

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

โดยทั่วไปประชากรที่สนใจศึกษามักจะมีขนาดใหญ่ การรวบรวมข้อมูลจากทุกหน่วยของประชากรย่อมต้องใช้ทรัพยากรเป็นจำนวนมาก ทั้งเวลา กำลังคน และงบประมาณ ซึ่งบ่อยครั้งเป็นสิ่งที่เป็นไปไม่ได้ภายใต้ข้อจำกัดที่มีอยู่ ดังนั้น หากสามารถเลือกหน่วยที่ให้ข้อมูลจำนวนหนึ่งมาเป็นตัวอย่างได้ ย่อมทำให้การศึกษาในเรื่องนั้นๆ เป็นไปได้ภายใต้กรอบของข้อจำกัดด้านต่างๆ และยังสามารถได้ประโยชน์ด้านการจัดการควบคุมคุณภาพของข้อมูลอันเกิดจากปริมาณงานที่น้อยลง เมื่อเปรียบเทียบข้อดีของการสำรวจตัวอย่างกับการสำมะโน จะเห็นว่าการสำรวจตัวอย่างมีข้อดีมากกว่าหลายประเด็น ที่สำคัญคือ การสำรวจตัวอย่างใช้เวลาและค่าใช้จ่ายน้อยกว่าการสำมะโน เนื่องจากขนาดของข้อมูลที่รวบรวมลดลง ทำให้ในขั้นตอนการรวบรวมข้อมูล และการวิเคราะห์ข้อมูลใช้เวลาเฉลี่ยน้อยลงกว่าคือ ได้ผลการศึกษาเร็วขึ้น และเมื่อปริมาณการรวบรวมข้อมูลลดลง การควบคุมคุณภาพในการรวบรวมข้อมูลย่อมทำได้ง่ายขึ้น ทำให้ข้อมูลที่ได้อาจมีความถูกต้องเชื่อถือมากขึ้นกว่ากรณีที่ต้องรวบรวมจากทุกหน่วยในประชากร

หลักการพื้นฐานของการสำรวจตัวอย่าง คือ การสุ่มตัวอย่างที่เป็นตัวแทนที่ดีของประชากร โดยคำนึงถึงลักษณะของประชากร เงื่อนไขข้อจำกัดต่างๆ โดยมีจุดประสงค์ที่จะสรุปผลหรือหาความรู้เกี่ยวกับประชากร ลักษณะสำคัญของประชากร (Population Characteristic) ที่นำมาใช้อธิบายลักษณะประชากรภายใต้กรอบความสนใจมี 4 ประเภท คือ ค่าเฉลี่ยประชากร (Population Mean) ยอดรวมประชากร (Population Total) สัดส่วนประชากร (Population Proportion) อัตราส่วนประชากร (Population Ratio)

วิธีการสุ่มตัวอย่างที่นับว่าเป็นแม่บทในการสุ่มตัวอย่าง จำแนกออกได้เป็น 4 วิธีคือ แผนแบบการสุ่มตัวอย่างแบบง่าย (Simple Random Sampling) แผนแบบการสุ่มตัวอย่างแบบมีระบบ (Systematic Sampling) แผนแบบสุ่มตัวอย่างแบบมีชั้นภูมิ (Stratified Sampling) แผนแบบการสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling) วิธีการสุ่มตัวอย่างแต่ละวิธีจะมีความเหมาะสมแตกต่างกันไป ขึ้นอยู่กับลักษณะประชากรที่สนใจอยู่ มีลักษณะแบบไหน เหมาะสมกับวิธีการเลือกแบบใด ถ้าหากเลือกแผนแบบการสุ่มตัวอย่างที่ไม่เหมาะสมกับลักษณะประชากร อาจส่งผลให้การประมาณค่าประชากรมีความคลาดเคลื่อน การนำค่าประมาณที่ได้ไปใช้ในการตัดสินใจมีความเสี่ยงมากขึ้น

โดยปกติแผนแบบการสุ่มตัวอย่างที่นิยมใช้กันมากคือ แผนแบบการสุ่มตัวอย่างแบบง่าย เนื่องจากเป็นแผนแบบการสุ่มตัวอย่างที่ง่ายและสะดวกต่อการนำไปใช้ การประมวลผลทำได้ง่าย

สามารถคำนวณได้โดยไม่มีความยุ่งยากมากนัก และสามารถปรับวิธีการประมาณได้ เช่น ถ้ารวบรวมข้อมูลได้เพียง n' หน่วย ($n' < n$) ก็สามารถ ใช้ข้อมูล n' หน่วยไปคำนวณค่าประมาณได้ แผนแบบการสุ่มตัวอย่างแบบง่ายเหมาะสำหรับการสุ่มตัวอย่างจากประชากรที่ประกอบด้วยหน่วยตัวอย่างที่มีลักษณะคล้ายคลึงกัน แต่แผนแบบการสุ่มตัวอย่างแบบง่ายก็ไม่เหมาะสมกับการสำรวจข้อมูลจากประชากรที่มีความแตกต่างกันในลักษณะที่ต้องการศึกษาอยู่มากๆ เช่น ในการสำรวจคลังสินค้าข้าวจากโรงสีที่มีขนาดแตกต่างกัน การสำรวจภาวะทางอุตสาหกรรมจากโรงงานอุตสาหกรรมที่แตกต่างกันทั้งประเภทและขนาดของการสำรวจ การประมาณจำนวนประชากรจากประชากรและหมู่บ้านที่มีขนาดและความหนาแน่นแตกต่างกันมากๆ การสำรวจระดับรายได้รายจ่ายของประชากรในท้องที่ต่างๆกัน และประชากรมีความแตกต่างกันในอาชีพและระดับการศึกษา ฯลฯ ถ้าหากใช้แผนการสุ่มแบบง่ายกับการสำรวจดังกล่าวจะทำให้ได้ค่าประมาณที่มีความแปรปรวนสูง ทำให้ผู้ที่ใช้ข้อมูลไม่มีความมั่นใจในผลการสำรวจ อีกทั้งแผนแบบการสุ่มตัวอย่างแบบง่ายยังเป็นแผนแบบการสุ่มตัวอย่างที่อาจจะทำให้เสียค่าใช้จ่ายในการดำเนินงานสูง เพราะอาจจะต้องใช้ขนาดตัวอย่างที่ใหญ่หรือมากเพื่อควบคุมขนาดของความคลาดเคลื่อนของค่าประมาณให้อยู่ในระดับที่เป็นที่พอใจของทุกฝ่าย ในงานสำรวจนั้นอาจพบปัญหาในทางปฏิบัติ เช่น หน่วยตัวอย่างกระจัดกระจายทั่วไปทำให้การควบคุมทำได้ยาก บางครั้งเมื่อใช้แผนแบบการสุ่มตัวอย่างแบบง่ายเป็นแผนแบบการสุ่มตัวอย่างพบว่า กลุ่มตัวอย่างที่รวบรวมมาได้ไม่ใช่ตัวแทนที่ดีพอของประชากรเพราะบางส่วนของประชากรไม่ได้รับการเลือก หรือได้รับเลือกเข้ามาในปริมาณที่ต่ำทำให้ค่าประมาณที่ได้มีค่าต่ำหรือสูงมากเกินไป และ บางครั้งผู้ศึกษาต้องการศึกษาประชากรย่อยบางกลุ่ม แต่ข้อมูลที่รวบรวมได้ไม่เพียงพอ กล่าวคือ หน่วยการสำรวจจากกลุ่มประชากรย่อยที่สนใจ ได้รับเลือกมาเป็นตัวอย่างน้อยเกินไปหรือไม่ปรากฏเลย เหล่านี้เป็นปัญหาพื้นฐานที่นำไปสู่การหาทางพัฒนาแผนแบบการสุ่มตัวอย่างใหม่ขึ้นมาเพื่อที่จะสามารถแก้ปัญหาเหล่านี้พร้อมทั้งเพิ่มระดับความแม่นยำของค่าประมาณให้สูงขึ้นด้วย จากเหตุผลที่กล่าวมาจึงได้มีการพัฒนาแผนแบบการสุ่มตัวอย่างแบบต่างๆขึ้นมาเพื่อใช้ให้เหมาะสมกับลักษณะประชากรแต่ละแบบ เช่น แผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling) McIntyre ได้ทำการศึกษาและได้ข้อสรุปว่าแผนแบบการสุ่มตัวอย่างอันดับ จะประมาณค่าเฉลี่ยของประชากรและให้ความแม่นยำได้ดีกว่าแผนแบบการสุ่มตัวอย่างอย่างง่าย (Simple Random Sampling) เมื่อประชากรมีลักษณะแตกต่างกันมากๆ ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาแนวคิดกระบวนการของแผนแบบการสุ่มตัวอย่างอันดับเพื่อให้เกิดความเข้าใจในกระบวนการ เจือจางของการใช้แผนแบบการสุ่มตัวอย่างชนิดนี้ เพื่อให้มีความพร้อมที่จะนำไปใช้ประโยชน์ในอนาคตต่อไป

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาแนวคิดของแผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling)
2. เพื่อศึกษากระบวนการของแผนแบบการสุ่มตัวอย่างของแผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling)
3. เพื่อศึกษาเปรียบเทียบประสิทธิภาพของแผนแบบการสุ่มตัวอย่างการแบบง่าย (Simple Random Sampling) และ แผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling)

1.3 ขอบเขตของการวิจัย

ศึกษาแผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling) กรณีที่เป็นแผนแบบการสุ่มตัวอย่างอันดับชนิดสมดุล (Balanced Ranked Set Sampling) และเป็นการสุ่มตัวอย่างแบบไม่ใส่คืน (Without Replacement) และการประมาณค่าเฉลี่ยของประชากร (Population Mean)

1.4 คำจำกัดความ

แผนแบบการสุ่มตัวอย่างแบบง่าย (Simple Random Sampling) หมายถึง วิธีการเลือกตัวอย่างที่กำหนดให้ตัวอย่างในขนาดที่กำหนดทุกตัวอย่างที่เป็นไปได้มีความน่าจะเป็นที่จะถูกเลือกเท่าๆกัน

แผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling) หมายถึง วิธีการเลือกตัวอย่างที่จะพิจารณาอันดับในการเลือกตัวอย่าง

ความไม่เอนเอียงของค่าประมาณ (Unbiased Estimator) หมายถึง การที่ค่าคาดหวัง (Expectation) ของค่าประมาณพารามิเตอร์มีค่าเท่ากับค่าพารามิเตอร์ หรือค่าเฉลี่ยของฟังก์ชันที่ขึ้นอยู่กับตัวอย่าง (Sampling Distribution) ของตัวประมาณมีค่าเท่ากับค่าจริง

ประชากร (Population) หมายถึง กลุ่มของข้อมูลทั้งหมดที่สนใจศึกษา

ตัวอย่าง (Sample) หมายถึง กลุ่มย่อยของประชากรหรือกลุ่มของข้อมูลบางส่วนของประชากร โดยกลุ่มย่อยนี้ประกอบไปด้วย หน่วยบางหน่วยของประชากร

ตัวอย่างสุ่ม (Random Sample) หมายถึง กลุ่มของตัวแปรสุ่มขนาด n ที่ได้มาจากประชากรเดียวกันและเป็นอิสระซึ่งกันและกัน

พารามิเตอร์ (Parameter) หมายถึง ค่าที่แสดงลักษณะหรือคุณสมบัติของประชากร

ตัวประมาณค่า (Estimator) หมายถึง ฟังก์ชันที่ขึ้นอยู่กับตัวอย่างสุ่ม ใช้สำหรับประมาณค่าพารามิเตอร์

ความแม่นยำหรือคงเส้นคงวา (Consistency) หมายถึง คุณสมบัติของค่าประมาณจะมีค่าลู่เข้าหาค่าพารามิเตอร์เมื่อขนาดตัวอย่างมากขึ้น และมีค่าเท่ากับค่าพารามิเตอร์ เมื่อขนาดตัวอย่างมีค่าเท่ากับขนาดของประชากร

ความแปรปรวน (Variance) หมายถึง ค่าคาดหวังกำลังสองของผลต่างระหว่าง ค่าประมาณพารามิเตอร์กับค่าคาดหวังของค่าประมาณพารามิเตอร์

ประชากรที่มีมูลค่าสูง หมายถึง ประชากรที่สนใจศึกษามีราคาแพง หาได้ยากหรือมีคุณค่าทางจิตใจสูง

1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เกิดความเข้าใจที่ชัดเจนในทฤษฎีของแผนการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling)
2. เพื่อที่จะหลักการทางทฤษฎีของแผนแบบการสุ่มตัวอย่างแบบอันดับไปใช้ได้เหมาะสม กับลักษณะประชากร
3. เพื่อเป็นแนวทางในพัฒนาแผนแบบการสุ่มตัวอย่างแบบอันดับในมิติอื่นๆต่อไป

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย