บทที่ 3
วิธีดำเนินการศึกษา

การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อให้เกิดความรู้ความเข้าใจถึงกระบวนการพัฒนาทางสถิติที่เกี่ยวข้อง เพื่อเป็นแนวทางในการสร้างช่วงความเชื่อมั่นให้มีความถูกต้องเหมาะสม และมีความน่าเชื่อถือภายใต้ทฤษฎีการสำรวจตัวอย่าง และทฤษฎีการอนุมานทางสถิติ การศึกษาแบ่งออกเป็น 3 ส่วนหลักคือ การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่ายกรณีที่ขนาดประชากรเท่ากับ 10 ถึง300 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่ายกรณีที่ประชากรมีขนาดใหญ่ และการตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยภายใต้แผนการสุ่มแบบมีชั้นภูมิอย่างง่ายโดยแต่ละส่วนมีรายละเอียดดังนี้

3.1 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่าย กรณีที่ขนาดประชากรเท่ากับ 10 ถึง 300

ในส่วนนี้จะทำการศึกษาการตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่ายเมื่อขนาดประชากรเท่ากับ 10 ถึง 300 ซึ่ งจะเริ่มจากการจำลองหน่วยประชากรด้วยโปรแกรม Minitab ประชากรที่ได้จึงเป็นประชากรอันตะที่นับจำนวนหน่วยประชากรได้อย่างครบถ้วนแล้วจึงและทำการสุ่มตัวอย่างภายใต้แนวทางของการสำรวจตัวอย่างสำหรับประชากรอันตะเพื่อให้ได้ชุดตัวอย่างทั้งหมดที่เป็นไปได้ โดยใช้โปรแกรม VB (Visual Basic) โดยมีรายละเอียดการศึกษาดังนี้

3.1.1 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยประชากร

การศึกษาเริ่มจากการจำลองข้อมูลโดยใช้โปรแกรม Minitab เพื่อให้ได้ข้อมูลของหน่วยประชากร โดยมีขอบเขตของการศึกษาตามที่กำหนดต่อไปนี้

- จำลองข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยประชากรเท่ากับ 0 ความแปรปรวนประชากรเท่ากับ 1 โดยกำหนดขนาดประชากรมีค่าตั้งแต่ 10 ถึง 300
- จำลองข้อมูลที่มีการแจกแจงแบบปัวซง ที่มีค่าเฉลี่ยประชากรเท่ากับ 10 โดยกำหนดขนาด

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ประชากรมีค่าตั้งแต่ 10 ถึง 300

- จำลองข้อมูลที่มีการแจกแจงแบบเอกซ์โพเนนเชียลที่มีค่าเฉลี่ยประชากรเท่ากับ 10 โดยกำหนดขนาดประชากรมีค่าตั้งแต่ 10 ถึง 300

เมื่อได้ข้อมูลหน่วยประชากรทั้งหมดที่มีขนาดประชากร (N) ตามสถานการณ์ที่สนใจแล้วทำการสุ่มตัวอย่างซึ่งมีลักษณะการสุ่มตัวอย่างเป็นแบบไม่ใส่คืน จากนั้นทำการประมาณช่วงความเชื่อมั่นของค่าเฉลี่ยโดยใช้วิธีต่างๆ ดังนี้ (Steven, 1992)

วิธีที่ 1 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวอย่าง

$$\overline{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$$

วิธีที่2 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวประมาณ

$$\overline{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(1-\frac{n}{N}\right)\frac{s^2}{n}}$$

จากนั้นทำการวัดความคลาดเคลื่อนของการประมาณ โดยพิจารณาจากค่าครอบคลุมความน่าจะเป็น เทียบกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด โดยค่าความน่าจะเป็นของการครอบคลุมที่ได้จากตัวประมาณแบบช่วงทั้ง 2 วิธีสามารถทำได้โดยการนับจำนวนช่วงความเชื่อมั่นที่ค่าเฉลี่ยประชากร $\left(\overline{Y}\right)$ ตกอยู่ในช่วงที่สร้างขึ้นแล้วนำมาหารด้วยจำนวนชุดตัวอย่างทั้งหมดที่เป็นไปได้ซึ่งมีค่าเท่ากับ ${}^{N}C_{n}$ โดยการศึกษาครั้งนี้กำหนดระดับความเชื่อมั่น เท่ากับ 0.95

3.1.2 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับสัดส่วนประชากร

การศึกษาเริ่มจากการจำลองข้อมูลโดยใช้โปรแกรม Minitab เพื่อให้ได้ข้อมูลของหน่วยประชากรมีการแจกแจงแบบแบร์นูลลี ที่มีสัดส่วนประชากรเท่ากับ 0.5 โดยกำหนดขนาดประชากรมีค่าตั้งแต่ 10 ถึง 300 และทำการสุ่มตัวอย่างซึ่งมีลักษณะการสุ่มตัวอย่างเป็นแบบไม่ใส่คืน จากนั้นทำการประมาณช่วงความเชื่อมั่นของค่าเฉลี่ยโดยใช้วิธีต่างๆ ดังนี้ (Steven, 1992)

วิธีที่ 1 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวอย่าง

$$p \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$$

วิธีที่2 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวประมาณ

$$p \pm Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{N-n}{N-1}\right)\frac{p(1-p)}{n}}$$

จากนั้นทำการวัดความคลาดเคลื่อนของการประมาณ โดยพิจารณาจากค่าครอบคลุมความน่าจะเป็น เทียบกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด โดยค่าความน่าจะเป็นของการครอบคลุมที่ได้จากตัวประมาณแบบช่วงทั้ง 2 วิธีสามารถทำได้โดยทำการนับจำนวนช่วงความ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

เชื่อมั่นที่สัดส่วนประชากร (P) ตกอยู่ในช่วงที่สร้างขึ้นแล้วนำมาหารด้วยจำนวนชุดตัวอย่างทั้งหมดที่เป็นไปได้ซึ่งมีค่าเท่ากับ ${}^{N}C_{n}$ โดยการศึกษาครั้งนี้กำหนดระดับความเชื่อมั่นเท่ากับ 0.95

3.2 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่ายกรณีที่ประชากรมีขนาดใหญ่

ในส่วนนี้จะทำการศึกษาการตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยและสัดส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่ายเมื่อประชากรมีขนาดใหญ่ ซึ่งจะเริ่มจากการจำลองหน่วยประชากรและทำการสุ่มตัวอย่างจำนวน 1,000 รอบ ด้วยโปรแกรม Minitab จากนั้นนำข้อมูลมาทำการศึกษาตามขั้นตอนที่กำหนดไว้ โดยมีรายละเอียดดังนี้

3.2.1 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยประชากร

การศึกษาเริ่มจากการจำลองข้อมูลโดยใช้โปรแกรม Minitab เพื่อให้ได้ข้อมูลของหน่วยประชากร โดยมีขอบเขตของการศึกษาตามที่กำหนดต่อไปนี้

- จำลองข้อมูลที่มีการแจกแจงแบบปกติโดยมีค่าเฉลี่ยประชากรเท่ากับ 0 ความแปรปรวนประชากรเท่ากับ 1 และขนาดประชากรที่ใช้ศึกษากำหนดให้มีขนาดเท่ากับ 100, 500, 1,000, 2,000 และ 3,000
- จำลองข้อมูลที่มีการแจกแจงแบบปัวซง โดยมีค่าเฉลี่ยประชากรเท่ากับ 10 และขนาดประชากรที่ใช้ศึกษากำหนดให้มีขนาดเท่ากับ 100, 500, 1,000, 2,000 และ 3,000
- จำลองข้อมูลที่มีการแจกแจงแบบเอกซ์โพเนนเชียล โดยมีค่าเฉลี่ยประชากรเท่ากับ 10 และขนาดประชากรที่ใช้ศึกษากำหนดให้มีขนาดเท่ากับ 100, 500, 1,000, 2,000 และ 3,000

เมื่อได้ข้อมูลหน่วยประชากรทั้งหมดที่มีขนาดประชากร ตามสถานการณ์ที่สนใจแล้ว จากนั้นทำการสุ่มตัวอย่าง โดยกำหนดขนาดตัวอย่างเป็นร้อยละ 1, 3, 5 และ 10 ของประชากร ลักษณะการสุ่มตัวอย่างเป็นแบบใส่คืนและไม่ใส่คืน และทำการประมาณค่าเฉลี่ยของประชากรและทำการตรวจสอบข้อมูลเบื้องต้นจาก 3 วิธีคือ Stem-and Leaf, Normal Q-Q plot และ Descriptive Statistics จากนั้นทำการทดสอบคุณสมบัติการแจกแจงแบบปกติของค่าเฉลี่ยว่ามีการแจกแจงแบบปกติหรือไม่ โดยใช้สถิติทดสอบของ Kolmogorov-Smirnov ซึ่งมีสมมติฐานการทดสอบคือ

H_0 : ค่าเฉลี่ยตัวอย่างมีรูปแบบการแจกแจงแบบปกติ

H_1 : ค่าเฉลี่ยตัวอย่างไม่มีรูปแบบการแจกแจงแบบปกติ

โดยกำหนดระดับนัยสำคัญในการทดสอบเท่ากับ 0.05 จากนั้นทำการประมาณช่วงความเชื่อมั่นของค่าเฉลี่ยโดยใช้วิธีต่างๆ ดังนี้ (Steven, 1992)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

วิธีที่ 1 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวอย่าง

กรณีสุ่มแบบใส่คืนและไม่ใส่คืน $\bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$

วิธีที่2 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวประมาณ

กรณีสุ่มแบบใส่คืน $\bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(1-\frac{1}{N}\right)\frac{s^2}{n}}$

กรณีสุ่มแบบไม่ใส่คืน $\bar{X} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(1-\frac{n}{N}\right)\frac{s^2}{n}}$

วิธีที่3 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติเมื่อ n มีขนาดใหญ่ของ M.E. Thompson

กรณีสุ่มแบบใส่คืน $\bar{X} \pm \frac{1}{2n} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(1-\frac{1}{N}\right)\frac{s^2}{n}}$

กรณีสุ่มแบบไม่ใส่คืน $\bar{X} \pm \frac{1}{2n} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(1-\frac{n}{N}\right)\frac{s^2}{n}}$

จากนั้นทำการวัดความคลาดเคลื่อนของการประมาณ โดยพิจารณาจากค่าครอบคลุมความน่าจะเป็น เทียบกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด โดยค่าความน่าจะเป็นของการครอบคลุมที่ได้จากตัวประมาณแบบช่วงทั้ง 3 วิธีสามารถทำได้โดยการนับจำนวนช่วงความเชื่อมั่นที่ค่าเฉลี่ยประชากร $(\bar{Y})$ ตกอยู่ในช่วงที่สร้างขึ้นแล้วนำมาหารด้วยจำนวนรอบของการทำซ้ำนั่นคือ 1,000 รอบ โดยการศึกษาครั้งนี้กำหนดระดับความเชื่อมั่น เท่ากับ 0.95

3.2.2 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับสัดส่วนประชากร

การศึกษาเริ่มจากการจำลองข้อมูลโดยใช้โปรแกรม Minitab เพื่อให้ได้ข้อมูลของหน่วยประชากร ที่มีการแจกแจงแบบแบร์นูลลี โดยกำหนดค่าสัดส่วนประชากรมีค่าตั้งแต่ 0.05 ถึง 0.95 โดยทำการเพิ่มค่าทีละ 0.05 และขนาดประชากรที่ใช้ศึกษา กำหนดให้มีขนาดเท่ากับ 100, 500, 1,000, 2,000 และ 3,000

จากนั้นทำการสุ่มตัวอย่างโดยกำหนดขนาดตัวอย่างเป็นร้อยละ 1, 3, 5 และ 10 ของประชากร ลักษณะการสุ่มตัวอย่างเป็นแบบใส่คืนและไม่ใส่คืน และทำการประมาณสัดส่วนของประชากรและทำการตรวจสอบข้อมูลเบื้องต้นจาก 3 วิธีคือ Stem-and Leaf, Normal Q-Q plot และ Descriptive Statistics จากนั้นทำการทดสอบคุณสมบัติการแจกแจงแบบปกติของสัดส่วน ว่ามีการแจกแจงแบบปกติหรือไม่ โดยใช้สถิติทดสอบของ Kolmogorov-Smirnov ซึ่งมีสมมติฐานการทดสอบคือ

H_0 : สัดส่วนตัวอย่างมีรูปแบบการแจกแจงแบบปกติ

H_1 : สัดส่วนตัวอย่างไม่มีรูปแบบการแจกแจงแบบปกติ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

โดยกำหนดระดับนัยสำคัญในการทดสอบเท่ากับ 0.05 จากนั้นทำการประมาณช่วงความเชื่อมั่นของสัดส่วนโดยใช้วิธีต่างๆ ดังนี้ (Steven, 1992)

วิธีที่ 1 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวอย่าง

กรณีสุ่มแบบใส่คืนและไม่ใส่คืน $p \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$

วิธีที่2 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติโดยใช้ความแปรปรวนของตัวประมาณ

กรณีสุ่มแบบใส่คืน $p \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$

กรณีสุ่มแบบไม่ใส่คืน $p \pm Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{N-n}{N-1}\right)\frac{p(1-p)}{n}}$

วิธีที่3 ประมาณช่วงความเชื่อมั่นด้วยการแจกแจงแบบปกติเมื่อ n มีขนาดใหญ่ของ M.E. Thompson

กรณีสุ่มแบบใส่คืน $p \pm \frac{1}{2n} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$

กรณีสุ่มแบบไม่ใส่คืน $p \pm \frac{1}{2n} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{N-n}{N-1}\right)\frac{p(1-p)}{n}}$

จากนั้นทำการวัดความคลาดเคลื่อนของการประมาณ โดยพิจารณาจากค่าครอบคลุมความน่าจะเป็น เทียบกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด โดยค่าความน่าจะเป็นของการครอบคลุมที่ได้จากตัวประมาณแบบช่วงทั้ง 3 วิธีสามารถทำได้โดยทำการนับจำนวนช่วงความเชื่อมั่นที่สัดส่วนประชากร (P) ตกอยู่ในช่วงที่สร้างขึ้นแล้วนำมาหารด้วยจำนวนรอบของการทำซ้ำ นั่นคือ 1,000 รอบ โดยการศึกษาครั้งนี้กำหนดระดับความเชื่อมั่นเท่ากับ 0.95

3.3 การตรวจสอบความเหมาะสมของช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยภายใต้แผนการสุ่มแบบมีชั้นภูมิอย่างง่าย

เพื่อแสดงให้เห็นถึงการใช้ตัวประมาณแบบช่วงที่ถูกต้องตามแผนการสุ่มตัวอย่าง ซึ่งนอกจากการศึกษาภายใต้การสุ่มตัวอย่างแบบง่ายแล้ว ได้เพิ่มเติมการศึกษาภายใต้การสุ่มตัวอย่างแบบมีชั้นภูมิอย่างง่ายเพื่อเป็นการพัฒนาความเข้าใจในขั้นที่สูงขึ้น การศึกษาในส่วนนี้กำหนดให้จำนวนชั้นภูมิเท่ากับ 3 โดยเริ่มจากการจำลองข้อมูลของหน่วยประชากรซึ่งมีการแจกแจงแบบปกติ การแจกแจงแบบปัวซง และการแจกแจงแบบเอกซ์โพเนนเชียล โดยมีการกำหนดพารามิเตอร์ในการศึกษาเป็นดังนี้

- กรณีที่ประชากรมีการแจกแจงแบบปกติ กำหนดให้

1. ทุกชั้นภูมิมีค่าเฉลี่ยคงที่แต่มีความแปรปรวนที่ต่างกัน นั่นคือ

$$\mu_1 = \mu_2 = \mu_3 = 0$$

$$\sigma_1^2 = 2\sigma_2^2 = 4\sigma_3^2$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2. แต่ละชั้นภูมิมีค่าเฉลี่ยที่ต่างกัน แต่มีความแปรปรวนคงที่ นั่นคือ

$$\mu_1 = 2\mu_2 = 4\mu_3$$

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$$

- กรณีที่ประชากรมีการแจกแจงแบบปัวซง กำหนดให้ แต่ละชั้นภูมิมีค่าเฉลี่ยแตกต่างกัน นั่นคือ $\lambda_1 = 2\lambda_2 = 4\lambda_3$
- กรณีที่ประชากรมีการแจกแจงเอกซ์โพเนนเชียล กำหนดให้ แต่ละชั้นภูมิมีค่าเฉลี่ยแตกต่างกัน นั่นคือ $\theta_1 = 2\theta_2 = 4\theta_3$

กำหนดขนาดประชากร เท่ากับ 1,500 และ 3,000 โดยขนาดของชั้นภูมิมี 2 ลักษณะ คือ มีขนาดเท่ากันทุกชั้นภูมิ นั่นคือ $N_1 = N_2 = N_3$ และแต่ละชั้นภูมิมีขนาดที่แตกต่างกัน นั่นคือ $N_1 = 2N_2 = N_3$

เมื่อได้ข้อมูลหน่วยประชากรทั้งหมดที่มีขนาดประชากรตามสถานการณ์ที่สนใจแล้ว จากนั้นทำการสุ่มตัวอย่างจำนวน 1,000 รอบ ด้วยโปรแกรม Minitab โดยมีลักษณะการสุ่มตัวอย่างเป็นแบบไม่ใส่คืน กำหนดขนาดตัวอย่างในรูปร้อยละของประชากร 4 ระดับ คือ 1, 3, 5 และ 10 โดยขนาดตัวอย่างชั้นภูมิแบ่งเป็น 2 ลักษณะ คือ สัดส่วนโดยตรงกับขนาดของชั้นภูมิและขนาดเท่ากันทุกชั้นภูมิ และทำการประมาณค่าเฉลี่ยของประชากรโดยใช้การประมาณ 2 วิธีคือ

1. ตัวประมาณแบบ Simple Random Sampling $\bar{y} = \frac{1}{n}\sum_{h=1}^{L}\sum_{i=1}^{n} y_{hi}$

2. ตัวประมาณแบบ Stratified Random Sampling $\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$

จากนั้นนำตัวประมาณที่ได้จากทั้ง 2 วิธี ไปประมาณช่างความเชื่อมั่นสำหรับค่าเฉลี่ยประชากร โดยใช้การประมาณด้วยการแจกแจงแบบปกติของค่าลักษณะประชากร โดยใช้ความแปรปรวนของตัวอย่าง และทำการเปรียบเทียบค่าความน่าจะเป็นของการครอบคลุมที่ได้จากการประมาณทั้งสองวิธีว่าวิธีใดให้ค่าครอบคลุมความน่าจะเป็นใกล้เคียงกับค่าครอบคลุมความน่าจะเป็นที่กำหนดมากกว่ากัน โดยค่าความน่าจะเป็นของการครอบคลุมสามารถคำนวณได้โดยการนับจำนวนช่วงความเชื่อมั่นที่ค่าเฉลี่ยประชากร $(\bar{Y})$ ตกอยู่ในช่วงที่สร้างขึ้นแล้วนำมาหารด้วยจำนวนรอบของการทำซ้ำนั่นคือ 1,000 รอบ โดยการศึกษาครั้งนี้กำหนดระดับความเชื่อมั่น เท่ากับ 0.95

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved