกำหนดค่าเริ่มต้นแล้ว ค่า X_i ตัวต่อไปก็จะเป็นไปตามตัวแบบในสมการ (3.1) และทุกครั้งที่เริ่มต้นด้วย X_0 ค่าเดิม (โดยที่ a และ m ไม่เปลี่ยนแปลง) จะได้ X_i เป็นเลขชุดเดิม

ในงานวิจัยครั้งนี้ผู้วิจัยได้กำหนดให้ a = 7^5= 16807 และ m = $2^{31} - 1$ ซึ่งเป็นตัวแบบจำลองสมภาคการคูณที่ใช้กันมากตัวแบบหนึ่งที่ได้ผ่านการตรวจสอบคุณบัติแล้วอย่างกว้างขวาง

3.5 การสร้างข้อมูลความคลาดเคลื่อน

การสร้างความคลาดเคลื่อนที่ใช้ในงานวิจัยแบ่งออกเป็น 2 ส่วนดังนี้

3.5.1 ความคลาดเคลื่อนในระดับที่ 1 ($\underline{r}_j$) สร้างให้มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย $\underline{0}$ และเมตริกซ์ความแปรปรวนร่วม $\sigma^2 \cdot I_{(n)}$ โดยมีสูตรการสร้างดังต่อไปนี้

$$NORMAL_1 = \mu + \sigma Z_1$$
$$NORMAL_2 = \mu + \sigma Z_2$$

3.5.2 ความคลาดเคลื่อนในระดับที่ 2 ($\underline{u}_j = (u_{0j}, u_{1j})^T$) สร้างให้มีการแจกแจงแบบเอกโพเนนเชียลพาวเวอร์แบบไม่สมมาตรสองตัวแปรที่มีค่าเฉลี่ย $\underline{0}$ และเมตริกซ์ความแปรปรวนร่วม Σ พารามิเตอร์ความโด่ง α และพารามิเตอร์ความเบ้ κ โดยมีสูตรสร้างต่อไปนี้

$$EXPO = \theta + \sigma IW^{1/\alpha}$$

3.6 การสร้างข้อมูลตัวแปรอิสระ (X)

ในขั้นตอนนี้ สร้างตัวแปรอิสระ(x_{ij}) ให้มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย 3 และมีความแปรปรวน 16

3.7 การสร้างข้อมูลของตัวแปรตาม (y)

การสร้างตัวแปรตาม (y_{ij}) จากตัวแบบความถดถอยเชิงลำดับชั้นซึ่งอยู่ในรูปแบบดังนี้

$$y_{ij} = \gamma_{00} + \gamma_{01} x_{ij} + u_{0j} + u_{1j} x_{ij} + r_{ij} \qquad ;i = 1,2,...,n \quad j = 1,2,...,J$$

3.8 การประมาณค่าพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐานด้วยวิธีประมาณแต่ละวิธี

เมื่อสร้างข้อมูลดังในข้างต้นได้แล้ว ขั้นตอนต่อไปคือการนำข้อมูลไปประมาณค่าพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐาน ซึ่งในการวิจัยนี้ได้ทำการประมาณค่าพารามิเตอร์ด้วยวิธีประมาณ 2 วิธี และประมาณค่าคลาดเคลื่อนมาตรฐานจากวิธีประมาณแต่ละวิธี และจากวิธีประมาณค่าคลาดเคลื่อนมาตรฐานที่มีความแกร่งของฮูเบอร์ไวท์ ซึ่งรายละเอียดของวิธีประมาณดังกล่าวได้กล่าวไว้แล้วในบทที่ 2

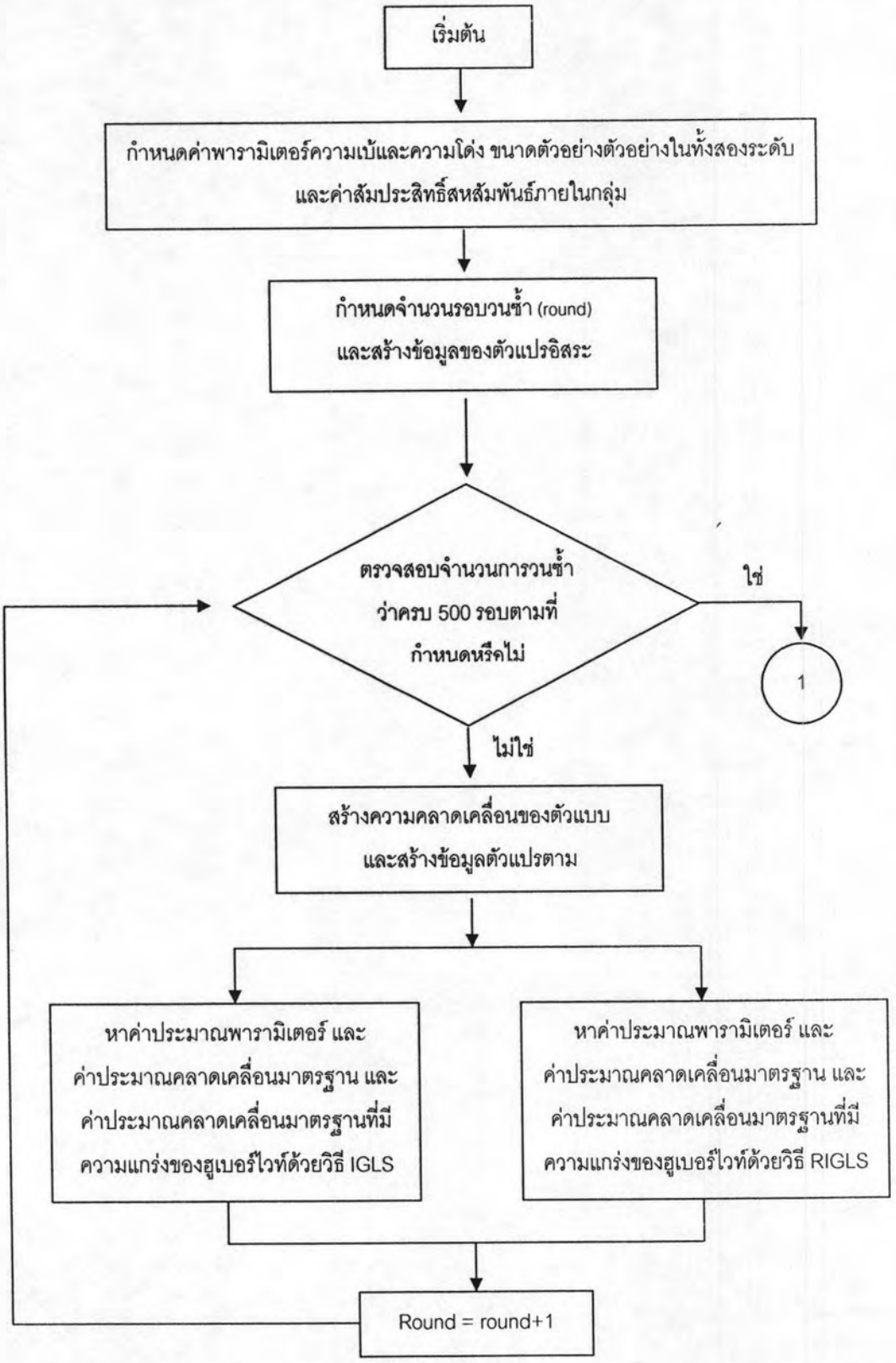
3.9 การหาค่าเกณฑ์การพิจารณา

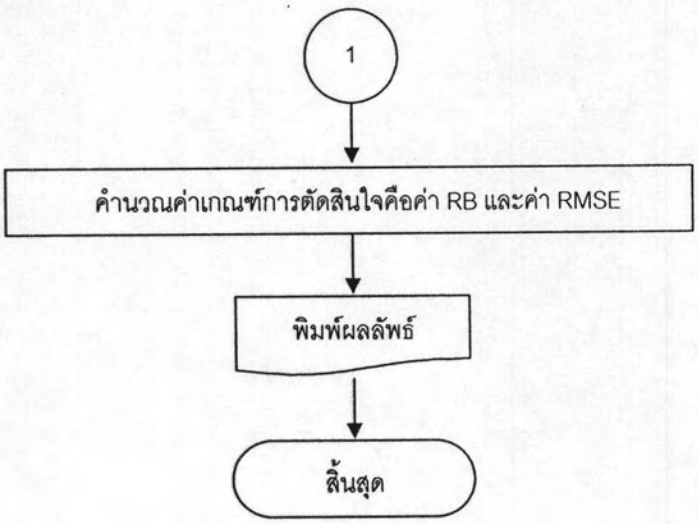
เมื่อได้ค่าประมาณของพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐานจากแต่ละวิธีแล้ว จะนำค่าประมาณดังกล่าวมาคำนวณค่าเกณฑ์การพิจารณา โดยค่าประมาณของพารามิเตอร์จะนำมาคำนวณค่าเกณฑ์ความเอนเอียงสัมพัทธ์ และค่าประมาณค่าคลาดเคลื่อนมาตรฐานจะนำมาคำนวณค่าเกณฑ์อัตราส่วนของค่าสัมบูรณ์ของค่าผลต่างระหว่างค่าเฉลี่ยของค่าประมาณค่าคลาดเคลื่อนมาตรฐานกับค่าคลาดเคลื่อนมาตรฐานแบบมอนติคาร์โลสัมพัทธ์ โดยในการ

ประมาณค่าพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐานดังกล่าวจะทำการประมาณซ้ำจนครบ 500 ครั้ง ซึ่งรายละเอียดของการคำนวณค่าเกณฑ์การพิจารณาทั้งสอง ได้กล่าวไว้อย่างละเอียดแล้วในบทที่ 1

จากนั้นจึงทำการเปรียบเทียบค่าเกณฑ์ที่ได้ของตัวประมาณในแต่ละวิธี เพื่อศึกษาถึงผลกระทบ และประสิทธิภาพในการประมาณค่าพารามิเตอร์ และค่าคลาดเคลื่อนมาตรฐาน เมื่อการแจกแจงความคลาดเคลื่อนสุ่มในระดับที่ 2 ไม่ได้มีการแจกแจงแบบปกติ โดยทำการเปลี่ยนระดับของความเบ้ และความโด่ง ขนาดตัวอย่าง และค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่ม จนกระทั่งครบทุกรูปแบบของสถานการณ์ที่ต้องการศึกษา และเพื่อง่ายแก่การเข้าใจจึงได้แสดงผังงานซึ่งแสดงขั้นตอนการวิจัยทั้งหมดในรูปที่ 3.1 ดังนี้

แผนภูมิที่ 3.1 แสดงผังงานสำหรับการหาค่าเกณฑ์การพิจารณาของวิธีประมาณพารามิเตอร์ในตัวแบบความถดถอยเชิงลำดับชั้น




1
คำนวณค่าเกณฑ์การตัดสินใจคือค่า RB และค่า RMSE
พิมพ์ผลลัพธ์
สิ้นสุด