

บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้มีวิธีดำเนินการวิจัยโดยแบ่งเป็นขั้นตอนดังนี้

- ขั้นตอนที่ 1 ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องกับเนื้อหาที่ต้องการศึกษา
 - ขั้นตอนที่ 2 เสนอตัวประมาณค่าเฉลี่ยประชากร
 - ขั้นตอนที่ 3 หารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าที่นำเสนอ
 - ขั้นตอนที่ 4 เปรียบเทียบประสิทธิภาพของตัวประมาณค่าที่นำเสนอกับตัวประมาณค่าอื่น ๆ
 - ขั้นตอนที่ 5 แสดงตัวอย่างการคำนวณเพื่อสนับสนุนผลลัพธ์ในเชิงทฤษฎีที่ได้
- สำหรับรายละเอียดของแต่ละขั้นตอนเป็นดังนี้

3.1 ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับข้อมูลสุ่มหา รูปแบบของข้อมูลสุ่มหา ขนาดของข้อมูลสุ่มหา เทคนิคใช้ในการประมาณค่าข้อมูลสุ่มหา การประมาณค่าเฉลี่ยประชากร การจำลองสถานการณ์ การประมาณค่าในการสำรวจตัวอย่างเชิงสุ่ม และศึกษาการประมาณค่าแบบอื่น ๆ เพิ่มเติม

3.2 เสนอตัวประมาณค่าเฉลี่ยประชากรเมื่อมีข้อมูลสุ่มหา

จากการศึกษางานวิจัยซึ่งกล่าวไว้ในบทที่ 2 ผู้วิจัยจึงมีความสนใจที่จะพัฒนาตัวประมาณค่าเฉลี่ยประชากร โดยตัวประมาณค่าที่ผู้วิจัยเสนอเป็นตัวประมาณค่าที่ประยุกต์มาจากแนวความคิดของ Singh et al. (2010) ซึ่งอยู่ในรูป

$$T_{new} = \lambda T_{d1} \quad (3.1)$$

เมื่อ T_{d1} คือ ตัวประมาณค่าเฉลี่ยประชากรของ Singh et al. (2010)

และ λ คือ ค่าคงตัวที่ทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของ T_{new} มีค่าต่ำสุด

โดย λ คำนวณจากการหาอนุพันธ์ (Differential) ของค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่เหมาะสมของตัวประมาณค่า T_{d1} เทียบกับค่า λ และให้เท่ากับศูนย์

3.3 การหาค่าที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าที่น่าเสนอ

ในการหาค่าที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าที่น่าเสนอผู้วิจัยหาค่าที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าที่น่าเสนอ T_{new}

3.4 เปรียบเทียบประสิทธิภาพของตัวประมาณค่าที่น่าเสนอกับตัวประมาณค่าอื่น ๆ

เปรียบเทียบประสิทธิภาพของตัวประมาณค่าที่น่าเสนอ และตัวประมาณค่า T_{d1} T_{d2} ตัวประมาณค่าเมื่อประมาณค่าข้อมูลสูญหายด้วยวิธีการประมาณข้อมูลสูญหายด้วยค่าเฉลี่ย วิธีการประมาณข้อมูลสูญหายด้วยค่าอัตราส่วน วิธีการประมาณข้อมูลสูญหายด้วยค่าคอมโพรไมซ์และวิธีการประมาณข้อมูลสูญหายด้วยค่าถดถอยของค่าเข้าใกล้สุดแบบถ่วงน้ำหนักโดยใช้ค่าที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (ซึ่งกล่าวว่าตัวประมาณค่า $\hat{\theta}_1$ เป็นตัวประมาณค่าที่มีประสิทธิภาพสูงกว่า $\hat{\theta}_2$ ก็ต่อเมื่อ $RMSE(\hat{\theta}_1) < RMSE(\hat{\theta}_2)$)

3.5 ตัวอย่างการคำนวณเชิงตัวเลข

ในหัวข้อนี้เป็นการศึกษาหรือทดสอบว่าผลลัพธ์เชิงตัวเลขมีความสอดคล้องกับผลลัพธ์ในเชิงทฤษฎีหรือไม่ ดังนั้นผู้วิจัยจึงนำเสนอตัวอย่างเชิงตัวเลขเพื่อแสดงให้เห็นว่าผลลัพธ์เชิงตัวเลขสนับสนุนผลการศึกษาในเชิงทฤษฎีที่ได้ในหัวข้อที่ผ่านมา ด้วยการสร้างแบบจำลองสถานการณ์ด้วยโปรแกรม R ซึ่งดำเนินการดังนี้

1. จำลองข้อมูลประชากรขนาด N เท่ากับ 50,000 ซึ่งประกอบด้วยตัวแปรช่วย X ให้มีการแจกแจงเอกรูป (Uniform)
2. สร้างตัวแปรที่สนใจ Y ที่มีความสัมพันธ์กับตัวแปรช่วย X ที่สร้างขึ้น ณ ระดับค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) เท่ากับ 0.5 0.75 และ 0.9 (Chaimongkol & Suwatee, 2005)
3. ทดสอบระดับค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรที่สนใจ Y และตัวแปรช่วย X ว่ามีระดับความสัมพันธ์ตามที่ระบุไว้หรือไม่
4. ชักตัวอย่างขนาด n เท่ากับ 0.2% 0.6% 1% และ 2% ของขนาดประชากร N ด้วยวิธีการชักตัวอย่างแบบสุ่มเชิงเดียวชนิดไม่คืนที่ (Simple random sampling without replacement หรือ $SRSWOR$)
5. สุ่มตำแหน่งการสูญหายของข้อมูลให้กับตัวแปรที่สนใจ Y โดยกำหนดขนาดของข้อมูลสูญหาย m เท่ากับ 5% 10% 15% และ 20% ของขนาดตัวอย่าง n (Little & Rubin, 1987)

6. หาค่า d ที่เหมาะสม จากสมการ

$$-\rho_{yx} \frac{C_y}{C_x} = \frac{\frac{n}{N}[(d-1)(d-4)] - [(d-2)(d-3)(d-4)]}{(d-1)(d-4) + \frac{n}{N}[(d-1)(d-4)] + [(d-2)(d-3)(d-4)]} \quad (3.2)$$

ด้วยโปรแกรม maple แล้วนำค่า d ที่ได้ทั้ง 3 ค่า แทนลงในค่า $A=(d-1)(d-2)$,
 $B=(d-1)(d-4)$ และ $C=(d-2)(d-3)(d-4)$

7. นำค่า A B และ C ที่ได้แทนค่าลงในตัวประมาณค่า T_{d1} T_{d2} และตัวประมาณค่า T_{new}
 ประมาณค่าข้อมูลสูญหายด้วยตัวประมาณค่า T_{d1} T_{d2} และตัวประมาณค่า T_{new}

8. เปรียบเทียบประสิทธิภาพของรอกทั้งสองของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัว
 ประมาณค่าที่นำเสนอ ตัวประมาณค่า T_{d1} และ T_{d2} ตัวประมาณค่าเมื่อประมาณค่าข้อมูลสูญหาย
 ด้วยวิธีการประมาณข้อมูลสูญหายด้วยค่าเฉลี่ย การประมาณข้อมูลสูญหายด้วยค่าอัตราส่วน การ
 ประมาณข้อมูลสูญหายด้วยค่าคอมโพรไมซ์และวิธีการประมาณข้อมูลสูญหายด้วยค่าถดถอยของ
 ค่าเข้าใกล้สุดแบบถ่วงน้ำหนัก