บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้ศึกษาการประมาณค่าสัมประสิทธิ์การถดถอยพหุ เมื่อค่าความผิดพลาดมีการแจกแจงแบบปกติปลอมปนแบบหางยาวกว่าการแจกแจงปกติ การแจกแจงแบบเบ้และการแจกแจงแบบดับเบิลเอกซ์โปเนนเชียลโดยจะพิจารณาเปรียบเทียบประสิทธิภาพของการประมาณค่าสัมประสิทธิ์การถดถอยของวิธีกำลังสองน้อยที่สุด วิธีบูตสแตรปและวิธีตัวประมาณชนิด M โดยใช้ค่าเฉลี่ยความผิดพลาดกำลังสอง(MSE)เป็นตัวเปรียบเทียบ โดยมีวิธีการและแผนดำเนินงานดังต่อไปนี้

<u>3.1 วิธีมอนติคาร์โล</u> (Monte Carlo method)

เทคนิคการจำลองตัวแบบทางคณิตศาสตร์ที่นิยมใช้วิธีหนึ่ง คือ วิธีมอนติคาร์โล ซึ่งเป็นการจำลองตัวเลขสุ่ม (random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษาโดยการสร้างข้อมูลให้มีสภาพการแจกแจงตามที่ต้องการ ซึ่งขั้นตอนของวิธีมอนติคาร์โลที่ใช้กันอยู่ในปัจจุบันแบ่งเป็น 3 ขั้นตอนดังนี้

3.1.1 การสร้างตัวเลขสุ่ม การใช้ตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล เพราะว่าหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหาซึ่งลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบสม่ำเสมอ (uniform distribution) ในช่วง (0,1)สำหรับวิธีสร้างตัวเลขสุ่มนั้นมีผู้เสนอไว้หลายวิธี แต่วิธีหนึ่งที่ดีก็คือ วิธีของไวท์และชมิดท์ (White and Schemidt 1795 : 421) ที่ได้เสนอไว้กล่าวคือ ลักษณะของตัวเลขสุ่มที่ได้มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) และเป็นอิสระซึ่งกันและกัน ซึ่งรายละเอียดเสนอไว้ในภาคผนวก ก.

3.1.2 การประยุกต์ปัญหาที่ต้องการศึกษาเพื่อใช้กับตัวเลขสุ่ม ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจไม่ใช้ตัวเลขสุ่มโดยตรง แต่จะมีขั้นตอนอื่นอีกหลาย ๆ ขั้นตอนซึ่งบางครั้งต้องใช้ตัวเลขสุ่ม

3.1.3 การทดลองกระทำซ้ำ (replication) เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นต่อไปก็คือการทดลองโดยกระบวนการของเลขสุ่ม (random process) มากระทำในลักษณะที่กระทำซ้ำ ๆ กันเพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.2 แผนดำเนินการวิจัย

การวิจัยในครั้งนี้กำหนดสถานการณ์ต่าง ๆ ที่ต้องการศึกษาโดยสุ่มตัวอย่างจากประชากรที่มีการแจกแจงแบบเดียวกัน ซึ่งลักษณะการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติที่ต้องการศึกษาคือ แบบปกติปลอมปน สำหรับลักษณะการแจกแจงแบบเบ้ที่ต้องการศึกษา คือ ลอกนอร์มอล แกมมา และการแจกแจงแบบดับเบิลเอกซ์โปเนนเชียล

สำหรับการแจกแจงแบบปกติปลอมปนจะกำหนดเปอร์เซนต์การปลอมปนเป็น 1%,5%, 10% และ25% ส่วนสเกลแฟคเตอร์มี 2 ระดับ คือ 3 และ 10 ถ้าสเกลแฟคเตอร์สูงจะทำให้เกิดค่าผิดปกติสูง ซึ่งได้แสดงรายละเอียดในตารางที่ 3.1

ตารางที่ 3.1 แสดงค่าสเกลแฟคเตอร์และเปอร์เซนต์การปลอมปนทั้งหมดที่ใช้ในการวิจัยครั้งนี้

c \ p	1	5	10	25
3	(3,1)	(3,5)	(3,10)	(3,25)
10	(10,1)	(10,5)	(10,10)	(10,25)

สำหรับการแจกแจงแบบลอกนอร์มอล สนใจศึกษาเมื่อ ค่า $u = 0$, $\sigma^2 = 1$ ซึ่งเป็นค่าเฉลี่ยและค่าความแปรปรวนของการแจกแจงแบบปกติ

สำหรับการแจกแจงแบบแกมมาจะสนใจศึกษาเมื่อ $\alpha = 1, 2,$ และ 3 เมื่อ $B = 1$ โดยมีสัมประสิทธิ์ความแปรปรวนเป็น 100%,70% และ 58% ตามลำดับ

สำหรับการแจกแจงแบบดับเบิลเอ็กซ์โปแนนเชียลสนใจศึกษาเมื่อ

$\alpha = 0$ $\beta = 5,10$

ทุก ๆ การแจกแจงจะใช้ตัวแปรอิสระและขนาดตัวอย่างดังนี้

ตารางที่ 3.2 แสดงขนาดตัวอย่างและจำนวนตัวแปรอิสระทั้งหมดที่ใช้ในการวิจัยดังนี้

จำนวนตัวแปรอิสระ / ขนาดตัวอย่าง	2	3	5
4	(4,2)	-	-
5	-	(5,3)	-
10	-	(10,3)	-
20	-	(20,3)	-
50	-	-	(50,5)
100	-	-	(100,5)

3.3 ขั้นตอนของการดำเนินการวิจัย

การวิจัยนี้สนใจศึกษาการประมาณค่าสัมประสิทธิ์การถดถอยพหุ เมื่อค่าประมาณสเกลเปลี่ยนไป ในที่นี้พิจารณาเมื่อความผิดพลาดของข้อมูลมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติแบบดับเบิลเอกซ์โปเนนเชียลและแบบเบ้ โดยการสร้างโปรแกรม สำหรับการแจกแจงที่มีหางยาวกว่าการแจกแจงแบบปกติ แบบดับเบิลเอกซ์โปเนนเชียลและการแจกแจงแบบเบ้ ด้วยภาษาฟอร์แทน IV กับเครื่อง Vax-11/780 ซึ่งแต่ละโปรแกรมมีขั้นตอนในการดำเนินการดังนี้

3.3.1 กำหนดลักษณะการแจกแจงของความผิดพลาด (II) ขนาดตัวอย่างและจำนวนตัวแปรอิสระ เพื่อนำไปสร้างตัวแปรตาม (Y)

3.3.2 การสร้างข้อมูล คือตัวแปรตาม (Y) และตัวแปรอิสระ (X) เมื่อการแจกแจงของความผิดพลาดเป็นแบบหางยาวกว่าการแจกแจงปกติและแบบดับเบิ้ลเอกซ์โปเนนเชียล ตัวแปรตาม (Y) สร้างจากตัวแปรอิสระ (x) ซึ่งเป็นค่าคงที่และค่าความผิดพลาดมีลักษณะ

การแจกแจงตามที่ต้องการศึกษา โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระ ส่วนกรณีที่การแจกแจงของความผิดพลาดเป็นแบบเบ้ ตัวแปรตาม (Y) สร้างให้มีลักษณะการแจกแจงตาม การแจกแจงของค่าความผิดพลาด

3.3.3 ทดสอบความเหมาะสมของตัวแปรอิสระ (X) โดยทดสอบความเป็นอิสระซึ่งกันและกันหรือไม่

3.3.4 การประมาณค่าสัมประสิทธิ์การถดถอยพหุเชิงเส้น เมื่อความผิดพลาดมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ โดยใช้วิธีกำลังสองน้อยที่สุด วิธีบูตสแตรป วิธีตัวประมาณชนิด M เมื่อใช้สัดส่วนค่าสัมบูรณ์ของมัธยฐาน (MAD) เป็น ตัวประมาณสเกล ในที่นี้ใช้เกณฑ์ความแกร่งของ Huber ที่ b = 2

3.3.5 การประมาณค่าสัมประสิทธิ์การถดถอยพหุเชิงเส้น เมื่อความผิดพลาดมีการแจกแจงแบบเบ้จะอาศัยการแปลงที่อยู่ในรูปยกกำลังของ Box และ Cox แปลงข้อมูลให้อยู่ในภาวะปกติ แล้วจึงประมาณค่าสัมประสิทธิ์การถดถอยพหุเชิงเส้นโดยใช้วิธีกำลังสองน้อยที่สุด วิธีบูตสแตรป และ วิธีตัวประมาณชนิด M เมื่อใช้ สัดส่วนค่าสัมบูรณ์ของมัธยฐาน (MAD) เป็นตัวประมาณสเกล ใช้เกณฑ์ความแกร่งของ Humber ที่ $\beta = 2$

3.3.6 คำนวณหาค่าเฉลี่ยของความผิดพลาดกำลังสอง ของการประมาณค่า - สัมประสิทธิ์การถดถอย (mean square error)

3.3.7 สรุปผลในรูปของตาราง

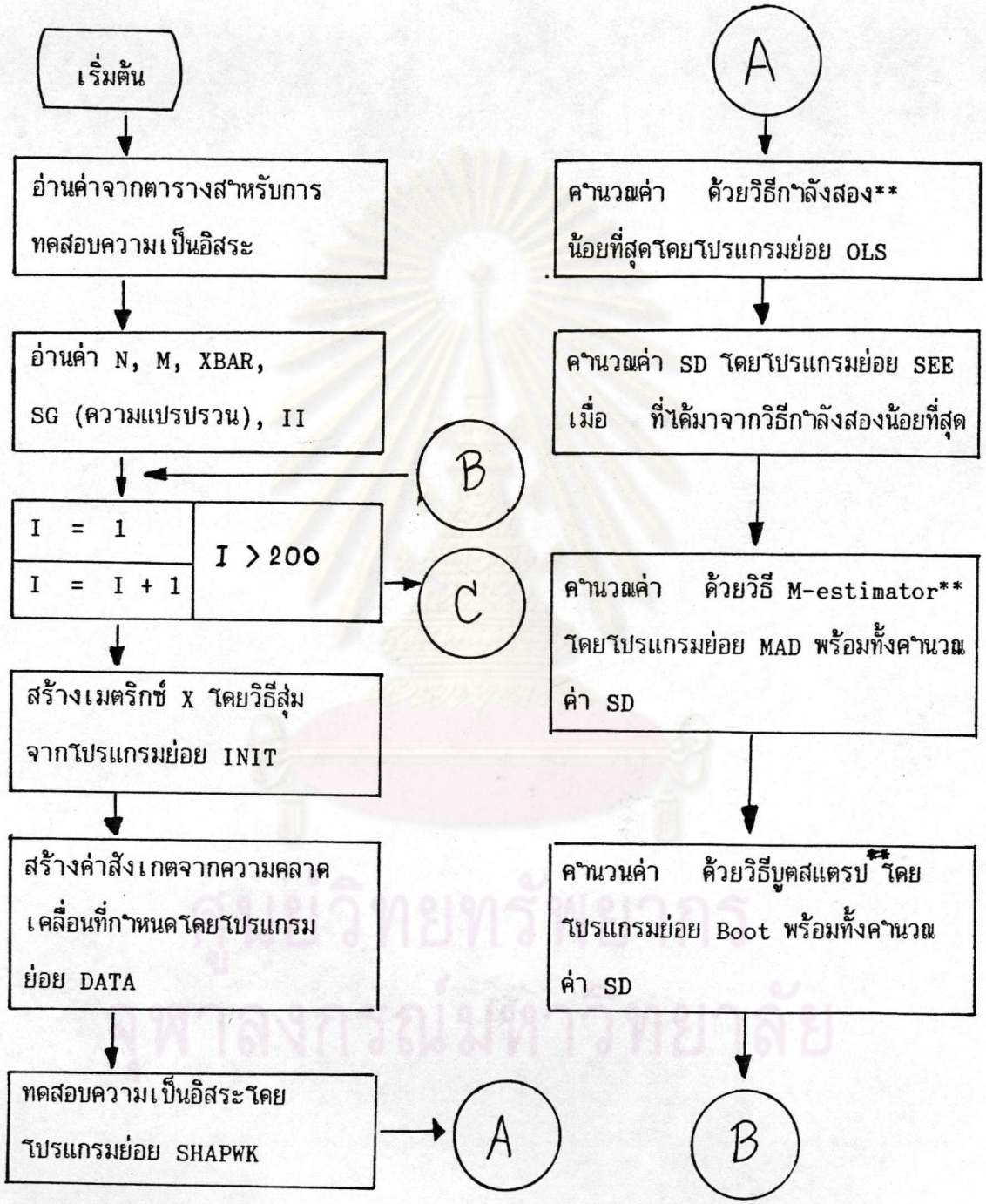
3.4 ผังงานวิธีดำเนินการวิจัย

ขั้นตอนในการดำเนินงานสามารถแสดงเป็นผังงานได้ ดังรูปที่ 3.1 ส่วนรายละเอียดของโปรแกรม แสดงไว้ในภาคผนวก ข. และความหมายของสัญลักษณ์ที่ใช้ในผังงานมีดังนี้

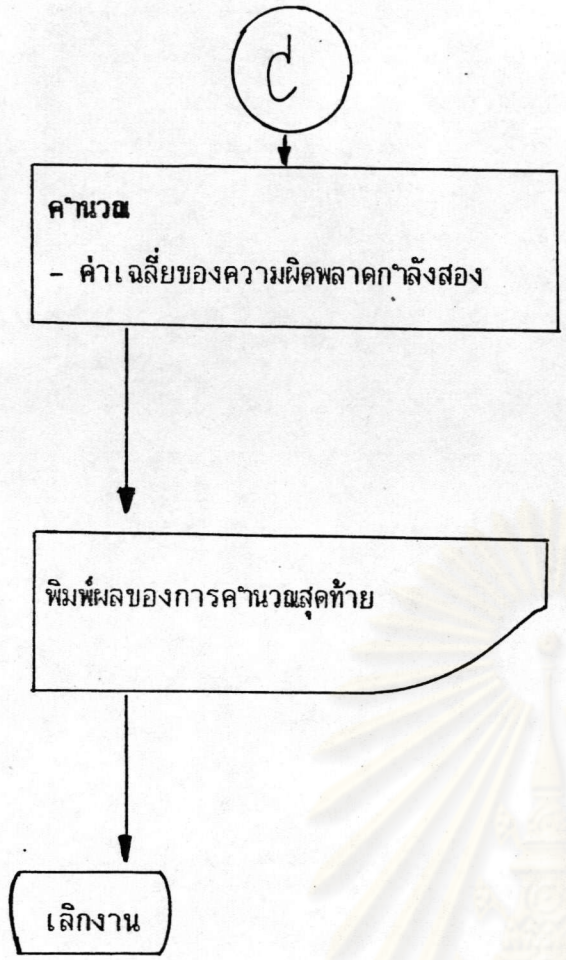
II	รูปแบบของการแจกแจงของความผิดพลาด
N	จำนวนค่าสังเกต
M	จำนวนพารามิเตอร์
IC	สเกลแฟคเตอร์
P	เปอร์เซนต์การปลอมปน
MSE	ค่าเฉลี่ยของความผิดพลาดกำลังสอง

i17191069

รูปที่ 3.1 แสดงผังงานสำหรับการประมาณค่าสัมประสิทธิ์การถดถอยพหุ เมื่อค่าประมาณสเกลเปลี่ยนไป



**กรณีที่ค่าความผิดพลาดมีการแจกแจงแบบเบ้จะต้องทำการแปลงค่าความผิดพลาดโดยใช้เทคนิคที่อยู่ในรูปยกกำลังของการแปลง Box และ Cox ก่อนการประมาณค่าสัมประสิทธิ์การถดถอยพหุ ทุกวิธี



จากรูปที่ 3.1 สามารถแยกอธิบายแต่ละขั้นตอนได้ดังนี้

3.4.1 เริ่มต้นระบบโดยการอ่านค่าจากตารางสำหรับการทดสอบความเป็นอิสระได้แก่ ตารางค่าสัมประสิทธิ์การทดสอบของ Shapiro และ Wilk และตาราง Percentage point ของการทดสอบ Shapiro และ Wilk ทั้งนี้พิจารณาตามขนาดตัวอย่างค่าสังเกต n = 4,5,10,20,50,100

3.4.2 อ่านค่าขนาดตัวอย่าง (N) จำนวนตัวแปรอิสระ (M) ค่าเฉลี่ย (XBAR) ค่าความแปรปรวน (SG) ของการแจกแจงแบบปกติและลักษณะการแจกแจงของประชากร(II)

3.4.3 สร้างเมตริกซ์ X ที่มีการแจกแจงแบบปกติมีขนาด n x m โดยวิธีสุ่มพร้อมทั้งสร้างเมตริกซ์ $(XX)^{-1}$ จากโปรแกรมย่อย INIT

3.4.4 สร้างค่าสังเกตจากความคลาดเคลื่อนที่กำหนดตาม II จากโปรแกรมย่อย DATA ได้แก่

3.4.4.1 การแจกแจงปกติ ฟังก์ชันที่ใช้ในการแจกแจงแบบปกติจะใช้วิธีของ BOX และ Muller (1958) ซึ่งเป็นวิธีที่สร้างการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น u และความแปรปรวนเป็น O^2 โดย DMEAN และ $(SIGMA)^2$ คือค่าเฉลี่ย และความแปรปรวนที่กำหนดรายละเอียดแสดงไว้ในภาคผนวก ก. การใช้ฟังก์ชันนี้ใช้คำสั่ง NORMAL (DMEAM, SIGMA) ผลลัพธ์คือ NORMAL ที่มีค่าเฉลี่ยและความแปรปรวนที่กำหนด

3.4.4.2 การแจกแจงแบบปกติปลอมปน สร้างโดยการใช้วิธีการแปลงข้อมูลมาจากการแจกแจงแบบปกติที่ได้มาจาก NORMAL (DMEAN,SIGMA) ด้วยความน่าจะเป็น (1-P) เมื่อ DMEAN SIGMA และ P เป็นค่าเฉลี่ย ความแปรปรวนและเปอร์เซนต์การปลอมปน และได้มาจาก NORMAL (DMEAN,SG2) ด้วยความน่าจะเป็น P เมื่อ SG2 เป็นผลคูณของ IC กับ SIGMA IC เป็นสเกลแฟกเตอร์สำหรับความแปรปรวนโดยที่รายละเอียดแสดงไว้ในภาคผนวก ก.

3.4.4.3 การแจกแจงแบบลอกนอร์มอล หาได้จากการหาเอกซ์โปเนนเชียล ของฟังก์ชัน NORMAL (DMEAN,SIGMA) เมื่อ DMEAN และ SIGMA เป็นค่าเฉลี่ยและความแปรปรวนของการแจกแจงแบบปกติสำหรับรายละเอียดแสดงไว้ในภาคผนวก ก.

3.4.4.4 การแจกแจงแบบแกมมา ฟังก์ชันที่ใช้คือ GAMMA1(ALPHA1, BETA1) ฟังก์ชันนี้สร้างโดยวิธี reproductive property ซึ่งรายละเอียดแสดงไว้ในภาคผนวก ก. ฟังก์ชัน GAMMA1 จะกำหนดค่า ALPHA1 และ BETA1 เพื่อให้ได้ค่าเฉลี่ยและความแปรปรวนตามที่ต้องการ ผลลัพธ์คือค่า GAMMA1 จะมีการแจกแจงแบบแกมมามีค่าเฉลี่ยเป็น BETA1 x ALPHA1 และความแปรปรวนเป็น $BETA1^2$ x ALPHA1

3.4.4.5 การแจกแจงแบบดับเบิ้ลเอกซ์โปเนนเชียล ฟังค์ชั่นที่ใช้ คือ DOUBLE (ALPHA, BETA,X) ฟังก์ชั่นนี้สร้างโดยวิธีการแปลงข้อมูลแบบผกผัน (Inverse Transformation) เมื่อกำหนดค่า ALPHA, BETA และ X เพื่อให้ได้ค่าเฉลี่ยและความแปรปรวนตามที่ต้องการ รายละเอียดแสดงไว้ในภาคผนวก ก.

3.4.5 การทดสอบความเหมาะสมของตัวแปรอิสระ X โดยทดสอบความเป็นอิสระซึ่งกันและกัน เนื่องจากตัวแปรอิสระเกิดภาวะความไม่เป็นอิสระซึ่งกันและกันแล้ว จะทำให้การประมาณค่าสัมประสิทธิ์การถดถอยพหุด้วยวิธีกำลังสองน้อยที่สุด วิธีบูตสแตรป และวิธีตัวประมาณ

ชนิด M ลดประสิทธิภาพลง Mongomery (ค.ศ.1982 หน้า 364-381) Mason และ Gunst (ค.ศ.1985 หน้า 401-407) และผลของการที่ตัวแปรอิสระมีความสัมพันธ์กันหรือเกิดภาวะความไม่เป็นอิสระซึ่งกันและกันแล้ว จะทำให้เมตริกซ์ x เกือบจะเป็น เมตริกซ์เดียว นั่นคือ ค่าดีเทอร์มิเนนท์ มีค่าน้อยและ เมตริกซ์ผกผันมีค่าสูงซึ่งจะทำให้ความแปรปรวนของสมการถดถอยมีค่าสูง และมีผลทำให้ยอมรับ H_o บ่อยครั้งดังนั้นการสรุปความสัมพันธ์ของ $\underset{\sim}{Y}$ และ X อาจเกิดความผิดพลาดได้

การทดสอบว่าเกิดภาวะความไม่เป็นอิสระซึ่งกันและกัน (multicollinearity) หรือไม่อาจทำการทดสอบสมมติฐานดังนี้

สมมติฐานของการทดสอบ H_o : ไม่เกิด multicollinearity

H_a : เกิด multicollinearity

ณ. ระดับสำคัญ α = 0.05 และ 0.01

ตัวสถิติทดสอบ $F_{(n-m+1)}^{(m-2)} = \frac{R_i^2/(m-2)}{(S-R_i^2)/(n-m+1)}$ (ไม่รวมจุดตัด (intercept))

โดยที่ R_i^2 เป็นสหสัมพันธ์ (correlation) ของ การถดถอย (regression) ที่ X_i บน x_j $(i \neq j)$; i,j, = 1,.....,m

ขอบเขตวิกฤต F จะหาได้จากตาราง ณ. ระดับนัยสำคัญ 0.05 และ 0.01

การตัดสินใจ จะปฏิเสธ H_o เมื่อ F (คำนวณ) > F (ตาราง)

3.4.6 การประมาณค่าสัมประสิทธิ์การถดถอยพหุด้วยวิธีกำลังสองน้อยที่สุด

3.4.6.1 เมื่อความผิดพลาดมีการแจกแจงเป็นแบบหางยาวว่าการแจกแจงแบบปกติ โปรแกรมย่อยที่ใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยคือ OLS (B) โดยที่ B เป็นเวกเตอร์ของสัมประสิทธ์การถดถอยพหุ ซึ่งเมตริกของ x จะใช้ร่วมกับโปรแกรมหลัก เมตริกซ์ x ประกอบด้วยโปรแกรมย่อยอีกดังนี้

ก) โปรแกรมย่อยสำหรับหาเมตริกซ์ผกผัน (inverse matrix) คือ โปรแกรม INVS(M,A) เมื่อ M เป็นขนาดของเมตริกซ์ผกผันที่ต้องการหาและ A เป็นเมตริกซ์ผกผันที่ต้องการ

ข) โปรแกรมย่อยสำหรับประมาณค่าผิดพลาด คือโปรแกรม YRESID (X,N,M, YHAT,YRES,β) โดยที่ x เป็นเมตริกซ์ของตัวแปรอิสระและตัวแปรตาม N คือจำนวนข้อมูล M คือจำนวนตัวแปรอิสระ YHAT คือค่าประมาณ ε_i YRES คือค่าผิดพลาด และ β คือสัมประสิทธิ์การถดถอยพหุ

3.4.6.2 เมื่อการแจกแจงของค่าผิดพลาดมีการแจกแจงแบบเบ้ จะอาศัยการแปลงที่อยู่ในรูป BOX และ COX โดยโปรแกรมย่อย BOXCOX (B,FL1) ซึ่ง B คือค่าสัมประสิทธิ์การถดถอยพหุ FL1 คือ $\hat{\lambda}$ ซึ่งเป็นค่าประมาณ λ ที่อยู่ในรูปยกกำลัง BOX และ COX โดยมีขั้นตอนการทำงานดังนี้

ขั้นตอน	ความหมาย
1	คำนวณค่าเฉลี่ยเรขาคณิต (geometric mean (G)) ของเวกเตอร์ Y โดย Y_i มีค่ามากกว่า 0 แล้วคำนวณค่า $y_i = y/g$
2	กำหนดค่าเริ่มต้น st ค่าสุดท้าย Fin เป็นขอบเขตในการหาค่า $\hat{\lambda}$ หรือ FL1 โดยใช้เทคนิคของการหา ที่มีค่าผลบวกกำลังสองของค่าผิดพลาดหรือที่มีค่า $\hat{\sigma}^2_{\varepsilon \mid \lambda}$ ต่ำสุด

3.4.6.3 การหาค่าเฉลี่ยความผิดพลาดกำลังสอง(Mean Square Error-(MSE)) โดยโปรแกรมย่อย SEE(B,SE) เมื่อ B ที่ใช้มาจากวิธีกำลังสองน้อยที่สุดโดยการหาผลรวมของ MSE ของ $\hat{\underset{\sim}{\beta}}$ ที่ได้จากการประมาณเทียบกับ $\underset{\sim}{\beta}$ ของประชากรทุกค่าของ ; i = 1...,m แล้วหารด้วย M

$$\text{MSE ของวิธีกำลังสองน้อยที่สุด} = \frac{\sum_i^m \left(\hat{\beta}_i(OLS) - \beta_i\right)^2}{M}$$

เมื่อ $\hat{\beta}_{(OLS)}$ เป็น $\hat{\beta}_i$ ของวิธีกำลังสองน้อยที่สุด

3.4.7 การประมาณค่าสัมประสิทธิ์การถดถอยพหุด้วยวิธีตัวประมาณชนิด M โปรแกรมที่เรียกใช้คือ MEST ซึ่งพิจารณาตามการแจกแจงของค่าผิดพลาดดังนี้

3.4.7.1 เมื่อความผิดพลาดเป็นแบบหางยาวกว่าการแจกแจงแบบปกติ จะเรียกโปรแกรมย่อยในการประมาณค่าสัมประสิทธิ์การถดถอยคือ

โปรแกรมย่อย MAD (B1,BMMAD,SIGMA) เมื่อ BMAD คือค่า $\hat{B}$ ที่ได้จากการประมาณค่าโดยวิธีตัวประมาณชนิด M เมื่อใช้สัดส่วนค่าสัมบูรณ์ของค่ามัธยฐาน (median absolute deviation) เป็นค่าประมาณสเกล และใช้เกณฑ์ความแกร่งของ Huber เมื่อ b=2 และ SIGMA คือค่า ประมาณสเกล

3.4.7.2 เมื่อความผิดพลาดเป็นแบบเบ้ จะอาศัยการแปลงที่อยู่ในรูปยกกำลังของ Box และ COX และประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยใช้โปรแกรมย่อยคือ

ก) การประมาณค่าสัมประสิทธิ์การถดถอยพหุ โดยวิธีตัวประมาณชนิด M ที่ใช้ค่าสัดส่วนค่าสัมบูรณ์ของมัธยฐาน เป็นตัวประมาณสเกล จากเทคนิคการแปลงที่อยู่ในรูปยกกำลังของ Box และ Cox โปรแกรมย่อยที่เรียกใช้คือ BOXC2 (B,FL1,SIGMA) ซึ่ง β เป็นค่าสัมประสิทธิ์ของการถดถอยพหุ FL1 คือ $\hat{\lambda}$ ซึ่งเป็นค่าประมาณที่อยู่ในรูปยกกำลังของ BOX และ COX สำหรับโปรแกรมย่อย BOXC2 (B,FL1,SIGMA) นี้ใช้คำตอบที่ได้จากโปรแกรม BOXCOX (B,FL1) ในการประมาณสัมประสิทธิ์การถดถอยพหุ และ $\hat{\lambda}$ ซึ่งเป็นค่าประมาณของการแปลง

ข) โปรแกรมย่อยสำหรับการคำนวณค่าผลบวกกำลังสองของค่าผิดพลาดโปรแกรมที่เรียกใช้คือ SUMMAD (FLX,SSE,K,SIGMA) โดยที่ FLX เป็นค่า $\hat{\lambda}$ที่ต้องการใช้ในการแปลงข้อมูล $\underset{\sim}{Y}$ ให้อยู่ในรูปตัวแบบเชิงเส้นดังนี้

$$(3.4.7.2.1)\quad \underset{\sim}{Y}^{(\lambda)} = \underset{\sim}{X}\underset{\sim}{B} + \underset{\sim}{\varepsilon}$$

SEE คือ $\hat{\sigma}_{\varepsilon|\lambda}$ ที่คำนวณได้จากตัวแบบเชิงเส้นนี้ K เป็นค่าใช้สำหรับบอกตำแหน่งของ SEE ว่าเป็นค่าของจุดเริ่มต้น ค่ากลาง หรือค่าสุดท้าย

ค) โปรแกรมย่อยสำหรับการแปลงข้อมูลโดยเทคนิคการแปลงที่อยู่ในรูปยกกำลังของ BOX และ COX โปรแกรมย่อยที่ใช้คือ BCOX(FL1) ซึ่ง FL1 คือค่า $\hat{\lambda}$ สำหรับการแปลงข้อมูล หรือ Y1 ให้อยู่ในรูปตัวแบบเชิงเส้น โดยที่ Y1 ใช้ข้อมูลร่วมกับโปรแกรม SUMMAD และ

ผลลัพธ์จะถูกแทนเมตริกซ์ X ในแนวตั้งที่ M2 เพื่อใช้ในการประมาณสัมประสิทธิ์การถดถอยพหุต่อไป

หลังจากการประมาณค่าสัมประสิทธิ์การถดถอยพหุแล้ว อันดับต่อไปคือการหาค่าเฉลี่ยความผิดพลาดกำลังสองของการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธี MAD โดยการหาผลรวมของ MSE ของ $\hat{\beta}_i$ ที่ได้จากการประมาณเทียบกับ β_i ของประชากรทุกค่าของ $\beta i; i = 1, \ldots\ldots, M$ แล้วหารด้วย M

3.4.8 การประมาณค่าสัมประสิทธิ์การถดถอยพหุ ด้วยวิธี Bootstrap โปรแกรมที่เรียกใช้คือ Boot ซึ่งมีขั้นตอนการทำงานดังนี้

3.4.8.1 กำหนดรูปแบบการแจกแจงของค่าความผิดพลาด และกระทำการสุ่มข้อมูลแบบใส่คืนซ้ำ ๆ กัน 100 ครั้ง ด้วยโปรแกรมย่อย WR

3.4.8.2 คำนวณการประมาณค่าสัมประสิทธิ์การถดถอยพหุ, β และค่า MSE ด้วยโปรแกรมย่อย Boot

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย