

บทที่ 1

บทนำ

ความเป็นมาของปัญหา

เทคนิคการสุ่มตัวอย่างเป็นระเบียบวิธีการทางสถิติอย่างหนึ่งที่มีคามสำคัญอย่างยิ่งสามารถนำไปใช้ประโยชน์ได้ทุกหน่วยงานไม่ว่าจะเป็นหน่วยงานในภาครัฐ ภาคเอกชน องค์กรระหว่างประเทศหรือในแวดวงการศึกษา และสื่อมวลชน เป็นต้น ดังเช่นในสำนักงานสถิติแห่งชาติได้ดำเนินการสำรวจการทำงานของคนไทย เพื่อเป็นข้อมูลพื้นฐานในการแสดงภาวะการมีงานทำ อัตราการว่างงาน หรือโครงการสำรวจความคิดเห็นของประชาชนเกี่ยวกับสถานการณ์การแพร่ระบาดของยาเสพติด (สำนักงานสถิติแห่งชาติ, 2554) และโครงการสำรวจทรัพยากรธรณีในทะเล (กรมทรัพยากรธรณี, 2550) เป็นต้น ทั้งนี้โครงการสำรวจด้วยตัวอย่างจะมีจุดเด่นที่ดีกว่าการสำรวจข้อมูลที่ใช้ทั้งประชากรหรือที่เรียกว่า การทำสำมะโนเนื่องจาก 1) เสียงบประมาณน้อยกว่า 2) ใช้เวลาน้อยกว่า 3) ใช้อัตรากำลังคนน้อยกว่า และ 4) การศึกษาจะมีคุณภาพที่ดีกว่า (สุชาติ กิระนันท์, 2542, หน้า 5)

การสำรวจด้วยตัวอย่างเป็นวิธีการคัดเลือกหน่วยตัวอย่างเพียงบางส่วนจากประชากร และการที่จะได้หน่วยตัวอย่างที่เป็นตัวแทนที่ดีของประชากรจะต้องพิจารณาความเหมาะสมของคุณลักษณะประชากร โดยการเลือกใช้แผนการสุ่มตัวอย่างที่ดี ควรรู้จักกับโครงสร้างของประชากรซึ่งจะทำให้เลือกแผนการสุ่มตัวอย่างได้เหมาะสม นอกจากนี้การใช้วิธีการประมาณค่าที่เหมาะสมกับสถานการณ์ที่สนใจศึกษาจะทำให้สามารถผลิตสถิติที่มีคุณภาพ เช่น โครงการสำรวจภาวะการมีงานทำของประเทศไทย ได้ใช้แผนการสุ่มตัวอย่างแบบมีชั้นภูมิ 3 ชั้นตอน (Stratified Three-stage Cluster Sampling) และใช้ตัวประมาณแบบอัตราส่วนในการประมาณค่า (สำนักงานสถิติแห่งชาติ, 2542) ซึ่งวิธีดังกล่าวเป็นการวางแผนที่ประสบความสำเร็จ และสถิติที่ได้จะมีความน่าเชื่อถือ ซึ่งแบบแผนนี้ใช้กันอย่างกว้างขวางทั่วโลก แต่แบบแผนนี้อาจไม่ประสบความสำเร็จเมื่อนำไปใช้กับสิ่งที่สนใจศึกษาที่มีลักษณะหายาก และอยู่รวมกันเป็นกลุ่ม ดังนั้นจึงมีการปรับแผนการสุ่มตัวอย่างให้เหมาะสมกับสถานการณ์มากยิ่งขึ้น (ประชุม สุวดี, 2552, หน้า 445)

แผนการสุ่มตัวอย่างแบบกลุ่มปรับ (Adaptive Cluster Sampling) เป็นวิธีการที่ดัดแปลงมาจากแผนการสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling) เพื่อให้สำรวจสิ่งที่หายากที่อยู่รวมกันเป็นกลุ่ม เช่น ต้องการประมาณจำนวนประชากรสัตว์ป่าที่หายากในพื้นที่แห่งหนึ่ง โดยที่หน่วยตัวอย่างถูกสุ่มด้วยแผนการสุ่มตัวอย่างความน่าจะเป็นขั้นพื้นฐาน (Conventional Probability

