หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการเปรียบเทียบค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของวิธีการประมาณค่าพารามิเตอร์ในการวิเคราะห์ความถดถอยพหุนามกรณีที่มีตัวแปรอิสระ 1 ตัวและกำลังสูงสุดของตัวแปรอิสระไม่เกิน 6 จากวิธีการประมาณค่าพารามิเตอร์ 3 วิธี ซึ่งมีดังนี้

ก) การประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยสุดสามัญ (Ordinary Least Squares Method (OLS))

ข) การประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยสุดปรับปรุง (Adjusted Least Squares Method (ALS))

ค) การประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (Weighted Least Squares Method (WLS))

3.1 การหาข้อสรุปโดยใช้การจำลองของข้อมูลด้วยวิธีมอนติคาร์โล [4] (Simulation by Monte Carlo Method)

วิธีมอนติคาร์โลเป็นเทคนิคในการจำลองตัวแบบทางคณิตศาสตร์ ซึ่งหลักการเบื้องต้นนั้นต้องจำลองตัวเลขสุ่ม (random number) เพื่อช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา โดยในการศึกษาวิจัยนี้จะใช้เทคนิคมอนติคาร์โลในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการศึกษา ขั้นตอนที่สำคัญของการจำลองข้อมูลด้วยวิธีมอนติคาร์โลมี 3 ขั้นตอนคือ

ขั้นตอนที่ 1 การสร้างตัวเลขสุ่ม

การสร้างตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล ทั้งนี้ก็เพราะว่าหลักการของมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา ลักษณะของตัวเลขสุ่มที่ดีจะมีการแจกแจงสม่ำเสมอ (Uniform Distribution) ในช่วง $[0,1]$ และเป็นอิสระซึ่งกันและกัน

---

[4] จะเด็จ สวรรค์ตรานนท์,"การเปรียบเทียบวิธีที่ใช้สำหรับการคัดเลือกสมการถดถอยที่ดีที่สุด,"(วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย,2531),หน้า36.

ขั้นตอนที่ 2 การประยุกต์ของปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม

ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจใช้ตัวเลขสุ่มโดยตรง แต่บางปัญหาอาจใช้ตัวเลขสุ่มเพียงบางขั้นตอนของปัญหาเท่านั้น

ขั้นตอนที่ 3 การทดลองกระทำ

เมื่อประยุกต์ปัญหาเพื่อใช้กับตัวเลขสุ่มได้แล้วขั้นตอนต่อไปก็คือการทดลองโดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะที่ซ้ำ ๆ กัน (replicaton) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.2 แผนการทดลอง

ผู้วิจัยกำหนดสถานการณ์ในการหาค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของตัวแบบทั้ง 3 วิธีดังนี้

1) เลือกตัวอย่างอย่างสุ่มของ ε_i^* จากประชากรที่มีการแจกแจงเดียวกัน การแจกแจงที่สนใจศึกษาคือ การแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และส่วนเบี่ยงเบนมาตรฐาน $0.1, 0.3, 0.5, 0.7$ และ 1.0

2) กำหนดขนาดตัวอย่างที่ศึกษาคือ $15, 30, 50, 100$ และ 200

3) ตัวแปรอิสระเริ่มต้นที่ใช้ในการศึกษาครั้งนี้เท่ากับ 1 ซึ่งสร้างจากตัวแปรแฝง (x_i^*) ที่มีการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 1 และความคลาดเคลื่อนในตัวแปรอิสระ (δ_i) ที่มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย 0 ความแปรปรวน $0.1, 0.3, 0.5$ และ 0.7

4) กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (highest degree of independent variables for dependent variable building in model)(MB)) คือ $2, 3, 4, 5$ และ 6

3.3 ขั้นตอนในการศึกษาวิจัย

ขั้นตอนในการศึกษาวิจัยมีดังนี้

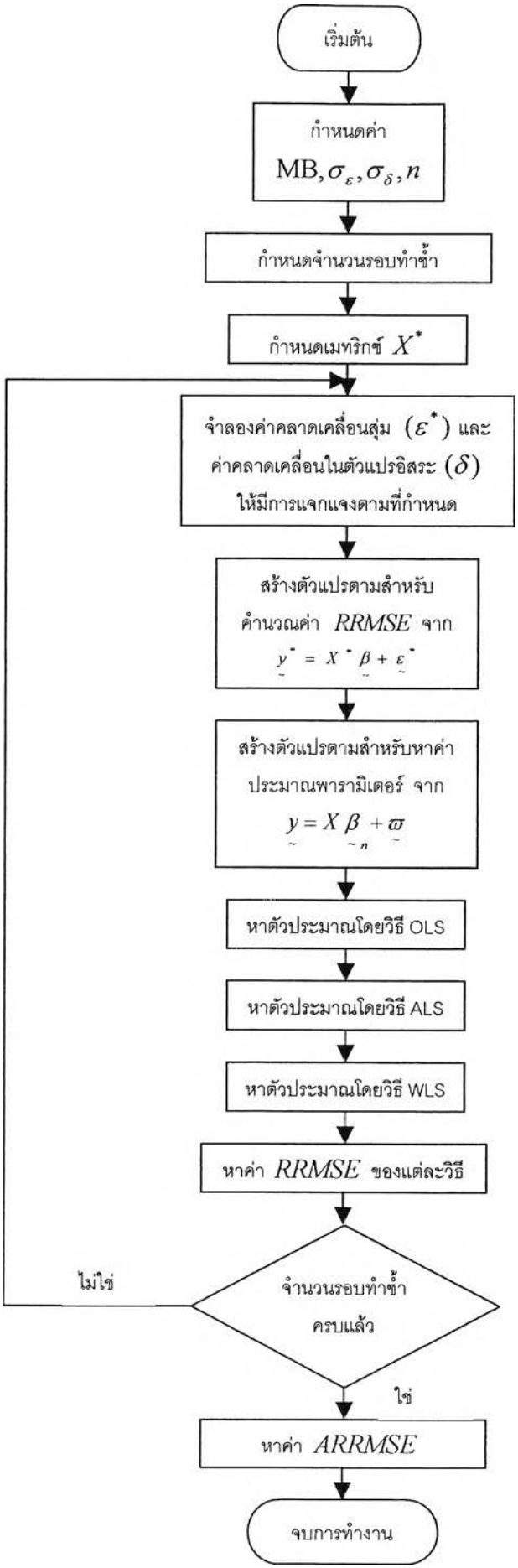
1. กำหนดลักษณะการแจกแจงของความคลาดเคลื่อนสุ่มในตัวแปรตาม (ε_i^*) และความคลาดเคลื่อนในตัวแปรอิสระ (δ_i) ขนาดตัวอย่าง (n) กำลังสูงสุดของตัวแปรอิสระที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB)

2. การสร้างข้อมูลตัวแปรตาม (y^*) ที่ใช้สำหรับคำนวณค่ารากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ จากตัวแปรแฝง(พจน์พหุนามของ x^*) และความคลาดเคลื่อนสุ่มในตัวแปรตามที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา (ε^*) โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรแฝง และสร้างข้อมูลตัวแปรแฝงที่ใช้สำหรับสร้างตัวแบบเริ่มต้น

3. การสร้างข้อมูลตัวแปรตาม (y) ที่ใช้สำหรับหาตัวประมาณค่าพารามิเตอร์ จากตัวแปรอิสระ(พจน์พหุนามของ x) และความคลาดเคลื่อนที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา (ϖ) โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระดังที่แสดงในหัวข้อ 2.1 และสร้างข้อมูลตัวแปรอิสระที่ใช้สำหรับสร้างตัวแบบเริ่มต้นจากตัวแปรแฝง (x^*) และความคลาดเคลื่อนในตัวแปรอิสระ (δ)

4. การประมาณค่าพารามิเตอร์โดยวิธีทั้ง 3 วิธี
   ก) วิธีกำลังสองน้อยสุดสามัญ
   ข) วิธีกำลังสองน้อยสุดปรับปรุง
   ค) วิธีกำลังสองน้อยสุดถ่วงน้ำหนัก
5. การหาค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์
6. สรุปผลในรูปของตาราง

ผังงานแสดงขั้นตอนในงานวิจัยแสดงดังรูปที่ 3.1



รูปที่ 3.1 แสดงผังงานของขั้นตอนในงานวิจัย

3.4 รายละเอียดของแต่ละขั้นตอนเป็นดังนี้

ขั้นตอนที่ 1

กำหนดลักษณะการแจกแจงของค่าคลาดเคลื่อนสุ่มในตัวแปรตาม (ε^*) และค่าคลาดเคลื่อนในตัวแปรอิสระ (δ) ขนาดตัวอย่าง (n) จำนวนตัวแปรอิสระสูงสุดที่ใช้สำหรับการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB) ซึ่งกำหนดตามแผนการทดลองที่ได้กล่าวมาแล้วข้างต้น

ขั้นตอนที่ 2

การสร้างข้อมูลที่มีลักษณะการแจกแจงที่ต้องการศึกษา จะใช้โปรแกรมภาษาฟอร์แทรน 4.0 (Fortran 4.0) บนเครื่อง pc computer ซึ่งการสร้างข้อมูลที่มีลักษณะการแจกแจงแบบต่าง ๆ นั้นจะต้องใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้างฟังก์ชันที่ใช้ในการสร้างตัวเลขสุ่ม คือ RD(IX) เมื่อสร้างตัวเลขสุ่มแล้วจะนำตัวเลขสุ่มที่ได้มาสร้างข้อมูลที่มีการแจกแจงแบบปกติมีค่าเฉลี่ย 0 และความแปรปรวน σ^2 โดยสร้างจากโปรแกรมย่อย NORMAL(DMEAN,SIGMA,X)[5] ข้อมูลที่สร้างได้แก่

1) ตัวแปรแฝงเริ่มต้น (x^*) สร้างจากการแจกแจงแบบปกติที่มีค่าเฉลี่ย 10 ความแปรปรวน 1
2) จำนวนตัวแปรแฝงสูงสุดที่ใช้ในการสร้างตัวแปรตามในตัวแบบถดถอยพหุนาม (MB) เป็นไปตามที่กำหนดไว้ในขั้นตอนที่ 1
3) สร้างตัวแปรตาม (y^*) ที่ใช้สำหรับคำนวณค่า $RRMSE$ โดยให้มีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรแฝง (X^*) และมีลักษณะการแจกแจงค่าคลาดเคลื่อนสุ่มที่เกิดจากตัวแปรตาม (ε^*) ตามที่กำหนดในตอนต้น ซึ่งมีรูปแบบดังนี้

$$y^* = X^* \beta + \varepsilon^*$$

เมื่อ y^* แทนตัวแปรตาม (เมื่อตัวแปรอิสระปราศจากความคลาดเคลื่อน)

X^* แทนเมทริกซ์ของตัวแฝง

β แทนพารามิเตอร์ที่กำหนด

และ ε^* แทนค่าคลาดเคลื่อนที่กำหนด ซึ่ง $\varepsilon_i^* \sim N(0, \sigma_\varepsilon^2)$

---

[5] ปราณี รัตนัง."การประมาณสัมประสิทธิ์การถดถอยพหุเมื่อความผิดพลาดมีการแจกแจงแบบเบ้และมีการแจกแจงแบบหางยาวกว่าการแจกแจงปกติ."(วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ จุฬาลงกรณ์มหาวิทยาลัย,2531) หน้า117

ขั้นตอนที่ 3

การสร้างข้อมูลสำหรับหาค่าประมาณพารามิเตอร์ ข้อมูลที่สร้างได้แก่

1) สร้างข้อมูลตัวแปรอิสระ (x_i) โดยสร้างจากตัวแปรแฝง (x_i^*) ที่ได้จากขั้นตอนที่ 2 และความคลาดเคลื่อนในตัวแปรอิสระ (δ_i) โดยมีลักษณะการแจกแจงตามที่กำหนดในตอนต้น ซึ่งมีรูปแบบดังนี้ $x_i = x_i^* + \delta_i$

2) สร้างข้อมูลตัวแปรตาม (y) ที่ใช้สำหรับหาค่าประมาณพารามิเตอร์ โดยให้มีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระ (x) และมีลักษณะการแจกแจงค่าคลาดเคลื่อนสุ่ม (ϖ_i) ตามที่ได้กล่าวไว้ในหัวข้อ 2.1 ซึ่งมีรูปแบบดังนี้

$$y = X\beta + \varpi$$

เมื่อ y แทนตัวแปรตาม (เมื่อเกิดความคลาดเคลื่อนในตัวแปรอิสระ)

X แทนเมทริกซ์ของตัวอิสระ

β แทนพารามิเตอร์ที่ได้จากการจัดรูปแบบสมการที่ได้กล่าวไว้ในหัวข้อ 2.1

และ ϖ แทนค่าคลาดเคลื่อนที่สุ่มที่ได้จากการจัดรูปแบบสมการที่ได้กล่าวไว้ในหัวข้อ 2.1

ขั้นตอนที่ 4

หลังจากที่มีข้อมูลพร้อมแล้วทำการหาตัวประมาณค่าพารามิเตอร์ที่เหมาะสมจากทั้ง 3 วิธี ด้วยโปรแกรมย่อย ดังต่อไปนี้

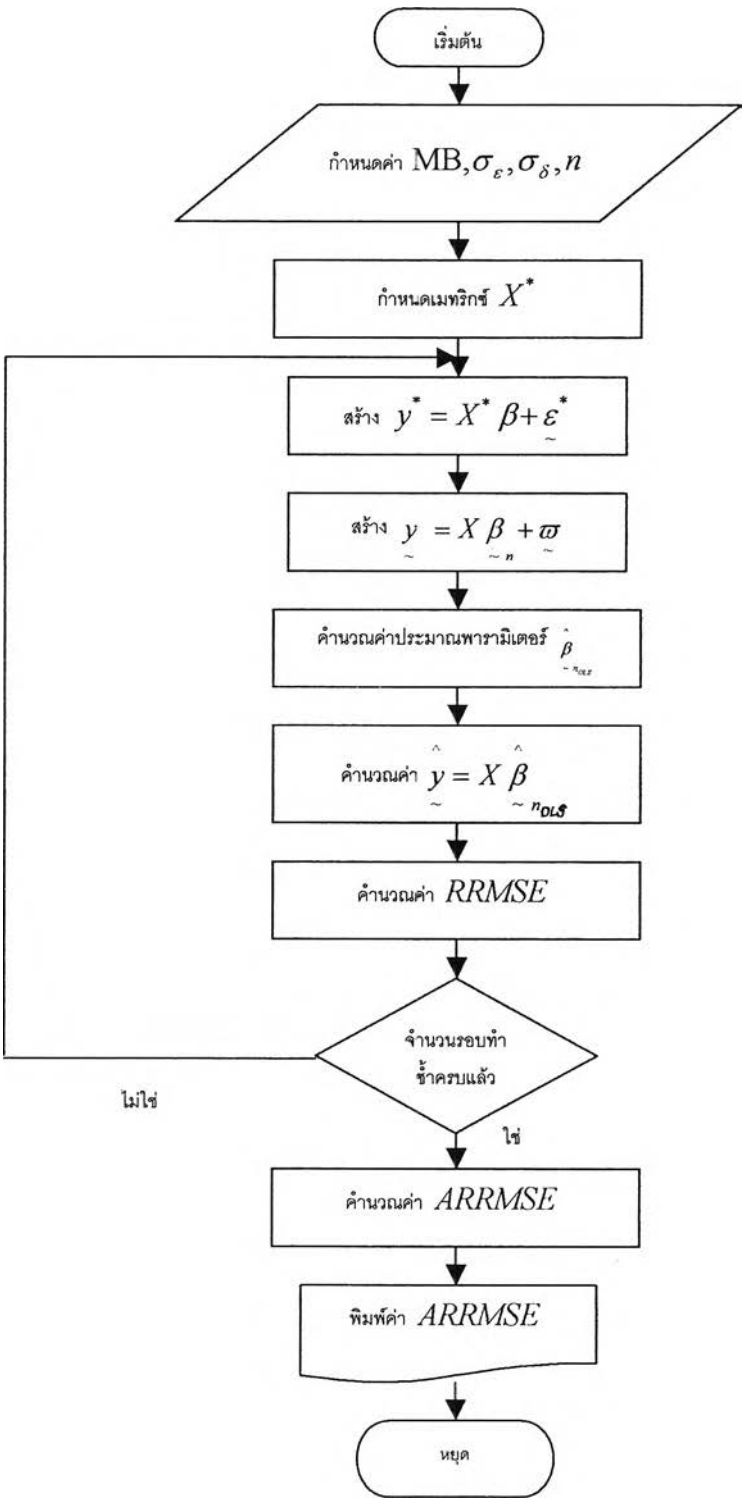
ก) โปรแกรมย่อยของวิธี OLS คือ OLS(N,COL,XE,Y,YS,DEL,RMSEO,TRMSEO)

ข) โปรแกรมย่อยของวิธี ALS คือ ALS(N,COL,XE,Y,YS,DEL,RMSEA,TRMSEA)

ค) โปรแกรมย่อยของวิธี WLS คือ WLS(N,COL,XE,Y,YS,DEL,RMSEW,TRMSEW)

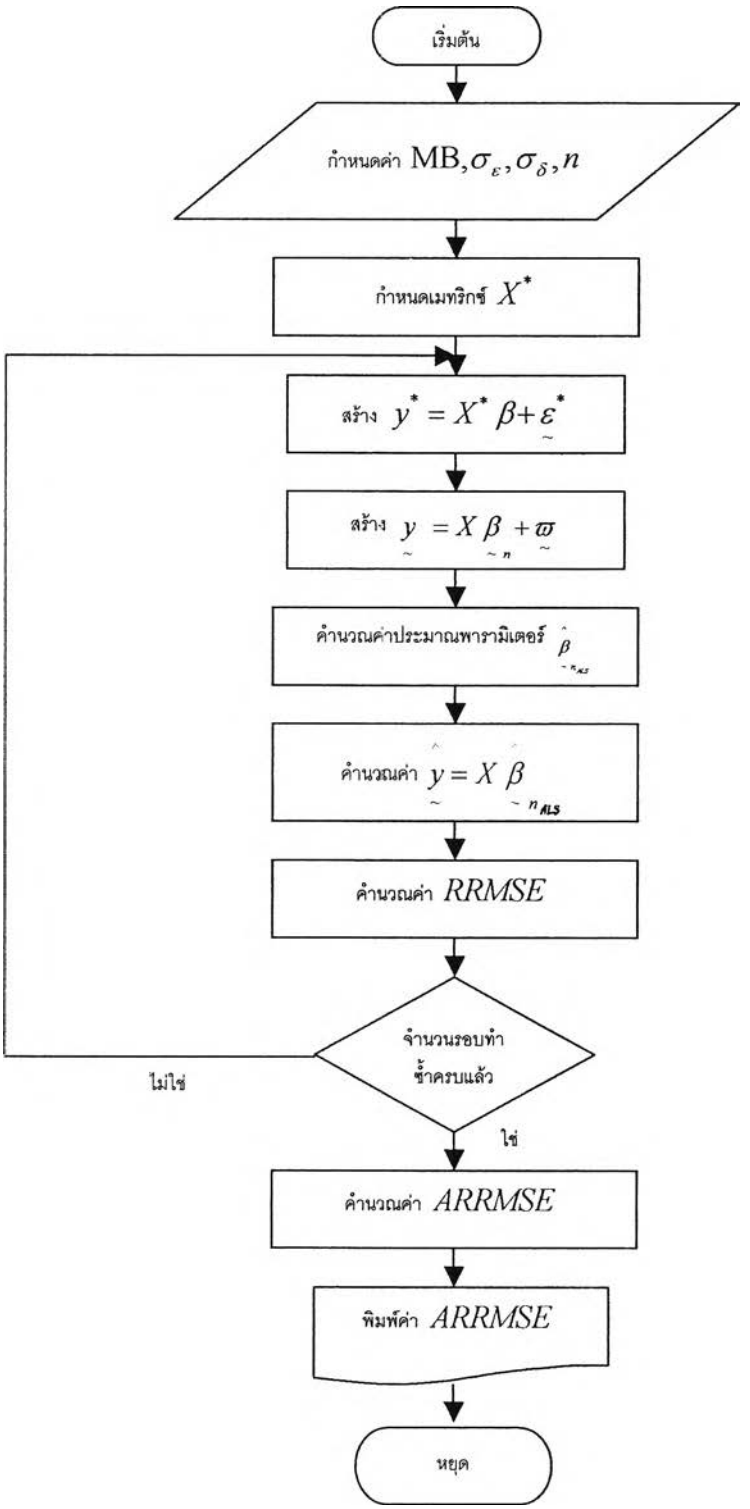
วิธีการประมาณค่าพารามิเตอร์ที่เหมาะสมแต่ละวิธีมีรายละเอียดดังนี้

ก) การประมาณค่าพารามิเตอร์โดยวิธี OLS เป็นการสร้างตัวแบบโดยไม่มีการตัดหรือเพิ่มตัวแปรอิสระเข้าสู่สมการ ซึ่งตัวประมาณที่ได้จากวิธีนี้คำนวณจาก $\hat{\beta}_{OLS} = (X'X)^{-1}(X'y)$ โดยจะคำนวณสัมประสิทธิ์การถดถอยจากโปรแกรมย่อย BETAH(N,COL,XX,Y,B) และผังงานแสดงขั้นตอนการประมาณค่าพารามิเตอร์ โดยวิธี OLS แสดงดังรูปที่ 3.2



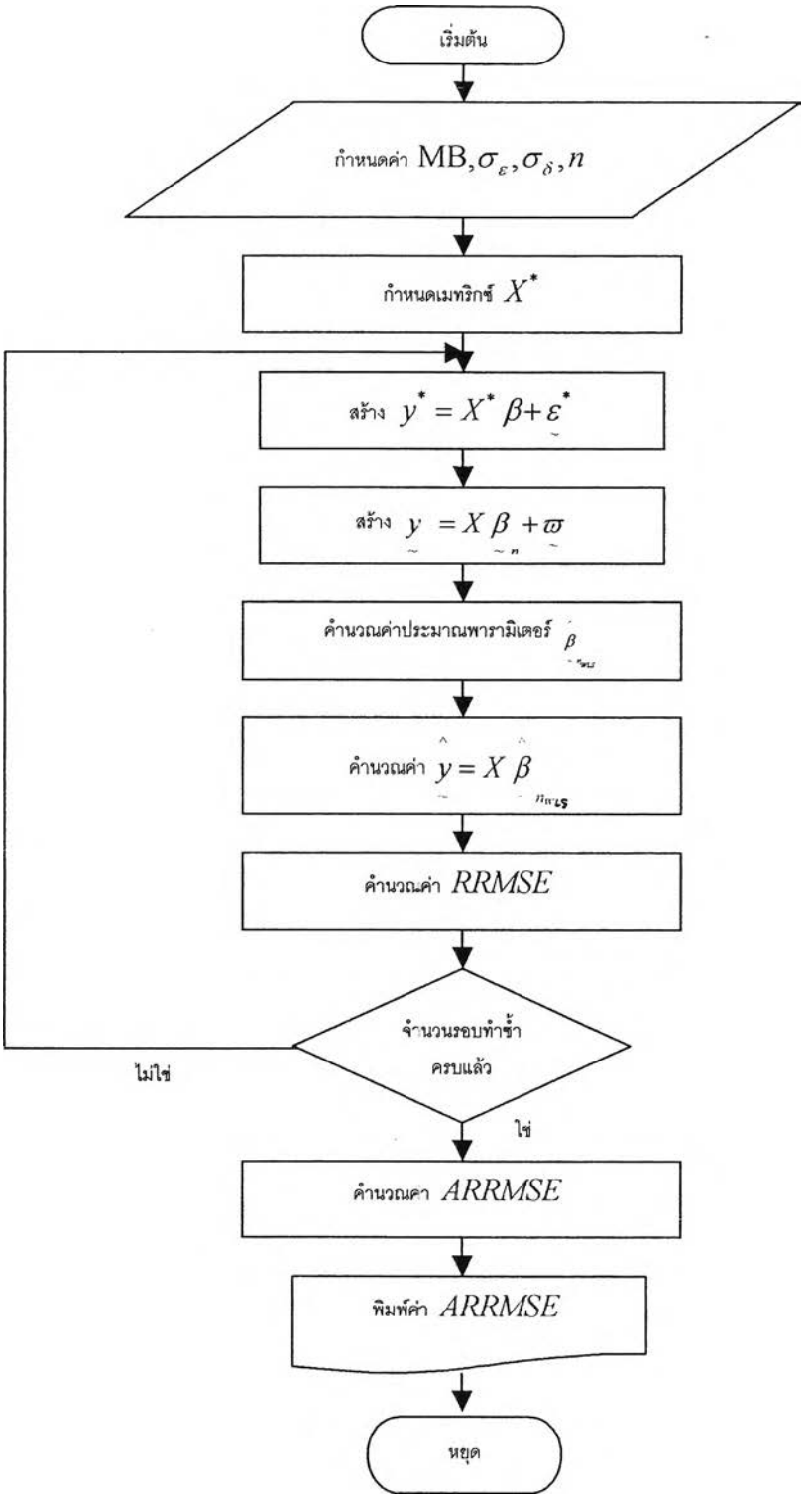
รูปที่ 3.2 แสดงผังงานขั้นตอนการประมาณค่าพารามิเตอร์โดยวิธี OLS

ข) การประมาณค่าพารามิเตอร์โดยวิธี ALS เป็นการสร้างตัวแบบโดยไม่มีการตัดหรือเพิ่มตัวแปรอิสระเข้าสู่สมการและเป็นวิธีที่พัฒนามาจากวิธีกำลังสองน้อยสุดสามัญ(OLS) ดังนั้นหลักการหาตัวประมาณพารามิเตอร์จึงเหมือนกันกับวิธี OLS แต่จะต่างกันที่ค่าของตัวประมาณพารามิเตอร์ ซึ่งตัวประมาณที่ได้จากวิธีนี้คำนวณจาก $\underset{\sim}{\hat{\beta}}_{ALS} = H^{-1}\underset{\sim}{h}$ และผังงานแสดงขั้นตอนการประมาณค่าพารามิเตอร์ โดยวิธี ALS แสดงดังรูปที่ 3.3



รูปที่ 3.3 แสดงผังงานขั้นตอนการประมาณค่าพารามิเตอร์โดยวิธี ALS

ค) การประมาณค่าพารามิเตอร์โดยวิธี WLS เป็นการสร้างตัวแบบโดยไม่มีการตัดหรือเพิ่มตัวแปรอิสระเข้าสู่สมการและเป็นวิธีที่พัฒนามาจากวิธีกำลังสองน้อยสุดสามัญ(OLS) เช่นเดียวกับวิธี ALS ดังนั้นหลักการหาตัวประมาณพารามิเตอร์จึงเหมือนกันกับวิธี OLS และ ALS แต่จะต่างกันที่ค่าของตัวประมาณพารามิเตอร์ ซึ่งวิธี WLS จะมีการนำค่า $Var(y_i \mid x_i)$ มาใช้ในการถ่วงน้ำหนักในขั้นตอนของการหาค่าประมาณพารามิเตอร์ ดังนั้นตัวประมาณที่ได้จากวิธีนี้คำนวณได้จาก $\underset{\sim}{\hat{\beta}}_{WLS} = (X'w^{-1}X)^{-1}(X'w^{-1}\underset{\sim}{y})$ และผังงานแสดงขั้นตอนการประมาณค่าพารามิเตอร์ โดยวิธี WLS แสดงดังรูปที่ 3.4



รูปที่ 3.4 แสดงผังงานขั้นตอนการประมาณค่าพารามิเตอร์โดยวิธี WLS

ขั้นตอนที่ 5

เมื่อได้ตัวแบบที่เหมาะสมในแต่ละวิธีแล้วคำนวณค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์จากโปรแกรมย่อยของแต่ละวิธีดังนี้

ก) โปรแกรมย่อยของวิธี OLS คือ RRMSEO(N,COL,YS,YH,RMSEO)

ข) โปรแกรมย่อยของวิธี ALS คือ RRMSEA(N,COL,YS,YH,RMSEA)

ค) โปรแกรมย่อยของวิธี WLS คือ RRMSEW(N,COL,YS,YH,RMSEW)

ขั้นตอนที่ 6

เมื่อได้ค่าเฉลี่ยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยสัมพัทธ์ของแต่ละวิธีเมื่อทำซ้ำครบตามจำนวนรอบของการทำซ้ำตามที่กำหนดแล้ว จะนำผลมาสรุปลงในตารางเพื่อแสดงการเปรียบเทียบหรือศึกษาแนวโน้มของแต่ละวิธี