หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้สิ่งที่ต้องการคือ การเปรียบเทียบวิธีที่ใช้เลือกตัวแปรอิสระเข้าสู่สมการถดถอยซึ่งประกอบด้วยวิธี การกำจัดตัวแปรแบบถอยหลัง การเลือกตัวแปรแบบไปข้างหน้า การถดถอยแบบขั้นบันได การถดถอยแบบขั้นตอน การกำจัดตัวแปรโดยใช้สัมประสิทธิ์สหสัมพันธ์ ซึ่งวิธีทั้ง 5 นี้จะมีการเลือกตัวแปรอิสระแตกต่างกันไป และผลที่ได้นั้นอาจจะเหมือนกันหรือแตกต่างกันออกไปถ้าได้กล่าวถึงรายละเอียดแล้วในบทที่ 2 ในการศึกษาเปรียบเทียบครั้งนี้จะใช้ตัวสถิติคือ ผลรวมของความคลาดเคลื่อนกำลังสอง (SSE) ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (MSE) และตัวสถิติของธีล (U) เป็นตัววัดวิธีไหนจะให้ผลดีที่สุดในการคัดเลือกเอาตัวแปรอิสระเข้าสู่สมการถดถอย โดยจะศึกษาเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ โลจิสติก ดับเบิ้ลเอ็กซ์โปเนนเชียล และปกติปลอมปน ทั้งนี้ก็เพราะว่า การแจกแจงแบบปกตินี้เป็นการแจกแจงที่มีลักษณะการกระจายแบบปกติเป็นพื้นฐานในการดูการเปรียบเทียบในแต่ละวิธีว่ามีแนวโน้มออกมาในลักษณะอย่างไรและเมื่อเปลี่ยนแปลงไปสู่การแจกแจงที่มีลักษณะการกระจายสูงปกติ หรือการกระจายต่ำกว่าปกติ หรือมีลักษณะการกระจายไปทางหางมาก หรือที่เรียกว่ามีหางยาวกว่าปกตินั่นเอง ว่าจะมีผลหรือมีแนวโน้มออกมาเหมือนหรือใกล้เคียงกับลักษณะการกระจายปกติหรือไม่ สำหรับรูปแบบของการแจกแจงปกติปลอมปน จะศึกษาเมื่อเปอร์เซ็นต์การปลอมปนเป็น 5% 10% และ 25% สเกลแฟคเตอร์ที่ใช้มี 2 ระดับคือ 3 และ 10 ซึ่งถ้าสเกลแฟคเตอร์มีค่าสูงจะทำให้เกิดค่าผิดปกติ และจากการทดลองกระทำในการวิจัยครั้งนี้พบว่า ถ้าสเกลแฟคเตอร์มีค่าน้อยกว่า 3 จะมีโอกาสของการเกิดค่าผิดปกตินอยมาก ส่วนสเกลแฟคเตอร์มีค่ามากกว่า 10 จะมีโอกาสของการเกิดค่าผิดปกติสูงมาก ฉะนั้นจึงทำการศึกษาเพียง 2 ระดับดังกล่าว ขนาดตัวอย่างที่สนใจศึกษาในการวิจัยครั้งนี้คือ 15 30 50 และ 100 ส่วนจำนวนตัวแปรอิสระที่สนใจศึกษาคือ 3 5 7 และ 9 และระดับนัยสำคัญมี 2 ระดับคือ 0.01 และ 0.05 สำหรับการวิจัยในครั้งนี้จะใช้วิธีมอนติคาร์โลในการสร้างข้อมูลให้มีสภาพการแจกแจงตามที่ต้องการ

เนื่องจากวิธีมอนติคาร์โลเป็นเทคนิคที่ใช้ในการศึกษาวิจัยครั้งนี้ ดังนั้นในตอนแรกของบทนี้จะกล่าวถึงวิธีการของมอนติคาร์โลก่อน แล้วจึงแสดงรายละเอียดของแผนการทดลอง ขั้นตอนของการวิจัย และโปรแกรมที่ใช้ในการวิจัยตามลำดับ ซึ่งรายละเอียดต่าง ๆ นั้นมีดังนี้

