บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อทำการเปรียบเทียบตัวสถิติทดสอบที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบแกมมา กรณีที่ข้อมูลถูกตัดปลายประเภทที่ 2 โดยอาศัยตัวสถิติทดสอบ 3 ตัว คือ

1) ตัวสถิติทดสอบ $K\text{-}S^C$
2) ตัวสถิติทดสอบ AD^C
3) ตัวสถิติทดสอบ CVM^C

โดยใช้การจำลองด้วยวิธีมอนติคาร์โล (Monte Carlo Simulation Method) ซึ่งเขียนโปรแกรมคอมพิวเตอร์ด้วยภาษาฟอร์แทรน (Fortran) ในการจำลองซ้ำ 1,000 รอบในแต่ละสถานการณ์การทดลองและขั้นตอนในการวิจัยจะนำเสนอเป็นลำดับดังนี้

3.1 แผนการทดลอง

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแกมมาเมื่อข้อมูลถูกตัดปลายประเภทที่ 2 การกำหนดสถานการณ์ต่างๆ สำหรับการทดลองมีดังต่อไปนี้

3.1.1 ในการวิจัยครั้งเป็นการศึกษาข้อมูลในลักษณะข้อมูลถูกตัดปลายประเภทที่ 2 จึงไม่สามารถทดสอบสมมติฐานโดยการเปิดตารางค่าวิกฤตของข้อมูลแบบสมบูรณ์ได้ ดังนั้นในการวิจัยจะเริ่มจากการสร้างตารางแสดงค่าวิกฤตของตัวสถิติทดสอบ AD^C และ CVM^C (กรณีของตัวสถิติทดสอบ $K\text{-}S^C$ เปิดตารางค่าวิกฤตจากภาคผนวก ข) เพื่อใช้ในการทดสอบสมมติฐานสำหรับข้อมูลตัดปลายประเภทที่ 2 ต่อไป

3.1.2 สำหรับการเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นประเภทที่ 1 กำหนดให้ประชากรมีการแจกแจงแกมมา โดยกำหนดพารามิเตอร์เพื่อทดสอบภายใต้สมมติฐานว่างดังต่อไปนี้

การกำหนดพารามิเตอร์ (α) จะกำหนดจากสัมประสิทธิ์ความเบ้และความโด่ง สำหรับค่าพารามิเตอร์ (λ) จะกำหนดให้เท่ากับ 1.0

ตารางที่ 3.1 ค่าพารามิเตอร์ α และ λ และค่าสัมประสิทธิ์ความเบ้และความโด่งต่างๆ เมื่อประชากรมีการแจกแจงแกมมา

ส.ป.ส. ความเบ้	ส.ป.ส. ความโด่ง	พารามิเตอร์		ค่าเฉลี่ย	ความแปรปรวน
		α	λ		
0.5	3.38	16.00	1.0	16.00	16.00
1.1	4.82	3.30	1.0	3.30	3.30
2.3	10.94	0.76	1.0	0.76	0.76

3.1.3 สำหรับการเปรียบเทียบอำนาจการทดสอบ กำหนดให้ประชากรมีการแจกแจงต่างๆ ได้แก่ การแจกแจงแกมมาที่มิใช่พารามิเตอร์เดียวกันกับสมมติฐานว่าง การแจกแจงล็อกนอร์มอล การแจกแจงไวบูลล์ และการแจกแจงไคกำลังสอง การกำหนดค่าพารามิเตอร์ในแต่ละการแจกแจง จะใช้ค่าสัมประสิทธิ์ความเบ้และความโด่งเป็นตัวกำหนดค่า ดังนี้

<u>กรณีที่ประชากรมีการแจกแจงแกมมาที่มิใช่พารามิเตอร์เดียวกันกับสมมติฐานว่าง</u>

กำหนดค่าพารามิเตอร์ α และ λ ดังตารางที่ 3.2

ตารางที่ 3.2 ค่าพารามิเตอร์ α และ λ และค่าสัมประสิทธิ์ความเบ้และความโด่งต่างๆ เมื่อประชากรมีการแจกแจงแกมมาที่มิใช่พารามิเตอร์เดียวกันกับสมมติฐานว่าง

ส.ป.ส. ความเบ้	ส.ป.ส. ความโด่ง	พารามิเตอร์		ค่าเฉลี่ย	ความแปรปรวน
		α	λ		
0.48	3.34	17.4	1.0	17.4	17.4
1.00	4.50	4.0	1.0	4.0	4.0
1.90	8.42	1.1	1.0	1.1	1.1

<u>กรณีที่ประชากรมีการแจกแจงล็อกนอร์มอล</u>

กำหนดค่าพารามิเตอร์ μ และ σ ดังตารางที่ 3.3

ตารางที่ 3.3 ค่าพารามิเตอร์ μ และ σ และค่าสัมประสิทธิ์ความเบ้และความโด่งต่างๆ เมื่อประชากรมีการแจกแจงล็อกนอร์มอล

ส.ป.ส. ความเบ้	ส.ป.ส. ความโด่ง	พารามิเตอร์		ค่าเฉลี่ย	ความแปรปรวน
		μ	σ		
0.49	3.43	2.8	0.16	16.67	7.58
1.02	4.91	1.3	0.32	3.86	1.61
2.04	11.25	0.0	0.56	1.17	0.50

<u>กรณีที่ประชากรมีการแจกแจงไวบูลล์</u>

กำหนดค่าพารามิเตอร์ α และ β ดังตารางที่ 3.4

ตารางที่ 3.4 ค่าพารามิเตอร์ α และ β และค่าสัมประสิทธิ์ความเบ้และความโด่งต่างๆ เมื่อประชากรมีการแจกแจงไวบูลล์

ส.ป.ส. ความเบ้	ส.ป.ส. ความโด่ง	พารามิเตอร์		ค่าเฉลี่ย	ความแปรปรวน
		α	β		
0.6	3.22	2.02	18.0	15.95	68.27
1.1	4.47	1.48	3.5	3.17	4.73
1.9	8.44	1.03	1.0	0.98	0.92

<u>กรณีที่ประชากรมีการแจกแจงไคกำลังสอง</u>

กำหนดค่าพารามิเตอร์ n ดังตารางที่ 3.5

ตารางที่ 3.5 ค่าพารามิเตอร์ n และค่าสัมประสิทธิ์ความเบ้และความโด่งต่างๆ เมื่อประชากรมีการแจกแจงไคกำลังสอง

ส.ป.ส ความเบ้	ส.ป.ส ความโด่ง	พารามิเตอร์ (n)	ค่าเฉลี่ย	ความแปรปรวน
0.6	3.63	19	19	38
1.3	5.40	5	5	10
2.0	9.00	2	2	4

3.1.4 กำหนดขนาดตัวอย่างที่ศึกษา คือ 10, 20, 25, 30, 40, 50 และ 60

3.1.5 กำหนดเปอร์เซ็นต์การตัดปลายข้อมูล 3 ระดับ คือ 10%, 20% และ 30% จำนวนข้อมูลที่ถูกตัดปลายในแต่ละกรณีจะพิจารณาดังตารางที่ 3.6

ให้ n คือ ขนาดตัวอย่าง

s คือ จำนวนข้อมูลที่ถูกตัดปลาย

ตารางที่ 3.6 แสดงจำนวนข้อมูลที่ถูกตัดปลาย

s \ n	10	20	25	30	40	50	60
10%	1	2	3	3	4	5	6
20%	2	4	5	6	8	10	12
30%	3	6	8	9	12	15	18

3.1.6 กำหนดระดับนัยสำคัญของการทดสอบ 3 ระดับ คือ 0.01, 0.05 และ 0.10

3.2 ขั้นตอนในการทดลอง

ขั้นตอนในการทดลองแบ่งออกเป็น 5 ขั้นตอน ดังนี้

1. สร้างข้อมูลให้มีการแจกแจงแกมมา การแจกแจงล็อกนอร์มอล การแจกแจงไวบูลล์ และการแจกแจงไคกำลังสอง ให้มีลักษณะที่ต้องการศึกษา

2. คำนวณหาค่าวิกฤตของตัวสถิติทดสอบ AD^C และ CVM^C

3. คำนวณหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบ กระทำโดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่างหารด้วยจำนวนรอบการทดลอง 1,000 รอบ ทำการเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 กับเกณฑ์ที่กำหนด

4. เมื่อสถานการณ์ใดที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ จะทำการคำนวณหาค่าอำนาจการทดสอบของตัวสถิติทดสอบ กระทำโดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่างหารด้วยจำนวนรอบการทดลอง 1,000 รอบ

5. สรุปผลการทดลองในแต่ละสถานการณ์

3.3 การสร้างข้อมูลให้มีการแจกแจงตามลักษณะที่ต้องการศึกษา

การจำลองข้อมูลจากการแจกแจงของประชากรตามลักษณะที่กำหนดไว้ในการทดลอง สามารถทำได้โดยการจำลองด้วยเทคนิคมอนติคาร์โล โดยใช้โปรแกรมคอมพิวเตอร์ภาษาฟอร์แทน โดยใช้เลขสุ่มที่มีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) เป็นองค์ประกอบหลัก

การจำลองเลขสุ่มที่เป็นอิสระกันและมีการแจกแจงเอกรูปในช่วง[0,1][1]

วิธีการคณิตศาสตร์ในการจำลองเลขสุ่ม(เทียม) มีหลายวิธีการ สำหรับวิธีการที่ได้รับความนิยมใช้กันมากในปัจจุบัน คือ วิธีสมภาค(*Congruential Method*) ซึ่งมีสูตรหรือตัวแบบหนึ่งที่ใช้กันมาก คือ

$$X_i = (c + aX_{i-1}) \bmod m, \quad i = 1,2\ldots$$

---

[1] มานพ วราภักดิ์, <u>การจำลองเบื้องต้น</u>. (กรุงเทพฯ: ศูนย์ผลิตตำราเรียนสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2547), หน้า 43.

โดยที่ค่า c, a และ m เป็นค่าคงที่จำนวนเต็มค่าไม่เป็นลบ และความหมายของตัวแบบคือ X_i เป็นเศษเหลือที่เป็นจำนวนเต็มที่ได้จากการหาร $(c + aX_{i-1})$ ด้วย m นั่นคือ $X_i = (c + aX_{i-1}) - mk_i$ ซึ่ง $k_i = \lfloor (c + aX_{i-1})/m \rfloor$(หมายถึง จำนวนเต็มใหญ่ที่สุดที่น้อยกว่าหรือเท่ากับผลหาร $(c + aX_{i-1})/m$) ดังนั้นค่าเป็นไปได้ของ X_i คือ $0,1,\ldots,m-1$ และก่อนที่จะได้ค่าของ $X_1, X_2,\ldots$ ต้องกำหนดค่าของ c, a, m และ X_0 เราเรียก X_0 ว่า ซีด (*seed*) หรือ ค่าเริ่มต้น (*starting value*) จาก X_i ที่ได้จากการคำนวณนำมาหาค่า R_i ซึ่ง

$$R_i = \frac{X_i}{m}, \quad i = 1,2,\ldots$$

จะได้ R_i มีค่าอยู่ในช่วง [0,1) เรียก $R_1, R_2,\ldots$ ว่า เลขสุ่มเทียม หรือ เลขสุ่มคล้าย

ตัวแบบจำลองสมภาคแบบผลคูณที่ใช้กันมากตัวแบบหนึ่ง ซึ่งได้ผ่านการตรวจสอบคุณสมบัติแล้วหลายประการ คือ กำหนด $c = 0, m = 2^{31} - 1 = 2147483647, a = 7^5 = 16807$ และ X_0 เป็นจำนวนเต็มบวกที่เป็นเลขคี่ไม่เกิน m ฟังก์ชันการจำลองเลขสุ่มที่เป็นอิสระกันและมีการแจกแจงเอกรูปในช่วง [0,1] คือ SUBROUTINE URAND

การจำลองตัวแปรสุ่มที่มีการแจกแจงแกมมา[2]

การสร้างตัวแปรสุ่มให้มีการแจกแจงแกมมา สามารถแบ่งได้ออกเป็น 3 กรณี คือ

<u>กรณีที่ 1</u> $0 < \alpha < 1$

การจำลองตัวแปรสุ่มให้มีการแจกแจงแกมมาเมื่อพารามิเตอร์ α มีค่าในช่วง (0,1) โดยใช้เทคนิคการแปลง (*transformation technique*) ซึ่งเกี่ยวข้องกับการแจกแจงบีตา (*beta distribution*) ดังนี้

ให้ตัวแปรสุ่ม V และ W เป็นอิสระกัน และ V มีการแจกแจงเลขชี้กำลัง $Ex(1)$ และ W มีการแจกแจงบีตา $Be(\alpha, 1-\alpha)$, $0 < \alpha < 1$ ดังนั้น $X = \frac{VW}{\lambda}$ มีการแจกแจงแกมมา $G(\alpha, \lambda)$ จากเทคนิคดังกล่าวสามารถนำมาใช้ประโยชน์ในการจำลองตัวแปรสุ่มให้มีการแจกแจงแกมมาได้ตามขั้นตอนดังนี้

<u>ขั้นที่ 1</u> สร้างตัวแปรสุ่ม $V \sim Ex(1)$ และ $W \sim Be(\alpha, 1-\alpha)$

---

[2] เรื่องเดียวกัน.หน้า 157.

<u>ขั้นที่ 2</u> สร้างตัวแปรสุ่มให้มีการแจกแจงแกมมาที่มีพารามิเตอร์ α มีค่าอยู่ระหว่าง 0 และ 1 โดยนำตัวแปรสุ่มที่ได้จากขั้นที่ 1 มาแทนในสมการ

$$X = \frac{VW}{\lambda}$$

จะได้ว่า X มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α มีค่าอยู่ระหว่าง 0 และ 1

ฟังก์ชันการจำลองเลขสุ่มให้มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α มีค่าอยู่ระหว่าง 0 และ 1 คือ FUNCTION GAMMA1

<u>กรณีที่ 2</u> $\alpha = 1$

การแจกแจงแกมมาในกรณีที่ $\alpha = 1$ ก็คือ การแจกแจงเอกซโปแนนเชียล ดังนั้นในการจำลองเลขสุ่มที่มีการแจกแจงแกมมาในกรณีที่ $\alpha = 1$ จะใช้หลักการเดียวกันกับการจำลองเลขสุ่มที่มีการแจกแจงเอกซโปแนนเชียล โดยขั้นตอนการสร้าง ดังนี้

<u>ขั้นที่ 1</u> จำลองเลขสุ่ม R

<u>ขั้นที่ 2</u> สร้างตัวแปรสุ่มให้มีการแจกแจงแกมมาเมื่อพารามิเตอร์ $\alpha = 1$ จากสมการ

$$X = -\frac{1}{\lambda}\ln R$$

จะได้ว่า X มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α เท่ากับ 1

ฟังก์ชันการจำลองเลขสุ่มให้มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α เท่ากับ 1 คือ FUNCTION GAMMA2

<u>กรณีที่ 3</u> $\alpha > 1$

Cheng(1977) ได้นำเสนอวิธีวิธีการสร้างตัวแปรสุ่มให้มีการแจกแจงแกมมาเมื่อพารามิเตอร์มีค่ามากกว่า 1 โดยใช้วิธีการรับ-ปฏิเสธ (*acceptance-rejection*) โดยมีขั้นตอนวิธีสำหรับจำลองตัวแปรสุ่มแกมมาดังนี้

<u>ขั้นที่ 1</u> คำนวณค่าคงที่ต่างๆ จากสูตรต่อไปนี้

$$a = \sqrt{2\alpha - 1}$$

$$b = 2\alpha - \ln 4 + \frac{1}{a}$$

<u>ขั้นที่ 2</u> จำลองเลขสุ่ม R_1 และ R_2 จากโปรแกรมย่อย SUBROUTINE URAND

<u>ขั้นที่ 3</u> $Y = \alpha\left(\frac{R1}{1-R1}\right)^a$

<u>ขั้นที่ 4</u> ถ้า $Y > b - \ln(R_1^2 R_2)$ ให้กลับไปขั้นที่ 2

<u>ขั้นที่ 5</u> $X = Y / \lambda$

จะได้ว่า X มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α มากกว่า 1

ฟังก์ชันการจำลองเลขสุ่มให้มีการแจกแจงแกมมาที่มีพารามิเตอร์ α และ λ เมื่อพารามิเตอร์ α มากกว่า 1 คือ FUNCTION GAMMA3

การจำลองตัวแปรสุ่มที่มีการแจกแจงปกติ[3]

Marsaglia, Maclaren และ *Bray*(1964) ได้ดัดแปลงการจำลองตัวแปรสุ่มที่มีการแจกแจงปกติของ *Box* และ *Muller* โดยการหลีกเลี่ยงการคำนวณ *cosine* และ *sine* เรียกวิธีดังกล่าวว่า วิธีโพลาร์ ซึ่งมีตัวแบบการจำลองดังต่อไปนี้

$$Z_1 = V_1\sqrt{\frac{-2\ln(V_1^2 + V_2^2)}{V_1^2 + V_2^2}}$$

$$Z_2 = V_2\sqrt{\frac{-2\ln(V_1^2 + V_2^2)}{V_1^2 + V_2^2}}$$

ขั้นตอนการจำลองมีดังต่อไปนี้

<u>ขั้นที่ 1</u> จำลองเลขสุ่ม R_1 และ R_2

<u>ขั้นที่ 2</u> $V_1 = 2R_1 - 1$, $V_2 = 2R_2 - 1$ (จำลอง V_1, V_2 จาก $U(-1,1)$)

<u>ขั้นที่ 3</u> $S = V_1^2 + V_2^2$

---

[3] เรื่องเดียวกัน, หน้า 145.

<u>ขั้นที่ 4</u> ถ้า $S > 1$ กลับไปขั้นที่ 1

<u>ขั้นที่ 5</u> $W = \sqrt{\frac{-2\ln S}{S}}$

<u>ขั้นที่ 6</u> $Z_1 = V_1 W$, $Z_2 = V_2 W$

วิธีการจำลองดังกล่าวจะทำให้ได้ตัวแปรสุ่มปกติมาตรฐาน Z_1 และ Z_2 ซึ่งเป็นอิสระกัน เพื่อให้การเขียนโปรแกรมเรียกใช้ง่ายขึ้น ดังนั้นจึงได้เลือกใช้ Z_1 เพียงตัวเดียว

จากตัวแปรสุ่ม Z_1 สามารถแปลงให้เป็นตัวแปรสุ่มที่มีการแจกแจงปกติค่าเฉลี่ยเท่ากับ μ ความแปรปรวนเท่ากับ σ^2 ได้สมการดังต่อไปนี้

$$X_1 = \mu + \sigma Z_1$$

ฟังก์ชันการจำลองเลขสุ่มให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 คือ FUNCTION NORM

การจำลองตัวแปรสุ่มที่มีการแจกแจงล็อกนอร์มอล[4]

การสร้างตัวแปรสุ่มให้มีการแจกแจงแบบล็อกนอร์มอลที่มีพารามิเตอร์ μ และ σ^2 ใช้ความสัมพันธ์ระหว่างการแจกแจงปกติกับการแจกแจงล็อกนอร์มอล นำมาใช้ประโยชน์ในการจำลอง X โดยมีขั้นตอนดังต่อไปนี้

<u>ขั้นที่ 1</u> สร้างตัวแปรสุ่ม $Y = \mu + \sigma Z$ จาก FUNCTION NORM

<u>ขั้นที่ 2</u> สร้างตัวแปรสุ่มให้มีการแจกแจงแบบล็อกนอร์มอล โดยการนำตัวแปรสุ่มที่ได้จากขั้นที่ 1 แทนในสมการ

$$X = e^Y$$

จะได้ว่า X มีการแจกแจงแบบล็อกนอร์มอลที่มีพารามิเตอร์ μ และ σ^2

ฟังก์ชันของการจำลองเลขสุ่มให้มีการแจกแจงแบบล็อกนอร์มอลที่มีพารามิเตอร์ μ และ σ^2 คือ FUNCTION LOGNORM

---

[4] เรื่องเดียวกัน. หน้า 149.

การจำลองตัวแปรสุ่มที่มีการแจกแจงไวบูลล์[5]

การสร้างตัวแปรสุ่มให้มีการแจกแจงไวบูลล์ที่มีพารามิเตอร์ α และ β โดยใช้เทคนิคการแปลงผกผัน(*inverse transformation*) มีขั้นตอนดังต่อไปนี้

<u>ขั้นที่ 1</u> จำลองเลขสุ่ม R

<u>ขั้นที่ 2</u> $X = \beta[-\ln(1-R)]^{1/\alpha}$

จะได้ว่า X ที่มีการแจกแจงไวบูลล์ที่มีพารามิเตอร์ α และ β

ฟังก์ชันการจำลองเลขสุ่มให้มีการแจกแจงไวบูลล์ที่มีพารามิเตอร์ α และ β คือ FUNCTION WEI

การจำลองตัวแปรสุ่มที่มีการแจกแจงไคกำลังสอง[6]

การจำลองตัวแปรสุ่มไคกำลังสองใช้คุณสมบัติความสัมพันธ์ระหว่างตัวแปรสุ่มไคกำลังสองกับตัวแปรสุ่มแกมมา เนื่องจากการแจกแจงไคกำลังสอง $\chi^2(n)$ คือการแจกแจงแกมมา $G(\alpha,\lambda)$ที่มีพารามิเตอร์ $\alpha = \frac{n}{2}$ และ $\lambda = \frac{1}{2}$ ดังนั้น ถ้า n เป็นเลขคู่ เราสามารถจำลอง $X \sim \chi^2(n)$ ได้จากผลบวกของตัวแปรสุ่มเลขชี้กำลัง $Ex\left(\frac{1}{2}\right)$ จำนวน $\frac{n}{2}$ ตัว ได้ตัวแบบจำลอง

$$X = -2\ln\left(\prod_{i=1}^{n/2} R_i\right) \quad , \quad R_i \sim U(0,1) \quad , \quad n = \text{เลขคู่บวก}$$

แต่ถ้า n เป็นเลขคี่ เราสามารถจำลองตัวแปรสุ่มเลขชี้กำลัง $Ex\left(\frac{1}{2}\right)$ จำนวน $(n-1)/2$ และที่เหลืออีก 1 ระดับขั้นความเสรี จะจำลองจาก $N(0,1)$ ฉะนั้น ได้ตัวแบบจำลอง

$$X = -2\ln\left(\prod_{i=1}^{(n-1)/2} R_i\right) + Z^2 \quad , \quad R_i \sim U(0,1) \quad , \quad Z \sim N(0,1) \quad , \quad n = \text{เลขคี่บวก}$$

จะได้ว่า X มีการแจกแจงไคกำลังสองที่มีพารามิเตอร์ n

---

[5] Law A. W. and Kelton W. D., <u>Simulation Modeling and Analysis</u> (New York: McGraw-Hill, 1982), p. 258.

[6] มานพ วราภักดิ์, <u>การจำลองเบื้องต้น</u>. หน้า 165.

ฟังก์ชันของการจำลองเลขสุ่มให้มีการแจกแจงไคกำลังสองที่มีพารามิเตอร์ n คือ FUNCTION CHI

3.4 การหาค่าวิกฤตของตัวสถิติทดสอบ AD^C และ CVM^C

ในการวิจัยครั้งนี้จะทำการคำนวณหาค่าวิกฤตที่ระดับนัยสำคัญ (α) ต่างๆ โดยใช้หลักการหาค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ (*Percentile*) ที่ $100(1-\alpha)$ ซึ่งมีขั้นตอนดังต่อไปนี้

(1) กำหนดขนาดตัวอย่าง (n) และเปอร์เซ็นต์การตัดปลายข้อมูลที่ต้องการศึกษา

(2) จำลองข้อมูลให้ประชากรมีการแจกแจงแบบแกมมาตามพารามิเตอร์ที่กำหนดตามหัวข้อ 3.1.2 ตารางที่ 3.1

(3) จากจำลองในตามข้อ (2) จะได้ค่า $X_1, X_2, \ldots, X_n$

(4) นำค่า $X_1, X_2, \ldots, X_n$ มาจัดเรียงจากน้อยไปหามาก จากนั้นตัดปลายด้านขวาของข้อมูลตามจำนวนที่กำหนด จะได้ค่า $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$

(5) คำนวณค่าสถิติทดสอบ AD^C และ CVM^C

(6) ทำซ้ำข้อ (2) ถึงข้อ (5) จำนวน 50,000 ค่า เรียงลำดับค่าสถิติทดสอบจากน้อยไปหามาก หาค่าเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha)$ โดยมีวิธีการคำนวณดังต่อไปนี้

<u>การหาค่าเปอร์เซ็นต์ไทล์</u>[7]

กำหนด m คือ จำนวนข้อมูลทั้งหมด (m = 50,000 ตามข้อ (6))

y_α คือ ค่าเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha)$

การหาค่าเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha)$ คือ การคำนวณหาค่าของข้อมูลที่อยู่ในลำดับที่ $(1-\alpha)\times(m+1)$

ถ้าค่า $(1-\alpha)\times(m+1)$ เป็นเลขจำนวนเต็มเท่ากับ r ค่าเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha)$ คือ ค่าของข้อมูลที่อยู่ในลำดับที่ r

ถ้าค่า $(1-\alpha)\times(m+1)$ ไม่เป็นเลขจำนวนเต็ม คือถ้า $(1-\alpha)\times(m+1)=r.p$ โดยที่ r เป็นเลขหน้าจุดทศนิยม และ p เป็นเลขทศนิยม ($0< p <1$) ค่าเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha)$ จะคำนวณจากสูตร

$$y_\alpha = (1-p)\times(\text{ค่าของข้อมูลที่อยู่ในลำดับที่ } r) + p\times(\text{ค่าของข้อมูลที่อยู่ในลำดับที่ } r+1)$$

---

[7] เรื่องเดียวกัน, หน้า 375.

วิธีการตัดปลายข้อมูล

กำหนดขนาดตัวอย่างเท่ากับ n และกำหนดการตัดปลาย ($P\%$) ดังนั้นจำนวนข้อมูลที่ถูกตัดปลาย คือ $n \times P\%$ และจำนวนข้อมูลที่ต้องการคือ $r = n - (n \times P\%)$

เช่น กำหนดขนาดตัวอย่างเท่ากับ 100 และเปอร์เซ็นต์การตัดปลายเท่ากับ 10% จำนวนข้อมูลที่ถูกตัดปลาย คือ $100 \times 10\% = 10$ และจำนวนข้อมูลที่ต้องการคือ $r = 100 - 10 = 90$

3.5 การหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

ในการหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ขั้นตอนในการทำงานมีดังนี้

(1) กำหนดขนาดตัวอย่าง (n) และเปอร์เซ็นต์การตัดปลายข้อมูลที่ต้องการศึกษา

(2) จำลองข้อมูลให้ประชากรมีการแจกแจงแบบแกมมาตามพารามิเตอร์ที่กำหนดตามหัวข้อ 3.1.2 ตารางที่ 3.1

(3) จากจำลองในตามข้อ (2) จะได้ค่า $X_1, X_2, \ldots, X_n$

(4) นำค่า $X_1, X_2, \ldots, X_n$ มาจัดเรียงจากน้อยไปหามาก จากนั้นตัดปลายด้านขวาของข้อมูลตามจำนวนที่กำหนด จะได้ค่า $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$

(5) คำนวณค่าสถิติทดสอบ K-S^C AD^C และ CVM^C

(6) ทดสอบว่าปฏิเสธสมมติฐานว่างหรือไม่ และนับจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง

(7) ทำซ้ำข้อ (2) ถึงข้อ (6) จำนวน 1,000 รอบ

(8) คำนวณหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 คือ การนำจำนวนครั้งที่ปฏิเสธสมมติฐานว่างหารด้วยจำนวนรอบทั้งหมด (1,000 รอบ)

จากนั้นจะพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยการจะนำค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้ไปเปรียบเทียบกับเกณฑ์ที่กำหนด

เกณฑ์ในการทดสอบว่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในเกณฑ์ที่ควบคุมได้หรือไม่นั้นทำได้โดยการสอบทวินาม (*binomial test*)[8] โดยมีสมมติฐานของการทดสอบดังนี้

$$H_0 \quad : \quad \alpha \le \alpha_0$$

$$H_1 \quad : \quad \alpha > \alpha_0$$

---

[8] เรื่องเดียวกัน,หน้า 360

โดยใช้ทฤษฎีบทลิมิตสู่ส่วนกลาง (*Central Limit Theorem*) ความน่าจะเป็นที่จะยอมรับสมมติฐาน H_0 เท่ากับ

$$P\left[\frac{\hat{\alpha}-\alpha_0}{\sqrt{\frac{\alpha_0(1-\alpha_0)}{k}}} \le z_\gamma\right] = 1-\gamma$$

หรือ
$$P\left[\hat{\alpha} \le \alpha_0 + z_\gamma\sqrt{\frac{\alpha_0(1-\alpha_0)}{k}}\right] = 1-\gamma$$

ดังนั้น ช่วงของการยอมรับค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของตัวสถิติทดสอบ คือ

$$\left[0, \alpha_0 + z_\gamma\sqrt{\frac{\alpha_0(1-\alpha_0)}{k}}\right]$$

โดยที่ α คือ ระดับนัยสำคัญของการทดสอบ

α_0 คือ ระดับนัยสำคัญที่กำหนดในการศึกษา มี 3 ระดับ คือ 0.01, 0.05 และ 0.10

$\hat{\alpha}$ คือ ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง มีค่าเท่ากับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นจริง หารด้วยจำนวนครั้งในการทดลอง

k คือ จำนวนรอบของการจำลอง เท่ากับ 1,000 รอบ

γ คือ ระดับนัยสำคัญที่ใช้ในการทดสอบทวินาม

วิธีการคำนวณเกณฑ์การทดสอบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 คือ กำหนดให้มีการทดสอบทวินามที่ระดับนัยสำคัญที่ $\gamma = 0.05$ ดังนั้นจะได้ $z_\gamma = 1.645$

1. ระดับนัยสำคัญ $\alpha_0 = 0.10$

   $\hat{\alpha}$ จะอยู่ในช่วง $\left[0, 0.10 + 1.645\sqrt{\frac{0.10(0.90)}{1,000}}\right]$ = [0 , 0.116]

2. ระดับนัยสำคัญ $\alpha_0 = 0.05$

   $\hat{\alpha}$ จะอยู่ในช่วง $\left[0, 0.10 + 1.645\sqrt{\frac{0.05(0.95)}{1,000}}\right]$ = [0 , 0.061]

3. ระดับนัยสำคัญ $\alpha_0 = 0.01$

   $\hat{\alpha}$ จะอยู่ในช่วง $\left[0, 0.10 + 1.645\sqrt{\frac{0.01(0.99)}{1,000}}\right]$ = [0 , 0.015]

ดังนั้นตัวสถิติทดสอบจะสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ถ้า $\hat{\alpha}$ อยู่ในช่วงของการยอมรับ ดังต่อไปนี้

ระดับนัยสำคัญ	ช่วงของการยอมรับ
0.10	[0 , 0.116]
0.05	[0 , 0.061]
0.01	[0 , 0.015]

เมื่อผ่านเกณฑ์การควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 แล้ว จะทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบเป็นลำดับต่อไป

3.6 การหาค่าอำนาจการทดสอบ

การคำนวณหาค่าอำนาจการทดสอบ ขั้นตอนในการทำงานมีดังนี้

(1) กำหนดขนาดตัวอย่าง (n) และเปอร์เซ็นต์การตัดปลายข้อมูลที่ต้องการศึกษา

(2) จำลองข้อมูลให้ประชากรมีการแจกแจงอื่นที่ไม่ใช่การแจกแจงเดียวกับสมมติฐานว่าง ซึ่งการแจกแจงที่ใช้งานวิจัยนี้ คือ การแจกแจงแกมมาที่มิใช่พารามิเตอร์เดียวกันกับสมมติฐานว่าง การแจกแจงล็อกนอร์มอล การแจกแจงไวบูลล์ และการแจกแจงไคกำลังสอง โดยทำการศึกษาทดลองครั้งละการแจกแจง

การจำลองข้อมูลให้มีการแจกแจงแกมมาที่มิใช่พารามิเตอร์เดียวกันกับสมมติฐานว่าง จะกำหนดค่าพารามิเตอร์ตามหัวข้อ 3.1.3 ตารางที่ 3.2

การจำลองข้อมูลให้มีการแจกแจงล็อกนอร์มอล จะกำหนดค่าพารามิเตอร์ตามหัวข้อ 3.1.3 ตารางที่ 3.3

การจำลองข้อมูลให้มีการแจกแจงไวบูลล์ จะกำหนดค่าพารามิเตอร์ตามหัวข้อ 3.1.3 ตารางที่ 3.4

การจำลองข้อมูลให้มีการแจกแจงไคกำลังสอง จะกำหนดค่าพารามิเตอร์ตามหัวข้อ 3.1.3 ตารางที่ 3.5

(3) จากจำลองในตามข้อ (2) จะได้ค่า $X_1, X_2, \ldots, X_n$

(4) นำค่า $X_1, X_2, \ldots, X_n$ มาจัดเรียงจากน้อยไปหามาก จากนั้นตัดปลายด้านขวาของข้อมูลตามจำนวนที่กำหนด จะได้ค่า $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$

(5) คำนวณค่าสถิติทดสอบ K-S^C AD^C และ CVM^C

(6) ทดสอบว่าปฏิเสธสมมติฐานว่างหรือไม่ และนับจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง

(7) ทำซ้ำข้อ (2) ถึงข้อ (6) จำนวน 1,000 รอบ

(8) คำนวณหาค่าอำนาจการทดสอบ คือ การนำจำนวนครั้งที่ปฏิเสธสมมติฐานว่างหารด้วยจำนวนรอบทั้งหมด (1,000 รอบ)

3.7 ขั้นตอนในการทำงานของโปรแกรม

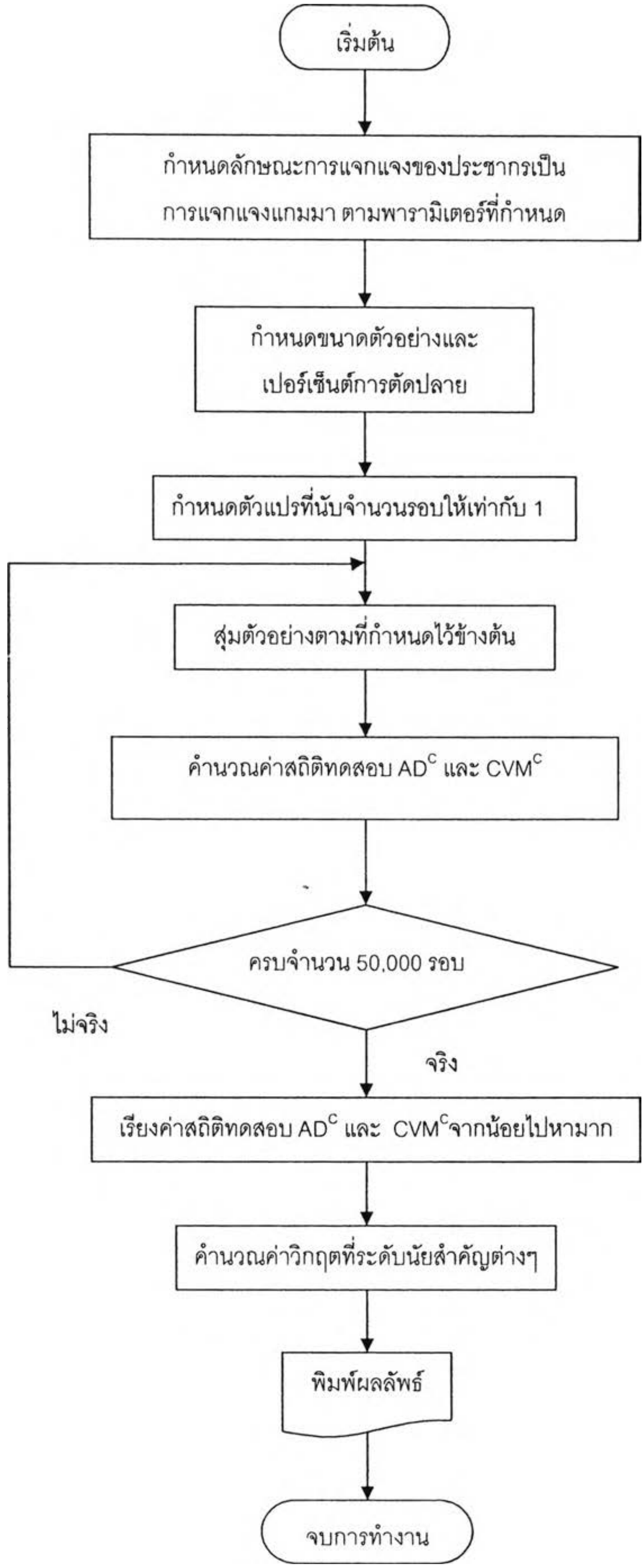
ขั้นตอนในการทำงานของโปรแกรมแบ่งออกเป็น 3 ส่วน ดังนี้

<u>ส่วนที่ 1</u> การคำนวณค่าวิกฤตของตัวสถิติทดสอบ AD^C และ CVM^C ได้นำเสนอไว้ในรูปที่ 3.1

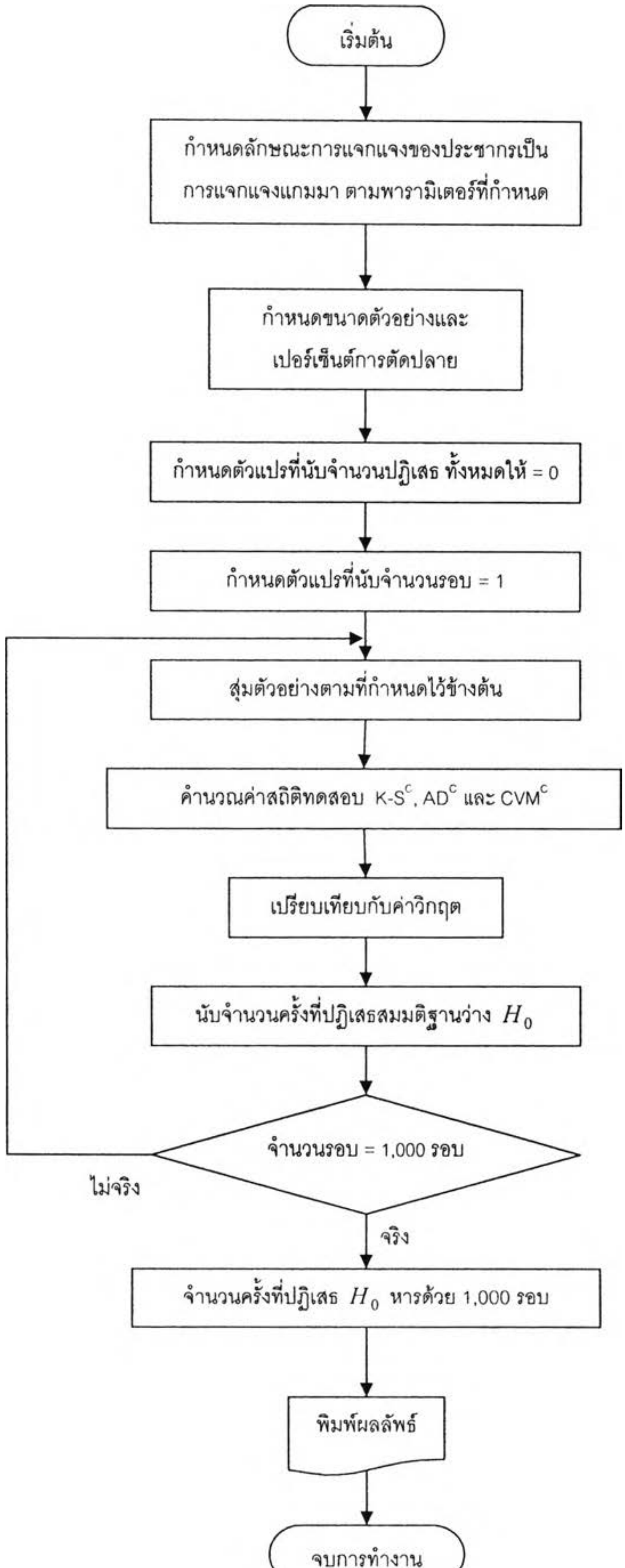
<u>ส่วนที่ 2</u> การคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้นำเสนอไว้ในรูปที่ 3.2

<u>ส่วนที่ 3</u> การคำนวณหาค่าอำนาจการทดสอบ ได้นำเสนอไว้ในรูปที่ 3.3

รูปที่ 3.1 แผนผังแสดงขั้นตอนการคำนวณหาค่าวิกฤตของสถิติทดสอบ AD^C และ CVM^C



รูปที่ 3.2 แผนผังแสดงขั้นตอนหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1



รูปที่ 3.3 แผนผังแสดงขั้นตอนหาค่าอำนาจการทดสอบ