หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้ทำการศึกษาถึงวิธีการแปลงข้อมูลให้มีการแจกแจงแบบปกติ ซึ่งข้อมูลที่นำมาทำการศึกษาส่วนหนึ่งเป็นข้อมูลที่ได้จากการจำลองข้อมูล อีกส่วนหนึ่งได้จากการเก็บรวบรวมข้อมูลที่เกิดขึ้นจริงในปัจจุบัน ซึ่งมีวิธีการดำเนินการวิจัยดังนี้

1. เตรียมข้อมูลเพื่อการศึกษา
   - เก็บรวบรวมข้อมูลที่เกิดขึ้นจริง
   - จำลองข้อมูล
2. หาค่าสถิติที่สนใจจากข้อมูลที่เตรียมไว้แล้วและทำการตรวจสอบข้อมูล
3. ทำการแปลงข้อมูล
4. ทดสอบการแจกแจงแบบปกติด้วยสถิติโคลโมโกรอฟ-สมินอฟ
5. สรุปผลการแปลงข้อมูล

1. เตรียมข้อมูลเพื่อการศึกษา

1.1 การเก็บรวบรวมข้อมูลที่เกิดขึ้นจริงในปัจจุบัน

ข้อมูลที่สนใจทำการศึกษาเป็นข้อมูลทุติยภูมิ คือ เป็นข้อมูลที่มีการเก็บรวบรวมไว้แล้ว แหล่งที่จะค้นหาได้คือ เอกสาร รายงานการวิจัย หรือหนังสือสถิติของหน่วยงานต่าง ๆ โดยข้อมูลที่เก็บรวบรวบมานั้นมีทั้งสิ้น 15 ชุดด้วยกัน รายละเอียดของแต่ละชุดข้อมูลแสดงไว้ในตารางที่ 3.1 หน้าที่

ตารางที่ 3.1 แสดงรายละเอียดของชุดข้อมูลที่ทำการศึกษาวิจัย

ชุดที่	เรื่อง	แหล่งที่มา
1	อุณหภูมิในบริเวณควนขี้เสียน เขตห้ามล่าสัตว์ป่าทะเลน้อยในเดือนตุลาคม พ.ศ. 2541	วิทยานิพนธ์เรื่อง " การศึกษาประชากรของนกน้ำและการประเมินความเหมาะสมของควนขี้เสียน เขตห้ามล่าสัตว์ป่าทะเลน้อย ในการจัดตั้งเป็นพื้นที่ชุ่มน้ำในทำเนียบ " ภาคสหสาขา วิชาวิทยาศาสตร์สภาวะแวดล้อม บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
2	จำนวนนกกาน้ำเล็กที่ควนขี้เสียนจังหวัดพัทลุงในเดือนเมษายน พ.ศ. 2541	
3	จำนวนนกกาน้ำเล็กที่ควนขี้เสียนจังหวัดพัทลุงในเดือนกรกฎาคม พ.ศ. 2541	
4	จำนวนนกกาน้ำเล็กที่ควนขี้เสียนจังหวัดพัทลุงในเดือนสิงหาคม พ.ศ. 2541	
5	จำนวนนกกระสาแดงที่ควนขี้เสียนจังหวัดพัทลุงในเดือนมิถุนายน พ.ศ. 2541	
6	จำนวนนกกระสาแดงที่ควนขี้เสียนจังหวัดพัทลุงในเดือนสิงหาคม พ.ศ. 2541	
7	จำนวนนกยางควายที่ควนขี้เสียนจังหวัดพัทลุงในเดือนกรกฎาคม พ.ศ. 2541	
8	ปริมาณปลาทรายแดงที่จับได้จากเครื่องประมงหลัก ระหว่างปี พ.ศ. 2514 ถึง 2535	รายงานการวิจัย เรื่อง " ทรัพยากรและการประมงทะเล ในเขตเศรษฐกิจจำเพาะประเทศไทย" กรมประมง
9	ปริมาณปลาตาโตที่จับได้จากเครื่องประมงหลัก ระหว่างปี พ.ศ. 2514 ถึง 2535	
10	ค่าเฉลี่ยระดับเสียงในช่วง 95 นาที ที่กรุงเทพมหานคร ปี พ.ศ. 2530	รายงานเรื่อง "การศึกษาปัญหามลพิษทางเสียง ปี พ.ศ. 2530" สำนักงานคณะกรรมการสิ่งแวดล้อม
11	อุณหภูมิของน้ำทะเลที่บริเวณน้ำขึ้นน้ำลง ชายฝั่งตะวันออกของอ่าวไทย ระหว่างปี พ.ศ.2530–2531	รายงานการวิจัยเรื่อง "การตรวจวัดระดับคุณภาพสิ่งแวดล้อมบริเวณเขตน้ำขึ้นน้ำลงของชายฝั่งทะเลด้านตะวันออกของอ่าวไทย" สถาบันวิจัยทรัพยากรทางน้ำ จุฬาลงกรณ์มหาวิทยาลัย

12	ปริมาณซิลิเกตที่ผิวของน้ำทะเล บริเวณชายฝั่งตะวันออกของอ่าวไทยในเดือนธันวาคม พ.ศ.2541	รายงานการวิจัยเรื่อง "การตรวจเฝ้าระวังมลพิษทางทะเล บริเวณชายฝั่งตะวันออกของอ่าวไทย ระยะที่ 5" สถาบันวิจัยทรัพยากรทางน้ำ จุฬาลงกรณ์มหาวิทยาลัย
13	ค่าความเค็มที่ระดับความลึก 1 เมตร บริเวณชายฝั่งตะวันออกของอ่าวไทยในเดือนมิถุนายน	
14	ค่า DO ที่ระดับความลึก 5 เมตร บริเวณชายฝั่งตะวันออกของอ่าวไทยในเดือนมิถุนายน	
15	จำนวนผู้ได้รับอนุมัติให้ก่อสร้างใหม่ในเขตเทศบาลในภาคตะวันออกเฉียงเหนือจำแนกเป็นรายจังหวัด	รายงานเรื่อง "การประมวลข้อมูลสถิติการก่อสร้าง" สำนักงานสถิติแห่งชาติ
16	ปริมาณฝนใช้การได้ของสถานีวัดน้ำฝน อำเภอเมือง จังหวัดยะลา ระหว่างปี ค.ศ. 1952 – 1991	โครงการศึกษาข้อมูลและศักยภาพการพัฒนาลุ่มน้ำปัตตานี กรมชลประทาน จังหวัดยะลา
17	อุณหภูมิเฉลี่ยจำแนกรายสถานีตรวจอากาศ ปี พ.ศ. 2540	สมุดสถิติรายปี บรรพ 45 สำนักงานสถิติแห่งชาติ

1.2 การจำลองข้อมูล

กระทำได้ด้วยการทำงานของเครื่องคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ซึ่งทำการจำลองข้อมูลให้มีการแจกแจงแบบเป้าตามที่ต้องการโดยอาศัยสมการการแจกแจงแลมดาร์ของตูกีร์ (Tukey ' s Lamda Distribution) และการจำลองข้อมูลจากการแจกแจงแบบไวบูลล์ ซึ่งการทำการจำลองข้อมูลจากการแจกแจงแบบไวบูลล์จะกำหนดที่ค่าความโด่งระดับต่ำกว่าที่การจำลองข้อมูลจากการแจกแจงแลมดาร์ของตูกีร์จะกระทำได้

1.2.1 การจำลองข้อมูลจากการแจกแจงแลมดาร์ของตูกีร์ ซึ่ง Ramberg และ Schmeiser ได้เสนอวิธีการสร้างตัวแปรสุ่มที่ขึ้นอยู่กับความเบ้และความโด่งเพื่อการจำลองข้อมูล โดยที่ตัวแปรสุ่มนั้นถูกกำหนดจากค่าพารามิเตอร์ 4 ค่าดังสมการ

$$X = R(p) = \lambda_1 + \frac{\left|p^{\lambda_3} - (1-p)^{\lambda_4}\right|}{\lambda_2} \quad ; \quad 0 < p < 1 \qquad (3.1)$$

เมื่อ

X คือ ตัวแปรสุ่มที่ทำการจำลองขึ้นมา

p คือ เลขสุ่มที่มีค่าระหว่าง 0 และ 1

λ_1 คือ พารามิเตอร์กำหนดตำแหน่ง (location parameter)

λ_2 คือ พารามิเตอร์มาตราส่วน (scale parameter)

λ_3 และ λ_4 คือ พารามิเตอร์สัณฐาน (shape parameter)

เราสามารถหาค่า $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ได้จากตารางของ Ramberg (แสดงในภาคผนวก)

เมื่อเรากำหนดค่าความเบ้และความโด่งต่าง ๆ โดยที่ λ_1 เป็นค่าเฉลี่ยมีค่าเท่ากับ 0 และ λ_2 เป็นค่าความแปรปรวนมีค่าเท่ากับ 1 แต่ถ้าต้องการค่าเฉลี่ยที่เป็นค่าใด ๆ μ และความแปรปรวนที่เป็นค่าใด ๆ σ^2 จะต้องแปลงค่า λ_1 และ λ_2 จากตารางเสียก่อน ซึ่งทำการแปลงโดยใช้สมการ

$$\lambda_1(\mu,\sigma^2) = \lambda_1(0,1)\sigma + \mu \qquad (3.2)$$

$$\lambda_2(\mu,\sigma^2) = \lambda_2(0,1) / \sigma \qquad (3.3)$$

สำหรับการจำลองข้อมูลในครั้งนี้กำหนดค่าเฉลี่ยเริ่มต้นเท่ากับ 100 และความแปรปรวนเริ่มต้นเท่ากับ 100 เพื่อหลีกเลี่ยงปัญหาค่าที่ติดลบ

ซึ่งการจำลองข้อมูลสามารถสรุปได้เป็นขั้นตอนดังนี้

1.2.1.1 กำหนดค่าพารามิเตอร์ของข้อมูลที่ต้องการ ประกอบไปด้วย

- ค่าเฉลี่ย
- ค่าความแปรปรวน
- ความเบ้
- ความโด่ง
- จำนวนข้อมูล

1.2.1.2 อ่านค่า $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ จากตาราง Ramberg ตามความเบ้และความโด่งที่กำหนด

1.2.1.3 แปลงค่า λ_1 และ λ_2 ตามสมการที่ 3.2 และ 3.3 เนื่องจากการจำลองข้อมูลในครั้งนี้ กำหนดให้ค่าเฉลี่ยเท่ากับ 100 และความแปรปรวนเท่ากับ 100

1.2.1.4 สร้างตัวเลขสุ่ม p ให้มีค่าอยู่ระหว่าง 0 ถึง 1

1.2.1.5 นำค่า $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ และ p ที่ได้ แทนในสมการที่ 3.1 จะได้ค่า X ซึ่งเป็นค่าข้อมูลตามที่ต้องการ

1.2.1.6 ทำซ้ำในข้อ 1.2.1.5 จนกว่าจะได้จำนวนข้อมูลครบตามต้องการ ซึ่งในการวิจัยครั้งนี้ กำหนดให้ขนาดตัวอย่างมี 5 ระดับด้วยกันคือ 20,30, 40, 50 และ 100

1.2.2 การจำลองข้อมูลจากการแจกแจงแบบไวบูลล์

การสร้างตัวแปรสุ่มให้มีการแจกแจงแบบไวบูลล์ อาศัยเทคนิคการแปลงผกผัน (Inverse Transformation) ซึ่งเป็นเทคนิคที่ใช้แปลงตัวแปรสุ่มที่มีลักษณะการกระจายของความน่าจะเป็นแบบสม่ำเสมอ (Uniform) ไปอยู่ในรูปของตัวแปรสุ่ม ที่มีลักษณะการกระจายของความน่าจะเป็นแบบอื่น ๆ สำหรับตัวแปรสุ่มที่มีลักษณะการกระจายความน่าจะเป็นแบบไวบูลล์มีขั้นตอนดังนี้

ขั้นที่ 1 กำหนดฟังก์ชันการแจกแจงสะสม

$$F(x) = 1 - exp[-(x/\beta)]^{\alpha}, \; x > 0 \qquad (3.4)$$

ขั้นที่ 2 ให้ $F(x) = 1 - exp[-(x/\beta)]^{\alpha} = p$ (3.5)

โดยที่ p คือตัวเลขสุ่มที่มีค่าอยู่ระหว่าง 0 และ 1

ขั้นที่ 3 หาค่าของ x ในเทอมของ p จะได้

$$x = \beta(-ln(p))^{\frac{1}{\alpha}} \qquad (3.6)$$

ซึ่งจากฟังก์ชันการแจกแจงแบบไวบูลล์ จะเห็นว่า α เป็นพารามิเตอร์สัณฐาน และ β เป็นพารามิเตอร์มาตราส่วน สำหรับงานวิจัยในครั้งนี้ได้กำหนดให้พารามิเตอร์ มีค่าแตกต่างออกไปตามความเบ้และความโด่งที่ต้องการ โดยการหาค่า α ที่เหมาะสมจากการคำนวณค่าความเบ้และความโด่งที่แต่ละค่าของ α โดยสามารถหาค่าความเบ้และความโด่งได้จากการสูตรโมเมนต์ที่ k นั่นคือ

$$m_k = E(x^k) = \beta^k \Gamma\left(1 + \frac{k}{\alpha}\right) \quad , \quad k > -\alpha$$

โดยที่

$$\mu_2 = m_2 - m_1^2$$

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$$

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

และสามารถหาค่าความเบ้และความโด่งได้จาก

$$\text{ความเบ้} = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{3/2}}$$

$$\text{ความโด่ง} = \frac{\mu_4}{\sigma^4} = \frac{\mu_4}{(\mu_2)^2}$$

เมื่อได้ค่า α ตามที่ต้องการ เราจะหาค่า β ได้จากการแก้สมการของความแปรปรวน ซึ่งจะขึ้นอยู่กับขนาดของความแปรปรวนที่ได้กำหนดไว้ในตอนต้นแล้ว โดยที่ความแปรปรวนอยู่ในรูป

$$V(x) = \frac{\beta^2}{\alpha}\left\{2\Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha}\left[\Gamma\left(\frac{1}{\alpha}\right)\right]^2\right\}$$

จากนั้นนำค่าพารามิเตอร์ทั้งหมดแทนค่าในสมการที่ 3.6 ทำการจำลองข้อมูลขึ้นมาตามขนาดตัวอย่างที่ต้องการ

2. หาค่าสถิติที่สนใจจากข้อมูลที่เตรียมไว้แล้วและทำการตรวจสอบข้อมูล

2.1 ค่าสถิติที่สนใจ

เมื่อได้ข้อมูลมาแล้วทั้งจากการเก็บรวบรวมและการจำลองขึ้นมา ทำการหาค่าสถิติจากข้อมูลที่ได้มานั้น โดยค่าสถิติที่สนใจคือ

2.1 ค่าความเบ้

2.2 ค่าความโด่ง

การหาค่าสถิติทั้งสองตัวนี้ จะทำให้สามารถจัดข้อมูลที่มีอยู่นั้นให้ตรงกับค่าความเบ้และค่าความโด่งที่กำหนดในการจำลองข้อมูล

2.2 ตรวจสอบข้อมูล

ในการตรวจสอบข้อมูลครั้งนี้ กระทำเพื่อให้ได้ข้อมูลที่เป็นไปตามความต้องการ ซึ่งข้อมูลที่ได้จากการจำลองข้อมูลกระทำ 2 ขั้นตอนด้วยกัน คือ

ขั้นตอนที่ 1 ทำการตรวจสอบที่ความเบ้และความโด่งของข้อมูล

ขั้นตอนที่ 2 ทำการตรวจสอบการแจกแจงปกติของข้อมูลด้วยสถิติโคลโมโกรอฟ-สมินอฟ

แต่ข้อมูลที่ได้จากการเก็บรวบรวมนั้น กระทำเพียงขั้นตอนที่ 2 เท่านั้น

2.2.1 การตรวจสอบที่ความเบ้และความโด่งของข้อมูล

ในการศึกษาวิจัยครั้งนี้ การจำลองข้อมูลนั้นถูกกำหนดด้วยค่าความเบ้และค่าความโด่งของข้อมูล ดังนั้น เมื่อทำการจำลองข้อมูลขึ้นมา จึงต้องทำการตรวจสอบว่ากลุ่มของข้อมูลที่ได้มานั้นมีค่าความเบ้และค่าความโด่งตามที่กำหนดหรือไม่ ซึ่งค่าความเบ้และค่าความโด่งที่กำหนดในการศึกษาวิจัยครั้งนี้ แสดงไว้แล้วตัวตารางที่ 3.5 หน้าที่ในหัวข้อ 1.2.2 การจำลองข้อมูล สูตรที่ใช้ในการหาค่าความเบ้ คือ

$$\text{ความเบ้} = \frac{\sqrt{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{3/2}}$$

นำค่าความเบ้ที่ได้มานั้นทำการเปรียบเทียบว่าเท่ากับค่าที่กำหนดหรือไม่ ถ้าไม่เท่ากันให้กระทำการจำลองข้อมูลชุดใหม่ขึ้นมาแทน และทำการเปรียบเทียบความเบ้อีกครั้งหนึ่ง กระทำจนกว่าจะได้ความเบ้ตามต้องการ

เมื่อได้ค่าความเบ้ตามที่ต้องการแล้ว ทำการหาค่าความโด่งจากสูตร

$$\text{ความโด่ง} = \frac{n\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^2}$$

ซึ่งค่าที่ได้มานั้น จะทำการเปรียบเทียบกับความโด่งที่กำหนดว่าเท่ากันหรือไม่ ถ้าไม่เท่ากันให้กระทำการจำลองข้อมูลชุดใหม่ขึ้นมาแทน และต้องทำการเปรียบเทียบค่าความเบ้อีกครั้งหนึ่ง การทำจนกว่าจะได้ค่าความโด่งตามต้องการ

จากการกระทำข้างต้น จะได้ข้อมูลที่มีความเบ้และความโด่งตามที่ต้องการ

2.2.2 การตรวจสอบการแจกแจงปกติของข้อมูลด้วยสถิติโคลโมโกรอฟ-สมินอฟ

การศึกษาวิจัยครั้งนี้ต้องการทดสอบการแปลงข้อมูลให้มีการแจกแจงปกติ ดังนั้น ข้อมูลเริ่มต้นที่จะนำมาทำการศึกษาต้องแน่ใจเสียก่อนว่าเป็นข้อมูลที่มีการแจกแจงไม่ปกติ ก่อนทำการแปลงข้อมูล เพื่อให้ได้ผลที่ถูกต้อง

การทดสอบเทียบความกลมกลืนที่เลือกใช้ในการศึกษาครั้งนี้คือ สถิติทดสอบโคลโมโกรอฟ-สมินอฟ กระทำได้โดยหาค่าความแตกต่างที่มากที่สุดระหว่างฟังก์ชันการแจกแจงสะสมของตัวอย่าง และฟังก์ชันการแจกแจงสะสมของการแจกแจงปกติ ซึ่งสามารถหาค่าสถิติได้จากสูตร

$$D = \max(D^+, D^-)$$

เมื่อ

$$D^+ = \max\left[{}^{i}/_{m} , F(x_i) \right]$$

$$D^- = \max\left[F(x_i) , {}^{(i-1)}/_{m} \right]$$

(ดังได้กล่าวไว้แล้วในบทที่ 2 หัวข้อ สถิติที่ใช้ในการทดสอบการแจกแจง หน้าที่ 16)

การทดสอบของสถิติโคลโมโกรอฟ-สมินอฟนั้น สามารถทำการทดสอบได้หลายการแจกแจง ซึ่งจะมีการปรับค่าสถิติ D ตามรูปแบบของการแจกแจงที่ทำการทดสอบนั้น โดยการตรวจสอบของการศึกษาวิจัยในขั้นตอนนี้มีสมมติฐานในการทดสอบคือ

H_0 : ข้อมูลมีการแจกแจงปกติ

H_1 : ข้อมูลมีการแจกแจงไม่ปกติ

และจะทำการปรับค่าสถิติ D ที่คำนวณได้ด้วยสูตร

$$Adjusted\ D = \left(\sqrt{n} - 0.01 + \left(0.85 / \sqrt{n}\right)\right) * D_n$$

โดยที่

n คือ จำนวนข้อมูลที่นำมาทำการตรวจสอบ

D_n คือ ค่าสถิติ D ที่คำนวณได้ในตอนแรก

นำค่า *Adjusted* D ไปเปรียบเทียบกับค่าวิกฤตตามระดับนัยสำคัญ (α) ที่กำหนด คือ

ระดับนัยสำคัญ (α)	ค่าวิกฤต D
0.01	1.035
0.05	0.895
0.10	0.819
0.15	0.775

จุดประสงค์ของการทดสอบในขั้นตอนนี้คือ ต้องการข้อมูลที่มีการแจกแจงไม่ปกติ จึงต้องการผลของการทดสอบเป็นปฏิเสธ H_0 นั่นคือ ค่า *Adjusted* D จะต้องมีค่ามากกว่าค่าวิกฤต D ที่ระดับนัยสำคัญ (α) ที่กำหนด

3. การแปลงข้อมูล

เมื่อแน่ใจว่าข้อมูล (x) ที่ได้มานั้นมีการแจกแจงไม่ปกติ (ผ่านขั้นตอนข้างต้นแล้ว) ทำการแปลงข้อมูล (x) ด้วยวิธีการต่าง ๆ ที่กำหนด คือ

3.1 วิธีการแปลงโดยใช้ล็อก (The Logarithm Transformation)

ได้แก่รูปแบบ $log(x)$, $ln(x)$

3.2 วิธีการแปลงโดยใช้รากที่สอง (The Square Root Transformation)

ได้แก่รูปแบบ $\sqrt{x}$

3.3 วิธีการแปลงโดยใช้การกลับเศษส่วน (The Recipocal Transformation)

ได้แก่รูปแบบ $1/x$

3.4 วิธีการแปลงโดยใช้เลขยกกำลัง (Power Transformations)

ได้แก่รูปแบบ x^p และ $\frac{x^p - 1}{p}$

เมื่อ p คือเลขจำนวนจริงใด ๆ

โดยที่วิธีการที่ 3.1 ถึงวิธีการที่ 3.3 สามารถกระทำได้เลยเมื่อมีข้อมูล แต่ในวิธีการที่ 3.4 จะต้องใช้เลขยกกำลัง (p) ที่มีค่าแตกต่างกันออกไป เพื่อให้ได้เลขยกกำลังที่สามารถแปลงข้อมูลให้มีการแจกแจงแบบปกติได้มากที่สุด จากการทดลองหาค่าเลขกำลัง (p) พบว่า

สำหรับการแจกแจงแบบเบ้ขวา : ค่า p ที่สามารถทำการแปลงข้อมูลได้จะมีค่าน้อยกว่า 1 และมีค่าลดลงจนถึง -4 ซึ่งเมื่อค่า p มีค่าน้อยกว่า –4 จะทำให้ข้อมูลนั้นมีค่าที่ใกล้เคียงกันมากจนทำให้การกระจายของข้อมูลมีค่าน้อยมาก และเป็นผลให้ค่าส่วนเบี่ยงเบนมาตรฐานมีค่าเข้าใกล้ 0 ไม่สามารถทำการทดสอบได้ด้วยสถิติโคลโมโกรอฟ-สมินอฟ ดังนั้นจึงกำหนดให้ค่า p มีค่าลดลงจาก 1 จนถึง –4 ด้วยอัตราลดลงทีละ 0.05

สำหรับการแจกแจงแบบเบ้ซ้าย : ค่า p ที่สามารถทำการแปลงข้อมูลได้จะมีค่ามากกว่า 1 และมีค่าเพิ่มจนถึง 6 ซึ่งเมื่อค่า มีค่ามากกว่า 6 จะทำให้ข้อมูลมีค่าใหญ่มาก เมื่อทำการหาค่าความโด่งและค่าความเบ้จะไม่สามารถกระทำได้ ดังนั้นจึงกำหนดให้ค่า p มีค่าเพิ่มขึ้นจาก 1 จนถึง 6 ด้วยอัตราเพิ่มขึ้นทีละ 0.05

4. ทดสอบการแจกแจงแบบปกติด้วยสถิติโคลโมโกรอฟ-สมินอฟ

เมื่อได้ข้อมูลที่ทำการแปลงด้วยวิธีการต่าง ๆ ที่กำหนดแล้ว ทำการทดสอบข้อมูลนั้นด้วยสถิติโคลโมโกรอฟ-สมินอฟอีกครั้ง โดยวิธีการทดสอบได้แสดงไว้แล้วในหัวข้อที่ 2.2.2 หน้าที่ แต่การทดสอบในขั้นตอนนี้กระทำเพื่อตรวจสอบว่าข้อมูลที่มีการแจกแจงไม่ปกติเมื่อได้รับการแปลงแล้ว การแจกแจงของข้อมูลนั้น ๆ เป็นการแจกแจงปกติหรือไม่ เนื่องจากจุดประสงค์ของการแปลงข้อมูล คือ เพื่อต้องการให้ข้อมูลที่ได้รับการแปลงแล้วมีการแจกแจงปกติ

การตรวจสอบด้วยสถิติโคลโมโกรอฟ-สมินอฟในขั้นตอนนี้ ถึงแม้ว่าจุดประสงค์ในการทดสอบแตกต่างกัน แต่เมื่อทำการทดสอบในหัวข้อที่ 2.2.2 ที่ระดับนัยสำคัญ (α) เท่าใด การทดสอบในขั้นตอนนี้ต้องทำการทดสอบที่ระดับนัยสำคัญ (α) เท่าเดิม ทั้งนี้เพราะการตรวจสอบว่าข้อมูลมีการแจกแจงปกติที่ระดับนัยสำคัญไม่เท่ากันทั้งสองขั้นตอน อาจทำให้เกิดความผิดพลาดได้ เนื่องจากถ้าข้อมูลปฏิเสธในขั้นตอนที่ 2.2.2 แล้ว แต่มีระดับนัยสำคัญ (α) น้อยกว่าระดับนัยสำคัญ (α) ที่ทำการทดสอบในขั้นตอนนี้ ก็อาจจะทำให้ยอมรับ H_0 ได้โดยไม่ต้องทำการแปลงข้อมูล

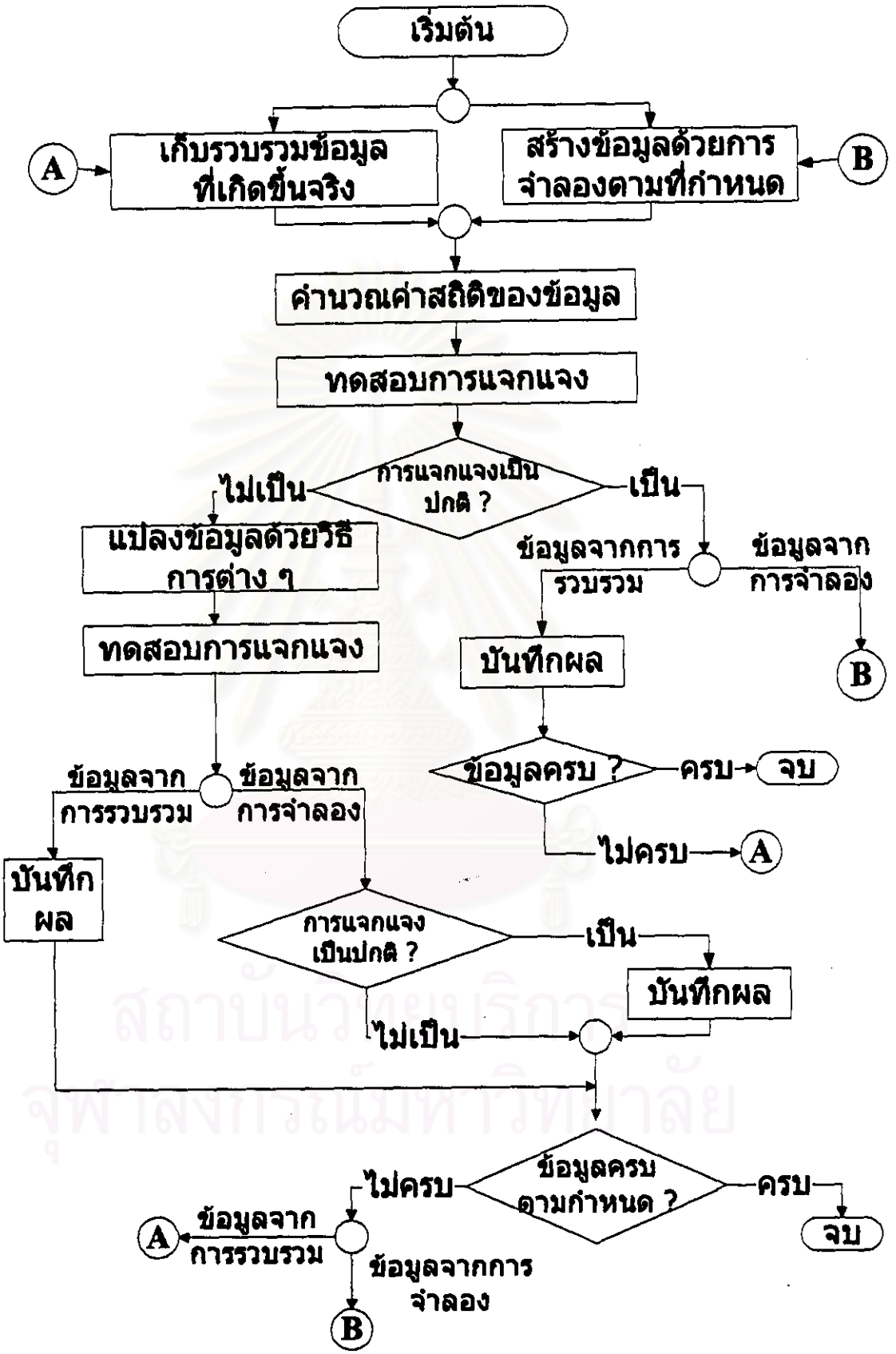
สำหรับข้อมูลที่ได้จากการจำลองข้อมูลจะกระทำซ้ำตั้งแต่ขั้นตอนที่ 1 จนกว่าจะได้ครบตามจำนวนรอบที่ต้องการ นั่นคือ แต่ละระดับนัยสำคัญ (α) ขนาดตัวอย่าง ความเบ้ และความโด่ง จะกระทำซ้ำเป็นจำนวน 200 รอบในแต่ละกลุ่ม ซึ่งจะทำการนับจำนวนชุดข้อมูลที่ได้ทำการตรวจสอบการแจกแจงในขั้นตอนที่ 4 แล้วได้ยอมรับว่าข้อมูลมีการแจกแจงปกติภายหลังจากการทำการแปลงข้อมูลด้วยวิธีการต่าง ๆ นำค่าที่นับได้นั้นมาหาเปอร์เซ็นต์ของการยอมรับ H_0: ข้อมูลมีการแจกแจงปกติ

5. สรุปผลการแปลงข้อมูล

การสรุปผลจะดูจากเปอร์เซ็นต์ของการยอมรับ H_0 : ข้อมูลมีการแจกแจงปกติ ที่ได้ในขั้นตอนที่ 4 มาเปรียบเทียบกันในแต่ละวิธีของการแปลงข้อมูล สำหรับการจำลองข้อมูลนั้นจะพิจารณาจากเปอร์เซ็นที่มากที่สุดสำหรับค่าความเบ้ ความโด่ง จำนวนข้อมูล และระดับนัยสำคัญ (α) นั้น ๆ และเสนอผลที่ได้ในรูปของตาราง แต่สำหรับข้อมูลที่ได้จากการเก็บรวบรวมนั้นไม่สามารถหาเปอร์เซ็นที่มากที่สุดได้ เนื่องจากแต่ละชุดตัวอย่างนั้นเมื่อทำการแปลงแล้วสามารถบอกได้เพียงยอมรับ H_0 หรือปฏิเสธ H_0 เท่านั้น ดังนั้นจึงพิจารณาผลของการแปลงข้อมูลในแต่ละชุดของข้อมูลและนำเสนอผลในรูปของตาราง

วิธีการดำเนินการวิจัยทั้งหมดที่ได้กล่าวมาแล้วข้างต้น สามารถแสดงได้ดังรูปที่ 3.1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.1 แสดงขั้นตอนการดำเนินงาน