I1030180

3.1 วิธีมอนติคาร์โล

เทคนิคในการจำลองตัวแบบทางคณิตศาสตร์นั้นมีอยู่หลายวิธีด้วยกัน วิธีของมอนติคาโลก็เป็นวิธีหนึ่งที่นิยมใช้กันแพร่หลายมากในปัจจุบัน ซึ่งหลักการของวิธีมอนติคาร์โลนั้นต้องจำลองตัวเลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ในการศึกษาวิจัยครั้งนี้จะใช้เทคนิคของมอนติคาร์โลดังกล่าวในการสร้างข้อมูลที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา ซึ่งขั้นตอนของวิธีมอนติคาร์โลที่ใช้กันอยู่ในปัจจุบันแบ่งได้เป็น 3 ขั้นตอนดังนี้

3.1.1 การสร้างตัวเลขสุ่ม การใช้ตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีของมอนติคาร์โล ทั้งนี้ก็เพราะว่าหลักการของมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา ซึ่งลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบยูนิฟอร์มในช่วง (0, 1) สำหรับวิธีการสร้างตัวเลขสุ่มนั้นมีผู้เสนอไว้หลายวิธีด้วยกัน แต่วิธีหนึ่งที่ดีก็คือวิธีที่ไวท์และสมิดท์ (White and Sehmidt 1975 : 421) เสนอไว้คือ ลักษณะของตัวเลขสุ่มที่เกิดขึ้นมีการแจกแจงแบบยูนิฟอร์มในช่วง (0, 1) และเป็นอิสระกัน ซึ่งรายละเอียดได้แสดงไว้ในภาคผนวก ก

3.1.2 การประยุกต์ของปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษาบางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจมีขั้นตอนอื่นอีกหลาย ๆ ขั้นตอนซึ่งขั้นตอนเหล่านี้มีบางขั้นตอนที่ต้องใช้ตัวเลขสุ่ม

3.1.3 การทดลองกระทำ เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นตอนต่อไปก็คือการทดลองโดยใช้กระบวนการสุ่ม (Random process) มากระทำในลักษณะที่ซ้ำ ๆ กัน (Replication) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.2 แผนการทดลอง

ในการศึกษาวิจัยครั้งนี้ ได้กำหนดสถานการณ์ต่าง ๆ ที่ต้องการศึกษาโดยสุ่มตัวอย่างจากประชากรที่มีการแจกแจงแบบเดียวกัน ซึ่งลักษณะการแจกแจงที่สนใจศึกษาคือ ปกติ โลจิสติก ดับเบิลเอ็กซ์โปเนนเชียล และปกติปลอมปน

สำหรับการแจกแจงแบบปกติปลอมปนกำหนดเปอร์เซ็นต์การปลอมปนเป็น 5% 10% และ 25% ส่วนสเกลแฟคเตอร์มี 2 ระดับคือ 3 และ 10 ซึ่งแสดงรายละเอียดดังตาราง 3.1

ตารางที่ 3.1 แสดงค่าเปอร์เซ็นต์การปลอมปนและสเกลแฟคเตอร์ทั้งหมดที่ใช้ในการศึกษาวิจัย.

(5, 3)	(5, 10)
(10, 3)	(10, 10)
(25, 3)	(25, 10)

นอกจากนี้ทุก ๆ การแจกแจงขนาดตัวอย่างที่สนใจศึกษาคือ 15 30 50 และ 100 ส่วนจำนวนตัวแปรอิสระที่สนใจศึกษาคือ 3 5 7 และ 9 ซึ่งแสดงรายละเอียดดังตาราง 3.2

ตารางที่ 3.2 แสดงขนาดตัวอย่างและจำนวนตัวแปรอิสระทั้งหมดที่ใช้ในการศึกษาวิจัย

(15, 3)	(15, 5)	(15, 7)	(15, 9)
(30, 3)	(30, 5)	(30, 7)	(30, 9)
(50, 3)	(50, 5)	(50, 7)	(50, 9)
(100, 3)	(100, 5)	(100, 7)	(100, 9)

3.3 ขั้นตอนในการศึกษาวิจัย

ขั้นตอนในการศึกษาวิจัยมีดังนี้คือ

1. การสร้างโปรแกรมย่อย (Subroutine) สำหรับสร้างการแจกแจงของความคลาดเคลื่อนตามที่ต้องการศึกษา

2. การสร้างข้อมูลคือตัวแปรตาม ($\underset{\sim}{Y}$) จากตัวแปรอิสระ (X) ซึ่งเป็นค่าคงที่ และความคลาดเคลื่อนที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระ

3. การแปลงข้อมูลให้มีหน่วยเป็นมาตรฐานเดียวกัน

4. ทำการคัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอยโดยวิธี

4.1 การกำจัดตัวแปรแบบถอยหลัง

4.2 การเลือกตัวแปรแบบไปข้างหน้า

4.3 การถดถอยแบบขั้นบันได

4.4 การถดถอยแบบขั้นตอน

4.5 การกำจัดตัวแปรโดยใช้สัมประสิทธิ์สหสัมพันธ์
เพื่อให้ได้สมการถดถอยออกมาในขั้นสุดท้าย

5. การประมาณค่าสัมประสิทธิ์ถดถอยเชิงเส้นโดยวิธีกำลังสองน้อยที่สุด

6. การหาค่าของผลรวมกำลังสองของความคลาดเคลื่อน

7. การหาค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน

8. หาค่าตัวสถิติของธีล

9. การหาค่าเฉลี่ยให้ได้ผลสัมพันธ์

สำหรับรายละเอียดของแต่ละขั้นตอนเป็นดังนี้

3.3.1 การสร้างโปรแกรมย่อยสำหรับการสร้างการแจกแจงของความคลาดเคลื่อนตามลักษณะการแจกแจงที่ต้องการศึกษา

การสร้างลักษณะการแจกแจงของความคลาดเคลื่อนทุกรูปแบบที่ต้องการศึกษานั้น ใช้โปรแกรมภาษาฟอร์แทรนโฟ (Fortran IV) โดยใช้กับเครื่อง IBM 370/3031 ซึ่งการสร้างลักษณะการแจกแจงแบบต่าง ๆ นั้นจะต้องใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้าง สำหรับรายละเอียดในการสร้างการแจกแจงแบบต่าง ๆ เป็นดังนี้

3.3.1.1 การแจกแจงแบบปกติ โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบปกติจะใช้วิธีของเกาส์ (Gauss) ซึ่งเป็นวิธีที่สร้างการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น μ และความแปรปรวนเป็น σ^2 โดย SMEAN และ $(SIGMA)^2$ คือค่าเฉลี่ย และความแปรปรวนที่กำหนด สำหรับรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ ใช้คำสั่ง Call Normal (SMEAN, SIGMA, X) โดยค่าเฉลี่ยและความแปรปรวนจะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์คือค่า X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น SMEAN และความแปรปรวนเป็น $(SIGMA)^2$

3.3.1.2 การแจกแจงแบบโลจิสติค โปรแกรมย่อยที่ใช้ในการสร้างลักษณะการแจกแจงแบบโลจิสติค ใช้วิธีการแปลงผกผัน (Inverse Transformation) รายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL LOGIST (ALPHA, BATA, X) โดยค่า ALPHA และ BETA เป็นค่าที่กำหนดขึ้นเพื่อให้ได้ค่าเฉลี่ย และความแปรปรวนตามที่

ต้องการ ค่า ALPHA และ BETA นี้จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบโลจิสติคที่มีค่าเฉลี่ยเป็น ALPHA และความแปรปรวนเป็น $\frac{1}{3}\ \pi^2\ (BETA)^2$

3.3.1.3 การแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียล โปรแกรมย่อยที่ใช้ในการสร้างลักษณะการแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลใช้วิธีการแปลงผกผัน ซึ่งรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL (ALPHA, BETA, X) โดยค่า ALPHA และ BETA เป็นค่าที่กำหนดขึ้นเพื่อให้ได้ค่าเฉลี่ยและความแปรปรวนตามที่ต้องการ ค่า ALPHA และ BETA นี้จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบดับเบิลเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเป็น ALPHA และความแปรปรวนเป็น $2(BETA)^2$

3.3.1.4 การแจกแจงแบบปกติปลอมปน โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบปกติปลอมปนใช้วิธีการแปลงข้อมูลมาจากการแจกแจงแบบปกติ โดยพิจารณาการแจกแจงแบบปกติปลอมปนของตัวแปร X ซึ่งมาจาก $F = (1-P)N(\mu,\sigma^2)+PN(\mu,c^2\sigma^2)$ เมื่อ P เป็นเปอร์เซ็นต์การปลอมปน C เป็นสเกลแฟคเตอร์ที่จะทำให้เกิดค่าผิดปกติ ($C>0$) ซึ่งหมายความว่าตัวแปรสุ่ม X จะมาจากการแจกแจงแบบ $N(\mu,\sigma^2)$ ด้วยความน่าจะเป็น (1-P) และมาจากการแจกแจงแบบ $N(\mu, c^2\sigma^2)$ ด้วยความน่าจะเป็น P

สำหรับรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL SCALE (C, P, SMEAN,SIGMA, X) C และ P เป็นค่าที่กำหนดสเกลแฟคเตอร์และเปอร์เซ็นต์การปลอมปน ส่วน SMEAN และ $(SIGMA)^2$ เป็นค่าเฉลี่ยและความแปรปรวนที่กำหนด ซึ่งค่า C, P, SMEAN และ SIGMA จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มาจากลักษณะการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น SMEAN และความแปรปรวนเป็น $(SIGMA)^2$ จำนวน (1-P) % ส่วนที่เหลือจะมาจากลักษณะการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น SMEAN และความแปรปรวนเป็น $(CSIGMA)^2$ จำนวน P %

3.3.2 การสร้างข้อมูลให้มีความสัมพันธ์เชิงเส้นตรง

ในการศึกษาวิจัยครั้งนี้จะทำการสร้างตัวแปรอิสระ X ซึ่งเป็นค่าคงที่ก่อนแล้วจึงสร้างตัวแปรตาม $\underset{\sim}{Y}$ ให้มีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระ X ให้มีลักษณะการแจกแจงความคลาดเคลื่อนที่กำหนดตามรูปแบบของความสัมพันธ์ดังนี้คือ $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ เมื่อ $\underset{\sim}{\beta}$ เป็นค่า

พารามิเตอร์ที่ถูกกำหนดขึ้นมา และ $\underset{\sim}{\varepsilon}$ เป็นความคลาดเคลื่อนที่มีรูปแบบการแจกแจงเป็นแบบต่าง ๆ ดังที่ได้กล่าวมาแล้วในตอนต้นของบท สำหรับการสร้าง $\underset{\sim}{Y}$ นั้นจะเริ่มจากการกำหนดขนาดตัวอย่าง และจำนวนตัวแปรอิสระที่ต้องการศึกษา เมื่อกำหนดพารามิเตอร์ $\underset{\sim}{\beta}$ ค่าเฉลี่ย ค่าความแปรปรวน และลักษณะการแจกแจงของความคลาดเคลื่อน รวมทั้งค่าคงที่ของตัวแปรอิสระ X แล้วจึงสร้างตัวแปรตาม $\underset{\sim}{Y}$ ตามรูปแบบดังกล่าว

3.3.3 การแปลงข้อมูลให้มีหน่วยเป็นมาตรฐานเดียวกัน

เมื่อสร้างข้อมูลที่มีความสัมพันธ์เชิงเส้นตามรูปแบบที่ต้องการได้แล้วในขั้นนี้ จะทำการแปลงข้อมูลคือ ตัวแปรตามและตัวแปรอิสระให้เป็นหน่วยเดียวกันหรือให้เป็นมาตรฐานเดียวกัน เพราะบางครั้งข้อมูลที่ได้จากการสังเกตมานั้นบางครั้งอาจจะเก็บคนละสถานที่หรือบางครั้งการเก็บข้อมูลอาจจะเก็บคนละเวลา จึงทำให้ข้อมูลที่ได้มานั้นมีคุณสมบัติแตกต่างกันก็เป็นได้ เช่น อาจจะเป็นข้อมูลเชิงปริมาณหรืออาจจะเป็นเชิงคุณภาพก็เป็นได้ ซึ่งเมื่อนำมาใช้การพยากรณ์แล้วอาจจะทำให้เกิดความคลาดเคลื่อนไปจากความจริงมาก ดังนั้นจึงมีการแปลงข้อมูลเสียก่อนโดยวิธีการใช้ลอการิทึมมาช่วยในการแปลงข้อมูลซึ่งมีวิธีการคือ นำค่าของตัวแปรตาม Y และค่าตัวแปรอิสระทุกตัวมาแปลงให้อยู่ในรูป ALOG(Y(I)+1) และ ALOG (X(I,J)+1)

3.3.4 การคัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอย

โปรแกรมย่อยที่เรียกใช้ในการคัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอยคือ CALL BACK (N, K, Kl, X, Y, U, SSE, SS) CALL FORN (N, K, Kl, X, Y, U, SSE, SS) CALL STEP (N, K, Kl, X, Y, U, SSE, SS) CALL STAGE(N, K, Kl, X, Y, U, SSE, SS) และ CALL CODEL (N, K, Kl, X, Y, U, SSE, SS) เป็นโปรแกรมย่อยของวิธี การกำจัดตัวแปรแบบถอยหลัง การเลือกตัวแปรแบบไปข้างหน้า การถดถอยแบบขั้นบันได การถดถอยแบบขั้นตอน และการกำจัดตัวแปรโดยใช้สัมประสิทธิ์สหสัมพันธ์ตามลำดับ ซึ่ง N คือ จำนวนข้อมูล k คือ จำนวนตัวแปรอิสระ kl คือ จำนวนตัวแปรอิสระบวก 1 X คือเมตริกซ์ของตัวแปรอิสระ Y คือเวกเตอร์ของตัวแปรตาม ส่วนผลลัพธ์ที่ได้จากโปรแกรมย่อยคือ U SSE และ SS ซึ่งเป็นค่าของตัวสถิติค่าผลรวมกำลังสองของความคลาดเคลื่อนและค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง ซึ่งได้มาจากโปรแกรมย่อยที่เรียกใช้ ดังจะกล่าวในหัวข้อต่อไป

3.3.5 การประมาณค่าสัมประสิทธิ์ถดถอยเชิงเส้นโดยวิธีกำลังสองน้อยที่สุด

โปรแกรมย่อยที่ใช้ในการประมาณค่าสัมประสิทธิ์ถดถอยเชิงเส้นจากข้อมูลจะใช้โปรแกรมย่อยที่ใช้คำสั่ง CALL BETA (N, K1, X, Y, B) โดย X คือเมตริกซ์ของตัวแปรอิสระ Y คือเวกเตอร์ของตัวแปรตาม N คือจำนวนตัวอย่าง k1 คือจำนวนตัวแปรอิสระบวก 1 B คือผลลัพธ์ซึ่งเป็นเวกเตอร์ของสัมประสิทธิ์ถดถอยเชิงเส้นที่มีขนาด (k1 x 1)

3.3.6 การหาค่าผลรวมกำลังสองของความคลาดเคลื่อน

โปรแกรมย่อยที่ใช้ในการหาค่าผลรวมกำลังสองของความคลาดเคลื่อนจากข้อมูลนั้น จะใช้โปรแกรมย่อยที่ใช้คำสั่ง CALL SMSQE (N, K, K1, X, y, B, SSE, SMSE) ซึ่ง N คือจำนวนตัวอย่าง K คือจำนวนตัวแปรอิสระ K1 คือจำนวนตัวแปรอิสระบวก 1 X คือเมตริกซ์ของตัวแปรอิสระ Y คือจำนวนตัวแปรตาม B คือ สัมประสิทธิ์ถดถอย SSE และ SMSE คือ ผลลัพธ์ซึ่งเป็นค่าผลรวมกำลังสองของความคลาดเคลื่อนและค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน

3.3.7 การหาค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน

โปรแกรมย่อยที่ใช้ในการหาค่าเฉลี่ยกำลังสองของความคลาดเคลื่อนจากข้อมูลนั้นรายละเอียดก็เหมือน 3.3.6

3.3.8 การหาค่าตัวสถิติของธีล

โปรแกรมย่อยที่ใช้ในการหาค่าของตัวสถิติU จากข้อมูลนั้นจะใช้โปรแกรมย่อยที่ใช้คำสั่ง CALL USTAT (N, Y, YP, U) โดย N คือจำนวนตัวอย่าง Y คือเวกเตอร์ของตัวแปรตาม YP คือ เวกเตอร์ของตัวแปรตามที่ประมาณค่าขึ้นมา โดยได้มาจากโปรแกรมย่อยที่ใช้คำสั่ง CALL PRE (N, K1, B, X, Y, YP, RES) ซึ่ง N คือขนาดตัวอย่าง K1 คือจำนวนตัวแปรอิสระบวก 1 B คือ สัมประสิทธิ์ถดถอย X คือเมตริกซ์ของตัวแปรอิสระ Y คือตัวแปรตาม YP เป็นผลลัพธ์ซึ่งเป็นเวกเตอร์ค่าประมาณของตัวแปรตาม และ RES เป็นผลลัพท์ ซึ่งเป็นค่าแตกต่างระหว่าง ค่าจริง Y และค่าประมาณ Y ส่วน U คือผลลัพท์ซึ่งเป็นตัวสถิติของธีล

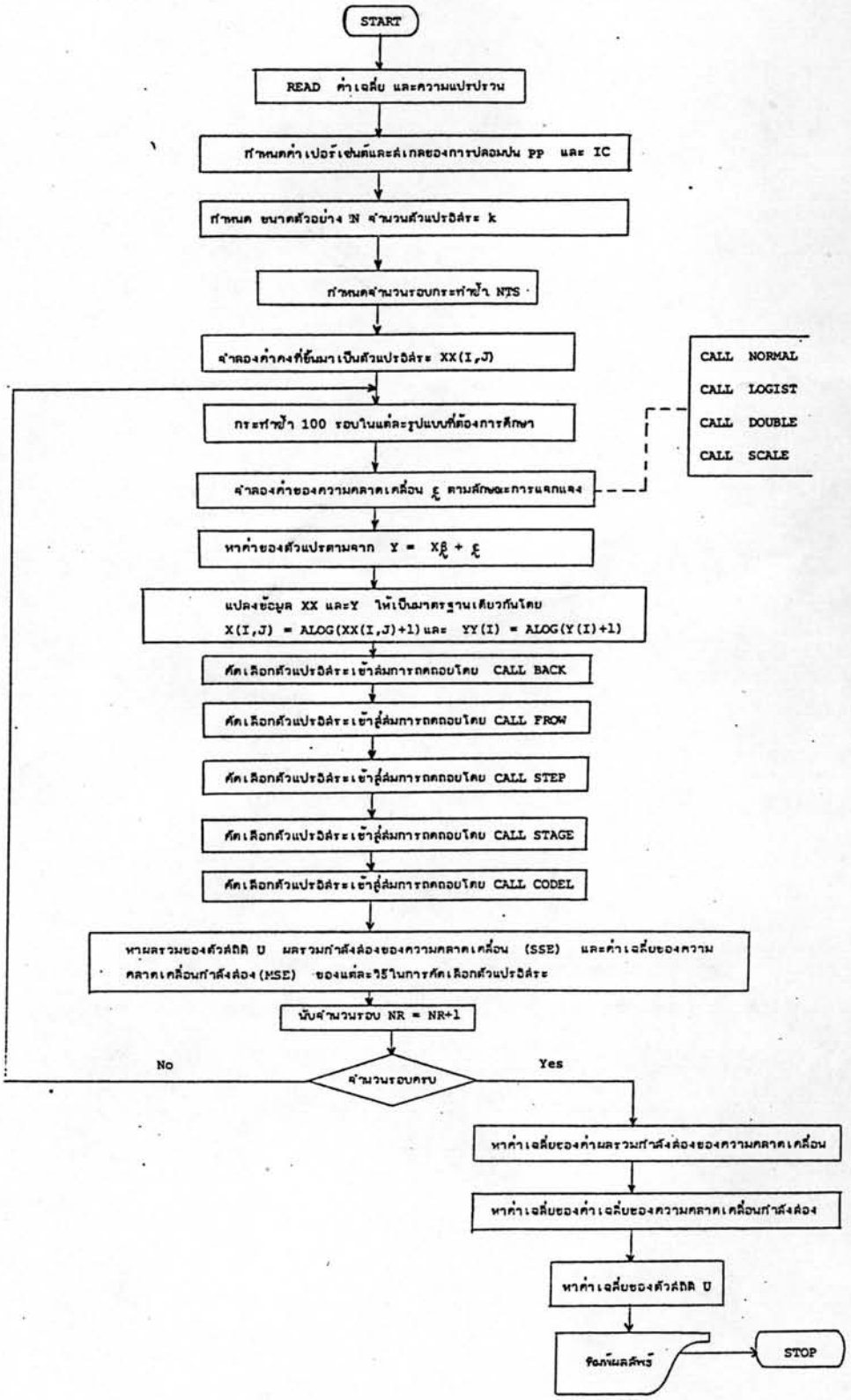
3.3.9 การหาค่าเฉลี่ยให้ได้ผลลัพธ์

ในการศึกษาแต่ละรูปแบบจะกระทำซ้ำกัน 100 ครั้ง ดังนั้นค่าของผลรวมกำลังสองของความคลาดเคลื่อน ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน และค่าของตัวสถิติ U นั้นจะมีทั้งหมด 100 ค่า ฉะนั้นจึงจำเป็นหาค่าเฉลี่ยเพื่อเอามาเป็นผลสรุปขั้นสุดท้าย ซึ่งสามารถทำได้

คือ นำเอาค่าที่ได้ทั้งหมดนำมารวมกันแล้วหารด้วย 100 เมื่อได้ขนาดตัวอย่างและจำนวนตัวแปรอิสระครบทุกรูปแบบที่ต้องการศึกษาแล้ว ขั้นต่อไปก็จะเปลี่ยนลักษณะการแจกแจงของความคลาดเคลื่อนจนครบทุกการแจกแจง โดยแต่ละการแจกแจงของความคลาดเคลื่อนจะใช้ขนาดตัวอย่างและจำนวนตัวแปรอิสระครบทุกรูปแบบที่ต้องการศึกษา

3.4 ผังงานของขั้นตอนในการวิจัย

รูปที่ 3.1 แสดงผังงานของขั้นตอนในการวิจัย



3.5 โปรแกรมที่ใช้ในการวิจัย

โปรแกรมที่ใช้ในการวิจัยทั้งหมดเขียนด้วยภาษาฟอร์แทรนโพ ซึ่งใช้กับเครื่อง IBM 370/3031 โดยลักษณะการทำงานของโปรแกรมแสดงในตาราง 3.3

ตาราง 3.3 แสดงลักษณะการทำงานของโปรแกรมทั้งหมดที่ใช้ในการวิจัย

ลำดับที่ของ ของโปรแกรม	ลักษณะการทำงาน	โปรแกรมย่อยที่เรียกใช้
1	สร้างตัวแปรตาม ตัวแปรอิสระ และความคลาดเคลื่อนเมื่อกำหนดขนาดของข้อมูลมีค่าเป็น 15 30 50 และ 100 จำนวนตัวแปรอิสระมีค่า 3, 5, 7 และ 9 โดยมีลักษณะการแจกแจงของความคลาดเคลื่อนคือ ปกติ โลจิสติค ดับเบิลเอ็กซ์โปเนนเชียล และปกติปลอมปน	โปรแกรมสร้างตัวเลขสุ่ม โปรแกรมสร้างลักษณะการแจกแจงของความคลาดเคลื่อน โปรแกรมแปลงข้อมูลให้อยู่ในมาตรฐานเดียวกัน
2	คัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอยโดยวิธีทั้ง 5	โปรแกรมการกำจัดตัวแปรแบบถอยหลัง โปรแกรมการเลือกตัวแปรแบบไปข้างหน้า โปรแกรมการถดถอยแบบขั้นบันได โปรแกรมการถดถอยแบบขั้นตอน โปรแกรมการกำจัดตัวแปรโดยใช้สัมประสิทธิ์สหสัมพันธ์ โปรแกรมการหาสัมประสิทธิ์สหสัมพันธ์บางส่วน โปรแกรมการหาค่าประมาณและค่าแตกต่างระหว่างค่าจริงกับค่าประมาณ โปรแกรมการหาอินเวอร์ส โปรแกรมการหาสัมประสิทธิ์ถดถอย โปรแกรมการหาค่าตัวสถิติ U

ลำดับที่ของ	ลักษณะการทำงาน	โปรแกรมย่อยที่เรียกใช้
3	คำนวณค่าเฉลี่ยของค่าตัวสถิติ U ค่าเฉลี่ยของค่าผลรวมกำลังสองของความคลาดเคลื่อนและหาค่าเฉลี่ยของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองจากการกระทำซ้ำ 100 รอบ ในแต่ละสถานการณ์	โปรแกรมการหาค่าผลรวมกำลังสองของความคลาดเคลื่อนและค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง
4	ทำการพิมพ์ผลลัพธ์ขั้นสุดท้าย	