Sampling) เช่น การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีชั้นภูมิ และการสุ่มตัวอย่างแบบมีระบบ เป็นต้น แล้วอาจจะไม่พบสัตว์ป่า หรืออาจจะพบสัตว์ป่าที่สนใจเพียงหน่วยเดียวหรือน้อยมาก โดยข้อมูลที่ได้อาจจะไม่เพียงพอต่อการคำนวณ และอาจส่งผลให้ค่าประมาณที่ได้ไม่เสถียรภาพพอ ซึ่งอาจก่อให้เกิดความเอนเอียงในการประมาณค่าได้ ดังนั้นการใช้แผนการสุ่มตัวอย่างแบบกลุ่มปรับเป็นแนวทางหนึ่งที่ยอมรับใช้ในการแก้ไขปัญหาดังกล่าว (Thompson, 1990, pp.1050-1059) โดยที่เมื่อใดก็ตามที่เราทำการสุ่มตัวอย่างแล้วพบสัตว์ป่าก็จะทำการพิจารณาหน่วยบริเวณใกล้เคียงถ้าเป็นไปได้ตามเงื่อนไขก็จะทำการขยายหน่วยถัดไปทางซ้าย ขวา บน และล่าง ทำเช่นนี้ไปเรื่อย ๆ จนกระทั่งไม่พบสัตว์ป่าที่สนใจ ข้อมูลทั้งหมดที่พบก็จะเพียงพอในการนำไปคำนวณ การพบสิ่งที่สนใจศึกษาโดยบังเอิญ และการเพิ่มตัวอย่างในหน่วยที่ใกล้เคียง จะทำให้มีอัตราการพบสิ่งที่ศึกษาในอัตราที่ค่อนข้างสูง (สุชาติดา กิระนันท์, 2542, หน้า 352) ดังนั้นจะเห็นได้ว่าแผนการสุ่มตัวอย่างแบบกลุ่มปรับเป็นแผนแบบการสุ่มตัวอย่างที่น่าสนใจ และเป็นประโยชน์อย่างยิ่งถ้านำไปใช้ให้เหมาะสมกับลักษณะประชากรที่มีลักษณะอยู่รวมกันเป็นกลุ่ม ๆ โดยผู้วิจัยได้ทำการรวบรวม และสรุปผลงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับแผนการสุ่มตัวอย่างแบบกลุ่มปรับ และสามารถสรุปแนวคิดได้ดังนี้

ศิริประภา มโนมัยย์ (2539) ได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพของแผนการสุ่มตัวอย่างแบบกลุ่มปรับเมื่อตัวอย่างชั้นต้นใช้วิธีการสุ่มแบบง่าย แบบมีชั้นภูมิ และแบบมีระบบ โดยใช้ร้อยละของอัตราส่วนความแปรปรวนของค่าเฉลี่ย เป็นเกณฑ์ในการเปรียบเทียบ จากการวิจัยพบว่าประสิทธิภาพของแผนการสุ่มตัวอย่างแบบมีระบบมีประสิทธิภาพดีที่สุด รองลงมาคือแผนการสุ่มตัวอย่างแบบมีชั้นภูมิและแผนการสุ่มตัวอย่างแบบง่ายตามลำดับ วิชาญ โชควิวัฒน์ (2546) และ Thompson (1990) ได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพตัวประมาณค่าเฉลี่ยภายใต้แผนการสุ่มตัวอย่างแบบกลุ่มปรับและแผนการสุ่มตัวอย่างแบบกลุ่มที่ไม่ปรับ (Nonadaptive) โดยใช้ค่าประสิทธิภาพสัมพัทธ์ และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบตามลำดับ ซึ่งผลการวิจัยสอดคล้องไปในทิศทางเดียวกัน โดยพบว่าตัวประมาณค่าเฉลี่ยแบบกลุ่มปรับมีประสิทธิภาพที่ดีกว่าตัวประมาณค่าเฉลี่ยจากแผนการสุ่มตัวอย่างแบบกลุ่มที่ไม่ปรับ นอกจากนี้ Thompson ยังพบว่าตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson มีความแปรปรวนต่ำกว่าตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz นอกจากการประมาณค่าดังกล่าวแล้วได้มีผู้ที่ศึกษาวิธีประมาณค่าโดยใช้ประโยชน์จากตัวแปรช่วยและมีการพัฒนาด้วยกระบวนการ Rao-Blackwell มาทำการศึกษา โดยในปี ค.ศ. 2005 Dryver และ Thompson ได้ทำการศึกษาและเปรียบเทียบ

ประสิทธิภาพตัวประมาณค่าเฉลี่ยที่ปรับปรุงโดยวิธีการ Rao-Blackwell กับตัวประมาณค่าเฉลี่ยแบบพื้นฐาน โดยที่ตัวประมาณค่าเฉลี่ยดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz และ Horvitz-Thompson ผลการวิจัยพบว่าตัวประมาณค่าเฉลี่ย ที่ปรับปรุงโดยวิธีการ Rao-Blackwell จะมีประสิทธิภาพที่ดีกว่าตัวประมาณค่าเฉลี่ยแบบพื้นฐาน และในปี ค.ศ.2007 Dryver และ Chao มีแนวคิดในการใช้ประโยชน์จากตัวแปรช่วยในการประมาณค่า ซึ่งได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพตัวประมาณค่าเฉลี่ยแบบอัตราส่วนกับตัวประมาณค่าเฉลี่ยแบบพื้นฐาน โดยที่ตัวประมาณค่าเฉลี่ยดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz และ Horvitz-Thompson ผลการวิจัยพบว่าตัวประมาณค่าเฉลี่ยแบบอัตราส่วน จะมีประสิทธิภาพที่ดีกว่าตัวประมาณค่าเฉลี่ย ต่อมาในปี ค.ศ.2008 Chao, Lin และ Chiang มีแนวคิดในการปรับปรุงตัวประมาณค่าเฉลี่ยแบบอัตราส่วนให้มีประสิทธิภาพเพิ่มมากขึ้น โดยใช้ทฤษฎี Rao-Blackwell มาทำการศึกษาและเปรียบเทียบประสิทธิภาพตัวประมาณค่าเฉลี่ยแบบอัตราส่วนที่ปรับปรุงโดยวิธีการ Rao-Blackwell กับตัวประมาณค่าเฉลี่ยแบบอัตราส่วน ที่ Dryver และ Chao ศึกษาในปี ค.ศ. 2007 โดยที่ตัวประมาณค่าเฉลี่ยดัดแปลงมาจากตัวประมาณ Hansen-Hurwitz และ Horvitz-Thompson และผลการวิจัยพบว่า ตัวประมาณค่าเฉลี่ยแบบอัตราส่วนที่ปรับปรุงโดยวิธีการ Rao-Blackwell จะเป็นฟังก์ชันของสถิติพอเพียงต่ำที่สุด (Minimal Sufficient Statistic) และมีประสิทธิภาพที่ดีกว่าตัวประมาณค่าเฉลี่ยแบบอัตราส่วน (ศิริประภา มโนมรรย์, 2539; Thompson, 1990; วิชาญ ไชควิวัฒน์, 2546; Dryver and Thompson, 2005; Dryver and Chao, 2007; Chao, Lin and Chiang, 2008)

จากการศึกษาแนวคิดและงานวิจัยที่เกี่ยวข้องจะเห็นได้ว่าการใช้ประโยชน์จากตัวแปรช่วยที่เหมาะสม จะทำให้การประมาณค่ามีความคลาดเคลื่อนลดลง และการใช้ทฤษฎีสถิติที่มีความพอเพียง จะทำให้ได้สถิติที่มีคุณภาพและน่าเชื่อถือ ดังนั้นผู้วิจัยจึงมีความสนใจที่จะทำการเปรียบเทียบวิธีประมาณค่าเฉลี่ยด้วยวิธีประมาณค่าอย่างง่าย วิธีประมาณค่าแบบอัตราส่วน และวิธีประมาณค่าแบบอัตราส่วนที่ปรับปรุงโดยวิธีการ Rao-Blackwell โดยใช้ ตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-Thompson

จุดมุ่งหมายของการศึกษา

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพวิธีประมาณค่าเฉลี่ย ภายใต้แผนการสุ่มตัวอย่างแบบกลุ่มปรับ ด้วยวิธีประมาณค่าอย่างง่าย วิธีประมาณค่าแบบอัตราส่วน และวิธีประมาณค่าแบบอัตราส่วนที่ปรับปรุงโดยวิธีการ Rao-Blackwell เมื่อสุ่มหน่วยตัวอย่างอย่างง่าย

ขอบเขตของงานวิจัย

การวิจัยครั้งนี้ผู้วิจัยได้กำหนดขอบเขต และขั้นตอนดำเนินการวิจัย ดังนี้

1. แบ่งพื้นที่ทำการศึกษาค้นคว้าออกเป็น 100 พื้นที่ย่อยดังภาพ 1
2. จำลองแบบประชากรที่มีลักษณะหายาก และอยู่รวมกันเป็นกลุ่ม 2 กรณีที่แตกต่างกัน ลงใน 100 พื้นที่ย่อยโดยอาศัยกระบวนการของคลัสเตอร์ ที่มีพารามิเตอร์ λ เป็น 30 โดยใช้โปรแกรม Minitab 14 ในการจำลองแบบประชากร ซึ่งลักษณะประชากร 2 กรณีประกอบด้วย

2.1 ประชากรกรณีที่ 1 แบ่งเป็น 3 ลักษณะ

2.1.1 หน่วยขอบแยกกัน 2 กลุ่ม

2.1.2 หน่วยขอบแยกกัน 3 กลุ่ม

2.1.3 หน่วยขอบแยกกัน 4 กลุ่ม

2.2 ประชากรกรณีที่ 2 แบ่งเป็น 3 ลักษณะ

2.2.1 หน่วยขอบรวมกัน 2 กลุ่ม

2.2.2 หน่วยขอบรวมกัน 3 กลุ่ม

2.2.3 หน่วยขอบรวมกัน 4 กลุ่ม

3. ตัวแปรในแต่ละสถานการณ์ประกอบด้วย

3.1 ตัวแปร Y คือ ตัวแปรที่ต้องการศึกษา เป็นข้อมูลเชิงปริมาณ

3.2 ตัวแปร X คือ ตัวแปรช่วยของสิ่งที่ต้องการศึกษา เป็นข้อมูลเชิงปริมาณในพื้นที่

ย่อย

3.3 กำหนดความสัมพันธ์ระหว่างตัวแปรเชิงปริมาณ Y กับตัวแปรเชิงปริมาณ X

ความสัมพันธ์ระดับต่ำ อยู่ในช่วง (0 , 0.4)

ความสัมพันธ์ระดับปานกลาง อยู่ในช่วง [0.4, 0.7)

ความสัมพันธ์ระดับสูง อยู่ในช่วง [0.7, 1.00)

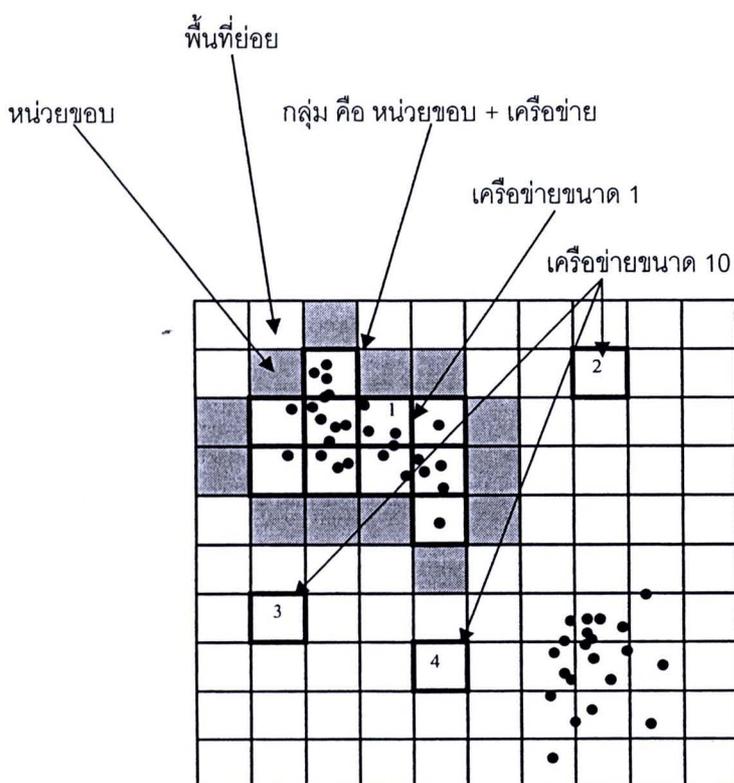
4. กำหนดขนาดพื้นที่ย่อยที่ใช้ในการวิจัยเท่ากับ 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 20

และ 24

5. คำนวณหาค่าเฉลี่ย และความแปรปรวนของวิธีประมาณค่าเฉลี่ย ภายใต้แผนการสุ่มตัวอย่างแบบกลุ่มปรับด้วยวิธีประมาณค่าอย่างง่าย วิธีประมาณค่าแบบอัตราส่วน และวิธีประมาณค่าแบบอัตราส่วนที่ปรับปรุงโดยวิธีการ Rao-Blackwell เมื่อสุ่มหน่วยตัวอย่างอย่างง่าย

6. เปรียบเทียบประสิทธิภาพของแต่ละวิธีโดยใช้เกณฑ์ของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

7. ทำซ้ำด้วยโปรแกรม MATLAB 1,000 รอบ



ภาพ 1 แสดงการสุ่มหน่วยตัวอย่างด้วยแผนการสุ่มตัวอย่างแบบง่ายที่มีขนาด 4 หน่วย

นิยามศัพท์เฉพาะ

แผนการสุ่มตัวอย่างแบบกลุ่มปรับ หมายถึง แผนการสุ่มตัวอย่างแบบที่ปรับกระบวนการสุ่มตัวอย่างให้สามารถนำลักษณะข้อมูลที่เก็บได้มากำหนดหน่วยตัวอย่างด้วยการขยายหน่วยตัวอย่างที่เป็นไปตามเงื่อนไขที่สนใจศึกษาทั้งด้าน บน ล่าง ขวา และซ้าย จนกระทั่งไม่พบหน่วยตัวอย่างที่เป็นไปตามเงื่อนไข

แผนการสุ่มตัวอย่างแบบที่ไม่ได้ปรับ หมายถึง แผนการสุ่มตัวอย่างแบบทั่วไปหรือแบบพื้นฐาน เช่น แผนการสุ่มตัวอย่างแบบง่าย แผนการสุ่มตัวอย่างแบบมีชั้นภูมิ หรือ แผนการสุ่มตัวอย่างแบบมีระบบ เป็นต้น

เงื่อนไขของสิ่งที่สนใจ หมายถึง สิ่งที่กำหนดเพื่อทำการขยายหน่วยถัดไป โดยงานวิจัยนี้ใช้เงื่อนไข $C = \{y : y \geq 3\}$ โดยที่ให้ C เป็นเซตของค่าที่มากกว่าหรือเท่ากับ 3 โดยถ้า y ซึ่งแทนสิ่งที่สนใจที่มีลักษณะหายาก และอยู่รวมกันเป็นกลุ่มที่อยู่ใน C มีค่ามากกว่าหรือเท่ากับ 3 ซึ่งเป็นไปตามเงื่อนไขที่กำหนดก็จะพิจารณาหน่วยบริเวณใกล้เคียงเข้าไปเป็นหน่วยตัวอย่าง

หน่วยขอบ (Edge Unit) หมายถึง หน่วยที่ไม่เป็นไปตามเงื่อนไขที่กำหนด แต่หน่วยที่ใกล้เคียงของหน่วยขอบนั้นสอดคล้องตามเงื่อนไข

เครือข่าย (Network) หมายถึง กลุ่มย่อยซึ่งประกอบด้วยหน่วยย่อยที่เป็นไปตามเงื่อนไขภายในกลุ่ม แต่ถ้าหน่วยใดที่ไม่เป็นไปตามเงื่อนไขแต่ถูกเลือกเป็นตัวอย่างจะให้เครือข่ายนั้นมีขนาดเป็น 1

กลุ่ม (Cluster) หมายถึง กลุ่มที่มีการเก็บหน่วยทั้งหมดที่ถูกสังเกตภายใต้แผนการสุ่มตัวอย่างแบบกลุ่มปรับของการสุ่มตัวอย่างในหน่วยที่ i ซึ่งประกอบด้วยหน่วยขอบรวมกับเครือข่ายพื้นที่ย่อย (Plot) หมายถึง พื้นที่เล็ก ๆ ที่รวมกันเป็นประชากร

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อทราบถึงประสิทธิภาพของ วิธีประมาณค่าเฉลี่ยทั้ง 3 วิธีการ ว่าเหมาะสมในสถานการณ์ใด
2. เพื่อเกิดความเข้าใจที่ชัดเจนในทฤษฎีการสุ่มตัวอย่างแบบกลุ่มปรับ และนำไปประยุกต์ใช้สำหรับการสำรวจตัวอย่างที่หน่วยตัวอย่างมีลักษณะที่หายาก และอยู่รวมกันเป็นกลุ่ม