บทที่ 3

วิธีดำเนินการวิจัย

ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โล โดยใช้โปรแกรมภาษาฟอร์แทรน 77 ซึ่งขั้นตอนและโปรแกรมที่ใช้ในการวิจัยมีรายละเอียดดังต่อไปนี้

แผนการทดลอง

การวิจัยครั้งนี้กำหนดสถานการณ์ต่าง ๆ ที่ต้องการศึกษาสำหรับการเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของวิธีการทั้ง 5 วิธีซึ่งประชากรที่ศึกษามี 2 กลุ่ม แต่ละกลุ่มมีการแจกแจงแบบปกติของตัวแปรพหุ ดังนี้

1. กำหนดจำนวนตัวแปรตาม (K) เท่ากับ 3 5 และ 7
2. กำหนดโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรตามเป็น 2 กรณี คือ
   2.1. กรณีสหสัมพันธ์เท่ากัน $\rho_{kk'} = \rho$ สำหรับ $k \neq k'$
   2.2. กรณีสหสัมพันธ์ไม่เท่ากัน $\rho_{kk'} = \rho^{|k-k'|}$ ($k,k'=1,2,...,K$)
   โดยให้ทั้ง 2 กรณี มีค่า $\rho = 0 \rightarrow 0.9$
3. กำหนดขนาดตัวอย่างเท่ากันทั้ง 2 กลุ่มประชากรเท่ากับ 10 30 และ 50
4. กำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.01

การดำเนินการวิจัย

ขั้นตอนในการดำเนินการวิจัย มีดังนี้

1. การสร้างข้อมูลให้มีลักษณะตามที่กำหนดในแผนการทดลอง
2. การคำนวณค่าสถิติตามขั้นตอนของวิธีการทั้ง 5 วิธี
3. การหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ

รายละเอียดของแต่ละขั้นตอนมีดังนี้

1. การสร้างข้อมูลให้มีการแจกแจงปกติของตัวแปรพหุนั้น จำเป็นต้องสร้างตัวเลขสุ่มที่มีการแจกแจงปกติมาตรฐานของตัวแปรพหุเสียก่อน โดยใช้วิธีการของ Box และ Muller (ค.ศ. 1978) ซึ่งจะได้ตัวเลขสุ่มที่มีการแจกแจงปกติมาตรฐานพร้อมกัน 2 ค่า ดังนี้

$$Z_1 = (-2 \ln (R_1))^{1/2} \cos (2\pi R_2)$$

$$Z_2 = (-2 \ln (R_1))^{1/2} \sin (2\pi R_2)$$

เมื่อ Z_1 และ Z_2 เป็นตัวเลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน

R_1 และ R_2 เป็นตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ

จากนั้นจึงทำการสร้างตัวเลขสุ่มที่มีการแจกแจงปกติของตัวแปรพหุจากสมการ

$$\underset{\sim}{y} = C\underset{\sim}{z} + \underset{\sim}{\mu}$$

เมื่อ $\underset{\sim}{y}' = (y_1, y_2, \ldots, y_K)$ มีการแจกแจงแบบปกติของ K ตัวแปรสุ่ม โดยมีเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu}$ และเมทริกซ์ความแปรปรวนร่วม $\sum$

$\underset{\sim}{z}' = (z_1, z_2, \ldots, z_K)$ มีการแจกแจงแบบปกติของ K ตัวแปรสุ่ม โดยมีเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{0}$ และ เมทริกซ์ความแปรปรวนร่วม คือ I_K , I_K เป็นเมทริกซ์เอกลักษณ์ขนาด $K \times K$ C เป็นเมทริกซ์สามเหลี่ยมล่างเพียงเมทริกซ์เดียว (unique lower triangular metrix) ที่ทำให้

$CC' = \sum$ และ $\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1} & \sigma_{K2} & \cdots & \sigma_{KK} \end{bmatrix}$

การหาเมทริกซ์ C มีสูตรในการหาดังนี้

$$c_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik} c_{jk}}{\left(\sigma_{jj} - \sum_{k=1}^{j-1} c_{jk}^2 \right)^{1/2}}$$

เมื่อ $\sum_{k=1}^{0} c_{ik} c_{jk} = 0$, $1 \le j \le i \le K$

นั่นคือ $c_{11} = \frac{\sigma_{11}}{\sqrt{\sigma_{11}}}$, $1 \le i \le K$,

$$c_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} c_{ik}^2}$$, $1 < i \le K$,

$$c_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{jk}}{c_{jj}} \quad , 1 < j < i \leq K ,$$

$$c_{ij} = 0 \quad , 1 \leq i < j \leq K$$

ตัวอย่างเช่น $K = 3$ เราสามารถหาเมทริกซ์ $C = [\, c_{ij} \,]_{\substack{i=1,\ldots,K \\ j=1,\ldots,K}}$ ได้ดังนี้

1) $c_{11} = \sqrt{\sigma_{11}}$, $c_{21} = \frac{\sigma_{21}}{\sqrt{\sigma_{11}}}$, $c_{31} = \frac{\sigma_{31}}{\sqrt{\sigma_{11}}}$

2) $c_{22} = \sqrt{\sigma_{22} - c_{21}^2}$

3) $c_{32} = \frac{\sigma_{32} - c_{31}c_{21}}{c_{22}}$

4) $c_{33} = \sqrt{\sigma_{33} - (c_{21}^2 + c_{32}^2)}$

5) $c_{12} = c_{13} = c_{23} = 0$

เนื่องจากประชากรที่ศึกษาในการวิจัยครั้งนี้มี 2 กลุ่ม ดังนั้นเราจะต้องสร้างข้อมูลที่มีการแจกแจงแบบปกติของตัวแปรพหุ 2 ชุด แทนด้วย Y1 และ Y2 สำหรับประชากรกลุ่มที่ 1 และ 2 ตามลำดับ ซึ่งเป็นเมทริกซ์ขนาด $n \times K$ (เมื่อ n คือ ขนาดตัวอย่าง และ K คือ จำนวนตัวแปร)

โปรแกรมย่อยสำหรับสร้างข้อมูลที่มีการแจกแจงแบบปกติของตัวแปรพหุดังกล่าว คือ SUBROUTINE GENY(Y1, Y2) ซึ่งผู้วิจัยได้แสดงไว้ในภาคผนวก

2. การคำนวณค่าสถิติตามขั้นตอนของวิธีการทั้ง 5 วิธี

เมื่อผู้วิจัยได้ข้อมูลที่มีลักษณะตามที่ต้องการศึกษาแล้ว ก็จะทำการคำนวณค่าสถิติตามขั้นตอนของวิธีการต่าง ๆ ซึ่งได้เสนอไปแล้วในบทที่ 2 เพื่อทดสอบสมมุติฐานที่สนใจศึกษา

3. การหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ

เมื่อได้ผลลัพธ์ของการทดสอบสมมุติฐานด้วยวิธีการต่าง ๆ ในขั้นตอนที่ 2 แล้ว เราสามารถหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และ อำนาจการทดสอบของขั้นแรกของการทดสอบได้โดยกระทำขั้นตอนที่ 1 และ 2 ซ้ำ ๆ กัน เป็นจำนวน 500 ครั้ง แล้วนับจำนวนครั้งที่ปฏิเสธสมมุติฐานว่างที่มีมิติสูงสุด จากนั้นนำค่าผลรวมที่ได้หารด้วย 500 ส่วนค่าอำนาจการทดสอบเฉลี่ยของแต่ละขั้นของการทดสอบ (ขั้นที่ 2 ถึงขั้นที่ K) ในการเปรียบเทียบพหุฤฒแบบขั้นลดลงนั้นจะหาได้ดังนี้

$$PO(J) = \frac{\sum_{i=1}^{500} RPR_i(J)}{FAC(K,K-J+1)\times 500} \quad , J = 2, \ldots, K$$

เมื่อ $PO(J)$ แทน อำนาจการทดสอบเฉลี่ยของขั้นที่ J

$RPR_i(J)$ แทน จำนวนสมมุติฐานที่ปฏิเสธในขั้นที่ J จากการทดสอบสมมุติฐานที่เป็นไปได้ทั้งหมดในขั้นที่ J ของรอบที่ i

และ $FAC(K,K-I+1) = \binom{K}{K-J+1}$ = จำนวนสมมุติฐานที่เป็นไปได้ทั้งหมดในขั้นที่ J

หมายเหตุ การทดสอบด้วยวิธี BON JAM และ WFY ซึ่งเป็นวิธีการที่ทดสอบสมมุติฐานเชิงเดียว (Single Hypothesis , H_0^k) นั้น จะได้ผลของทุกขั้นในการทดสอบครั้งเดียว กล่าวคือ เมื่อทำการทดสอบสมมุติฐานด้วยวิธีดังกล่าวแล้ว จะได้ผลของการทดสอบขั้นที่ K หรือ $RPR_i(K)$ ทันที และถ้า $RPR_i(K) > 0$ ก็แสดงว่ามีการปฏิเสธสมมุติฐานขั้นที่ 1 ในรอบที่ i ดังนั้นเราจะต้องทำการหา $RPR_i(J)$, $J = 2, \ldots, K-1$ ต่อไป ซึ่งสามารถหาได้จากการคำนวณ โดยอาศัยนิยามของการทดสอบแบบขั้นลดลง (ซึ่งแสดงไว้ในบทที่ 2) ดังนี้

$$RPR_i(J) = \begin{cases} \binom{K}{K-J+1} - \binom{K-RPR_i(K)}{K-J+1}, & RPR_i(K) < J \\ \binom{K}{K-J+1}, & RPR_i(K) \geq J \end{cases}$$

เมื่อ $J = 2, 3, \ldots, K-1$

ซึ่งการคำนวณค่าผลรวมของ $RPR_i(J)$; $J = 2, 3, \ldots, K-1$ จะใช้โปรแกรมย่อย SUBROUTINE POWER

ส่วนค่า $RPR_i(J)$; $J = 2, 3, \ldots, K$ ของวิธี OLS และ GLS จะหาได้โดยตรงจากการทดสอบสมมุติฐานในแต่ละขั้น (ขั้นที่ 2 ถึง ขั้นที่ K) ตามขั้นตอนของวิธีการ

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เปรียบเทียบพหุคูณแบบขั้นตอนลง (ถ้ามีการปฏิเสธสมมุติฐานในขั้นแรกของการทดสอบรอบที่ i)

ผู้วิจัยได้แสดงตารางลักษณะการทำงานของโปรแกรมทั้งหมดที่ใช้ในการวิจัย รวมทั้งแผนผังโปรแกรมสำหรับการคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และ ค่าอำนาจการทดสอบ โดยใช้สัญญลักษณ์ต่าง ๆ ดังต่อไปนี้

วิธีที่ 1 แทน วิธีบอนเฟอร์โรนี - โฮล์ม (BON)

วิธีที่ 2 แทน วิธีเจมส์ - โฮล์ม (JAM)

วิธีที่ 3 แทน วิธีการทดสอบแบบปิด OLS (OLS)

วิธีที่ 4 แทน วิธีการทดสอบแบบปิด GLS (GLS)

วิธีที่ 5 แทน วิธีเวสต์ฟอล - ยัง (WFY)

NRI แทน จำนวนครั้งที่ปฏิเสธสมมุติฐานขั้นที่ 1 ของวิธีที่ I (เมื่อสมมุติฐานเป็นจริง)

RPRI(J) แทน จำนวนสมมุติฐานที่ปฏิเสธในขั้นที่ J จากการทดสอบสมมุติฐานที่ที่เป็นไปได้ทั้งหมดในขั้นที่ J ของวิธีที่ I (เมื่อสมมุติฐานเป็นเท็จ)

CPI(J) แทน ผลรวมของจำนวนครั้งที่ปฏิเสธสมมุติฐานขั้นที่ 1 เมื่อ J = 1 หรือ ผลรวมของจำนวนสมมุติฐานที่ปฏิเสธในขั้นที่ J เมื่อ J = 2 , 3 ,... , K ของวิธีที่ I (เมื่อสมมุติฐานเป็นเท็จ) จากการทำซ้ำ 500 รอบ

THI แทน ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของวิธีที่ I

POI (J) แทน อำนาจการทดสอบเฉลี่ยของขั้นที่ J ของวิธีที่ I

FAC (a,b) แทน $\binom{a}{b} = \frac{a!}{(a-b)!b!}$

ECORR=Y แทน โครงสร้างสหสัมพันธ์เป็นแบบเท่ากัน

ECORR=N แทน โครงสร้างสหสัมพันธ์เป็นแบบไม่เท่ากัน

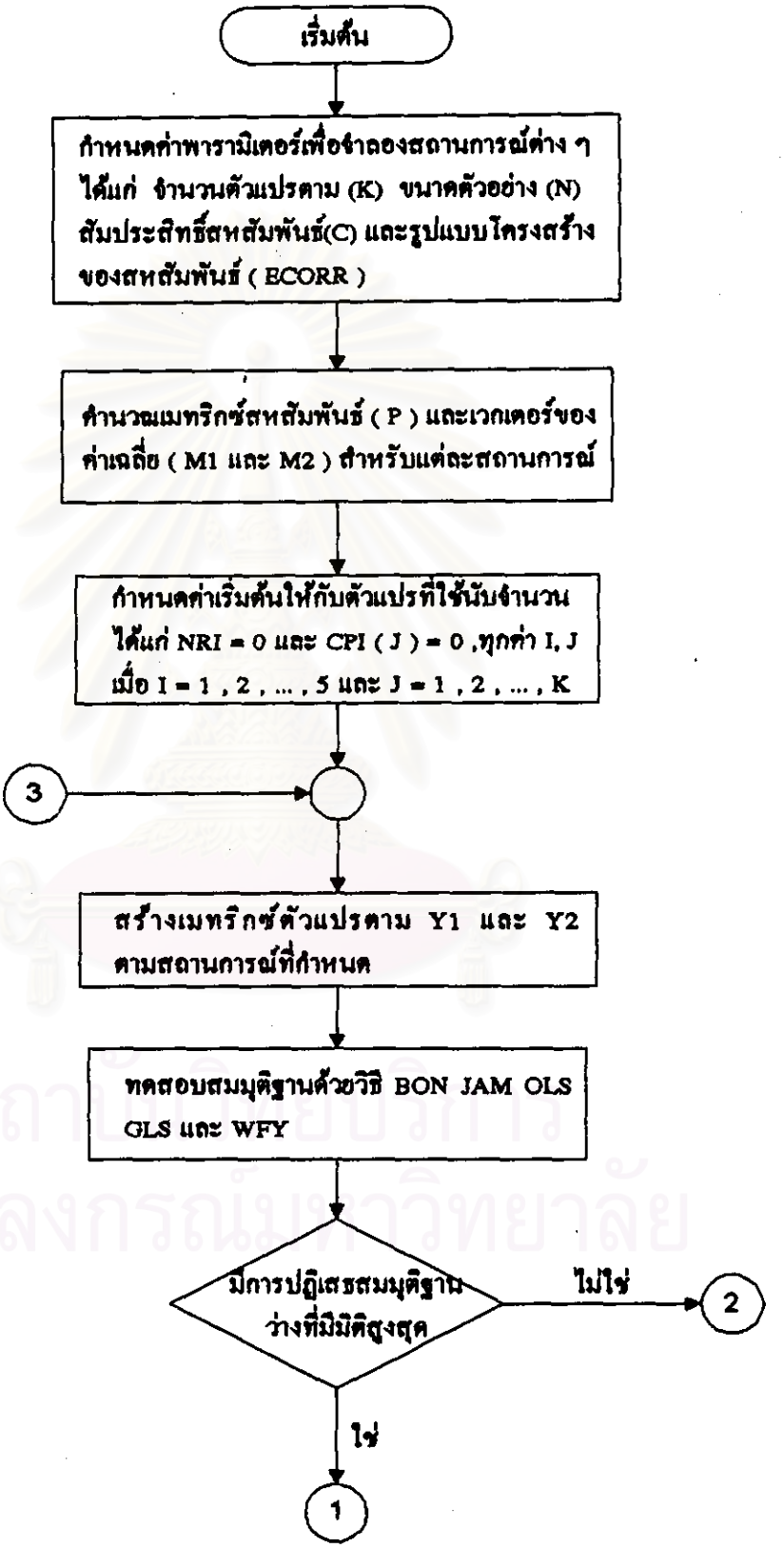
ตารางที่ 3.1 แสดงลักษณะการทำงานของโปรแกรมทั้งหมดที่ใช้ในการวิจัย

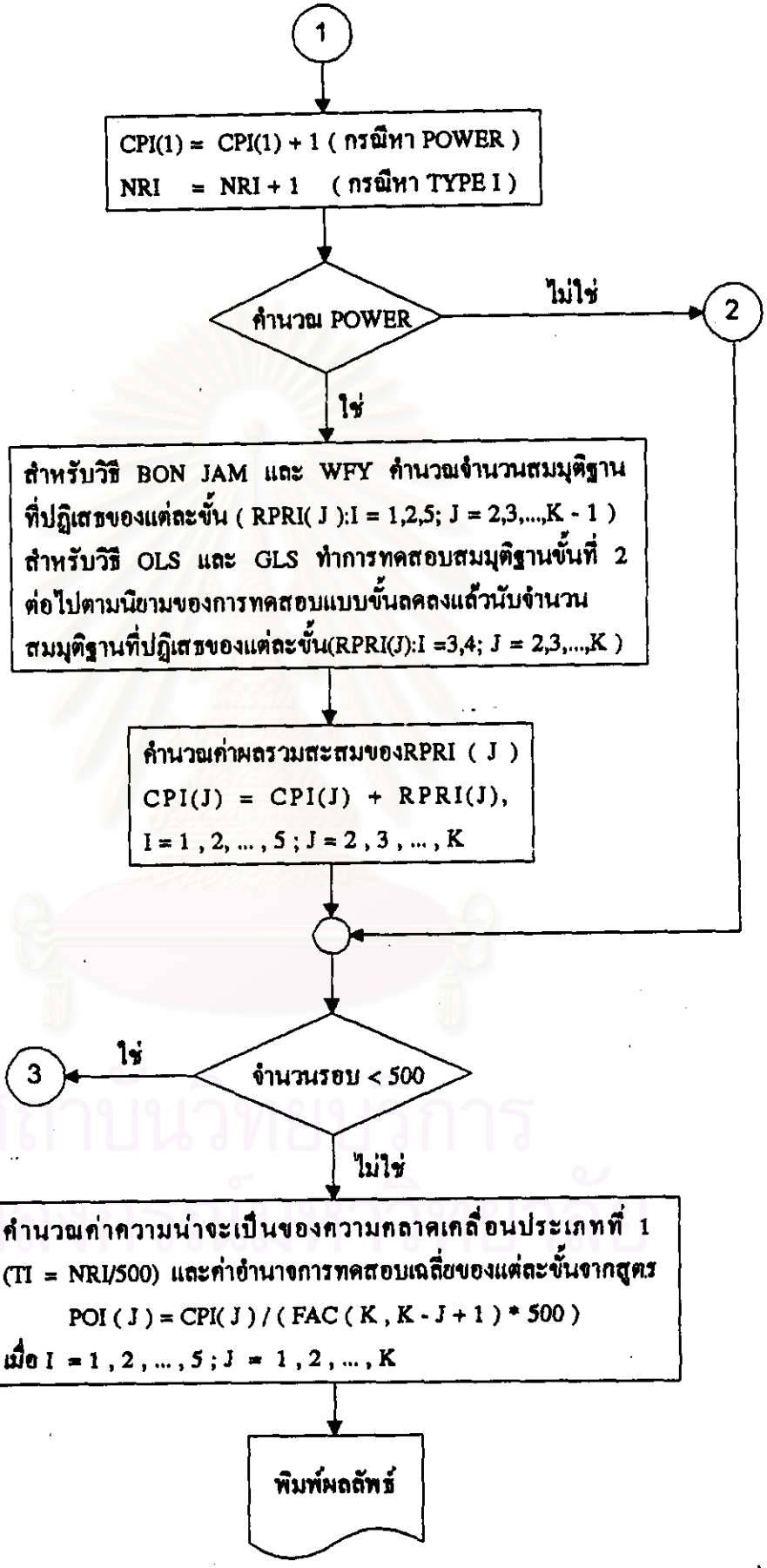
อันดับที่	ชื่อโปรแกรม	คุณสมบัติของโปรแกรม	ชื่อโปรแกรมที่เรียกใช้
โปรแกรมหลัก	MAIN	- อ่านค่าพารามิเตอร์ที่กำหนด - กำหนดค่าเริ่มต้นให้กับตัวแปรที่ใช้นับจำนวน - ทดสอบสมมุติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยระหว่าง 2 ประชากรบนพื้นฐานของตัวแปรตามพหุคูณด้วยวิธี BON JAM OLS GLS และ WFY - คำนวณค่าความน่าจะเป็นของความเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ - แสดงผล	GENY,CALPV,ASSORT, BON, JAM, OLS, GLS, WFY, POWER, CALPH, CALINP
SUBROUTINE			
1.	GENY	สร้างเมทริกซ์ตัวแปรตาม Y1 , Y2	NORMAL
2.	CALPV	คำนวณค่า P - value สำหรับวิธี BON JAM และ WFY	PROBST , SS , AVRG
3.	ASSORT	เรียงลำดับค่า P - value จากน้อยไปมาก	
4.	BON	ทดสอบสมมุติฐานด้วยวิธี BON	
5.	JAM	ทดสอบสมมุติฐานด้วยวิธี JAM	PPND , PHI
6.	WFY	ทดสอบสมมุติฐานด้วยวิธี WFY	SAMP , CALPV, MINP
7.	POWER	คำนวณค่า CPI (J) ; J = 1, ... , K สำหรับวิธี BON JAM และ WFY (I = 1 , 2 , 5)	FAC
8.	CALPH	คำนวณค่า $\hat{P}$ สำหรับวิธี OLS และ GLS	SKK , SS , AVRG
9.	OLS	ทดสอบสมมุติฐานด้วยวิธี OLS	FAC , INDEX , INSERT
10.	CALINP	คำนวณค่า $\hat{P}^{-1}$ สำหรับวิธี GLS	
11.	GLS	ทดสอบสมมุติฐานด้วยวิธี GLS	FAC , INDEX , INSERT

ตารางที่ 3.1 (ต่อ)

อันดับที่	ชื่อโปรแกรม	คุณสมบัติของโปรแกรม	ชื่อโปรแกรมที่เรียกใช้
12.	SAMP	สุ่มตัวอย่างแบบแทนที่สำหรับวิธีWFY	SRD
13.	RANDOM	สร้างเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ	
14.	SRD	สร้างเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอสำหรับการสุ่มซ้ำของวิธี WFY	
15.	INDEX	คำนวณดัชนีของสมมุติฐานที่เป็นไปได้ทั้งหมด ในแต่ละขั้นของการทดสอบสำหรับวิธี OLS และ GLS	FAC
16.	INSERT	คำนวณดัชนีของสมมุติฐานที่ต้องทดสอบในแต่ละขั้นของการทดสอบสำหรับวิธี OLS และ GLS	
FUNCTION			
1.	NORMAL	สร้างเลขสุ่มที่มีการแจกแจงแบบปกติ	
2.	PROBST	คำนวณค่าฟังก์ชันการแจกแจงสะสมของการแจกแจง t (F(t))	
3.	SS	คำนวณค่า SUM OF SQUARE	
4.	AVRG	คำนวณค่าเฉลี่ย	
5.	FAC	คำนวณค่า FAC (a,b)	
6.	PPND	คำนวณค่า Φ^{-1} (a) (ค่า z ที่ทำให้ฟังก์ชันการแจกแจงสะสมของ z หรือ Φ (z) มีค่าเท่ากับ a)	
7.	PHI	คำนวณค่า ϕ(a) ฟังก์ชันความหนาแน่นของการแจกแจงปกติ	
8.	SKK	คำนวณค่า $\hat{\sigma}_{kk'}$	
9.	MINP	คำนวณค่าต่ำสุดของ P - value	

รูปที่ 3.1 แผนผังโปรแกรมในการคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และ ค่าอำนาจการทดสอบ




1
CPI(1) = CPI(1) + 1 (กรณีหา POWER)
NRI = NRI + 1 (กรณีหา TYPE I)
คำนวณ POWER
ไม่ใช่
2
ใช่
สำหรับวิธี BON JAM และ WFY คำนวณจำนวนสมมุติฐาน ที่ปฏิเสธของแต่ละขั้น (RPRI(J):I = 1,2,5; J = 2,3,...,K - 1) สำหรับวิธี OLS และ GLS ทำการทดสอบสมมุติฐานขั้นที่ 2 ต่อไปตามนิยามของการทดสอบแบบขั้นลดลงแล้วนับจำนวน สมมุติฐานที่ปฏิเสธของแต่ละขั้น(RPRI(J):I =3,4; J = 2,3,...,K)
คำนวณค่าผลรวมสะสมของRPRI (J)
CPI(J) = CPI(J) + RPRI(J),
I = 1 , 2, ... , 5 ; J = 2 , 3 , ... , K
3
ใช่
จำนวนรอบ < 500
ไม่ใช่
คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (TI = NRI/500) และค่าอำนาจการทดสอบเฉลี่ยของแต่ละขั้นจากสูตร
POI (J) = CPI(J) / (FAC (K , K - J + 1) * 500)
เมื่อ I = 1 , 2 , ... , 5 ; J = 1 , 2 , ... , K
พิมพ์ผลลัพธ์