บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัย เรื่อง การเปรียบเทียบคุณสมบัติของตัวประมาณค่าพารามิเตอร์ที่ได้จากกลุ่มตัวอย่างสุ่มแบบหลายขั้นตอน ระหว่างวิธีสุ่มแบบง่ายกับแบบมีระบบนี้ มีวรรณคดีที่เกี่ยวข้องจากหนังสือเอกสารและรายงานการวิจัยต่าง ๆ โดยนำเสนอตามลำดับขั้น ดังนี้

ตอนที่ 1 การสุ่มตัวอย่างและขั้นตอนในการสุ่มตัวอย่าง
ตอนที่ 2 วิธีการสุ่มตัวอย่าง
ตอนที่ 3 การกำหนดขนาดของกลุ่มตัวอย่าง
ตอนที่ 4 ความคลาดเคลื่อนที่เกิดจากการสุ่มตัวอย่าง
ตอนที่ 5 การประมาณค่าและคุณสมบัติของตัวประมาณค่าที่ดี
ตอนที่ 6 การแจกแจงการสุ่มตัวอย่าง
ตอนที่ 7 การจำลองสถานการณ์และเทคนิคมอนติ คาร์โล
ตอนที่ 8 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 การสุ่มตัวอย่างและขั้นตอนในการสุ่มตัวอย่าง

การวิจัยในบางครั้งโดยเฉพาะการวิจัยที่มีกลุ่มประชากรขนาดใหญ่ จะไม่สามารถที่จะรวบรวมข้อมูลจากทุก ๆ หน่วยของสมาชิกในกลุ่มประชากร การเลือกสมาชิกจำนวนหนึ่งจากกลุ่มประชากรมาใช้ในการศึกษาวิจัย คือ การสุ่มตัวอย่าง การสุ่มตัวอย่างที่ดีจะทำให้ได้กลุ่มตัวอย่างที่เป็นตัวแทนที่ดีของประชากรหรือมีคุณสมบัติที่เหมือนกลุ่มประชากรที่ต้องการศึกษามากที่สุดเท่าที่จะสามารถทำได้

ขั้นตอนในการสุ่มตัวอย่าง

ขั้นตอนในการสุ่มตัวอย่าง อาจแบ่งได้ดังนี้ (สุชาดา กีระนันทน์, 2538 : 19-21)

1. กำหนดวัตถุประสงค์หรือเป้าหมายของการสำรวจตัวอย่างให้ชัดเจน ว่าการสำรวจตัวอย่างนี้จะทำเพื่อเก็บข้อมูลเอาไปใช้ประโยชน์อะไร ทั้งนี้เพื่อให้สามารถกำหนดได้ว่าจะเก็บข้อมูลอะไรบ้าง จากประชากรใด และข้อมูลต้องมีลักษณะอย่างไร

2. กำหนดประชากรที่จะศึกษา (Target Population) ประชากรประกอบขึ้นด้วยหน่วยทุกหน่วยที่จะให้ข้อมูลที่ต้องการตามเป้าหมายของการสำรวจประชากรที่จะศึกษาจะต้อง กำหนดให้เข้าใจได้ชัดแจ้ง ถ้าเป็นประชากรที่เปลี่ยนแปลงได้ง่าย เช่น ตามเวลาก็ต้องมีการกำหนดช่วงเวลาสำหรับอ้างอิงไว้ด้วย

3. หา หรือ สร้างกรอบสำหรับเลือกตัวอย่าง (Sampling Frame) หรือสำหรับแจงนับเนื่องจากประชากรที่กำหนดไว้ในข้อ 2. เป็นการกำหนดด้วยข้อความแต่ในการเลือกตัวอย่างนั้นจำเป็นจะต้องมีสิ่งที่สัมผัสได้ เพื่อทำการเลือกหน่วยตัวอย่าง กรอบตัวอย่างที่ดีจะต้องแสดงหน่วยทุกหน่วยในประชากรโดยไม่มีการซ้ำซ้อน และไม่มีหน่วยที่ไม่ได้อยู่ในประชากร ปรากฏในกรอบ

4. กำหนดวิธีการเลือกตัวอย่าง ว่าจะทำการเลือกโดยให้แต่ละหน่วยมีโอกาสถูกเลือกเช่นไร การเลือกจะเลือกจากหน่วยเล็กที่สุดในประชากรที่เป็นผู้ให้ข้อมูล หรือเลือกเป็นกลุ่มก่อนโดยใช้การเลือกแบบหลายขั้น เป็นต้น การกำหนดวิธีการเลือกตัวอย่างจะต้องสอดคล้องกับลักษณะของกรอบตัวอย่างที่มีและต้องพิจารณาถึงลักษณะของประชากรด้วย เพราะวิธีเลือกแต่ละวิธีก็มีความเหมาะสมต่าง ๆ กันไปสำหรับประชากรลักษณะต่าง ๆ

5. กำหนดวิธีการหาตัวประมาณที่ต้องการหรือวิธีการอนุมาน เพื่อประกอบในการพิจารณาคุณภาพของตัวประมาณต่อไป และเป็นการตรวจสอบความครบถ้วนของข้อมูลที่จะเก็บด้วย

6. กำหนดหรือสร้างสื่อที่ใช้ในการเก็บข้อมูล โดยปกติข้อมูลที่ต้องการเก็บหรือตัวแปรที่จะศึกษา จะถูกกำหนดด้วยวัตถุประสงค์ ขอบเขต และสมมุติฐานของการวิจัย ซึ่งผู้วิจัยต้องพิจารณาหาออกมาว่า การให้ได้สิ่งที่ตนต้องการนั้น จะต้องอาศัยข้อมูลอะไรบ้าง ข้อมูลหรือตัวแปรเหล่านี้ จะต้องนำมาแปลให้อยู่ในรูปของสื่อที่จะใช้เก็บข้อมูล เช่น แบบสอบถาม ซึ่งการสร้างแบบสอบถามก็ต้องคำนึงถึงลักษณะคำถาม วิธีการสร้าง ข้อควรระวังต่าง ๆ เป็นต้น

7. กำหนดขนาดตัวอย่าง ในการเลือกตัวอย่างตามความน่าจะเป็น คุณภาพของตัวประมาณส่วนหนึ่งจะขึ้นกับขนาดตัวอย่างในลักษณะที่ถ้าขนาดตัวอย่างใหญ่ขึ้น ค่าความแปรปรวนของตัวประมาณจะลดลง การกำหนดขนาดตัวอย่างจึงพิจารณาระหว่างคุณภาพและทรัพยากรที่จะต้องใช้ โดยอาจกำหนดคุณภาพที่ต้องการถ้าเปลี่ยนแปลงทรัพยากรได้ หรือถ้าทรัพยากรคงที่ก็ต้องหาวิธีที่จะให้ได้ตัวประมาณที่มีค่าแปรปรวนต่ำสุดภายใต้ขีดจำกัดของทรัพยากรที่มีในงานวิจัย ส่วนใหญ่ผู้วิจัยมักจะกังวลเกี่ยวกับขนาดตัวอย่างที่เหมาะสม เนื่องจากขนาดตัวอย่างจะขึ้นอยู่กับลักษณะการกระจายของตัวแปรที่จะศึกษาซึ่งผู้วิจัยไม่ทราบแน่ชัด

8. เลือกหน่วยตัวอย่าง การเลือกหน่วยตัวอย่างจะขึ้นอยู่กับแผนแบบการเลือกตัวอย่างที่กำหนดไว้และขนาดตัวอย่าง ในกรณีที่กรอบตัวอย่างสมบูรณ์อาจกำหนดหน่วยตัวอย่างได้เลย ในกรณีอื่น ๆ การกำหนดหน่วยตัวอย่างที่จะให้ข้อมูลอาจต้องผ่านขั้นตอนอื่น เช่น การนับจด (Listing) หน่วยที่ให้ข้อมูลในเขตที่ถูกเลือกเป็นตัวอย่าง เป็นต้น

9. กำหนดวิธีการเก็บข้อมูลจากหน่วยตัวอย่าง การเก็บข้อมูลอาจจะทำโดยใช้พนักงานแจงนับออกไปสังเกต (Observe) หรือวัดค่า (Measure) โดยไม่ต้องถามจากหน่วยที่ให้ข้อมูล หรือให้พนักงานแจงนับไปสัมภาษณ์ หรือจดจากหน่วยตัวอย่างโดยใช้แบบสอบถามหรือให้หน่วยตัวอย่างตอบแบบสอบถามเอง เช่น ส่งแบบสอบถามให้ทางไปรษณีย์ เป็นต้น การเลือกวิธีเก็บข้อมูลจากหน่วยตัวอย่าง จะต้องพิจารณาถึงลักษณะของหน่วยตัวอย่าง และสื่อที่ใช้หรือแบบสอบถามเป็นสำคัญ เพราะจะมีปัญหาเกี่ยวกับการไม่ตอบตามมาด้วย

10. **เตรียมงานสนาม** การเตรียมงานสนามจะต้องกำหนดช่วงเวลาที่จะออกงานสนาม จัดหาพนักงานสอบถามหรือสัมภาษณ์ในจำนวนที่ต้องการ และอบรมพนักงานสอบถามให้เข้าใจวัตถุประสงค์ของการสำรวจและเข้าใจแบบสอบถามที่จะใช้ อาจมีการทดสอบงานสนามเพื่อดูว่ามีปัญหาอะไรหรือไม่และจะแก้ไขปัญหาอย่างไร เนื่องจากพนักงานสอบถามมีผลอย่างมากต่อคุณภาพของข้อมูลที่ได้ การเตรียมพนักงานสอบถามสำหรับงานสนามจึงมีความสำคัญมาก รวมทั้งมีผู้ที่จะทำหน้าที่ควบคุมดูแลและแก้ปัญหาเฉพาะหน้าในงานสนามด้วย ถ้าเป็นกรณีที่ใช้การส่งแบบสอบถามทางไปรษณีย์ให้หน่วยตัวอย่างตอบเอง ก็ต้องมีการกำหนดระยะเวลา และวิธีการติดตามในกรณีไม่ตอบ เป็นต้น

11. **งานสนาม** เป็นการปฏิบัติการเก็บข้อมูลจากหน่วยตัวอย่างจริง ๆ ลักษณะของงานสนามจะแตกต่างกันไปตามวิธีการเก็บรวบรวมข้อมูลที่ใช้ในงานแต่ละงาน รวมทั้งลักษณะของหน่วยตัวอย่างและสภาพความเป็นอยู่ด้วย

12. **บรรณาธิกร** (Edit) **ลงรหัส** (Code)**และเตรียมข้อมูลสำหรับการประมวลผล** และการวิเคราะห์ที่จะทำต่อไป การบรรณาธิกร ควรกระทำโดยเร็วที่สุดต่อจากการเก็บข้อมูล ซึ่งถ้าสามารถทำการบรรณาธิกรข้อมูลที่เก็บมาได้วันต่อวัน ก็จะช่วยทำให้ข้อมูลมีความถูกต้องครบถ้วนมากขึ้น และจะมีแบบสอบถามที่ใช้ไม่ได้เต็มที่ลดลง ทั้งนี้เพราะผู้เก็บข้อมูลหรือพนักงานสัมภาษณ์ยังสามารถจะฟื้นความจำในเรื่องที่เป็นปัญหาได้ดีและการกลับไปสอบถามเพิ่มเติมจากหน่วยตัวอย่างอาจจะทำได้สะดวกกว่า ถ้าตัวอย่างมีขนาดใหญ่พอควร และมีตัวแปรที่ต้องการศึกษาจำนวนมาก การประมวลผลและการวิเคราะห์ข้อมูล มักจะต้องใช้เครื่องคอมพิวเตอร์ จึงต้องมีการลงรหัส เพื่อเปลี่ยนสภาพข้อมูลให้เป็นรหัสที่จะใช้คอมพิวเตอร์ทำงานขั้นต่อไปได้ แล้วจึงทำการเตรียมข้อมูลในรูปที่คอมพิวเตอร์สามารถรับได้ เช่น บันทึกข้อมูลลงในสื่อที่เป็นเทปแม่เหล็กหรือแผ่นดิสเกต เป็นต้น เพื่อให้คอมพิวเตอร์ทำการประมวลผลหรือวิเคราะห์ตามคำสั่งต่อไป

13. **วิเคราะห์ข้อมูลและเสนอผล** การวิเคราะห์ข้อมูลนับเป็นส่วนที่สำคัญมากสำหรับงานวิจัยหรือการศึกษาเรื่องใดเรื่องหนึ่ง การวิเคราะห์ข้อมูลจะต้องอาศัยทฤษฎีสถิติเพื่อทำการอนุมานจากตัวอย่างไปสู่ประชากร โดยทั่วไปส่วนของการวิเคราะห์ข้อมูลอาจจะต้องอยู่นอกเหนือขอบข่ายของการสำรวจตัวอย่าง และจะไม่กล่าวถึงในรายละเอียดในที่นี้

โดยสรุปแล้ว ในการสุ่มตัวอย่างที่ดี เพื่อให้ได้กลุ่มตัวอย่างซึ่งเป็นตัวแทนที่ดีของประชากร ควรพิจารณาความมุ่งหมายของการวิจัยหรือการเลือกกลุ่มตัวอย่างให้ชัดเจน กำหนดหน่วยตัวอย่าง ขนาดตัวอย่าง วิธีการสุ่มตัวอย่างและวิธีการประมาณค่าที่เหมาะสมกับลักษณะของประชากร นอกจากนั้นควรคำนึงถึงเวลา ค่าใช้จ่าย แรงงาน และทรัพยากรอื่น ๆ ในการรวบรวมข้อมูลจากกลุ่มตัวอย่างด้วย การสุ่มตัวอย่างที่ดีช่วยให้ผลการวิจัยมีคุณภาพ มีความถูกต้อง น่าเชื่อถือ

ตอนที่ 2 วิธีการสุ่มตัวอย่าง (Sampling Methods)

การเลือกกลุ่มตัวอย่างซึ่งเป็นตัวแทนที่ดีนั้น มีความสำคัญมากสำหรับการวิจัย สิ่งสำคัญประการหนึ่งนั้นขึ้นอยู่กับวิธีการสุ่มตัวอย่างด้วย สำหรับวิธีการสุ่มตัวอย่างสามารถแบ่งออกได้เป็น 2 ประเภท คือ

1. วิธีการสุ่มแบบไม่อิงความน่าจะเป็น (Nonproability Sampling)
2. วิธีการสุ่มแบบอิงความน่าจะเป็น (Proabability Sampling)

1. วิธีการสุ่มแบบไม่อิงความน่าจะเป็น (Nonprobability Sampling)

เป็นการสุ่มแบบไม่คำนึงถึงว่ากลุ่มตัวอย่างที่ได้รับเลือกมานั้นจะมีความน่าจะเป็นหรือโอกาสที่จะได้รับเลือกมาเป็นเท่าใด ไม่มีการเตรียมรายละเอียดของกลุ่มประชากร นักวิจัยเพียงแต่นิยามและกำหนดขอบเขตของกลุ่มประชากรเท่านั้น วิธีการสุ่มแบบไม่อิงความน่าจะเป็นมีหลายวิธี ได้แก่ (นงลักษณ์ วิรัชชัย, 2536)

1. วิธีการสุ่มแบบเจาะจงหรือตามวัตถุประสงค์ (Purposive or Judgement Sampling) เป็นวิธีการเลือกตัวอย่างที่นักวิจัยใช้วิจารณญาณตัดสินว่าเป็นหน่วยตัวอย่างที่มีลักษณะพิเศษตรงกับปัญหาการวิจัย และเจาะจงเลือกเป็นหน่วยตัวอย่างในการวิจัย วิธีการเลือกกลุ่มตัวอย่างแบบนี้นักวิจัยต้องกำหนดนิยาม และขอบเขตของกลุ่มประชากรเฉพาะก่อน แล้วศึกษาสมาชิกของกลุ่มประชากร เจาะจงเลือกเฉพาะสมาชิกที่มีลักษณะตัวแปรตรงกับปัญหาการวิจัย โดยมากนิยามเลือกสมาชิกที่มีลักษณะเด่นพิเศษกว่าสมาชิกหน่วยอื่น ๆ

2. วิธีการสุ่มตามโควต้า (Quota Sampling) เป็นวิธีการเลือกตัวอย่างที่นักวิจัยให้นิยามและกำหนดขอบเขตของกลุ่มประชากร รวมทั้งกำหนดขนาดของกลุ่มตัวอย่างที่ต้องการแล้วก็รวบรวมข้อมูลจากสมาชิกที่มีลักษณะตรงกับลักษณะกลุ่มประชากรจนกว่าจะได้ครบตามจำนวนที่ต้องการ หรือครบโควต้าที่กำหนดไว้ วิธีการเลือกตัวอย่างแบบนี้เป็นวิธีที่สะดวกรวดเร็วและประหยัดมากที่สุด ใช้มากในการวิจัยเกี่ยวกับการตลาดและการสำรวจความคิดเห็นของประชาชน

3. วิธีการสุ่มแบบบังเอิญหรืออุบัติการ (Accidental Sampling) เป็นวิธีการเลือกตัวอย่างที่มีลักษณะตรงตามลักษณะของกลุ่มประชากร เลือกสมาชิกที่อยู่ใกล้ตัวนักวิจัยมากที่สุด ไม่ต้องกำหนดขนาดของกลุ่มตัวอย่างไว้ก่อน เมื่อให้นิยามและกำหนดขอบเขตของกลุ่มประชากรเฉพาะแล้ว สามารถเลือกสมาชิกได้เลย วิธีนี้จึงสะดวก รวดเร็ว แต่นิยมใช้เฉพาะเหตุการณ์ อุบัติการที่ไม่เกิดขึ้นบ่อยครั้งหรือไม่อาจจัดกระทำได้

4. วิธีการสุ่มแบบลูกโซ่หรือแบบเชิงก้อนหิมะ (Chain or Snowball Sampling) เป็นวิธี การเลือกตัวอย่างที่ผู้วิจัยแสวงหาสมาชิกกลุ่มประชากรจำนวนหนึ่งเป็นหน่วยตัวอย่าง และอาศัยความร่วมมือจากสมาชิกกลุ่มนี้ช่วยหาสมาชิกต่อ ๆ ไปเป็นลูกโซ่จนกว่าจะได้จำนวนหน่วยตัวอย่างครบตามความต้องการ

การสุ่มตัวอย่างแบบไม่อิงความน่าจะเป็นในการสุ่มนั้น เป็นวิธีที่สะดวก รวดเร็ว และประหยัดมากที่สุด เป็นวิธีที่ใช้มากในการวิจัยเกี่ยวกับการตลาดและการสำรวจความคิดเห็นของประชาชน

ข้อเสียที่สำคัญของการสุ่มแบบไม่อิงความน่าจะเป็น คือ นักวิจัยไม่สามารถคำนวณหาค่าความน่าจะเป็นของการถูกเลือกเป็นกลุ่มตัวอย่าง และกลุ่มตัวอย่างขาดลักษณะการสุ่มเป็นผลให้ไม่สามารถอ้างอิงผลการวิจัยไปสู่กลุ่มประชากรได้ หากนักวิจัยต้องการอ้างอิงผลการวิจัยอาจจะแก้ไขได้โดยการแสดงให้เห็นชัดเจนว่า กลุ่มตัวอย่างที่เลือกมานั้นเป็นตัวแทนที่ดีของกลุ่มประชากรโดยใช้การเปรียบเทียบลักษณะต่าง ๆ ของกลุ่มตัวอย่างว่าใกล้เคียงกับลักษณะของกลุ่มประชากร อย่างไรก็ดีนักวิจัยมักมีปัญหาเนื่องจากขาดข้อมูลเกี่ยวกับลักษณะของกลุ่มประชากรและจำเป็นต้องใช้ข้อมูลจากการวิจัยที่มีผู้ทำไว้แล้ว เป็นเหตุให้ไม่สามารถยืนยันความเป็นตัวแทนที่ดีของกลุ่มตัวอย่างได้ ด้วยเหตุนี้หากสามารถทำได้นักวิจัยควรใช้กลุ่มตัวอย่างสุ่มหรือใช้การสุ่มแบบอิงความน่าจะเป็น เพื่อให้ผลการวิจัยมีความเที่ยงตรงภายนอกอยู่ในระดับสูง

2. วิธีการสุ่มแบบอิงความน่าจะเป็น (Probability Sampling)

เป็นการสุ่มตัวอย่างโดยคำนึงถึงความน่าจะเป็นของแต่ละหน่วยประชากรที่จะได้รับเลือก ซึ่งเป็นไปแบบสุ่ม เพื่อนำผลที่ได้ไปใช้สรุปอ้างอิงถึงประชากร วิธีการเลือกตัวอย่างที่นับว่าเป็นวิธีแม่บทในการสุ่มแบบอิงความน่าจะเป็น อาจแบ่งออกได้เป็น 4 วิธี คือ

1. วิธีสุ่มแบบง่าย (Simple Random Sampling)
2. วิธีสุ่มแบบมีระบบ (Systematic Sampling)
3. วิธีสุ่มแบบแบ่งชั้น (Stratified Sampling)
4. วิธีสุ่มแบบตามกลุ่ม (Cluster Sampling)

1. วิธีสุ่มแบบง่าย (Simple Random Sampling)

วิธีการสุ่มแบบง่าย เป็นวิธีการเลือกตัวอย่างที่เป็นไปตามเงื่อนไข 3 ประการของการสุ่ม คือ สมาชิกทุกหน่วยในกลุ่มประชากรมีความน่าจะเป็นหรือโอกาสในการถูกเลือกเท่ากัน หรือตัวอย่างแต่ละหน่วยมีความน่าจะเป็นในการถูกเลือกเป็นกลุ่มตัวอย่างเท่ากัน และการเลือกสมาชิกแต่ละหน่วยเป็นอิสระต่อกัน

วิธีดำเนินการ เมื่อกำหนดขอบเขตและนิยามกลุ่มประชากรเฉพาะการวิจัย และกำหนดขนาดของกลุ่มตัวอย่างที่ต้องการแล้ว สามารถดำเนินการสุ่มเลือกสมาชิกได้โดยวิธีการต่าง ๆ ดังนี้

1. การจับสลาก เป็นวิธีที่เหมาะในกรณีที่จะเลือกตัวอย่างจากประชากรที่มีจำนวนไม่มากนักโดยวิธีเขียนสลากให้หมายเลขทุกหน่วยในประชากร แล้วจับสลากเลือกตัวอย่างที่ต้องการ โดยคละสลากให้เข้ากันดีและหยิบอย่างไม่เจาะจงตามจำนวนที่ต้องการ

2. การใช้ตารางเลขสุ่ม (Random Numbers Table) ตารางเลขสุ่มเป็นตารางตัวเลขที่นักสถิติสร้างขึ้น โดยมีเลขจำนวนเต็มตั้งแต่ 0 ถึง 9 ปรากฏอยู่ด้วยความถี่เท่าเทียมกัน และความน่าจะเป็นของตัวเลขแต่ละตัวที่จะพิมพ์ปรากฏอยู่ในที่ต่าง ๆ ของตารางมีค่าเท่ากัน การใช้ตารางเลขสุ่มนี้ นักวิจัยต้องมีรายชื่อสมาชิกทุกหน่วยในประชากรและให้หมายเลขแก่สมาชิกเหล่านั้นไว้ล่วงหน้าก่อน จากนั้นจึงสุ่มเลือกแถว (Row) หรือ สดมภ์ (Column) หรืออ่านตามแนวทะแยงเรื่อย ๆ ไปจนกว่าจะได้สมาชิกครบตามจำนวนที่ต้องการ (นงลักษณ์ วิรัชชัย, 2536)

3. การใช้คอมพิวเตอร์ (Computer) อาจใช้โปรแกรมสำเร็จรูปเลือกตัวอย่างโดยมีโปรแกรมที่ใช้จะให้เลขสุ่ม (Random Numbers) ตั้งแต่ 1 ถึง N ให้กับหน่วยตัวอย่างแต่ละหน่วยโดยไม่ซ้ำกัน เพื่อสุ่มเลือกตัวอย่างจำนวน n ตามที่ต้องการ หน่วยใดตรงกับเลขสุ่มหน่วยนั้นจะตกเป็นตัวอย่าง

จุดเด่นของการสุ่มแบบง่าย มีดังนี้

1. มีขั้นตอนการปฏิบัติง่าย ไม่ซับซ้อน

2. เป็นวิธีสุ่มที่เหมาะกับประชากรที่มีลักษณะคล้ายคลึงกันหรือลักษณะเป็นเอกพันธ์ (Homogeneity)

3. มีวิธีการประมาณค่าที่ง่ายและสะดวกต่อการคำนวณ

จุดด้อยของการสุ่มแบบง่าย มีดังนี้

1. เป็นวิธีการสุ่มที่ไม่เหมาะกับประชากรที่มีลักษณะแตกต่างกันมาก หรือมีความเป็นวิวิธพันธ์ (Heterogeneity)

2. ใช้เวลาและเสียค่าใช้จ่ายสูง เพราะอาจต้องใช้ตัวอย่างขนาดใหญ่ เพื่อควบคุมความคลาดเคลื่อนจากการสุ่มตัวอย่าง โดยเฉพาะกับประชากรที่มีความแปรปรวนมาก

2. วิธีสุ่มแบบมีระบบ (Systematic Sampling)

วิธีการสุ่มแบบมีระบบ เป็นการเลือกหน่วยตัวอย่างอย่างมีระบบจากรายชื่อสมาชิกของกลุ่มประชากรที่มีการเรียงชื่อสมาชิกโดยการเรียงลำดับแบบสุ่ม (Random Order)

วิธีการสุ่มแบบนี้ผู้วิจัยต้องรู้กรอบของประชากร (Population Frame) ก่อนว่ามีจำนวนประชากรเท่าใดแล้วให้หมายเลขประชากรและกำหนดขนาดของกลุ่มตัวอย่างที่ต้องการเพื่อคำนวณหาช่วงการสุ่ม (Sampling Interval)

วิธีดำเนินการ เมื่อกำหนดขอบเขตและนิยามกลุ่มประชากรแล้ว วิธีสุ่มแบบมีระบบสามารถปฏิบัติได้หลายวิธีดังนี้ (ศิริลักษณ์ สุวรรณวงศ์ , 2538 : 60-66)

1. วิธีการสุ่มตัวอย่างแบบมีระบบชนิดเชิงเส้น (Linear Systematic Sampling)

เป็นวิธีการสุ่มตัวอย่างขนาด n จากประชากรขนาด N โดยนำหน่วยตัวอย่างของประชากรมาเรียงให้อยู่ในลักษณะเส้นตรงเป็นหน่วยตัวอย่างที่ 1, 2 , ... ,N ทำการสุ่มตัวอย่างขนาด n โดยสุ่มจุดเริ่มสุ่มที่มีค่าตั้งแต่ 1 ถึง k เมื่อ $k = N / n$ หรือเป็นจำนวนเต็มที่ใกล้เคียงกับ N / n จุดเริ่มสุ่ม r มีช่วงการสุ่ม k ดังนั้นเมื่อสุ่มจุดเริ่มสุ่ม r ได้ก็จะได้หน่วยตัวอย่างอื่น ๆ ที่ต้องสุ่มทันที และตัวอย่างที่สุ่มได้ ได้แก่ หน่วยตัวอย่างที่ r, r+k, ..., r+jk โดย $r+jk \leq N$ และ $j = 0, 1, 2, \ldots$ สำหรับ กรณี k เป็นจำนวนที่ใกล้เคียงกับ N / n อาจทำหน่วยตัวอย่างที่ได้ไม่ครบตามต้องการ ซึ่งจะใช้วิธีสุ่มแบบมีระบบชนิดอื่นแก้ปัญหา

2. วิธีการสุ่มตัวอย่างแบบมีระบบชนิดใช้ช่วงการสุ่มเป็นเศษส่วน

เป็นวิธีการสุ่มตัวอย่างที่แก้ปัญหาการสุ่มแบบมีระบบชนิดเชิงเส้น โดย $k = N/n$ เป็นเศษส่วน ซึ่งการเกิดปัญหาเมื่อช่วงการสุ่มเป็นเศษส่วน (Fractional Interval) นี้ทำได้โดยให้หน่วยตัวอย่างที่สุ่มได้เป็นหน่วยตัวอย่างที่มีค่าใกล้เคียงกับ r+jk มากที่สุดโดย $1 \leq r \leq k$

3. วิธีการสุ่มตัวอย่างแบบมีระบบชนิดเดินหน้าและถอยหลัง (Forward - Backward)

เป็นวิธีการสุ่มตัวอย่างแบบมีระบบเชิงเส้นในกรณีที่ $k = N/n$ ไม่ใช่จำนวนเต็ม วิธีปฏิบัติคือ เลือกจุดเริ่มสุ่ม r ที่ $1 \leq r \leq N$ ให้ k เป็นจำนวนเต็มที่มีค่าใกล้เคียงกับ N/n หน่วยตัวอย่างที่สุ่มได้คือ $r + jk$; $j = 0, 1, \ldots$ และ $1 \leq r + jk \leq N$ ดังรูป

4. วิธีการสุ่มตัวอย่างแบบมีระบบชนิดวงกลม (Circular Systematic Sampling) เป็นวิธีการสุ่มตัวอย่างแบบมีระบบในกรณีที่ $k = N/n$ ไม่ใช่จำนวนเต็ม วิธีปฏิบัติคือ นำหน่วยตัวอย่างของประชากรมาเรียงในลักษณะวงกลม เป็นหน่วยตัวอย่างที่ 1, 2, ..., N แล้วสุ่มตัวอย่างขนาด n โดยสุ่มจุดเริ่มสุ่มที่มีค่าตั้งแต่ 1 ถึง N เมื่อ $k = N/n$ หรือเป็นจำนวนเต็มที่ใกล้เคียงกับ N/n หน่วยตัวอย่างที่สุ่มได้ คือ $r + jk \bmod N$ (เศษของผลหารของ $r + jk$ ด้วย N) โดยที่ $j = 0, 1, \ldots, n - 1$

5. วิธีการสุ่มแบบมีระบบชนิด 2 มิติ (Two Dimensions Systematic Sampling)

เป็นวิธีการสุ่มตัวอย่างแบบมีระบบที่เกิดกับกรอบตัวอย่างเป็นกรอบแผนที่ กล่าวคือ หน่วยตัวอย่างไม่สามารถนำมาเรียงได้ เช่น ต้องการประมาณผลผลิตเฉลี่ยของพื้นที่เพาะปลูกตำบลหนึ่ง โดยมีกรอบตัวอย่างเป็นกรอบแผนที่ ซึ่งเป็นพื้นที่เพาะปลูก ดังนั้น จึงต้องทำการสุ่มตัวอย่างซึ่งเป็นแปลงเพาะปลูก โดยมีจุดเริ่มสุ่มในแกนราบ และแกนตั้ง ลักษณะตัวอย่างที่สุ่มได้จะเป็นคู่ลำดับที่ตั้งของหน่วยตัวอย่าง แต่การสุ่มตัวอย่างแบบมีระบบชนิดนี้ปัจจุบันไม่นิยมใช้ เพราะไม่สะดวกในงานสนาม

จุดเด่นของวิธีการสุ่มแบบมีระบบ คือ

1. เป็นวิธีการสุ่มที่ทำได้ง่าย สะดวก รวดเร็ว คือ เมื่อสุ่มได้จุดเริ่มสุ่มเพียงตัวเดียว ทำให้ได้หน่วยตัวอย่างที่ตกเป็นตัวอย่างทั้งหมด

2. สามารถนำไปประกอบกับวิธีสุ่มตัวอย่างแบบอื่น ๆ ได้

3. โดยทั่วไป ถ้าหน่วยต่าง ๆ ในประชากรเรียงกันอยู่อย่างสุ่ม ทำให้วิธีการสุ่มแบบมีระบบจะให้ผลใกล้เคียงหรือสูงกว่าวิธีการสุ่มแบบง่าย

จุดด้อยของวิธีการสุ่มแบบมีระบบ คือ

1. ประสิทธิภาพของตัวประมาณ ขึ้นอยู่กับลักษณะการเรียงหน่วยตัวอย่างในประชากร

2. ถ้าสมาชิกในกลุ่มประชากรถูกจัดเรียงไว้อย่างมีระบบ หรือเรียงลำดับแบบสุ่ม จะทำให้กลุ่มตัวอย่างไม่เป็นตัวแทนที่ดี มีความลำเอียง

3. วิธีการสุ่มแบบแบ่งชั้น (Stratified Sampling)

เป็นวิธีการสุ่มตัวอย่างที่แบ่งกลุ่มประชากรออกเป็นชั้นภูมิย่อย ๆเสียก่อน บนพื้นฐานของตัวแปรที่ส่งผลกระทบต่อตัวแปรตาม แล้วจึงสุ่มตัวอย่าง

วิธีดำเนินการ เมื่อกำหนดขอบเขตและนิยามกลุ่มประชากรเฉพาะการวิจัยแล้ว จากนั้นแบ่งกลุ่มประชากรออกเป็นชั้นภูมิ (Strata) เสียก่อน โดยใช้ตัวแปรที่สนใจศึกษาเป็นเกณฑ์ในการแบ่ง มีหลักการให้ประชากรกลุ่มย่อยมีความเป็นวิวิธพันธ์ระหว่างชั้นภูมิ(กลุ่ม)สูง หรือมีความแตกต่างกันมากที่สุดเท่าที่จะทำได้ และภายในชั้นภูมิแต่ละชั้นมีความเป็นเอกพันธ์หรือมีลักษณะเหมือนกันให้มากที่สุดเท่าที่จะทำได้ สำหรับสมาชิกแต่ละหน่วยในกลุ่มประชากรแต่ละชั้นภูมิอาจใช้วิธีการสุ่มแบบง่าย หรือวิธีการสุ่มแบบมีระบบ อย่างใดก็ได้ตามความเหมาะสม ถ้าสุ่มหน่วยตัวอย่างจากแต่ละชั้นภูมิด้วยวิธีการสุ่มตัวอย่างแบบง่ายทุกชั้นภูมิ จะเรียกว่า การสุ่มตัวอย่างแบบแบ่งช่วงชั้นอย่างง่าย (Stratified Simple Random Sampling) ถ้าสุ่มหน่วยตัวอย่างจากแต่ละช่วงชั้นด้วยวิธีการสุ่มแบบมีระบบ จะเรียกว่า การสุ่มตัวอย่างแบบ

แบ่งช่วงชั้นแบบมีระบบ (Stratified Systematic Sampling) (สำหรับการวิจัยครั้งนี้จะเรียกว่า วิธีการสุ่มแบบแบ่งชั้นและสุ่มตัวอย่างย่อยแบบง่าย และวิธีการสุ่มแบบแบ่งชั้นและสุ่มตัวอย่างย่อยแบบมีระบบ)

ตัวแปรที่จะใช้เป็นหลักในการแบ่งชั้นของกลุ่มประชากร ควรเลือกใช้ตัวแปรที่มีความสัมพันธ์สูงกับตัวแปรที่ศึกษา (ตัวแปรอิสระ และ/หรือ ตัวแปรตามของการวิจัยนั้น) ตัวแปรที่ใช้ในการแบ่งชั้นหรือชั้นภูมิ ที่สามารถแบ่งช่วงชั้นหรือทำให้ประชากรในระหว่างชั้นภูมิมีความแตกต่างกัน หรือมีความเป็นวิวิชพันธ์ระหว่างชั้นภูมิสูง ความเป็นเอกพันธ์ภายในกลุ่มหรือภายในชั้นภูมิสูง จะทำให้ตัวประมาณมีประสิทธิภาพสูงและความคลาดเคลื่อนของตัวอย่างมีค่าน้อยลง (นงลักษณ์ วิรัชชัย , 2536 ; บุญเรียง ขจรศิลป์ , 2534)

วิธีกำหนดขนาดตัวอย่างในแต่ละชั้นภูมิ

วิธีการสุ่มตัวอย่างแบบแบ่งชั้น นอกเหนือจากขนาดตัวอย่างทั้งหมดที่ผู้เลือกตัวอย่างต้องกำหนดแล้ว จะต้องทำการกำหนดขนาดตัวอย่างของแต่ละชั้นภูมิด้วย ทั้งนี้ขนาดตัวอย่างของชั้นภูมิมีความสำคัญในแง่ของการกำหนดคุณภาพของตัวประมาณการลดความแปรปรวนของค่าประมาณ

การกำหนดขนาดตัวอย่างของชั้นภูมิ อาจกระทำได้ในหลายหลักการกล่าวคือ หากไม่มีความต้องการใช้ขนาดตัวอย่างเป็นกลไกควบคุมความแปรปรวน ก็อาจกำหนดขนาดตัวอย่างของชั้นภูมิอย่างไรก็ได้ หรือ ตามความสะดวก หรือความพอใจ และยอมรับผลที่จะเกิดตามมา เช่น การกำหนดขนาดตัวอย่างตามความพอใจของผู้เลือกตัวอย่าง หรือกำหนดขนาดตัวอย่างให้เท่ากันทุกชั้นเพื่อความสะดวกเป็นต้น แต่ถ้าหากต้องการใช้ประโยชน์ขนาดตัวอย่างของชั้นภูมิ หรือ กำหนดอย่างมีหลักเกณฑ์ เพื่ออาจจะนำมาใช้เพิ่มคุณภาพตัวประมาณได้

ขนาดตัวอย่างของชั้นภูมิ ควรแปรตามระดับความแปรปรวนของชั้นภูมิ ขนาดของชั้นภูมิ และค่าใช้จ่ายต่อหน่วยในชั้นภูมิโดยเปรียบเทียบ ทั้งนี้จะต้องคำนึงถึงความเป็นไปได้ในเชิงของข้อมูลที่มีอยู่ หรือที่สามารถหาได้โดยไม่ต้องเสียค่าใช้จ่ายมากนัก (สุชาดา กีระนันทน์ , 2538)

การกำหนดขนาดตัวอย่างในแต่ละชั้นภูมิสามารถทำได้ดังนี้

1. การกำหนดขนาดกลุ่มตัวอย่างตามสัดส่วนของชั้นภูมิ(Proportional Allocation)

หลักของวิธีการนี้มีว่า ชั้นภูมิใดมีขนาดใหญ่กว่า ก็จะมีขนาดตัวอย่างมากกว่าชั้นภูมิที่มีขนาดเล็กกว่า ต้องการให้สัดส่วนตัวอย่างในทุกชั้นภูมิเท่ากันหมด ทั้งนี้ให้ขนาดตัวอย่างของชั้นภูมิเป็นสัดส่วนโดยตรงกับขนาดชั้นภูมิ ซึ่งสามารถเขียนในรูปสูตร ดังนี้

$$n_h = nN_h / N$$ โดยที่ h = 1, 2, ..., L

(Yamane, 1967 : 125)

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อ n_h คือ ขนาดของกลุ่มตัวอย่างในชั้นภูมิที่ h
n คือ ขนาดของกลุ่มตัวอย่าง
N_h คือ จำนวนประชากรในชั้นภูมิที่ h
N คือ จำนวนประชากร

สำหรับการกำหนดขนาดตัวอย่างของชั้นภูมิ เช่นนี้ จะเห็นได้ว่า ขนาดตัวอย่างของชั้นภูมิขึ้นอยู่กับขนาดชั้นภูมิเพียงประการเดียว ถึงแม้ว่าจะสะดวกในการกำหนด แต่ก็ไม่ได้นำลักษณะการกระจายของค่าในชั้นภูมิหรือความแตกต่างระหว่างชั้นภูมิ และค่าใช้จ่ายต่อหน่วยในแต่ละชั้นภูมิมาพิจารณาเลย ซึ่งในกรณีที่ชั้นภูมิต่าง ๆ มีความกระจายภายในชั้นภูมิต่างกันหรือค่าใช้จ่ายต่อหน่วยไม่เท่ากัน คุณภาพของตัวประมาณอาจไม่ดีเท่าที่ควร

2. การกำหนดขนาดตัวอย่างแบบอุตมะหรือเหมาะสมที่สุดภายใต้เงื่อนไข (Optimum Allocation)

วิธีนี้เป็นกรณีที่มีวัตถุประสงค์ที่จะใช้การกำหนดขนาดตัวอย่างของชั้นภูมิเป็นกลไกในการเพิ่มความแม่นยำของตัวประมาณให้สูงที่สุด

หลักการคือ ภายใต้ข้อจำกัดด้านทรัพยากรที่มี ซึ่งอาจพิจารณาในรูปของงบประมาณทั้งหมดที่จะใช้ในการเก็บข้อมูลตัวอย่างหรือขนาดตัวอย่างรวมทั้งหมด ผู้เลือกตัวอย่างจะจัดสรรขนาดตัวอย่างให้ชั้นภูมิอย่างไรจึงจะทำให้ค่าแปรปรวนของตัวประมาณต่ำที่สุด นั่นคือขนาดตัวอย่างจะแปรตามขนาดของชั้นภูมิกับค่าความแปรปรวนในชั้นภูมิ และแปรผกผันกับค่าใช้จ่ายต่อหน่วยในชั้นภูมิ กล่าวคือ ขนาดตัวอย่างจะมีขนาดใหญ่ขึ้น หากชั้นภูมินั้นมีขนาดใหญ่ขึ้น หรือค่าความแปรปรวนของชั้นภูมินั้นสูงขึ้น หรือค่าใช้จ่ายต่อหน่วยในชั้นภูมินั้นลดลงนั่นเอง

สำหรับสูตรในการคำนวณขนาดตัวอย่าง ของแต่ละชั้นภูมิ มีดังนี้

$$n_h = \frac{nN_hS_h/\sqrt{C_h}}{\sum_{h=1}^{L} N_hS_h/\sqrt{C_h}}$$

โดยที่ h = 1, 2, 3, ..., L

(Cochran, 1977 : 98)

เมื่อ n_h คือ ขนาดของกลุ่มตัวอย่างในชั้นภูมิที่ h
S_h คือ ส่วนเบี่ยงเบนมาตรฐานของประชากรย่อยในชั้นภูมิที่ h
C_h คือ ค่าใช้จ่ายต่อหน่วยในชั้นภูมิที่ h
L คือ จำนวนชั้นภูมิ
n คือ ขนาดของกลุ่มตัวอย่าง

ในกรณีค่าใช้จ่ายต่อหน่วย ในการสำรวจแต่ละชั้นภูมิไม่แตกต่างกัน ความแปรปรวนในแต่ละชั้นภูมิคงที่หรือไม่แตกต่างกัน จะพบว่าการกำหนดขนาดตัวอย่างแบบอุตมะนี้ คือ การกำหนดขนาดตัวอย่างตามสัดส่วนของชั้นภูมินั่นเอง

3. วิธีกำหนดขนาดตัวอย่างแบบเท่าเทียมกัน (Equal Allocation)

เป็นการกำหนดขนาดตัวอย่างออกเป็นส่วน ส่วนละเท่ากันในแต่ละชั้นภูมิ วิธีนี้เหมาะสมในกรณีที่ชั้นภูมิต่าง ๆ มีขนาดเท่ากันหรือใกล้เคียงกัน มีสูตรในการกำหนดขนาดตัวอย่าง ดังนี้

$$n_h = n / L \qquad \text{โดยที่ h} = 1,2,3,...,L$$

(Yamane, 1967 : 142)

n_h คือ ขนาดของกลุ่มตัวอย่างในชั้นภูมิที่ h

n คือ ขนาดของกลุ่มตัวอย่าง

L คือ จำนวนชั้นภูมิ

4. วิธีกำหนดขนาดตัวอย่างแบบเนย์แมน (Neyman Allocation)

เป็นวิธีการกำหนดขนาดตัวอย่าง ที่ถือว่าค่าใช้จ่ายต่อหน่วยในการสำรวจแต่ละชั้นภูมิ มีค่าใกล้เคียงกันหรือ ไม่แตกต่างกันอย่างเห็นได้ชัดเจน หรือเท่ากัน

การคำนวณขนาดตัวอย่างของแต่ละชั้นภูมิ มีดังนี้

$$n_h = \frac{nN_hS_h}{\sum_{h=1}^{L} N_hS_h} \qquad \text{โดยที่ h} = 1, 2, 3, ..., L$$

(Yamane, 1967 : 136)

เมื่อ n_h คือ ขนาดของกลุ่มตัวอย่างในชั้นภูมิที่ h

S_h คือ ส่วนเบี่ยงเบนมาตรฐานของประชากรย่อยในชั้นภูมิที่ h

L คือ จำนวนชั้นภูมิ

n คือ ขนาดของกลุ่มตัวอย่าง

จากสูตร จะพบว่า ขนาดของตัวอย่างในแต่ละชั้นภูมิจะแปรผันตามขนาดของชั้นภูมิและความแปรปรวนภายในชั้นภูมิ คือ ชั้นภูมิใดที่มีขนาดใหญ่กว่า (N_h) ชั้นภูมินั้นจะมีขนาดตัวอย่าง (n_h) มากกว่า และชั้นภูมิใดที่มีความแปรปรวนภายใน (S_h) มาก ชั้นภูมินั้นจะมีขนาดตัวอย่างที่มากกว่าชั้นภูมิที่มีความแปรปรวนภายในน้อยกว่า

สำหรับในการวิจัยครั้งนี้ เลือกใช้วิธีกำหนดขนาดตัวอย่างแบบเนย์แมน เนื่องจากผลการวิจัยของ ดวงใจ ปวีณอภิชาต (2535) พบว่า เป็นวิธีการกำหนดขนาดตัวอย่างที่มีประสิทธิภาพมากที่สุดในการประมาณค่ามัชฌิมเลขคณิตและประมาณค่าความแปรปรวนของคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของประชากร

จุดเด่นของวิธีการสุ่มแบบแบ่งชั้น

1. การจัดแบ่งชั้นภูมิที่ใช้ตัวแปรอิสระหลายตัวเป็นเกณฑ์ ในการจัดแบ่งชั้นภูมิ จะยิ่งช่วยลดความคลาดเคลื่อนในการประมาณค่า

2. การสุ่มตัวอย่างจากชั้นภูมิในแต่ละชั้น สามารถใช้วิธีต่างกันไม่จำเป็นต้องใช้วิธีเดียวกัน ซึ่งเป็นประโยชน์ในการปฏิบัติมาก เพราะในบางชั้นภูมิแต่ละชั้นมีลักษณะแตกต่างกันมาก สามารถที่จะใช้วิธีสุ่มตัวอย่างที่เหมาะสมกับแต่ละชั้นภูมิ

3.กลุ่มตัวอย่างที่ได้เป็นตัวแทนของประชากรทุกลักษณะเพราะหน่วยตัวอย่างแต่ละชั้นภูมิถูกสุ่มมาเป็นกลุ่มตัวอย่าง

4. ทำให้กลุ่มตัวอย่างสามารถนำมาวิเคราะห์ข้อมูล และเสนอผลวิจัยเพื่อตอบคำถามหรือวัตถุประสงค์ที่กำหนดไว้อย่างแน่นอน คือ สามารถเสนอผลตามตัวแปรอิสระต่างๆ ที่ใช้แบ่งชั้นภูมิ เช่น เพศ คณะ สาขาวิชาที่ศึกษา ฯลฯ

จุดด้อยของวิธีการสุ่มแบบแบ่งชั้น

1. การใช้ตัวแปรหลายตัวเป็นเกณฑ์ในการแบ่งชั้นภูมิอาจจะทำให้เกิดปัญหาที่ตามมาคือมีบางชั้นภูมิที่หาสมาชิกไม่ได้ หรือมีน้อยมาก

2. สูตรในการคำนวณหาขนาดกลุ่มตัวอย่าง สูตรในการประมาณค่า ฯลฯ มีความยุ่งยาก

3. เนื่องจากเป็นวิธีการที่มีการแบ่งชั้นภูมิ ฉะนั้น งบประมาณ ค่าใช้จ่าย แรงงาน จึงเพิ่มขึ้นเนื่องจากต้องเก็บข้อมูลตามตัวแปรที่ใช้แบ่งชั้น ยิ่งมีหลายตัวแปรก็ยิ่งมีความสิ้นเปลืองหรือมีงานเพิ่มขึ้น ทั้งในส่วนของการเก็บรวบรวมและการวิเคราะห์ข้อมูล

4. วิธีการสุ่มแบบตามกลุ่ม หรือ แบบแบ่งพื้นที่ (Cluster or Area Sampling)

เป็นวิธีการสุ่มกลุ่มประชากรมีลักษณะรวมกันเป็นกลุ่มก้อนหลายกลุ่ม โดยที่แต่ละกลุ่มมีลักษณะเป็นเอกพันธ์ระหว่างกลุ่มสูง แต่ลักษณะเอกพันธ์ระหว่างสมาชิกในกลุ่มต่ำ ลักษณะการจัดกลุ่มดังกล่าวนี้ ตรงกันข้ามกับลักษณะการแบ่งชั้นของกลุ่มประชากรในการเลือกกลุ่มตัวอย่างโดยวิธีสุ่มแบบแบ่งชั้น ซึ่งมีลักษณะเอกพันธ์ระว่างชั้นภูมิสูง แต่มีลักษณะเอกพันธ์

ระหว่างสมาชิกในชั้นสูง เนื่องจากชั้นของกลุ่มประชากรในการสุ่มแบบแบ่งชั้นมีความเป็นวิวิธพันธ์สูง จึงต้องเลือกจากทุกชั้นประชากรเป็นกลุ่มตัวอย่าง แต่ในการสุ่มแบบแบ่งกลุ่ม กำหนดวิธีการในการแบ่งกลุ่มประชากรให้มีความเป็นเอกพันธ์สูง ดังนั้น จึงไม่ต้องศึกษาประชากรทุกกลุ่ม แต่สุ่มเลือกมาศึกษาตามจำนวนที่ต้องการ และใช้สมาชิกทุกหน่วยในแต่ละกลุ่มเป็นหน่วยตัวอย่าง เนื่องจากสมาชิกในกลุ่มมีลักษณะเป็นวิวิธพันธ์สูง อย่างไรก็ดีอาจสุ่มเลือกสมาชิก (หน่วยตัวอย่าง) จากกลุ่มประชากรที่เลือกไว้โดยไม่ใช้สมาชิกทุกหน่วยในกลุ่มก็ได้

วิธีดำเนินการ เมื่อกำหนดขอบเขตและนิยามกลุ่มประชากรเฉพาะการวิจัยแล้ว จัดแบ่งประชากรออกเป็นกลุ่มย่อย ๆ ตามเกณฑ์อย่างใดอย่างหนึ่ง ยึดหลักให้สมาชิกภายในกลุ่มมีความเป็นวิวิธพันธ์ (Heterogeniety within group) สูง และมีความเป็นเอกพันธ์ระหว่างกลุ่ม (Homogeniety between proup) มากที่สุด

การแบ่งกลุ่มประชากรโดยทั่วไปนิยมแบ่งตามลักษณะเส้นที่ทางภูมิศาสตร์หรือเขตการปกครอง เช่น แบ่งเป็นภาค จังหวัด อำเภอ/เขต ตำบล/แขวง และหมู่บ้าน เป็นต้น การแบ่งกลุ่มประชากรในลักษณะดังกล่าว จะช่วยให้ประหยัดค่าใช้จ่ายในการเดินทางและการรวบรวมข้อมูลได้มากกว่าการแบ่งกลุ่มประชากรตามลักษณะอื่น ๆ แต่มิได้หมายความว่าการแบ่งกลุ่มประชากรจะตัดแบ่งตามลักษณะทางภูมิศาสตร์หรือเขตการปกครองเสมอไป (นงลักษณ์ วิรัชชัย , 2536 ; พวงรัตน์ ทวีรัตน์ , 2531)

จุดเด่นของวิธีการสุ่มแบบตามกลุ่ม

1. ประหยัดเวลา แรงงานและค่าใช้จ่าย เริ่มตั้งแต่ การเตรียมกรอบตัวอย่างตลอดจน การเก็บรวบรวมข้อมูล
2. ใช้ได้กับประชากรที่ไม่มีกรอบตัวอย่าง หรือไม่สะดวกในการจัดเตรียมกรอบตัวอย่าง (Population Frame)

จุดด้อยของวิธีการสุ่มแบบตามกลุ่ม

1. มีความยุ่งยากซับซ้อนในการประมาณค่าพารามิเตอร์
2. มีความยุ่งยากในการแบ่งกลุ่มให้แต่ละกลุ่มมีลักษณะเป็นวิวิธพันธ์และภายในกลุ่มมีลักษณะเป็นเอกพันธ์ ในกรณีที่กลุ่มไม่ได้แบ่งอยู่ตามธรรมชาติ
3. ความแปรปรวนของตัวประมาณมีหลายส่วน ซึ่งเท่ากับจำนวนขั้นตอนของการสุ่มตัวอย่าง อาจทำให้ความแปรปรวนมีมากซึ่งมีผลต่อประสิทธิภาพของตัวประมาณ (ศิริลักษณ์ สุวรรณวงศ์ , 2538)

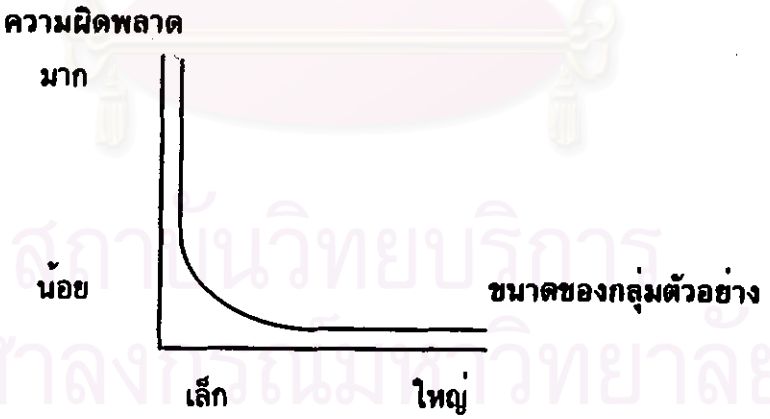
นอกจากนั้น ยังมีวิธีการสุ่มตัวอย่างที่เรียกว่า วิธีการสุ่มตัวอย่างแบบหลายขั้นตอน (Multi - Stage Sampling) อีกวิธีการหนึ่ง ซึ่งการเรียกโดยรวมของวิธีการสุ่มที่มีการเลือก หรือมีวิธีสุ่มมากกว่า 1 ครั้ง หรือหลายวิธีการ หลายขั้นตอน เช่น การสุ่มแบบตามกลุ่มที่มีหลายขั้นตอน, การสุ่มแบบแบ่งชั้นที่มีหลายขั้นตอน, การสุ่มตัวอย่างที่ขั้นตอนแรกใช้แบบแบ่งชั้น ขั้นตอนที่สองใช้แบบตามกลุ่ม หรือการสุ่มตัวอย่างที่มีลักษณะการจัดการประชากรเป็นแบบแบ่งชั้นหรือแบ่งกลุ่ม แล้วสุ่มกลุ่มประชากรย่อยเป็นแบบง่ายหรือแบบมีระบบ ดังเช่นการวิจัยครั้งนี้เป็นต้น

ตอนที่ 3 การกำหนดขนาดตัวอย่าง (Sample Size)

เมื่อเลือกวิธีสุ่มตัวอย่างแล้ว ต่อมาคือ การกำหนดขนาดกลุ่มตัวอย่างที่เหมาะสมเพื่ออ้างอิงไปสู่ประชากรแล้วมีความเชื่อถือได้มากที่สุด หรือมีความคลาดเคลื่อนหรือมีความผิดพลาดน้อยที่สุด

เคอร์ลิงเจอร์ (Kerlinger, 1972 : 61) ได้อธิบายถึงขนาดของกลุ่มตัวอย่างที่จะมีผลต่อความคลาดเคลื่อนหรือความผิดพลาดของผลวิจัยโดยแสดงเป็นกราฟไว้ดังนี้

แผนภาพที่ 1 ความสัมพันธ์ระหว่างขนาดตัวอย่างกับความผิดพลาดหรือความคลาดเคลื่อนของผลวิจัย



จากกราฟ อธิบายได้ว่า ถ้ากลุ่มตัวอย่างมีขนาดเล็ก ความผิดพลาดก็จะมีมาก ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ขึ้น ความผิดพลาดก็จะน้อยลง แสดงว่า ความผิดพลาดแปรผกผันกับขนาดของกลุ่มตัวอย่าง

ในการทำวิจัยจึงควรที่จะใช้กลุ่มตัวอย่างที่มีขนาดใหญ่ เพื่อให้เกิดความผิดพลาดน้อยที่สุด แต่ทั้งนี้การใช้กลุ่มตัวอย่างขนาดใหญ่เกินไป จะเป็นการสิ้นเปลืองงบประมาณ เวลา แรงงาน

ดังนั้น เมื่อมีความจำเป็นต้องใช้ขนาดตัวอย่างให้เหมาะสมโดยคำนึงถึงองค์ประกอบต่าง ๆ ร่วมกัน ทั้งนี้ McCall (1982 : 187-190 อ้างถึงใน สมชัย วงษ์นายะ, 2535) ได้สรุปเกี่ยวกับองค์ประกอบที่มีอิทธิพลต่อการกำหนดขนาดของกลุ่มตัวอย่าง ที่ผู้วิจัยควรคำนึงถึง คือ

1. ลักษณะประชากรและคุณลักษณะที่ต้องการวัด (The Population and Characteristics To Be Measured) องค์ประกอบนี้จะเป็นองค์ประกอบแรกที่ผู้วิจัยต้องพิจารณาในการกำหนดขนาดของกลุ่มตัวอย่าง ถ้าประชากรมีลักษณะเป็นวิวิธพันธ์ (Heterogeneous) มากเท่าไร ก็จำเป็นต้องใช้กลุ่มตัวอย่างมากขึ้น แต่ถ้าประชากรยิ่งมีความเป็นเอกพันธ์ (Homogeneous) มากเท่าไร ก็ทำให้สามารถใช้กลุ่มตัวอย่างได้น้อยลงมาก

นอกจากนี้จำนวนกลุ่มย่อย (Subgroup) ที่มุ่งศึกษา เช่น แยกเป็นกลุ่มเพศชาย เพศหญิง ซึ่งถ้ายิ่งมีจำนวนกลุ่มย่อยที่ต้องการศึกษามากเท่าไร ขนาดของกลุ่มตัวอย่างก็จำเป็นต้องมากขึ้นตาม

2. ข้อความรู้ที่นำมาใช้ (Summary Measures To Be Used) ในส่วนนี้จะพิจารณาถึงค่าสถิติที่ใช้แสดงคุณลักษณะของสิ่งที่สนใจศึกษาที่สำคัญ เช่น อาจจะใช้ค่าเฉลี่ย ผลรวม สัดส่วน เป็นต้น ซึ่งค่าต่าง ๆ ที่สนใจจะมีผลต่อการกำหนดขนาดของกลุ่มตัวอย่าง เพราะความคลาดเคลื่อนมาตรฐาน (Standard Error) เป็นพื้นฐานสำคัญต่อสูตรที่ใช้คำนวณหาขนาด ของกลุ่มตัวอย่าง

3. ความแม่นยำและความเชื่อมั่น (Precision and Confidence) ซึ่งผู้วิจัยต้องการในงานวิจัยครั้งนั้น

4. การวิเคราะห์ที่นำมาใช้ (Analyses To Be Made) การที่จะวิเคราะห์ข้อมูลหรือนำเสนอผลการวิเคราะห์ข้อมูลอย่างไร จะมีผลต่อการกำหนดขนาดของกลุ่มตัวอย่าง

5. ข้อพิจารณาอื่น ๆ ที่ไม่ใช่ทางสถิติ (Nonstatistical Considerations) เช่น งบประมาณ เวลา บุคลากร เครื่องมือในการเก็บรวบรวมข้อมูล และรายงานประชากรหรือกรอบตัวอย่าง (Sampling Frames)

วิธีการกำหนดขนาดของกลุ่มตัวอย่างนั้นมีหลายวิธี ซึ่งวิธีการสุ่มตัวอย่างแต่ละวิธีก็มีวิธีการประมาณขนาดของกลุ่มตัวอย่างด้วยสูตรที่แตกต่างกัน อย่างไรก็ตามสามารถสรุปวิธีกำหนดขนาดของตัวอย่างไว้ดังนี้ (อุทุมพร จามรมาน, 2537 : 25-45 ; นิยม ปุราคำ, 2517 : 117)

1. ใช้ตารางสำเร็จรูป

1.1 ตารางสำเร็จรูปของ Snedecor

เป็นตารางสำเร็จรูปที่เหมาะกับการใช้กำหนดขนาดของกลุ่มตัวอย่างที่เกี่ยวข้องกับข้อมูลประเภทแบ่งสอง (Dichotomous) ซึ่งมีลักษณะ เช่น เอา-ไม่เอา ชอบ-ไม่ชอบ อยู่-ไม่อยู่ หญิง-ชาย เป็นต้น วิธีใช้ตารางของ Snedecor ผู้ใช้จะต้องประมาณค่าพารามิเตอร์ของสัดส่วนก่อนว่าจะเป็นเท่าไร

1.2 ตารางสำเร็จรูปของ Yamane

ตารางสำเร็จรูปของ Yamane เหมาะสำหรับการประมาณสัดส่วน ตารางของ Yamane กำหนดระดับความเชื่อมั่นไว้ 2 ค่า คือ 95% กับ 99% ในการใช้ตารางนี้ ผู้ใช้ต้องประมาณขนาดของประชากร (N) และความคลาดเคลื่อนก่อน

2. ใช้สูตร

2.1 ถ้าข้อมูลเป็นแบบทวินาม (Dichotomous Data)

ใช้ในกรณีที่จุดมุ่งหมายของการวิจัยต้องการประมาณค่าสัดส่วน โดยมีสูตรดังนี้

$$n = (\frac{Z}{e})^2 (p)(1-p)$$

โดยที่ n คือ ขนาดของกลุ่มตัวอย่าง

Z คือ คะแนนมาตรฐานที่เกี่ยวข้องกับระดับความเชื่อมั่น (เปิดจากตารางโค้งปกติ)

p คือ สัดส่วนในประชากร

e คือ ความคลาดเคลื่อนในการสุ่ม

2.2 ถ้าข้อมูลเป็นแบบต่อเนื่อง (Continuous Data)

ในกรณีที่จุดมุ่งหมายของการวิจัยต้องการประมาณค่ามัชฌิมเลขคณิต ใช้สูตรดังนี้

$$n = (Z\sigma / e)$$

โดยที่ n คือ ขนาดของกลุ่มตัวอย่าง

σ คือ ส่วนเบี่ยงเบนมาตรฐานของประชากร

e คือ ความแตกต่างระหว่างค่ามัชฌิมเลขคณิตของกลุ่มตัวอย่างกับประชากร

สูตรนี้ใช้เมื่อต้องการประมาณค่ามัชฌิมเลขคณิตของประชากร โดยที่ประชากรมีขนาดใหญ่หรือประชากรมีจำนวนไม่จำกัด (Large or Infinite Population) และยังเหมาะสม

ในสถานการณ์อีก 2 ตัวอย่าง คือ 1) เมื่อประชากรที่มีขนาดเล็ก ได้รับการสุ่มแบบแทนที่ และ 2) เมื่อผู้วิจัยไม่สามารถประมาณขนาดของประชากรได้ สำหรับสูตรนี้ถึงแม้จะง่ายในการใช้แต่ก็จะทำให้ในกรณีที่ประชากรมีขนาดปานกลางและเล็ก จะได้ขนาดของกลุ่มตัวอย่างใหญ่กว่าความจำเป็น (McCall 1982:196-199)

2.3 ใช้สูตร ที (t)

$$n = (ts / e)$$

โดยที่ n คือ ขนาดของกลุ่มตัวอย่าง

t คือ ค่าสถิติทดสอบ t จากตาราง t ตามระดับความเชื่อมั่นที่กำหนด

s คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

e คือ ค่าความแตกต่างระหว่างค่ามัชฌิมเลขคณิตของกลุ่มตัวอย่างกับประชากรที่ยอมรับได้

2.4 ใช้สูตร ที (t) ตามวิธีของ Jaeger (1984)

สูตรนี้ใช้ในกรณีที่จุดมุ่งหมายของการวิจัยต้องการประมาณค่ามัชฌิมเลขคณิตของประชากรในวิธีการสุ่มอย่างง่าย โดยมีสูตรดังนี้

$$n = \frac{(ts / e)^2}{1 + (\frac{1}{N})(ts / e)^2}$$

โดยที่ n คือ ขนาดของกลุ่มตัวอย่าง

t คือ ค่าสถิติทดสอบ ที ตามระดับความเชื่อมั่นที่กำหนด

s คือ ค่าประมาณของส่วนเบี่ยงเบนมาตรฐานของประชากร

e คือ ความคลาดเคลื่อนของค่าสถิติจากค่าพารามิเตอร์

N คือ ขนาดของประชากร

สำหรับสูตรของ Jaeger นี้ ใช้ในกรณีที่ต้องการคำนวณหาขนาดของกลุ่มตัวอย่างที่ค่าพารามิเตอร์ คือ ค่ามัชฌิมเลขคณิตของประชากรในวิธีการสุ่มอย่างง่าย (Jaeger, 1984 : 44-45)

2.5 ใช้สูตรสัดส่วน (P)

ในกรณีที่ข้อมูลเป็นสัดส่วนใช้สูตรสัดส่วนเพื่อคำนวณขนาดของกลุ่มตัวอย่าง

$$n = \frac{Z^2 PQ}{e^2}$$

โดยที่ Z คือ ค่าสถิติทดสอบ ซี ที่ระดับความเชื่อมั่นที่กำหนด
P คือ สัดส่วนที่ได้จากข้อมูล
Q คือ 1-P
e คือ ความคลาดเคลื่อนของค่าสถิติจากค่าพารามิเตอร์

2.6 ใช้สูตร Casley และ Lury

$$n = k^2 v^2 / D^2$$

โดยที่ n คือ ขนาดของกลุ่มตัวอย่าง
k คือ Normal Deviate หรือค่า Z,t ที่ระดับความเชื่อมั่นที่กำหนด
V คือ ความแปรปรวนในประชากร ใช้สูตร V=PQ
เมื่อ P คือ ร้อยละของเหตุการณ์หรือคนที่ยอมรับ Q คือ 1-P
D คือ ร้อยละของระดับความมีนัยสำคัญ (α %)
หรือร้อยละของระดับความเชื่อมั่น (1-α %)

2.7ใช้สูตรคำนวณขนาดตัวอย่างในการสำรวจข้อมูลในระบบ Random Sampling

$$n = \frac{Nk^2\sigma_x^2}{NE^2 + k^2\sigma_x^2}$$ (Yamane , 1967)

เมื่อ n คือ ขนาดของกลุ่มตัวอย่างที่เหมาะสม
N คือ ขนาดของประชากร
k คือ ค่าคงที่ตามระดับความเชื่อมั่น
E คือ ค่าความคลาดเคลื่อนสูงสุดที่ยอมให้เกิดขึ้นได้ซึ่งเป็นความคลาดเคลื่อนระหว่างค่าสถิติที่ใช้เป็นตัวประมาณค่า (θ) กับค่าที่แท้จริง (θ) หรือ E = $|\theta - \theta|$
σ^2_x คือ ส่วนเบี่ยงเบนมาตรฐานของประชากร

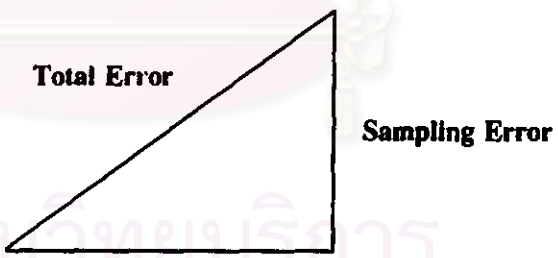
สูตรที่ใช้ในการคำนวณขนาดของกลุ่มตัวอย่าง โดยทั่วไปขึ้นอยู่กับว่า การวิจัยนั้นใช้วิธีการสุ่มตัวอย่างวิธีใด ต้องการประมาณค่าพารามิเตอร์ในส่วนใด เช่น ผลรวม ค่าเฉลี่ย สัดส่วน ร้อยละ ของประชากร เป็นต้น

ตอนที่ 4 ความคลาดเคลื่อนที่เกิดจากการสำรวจตัวอย่าง (Component of Sample Survey Error)

ในการวิจัยเรื่องหนึ่ง ๆ ผู้วิจัยมักมุ่งเน้นที่จะควบคุมคุณภาพของงานวิจัย โดยเฉพาะในงานสำรวจว่าผลการสำรวจมีความถูกต้องน่าเชื่อถือเพียงใด ซึ่งส่วนหนึ่งขึ้นอยู่กับว่างานวิจัยนั้นมีความคลาดเคลื่อนที่เกิดขึ้นจากการสำรวจตัวอย่างมากน้อยเพียงใด ในการสำรวจด้วยตัวอย่างจะมีความคลาดเคลื่อนที่อาจเกิดขึ้นได้ ความคลาดเคลื่อนทั้งหมด (Total Error) ประกอบด้วยความคลาดเคลื่อนที่เป็นอิสระจากกัน 2 แหล่ง คือ ความคลาดเคลื่อนจากการสุ่ม (Sampling Error) และความคลาดเคลื่อนที่ไม่ได้เกิดจากการสุ่ม (Non-Sampling Error) ดังรูป (ศิริชัย กาญจนวาสี, 2533)

แผนภาพที่ 2 ความคลาดเคลื่อนที่เกิดจากการสำรวจด้วยตัวอย่าง

Total Error of Estimate = Sampling Error + Non - Sampling Error



1. **ความคลาดเคลื่อนที่เกิดจากการสุ่มตัวอย่าง** (Sampling Error) (อภิชาติ พงษ์ศรีหดุลชัย, 2530) เป็นความคลาดเคลื่อนที่เกิดขึ้น เพราะไม่ได้ศึกษาจากทุกหน่วยของประชากร ดังนั้น ค่าที่ประมาณได้จากตัวอย่างจึงมักจะแตกต่างไปจากค่าที่แท้จริงของประชากรเสมอ ไม่มากก็น้อย

ความคลาดเคลื่อนประเภทนี้ อาจแบ่งตามสาเหตุและลักษณะของความคลาดเคลื่อนเป็น 2 ส่วน คือ (นิยม ปุราคำ, 2517 : 313 ; ศิริลักษณ์ สุวรรณวงค์, 2538 : 241)

1.1 ความคลาดเคลื่อนที่เกิดจากการใช้ตัวอย่างชุดหนึ่ง ๆ ประมาณค่าพารามิเตอร์

เนื่องจากการสุ่มตัวอย่างเป็นการหาตัวแทนเพียงชุดเดียว เพื่อประมาณค่าพารามิเตอร์ ซึ่งค่าสถิติที่คำนวณได้จากตัวอย่างนั้นจะมีค่าแตกต่างจากค่าพารามิเตอร์ ค่าประมาณที่ได้ (เช่น ค่าประมาณยอดรวมประชากร หรือค่าประมาณค่าเฉลี่ยประชากร) จากการสุ่มตัวอย่างเพียงชุดเดียวนี้จากโอกาสที่เป็นไปได้ทั้งหมดหลาย ๆ แบบก็ตาม ย่อมคลาดเคลื่อนจากค่าที่แท้จริง

สำหรับความคลาดเคลื่อนส่วนนี้วัดได้จาก ค่าความแปรปรวนจากการสุ่ม (Sampling Variance) หรือความแปรปรวนของตัวประมาณนั้น ๆ ความคลาดเคลื่อนส่วนที่วัดได้นี้จะมีมากหรือน้อยขึ้นอยู่กับ วิธีการเลือกตัวอย่างที่ใช้ ขนาดตัวอย่างที่ใช้ สูตรที่ใช้ในการประมาณค่า

ดังนั้น การควบคุมความคลาดเคลื่อนในส่วนนี้จึงอาศัยวิธีการเลือกตัวอย่างที่เหมาะสม รวมทั้งการกำหนดขนาดตัวอย่างที่เพียงพอ และเลือกใช้สูตรในการประมาณค่าที่ถูกต้องตามหลักวิชา (ศิริลักษณ์ สุวรรณวงศ์ , 2538 : 241-242)

1.2 ความคลาดเคลื่อนที่เกิดจากความมีอคติในตัวอย่าง (Sampling Bias)

การเกิดความมีอคติ หรือความเอนเอียง อาจเกิดจากการใช้กรอบตัวอย่างที่ไม่สมบูรณ์ หรือการใช้ตัวประมาณที่เอนเอียง (Biased Estimator)

การควบคุมความคลาดเคลื่อนในส่วนนี้ คือ การเตรียมกรอบตัวอย่างที่สมบูรณ์ทันสมัยและหลีกเลี่ยงการใช้ตัวประมาณที่เอนเอียง

2. ความคลาดเคลื่อนที่ไม่ได้เกิดจากการสุ่มตัวอย่าง (Non-Sampling Error)

เป็นความคลาดเคลื่อนที่เกิดขึ้นได้เสมอ อีกทั้งเกิดได้ในทุกขั้นตอนของการทำงานนับตั้งแต่ขั้นเตรียมงานจนถึงขั้นเสนอผล กล่าวคือ

1.1 ขั้นเตรียมการ เกิดจากการเตรียมการที่ไม่ดี เช่น แบบข้อถามบกพร่องคำสั่งชี้แจงไม่ชัดเจน พนักงานสนามด้อยคุณภาพ การควบคุมงานสนามหละหลวม

1.2 ขั้นเก็บรวบรวมข้อมูล เกิดจากผู้ตอบไม่ให้ข้อมูลที่ถูกต้อง (Response Error) ซึ่งส่วนมากมักเป็นคำถามเกี่ยวกับรายได้ หนี้สิน หรืออาจเกิดจากไม่ได้รับข้อมูลจากผู้ตอบ (Non-Respone Error) มักพบกับการสำรวจที่ใช้วิธีส่งแบบสอบถามทางไปรษณีย์ หรือการส่งพนักงานไปสัมภาษณ์แต่หาผู้ตอบไม่ได้เนื่องจากผู้ตอบไม่อยู่บ้าน หรือหาบ้านไม่พบซึ่งพบบ่อยในการเก็บรวบรวมข้อมูลในเขตกรุงเทพมหานคร หรือเมืองใหญ่

1.3 ขั้นการประมวลผล เช่น เกิดจากการบรรณาธิกรณ์ข้อมูล (Data Editing) การคำนวณ ถ้าไม่มีการวางแผนและการควบคุมคุณภาพการทำงานการประมวลผลจะทำให้ผลลัพธ์

ที่ได้คลาดเคลื่อนจากความจริงมากโดยเฉพาะการประมวลผลข้อมูลจำนวนมาก ๆ ซึ่งต้องใช้ทั้งคน และเครื่องคอมพิวเตอร์ประมวลผล

1.4 ขั้นการนำเสนอข้อมูล เกิดจากการแปลความหมาย หรือสรุปผลที่ได้ผิดพลาดรวมทั้งเกิดจากการพิมพ์รายงาน

ความคลาดเคลื่อนที่ไม่ได้เกิดจากการสุ่มตัวอย่างนี้ ถ้าไม่มีการควบคุมให้ดีจะเกิดได้ทั้งกรณีที่เราใช้วิธีการสุ่มตัวอย่างหรือไม่ใช้วิธีการสุ่มตัวอย่าง(การสำมะโน)ทางหนึ่งที่สามารถลดความคลาดเคลื่อนในส่วนนี้ คือ การใช้ระเบียบวิธีการสุ่มตัวอย่างซึ่งจะมีความคลาดเคลื่อนที่ไม่ได้เกิดจากการสุ่มตัวอย่างน้อยกว่าสำมะโน หรือการเก็บข้อมูลจากทุกหน่วยของประชากร เนื่องจากการสำมะโนต้องใช้คนและเครื่องจักรทำงานมากกว่า ดังนั้นการควบคุมคุณภาพของการทำงานย่อมยุ่งยาก และอาจทำให้ประสิทธิภาพด้อยกว่าการใช้ระเบียบวิธีการสุ่มตัวอย่าง

ตอนที่ 5 การประมาณค่าและคุณสมบัติของตัวประมาณค่าที่ดี

จุดมุ่งหมายประการสำคัญในการสุ่มกลุ่มตัวอย่างมาจากประชากร ก็คือ การประมาณค่าของพารามิเตอร์ อันเป็นลักษณะของประชากรที่ผู้วิจัยใช้ในการประมาณค่าของประชากร กระทำได้ 2 แบบ คือ การประมาณค่าแบบจุด (Point estimstion) และการประมาณค่าแบบช่วง (Interval Estimation) วิธีการประมาณค่าแบบจุดเป็นการประมาณค่าของประชากรที่สนใจศึกษาด้วยค่าเพียงค่าเดียวเท่านั้น เช่น ใช้ $\bar{x}$ ประมาณค่า μ หรือ s^2ประมาณค่า σ^2 เป็นต้น ส่วนวิธีการประมาณค่าแบบช่วงเป็นการประมาณค่าพารามิเตอร์ของประชากรที่สนใจศึกษาด้วยช่วงค่าช่วงหนึ่ง ซึ่งมีคุณสมบัติว่าค่าของประชากรที่แท้จริงจะตกอยู่ในช่วงค่าที่ประมาณนี้ ด้วยความเชื่อมั่นระดับหนึ่งโดยจะต้องอาศัย การประมาณค่าแบบจุด และการแจกแจงความน่าจะเป็นของตัวประมาณเป็นพื้นฐานในการคำนวณ การประมาณค่าทั้งสองแบบจะเหมาะสมกับกรณีการใช้งานที่ต่างกัน ในกรณีที่มีตัวประมาณค่า (Estimator) ซึ่งสามารถนำมาประมาณค่าพารามิเตอร์ได้หลายวิธี ได้มีการกำหนดคุณสมบัติของตัวประมาณค่าที่ดี ซึ่งควรมีคุณสมบัติ 4 ประการดังต่อไปนี้ (Hay, 1963; Yamane, 1967)

1. **ความไม่เอนเอียง** (Unbiasedness)

ถ้า $\hat{\theta}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของพารามิเตอร์ θ แล้วจะได้ว่า

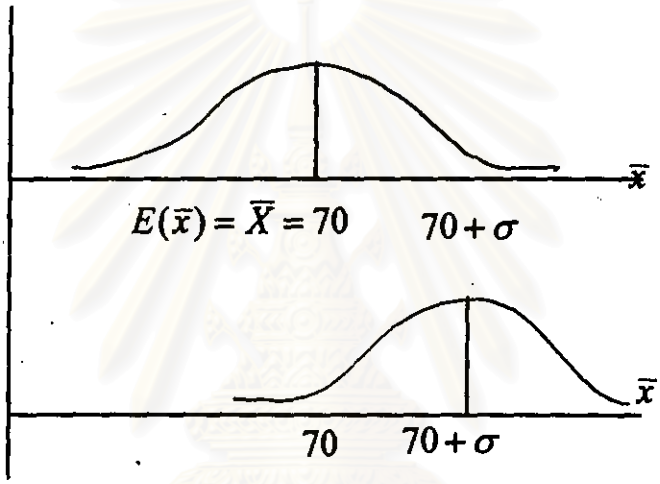
$$E(\hat{\theta}) = \theta$$

นั่นคือ ค่าคาดหวัง (Expected Value) ของตัวประมาณค่า θ มีค่าเท่ากับค่าของพารามิเตอร์ตัวที่ต้องการประมาณหรือกล่าวอีกนัยหนึ่งว่า ตัวประมาณค่านั้นไม่เอนเอียง ถ้า Expectation $\hat{\theta}$ มีค่าเท่ากับพารามิเตอร์ θ และตัวประมาณค่าที่ไม่เอนเอียงนั้นมีคุณสมบัติ

ที่ดีอยู่ว่า ถ้ามีชุดของค่าประมาณที่ไม่เอนเอียงที่เป็นอิสระต่อกันอยู่แล้วค่าเฉลี่ยเหล่านั้นย่อมไม่เอนเอียงด้วย และในทางตรงกันข้าม ค่าเฉลี่ยของค่าประมาณที่เอนเอียง ย่อมเอนเอียงด้วยไม่ว่าจะเฉลี่ยมากี่ค่าก็ตาม ส่วนค่าเอนเอียง(Bias)ของตัวประมาณ $\hat{\theta}$ คือ ค่าความแตกต่างระหว่างค่าที่คาดหวังของ $\hat{\theta}$ กับค่าพารามิเตอร์ θ กล่าวคือ

$$Bias(\theta) = E(\hat{\theta}) - \theta$$

แผนภาพที่ 3 การแจกแจงของตัวประมาณค่าที่ไม่เอนเอียงและตัวประมาณค่าที่เอนเอียง



2. ความคงเส้นคงวา (Consistency)

การประมาณค่าพารามิเตอร์ θ ด้วย $\hat{\theta}$ เมื่อขนาดของกลุ่มตัวอย่างใหญ่ขึ้น ตัวประมาณ $\hat{\theta}$ มีค่าเข้าใกล้พารามิเตอร์ θ มากขึ้นด้วย ซึ่งสามารถเขียนอยู่ในรูปประโยคทางคณิตศาสตร์ได้ว่า

$$\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| < E) = 1$$

เมื่อ E คือ ค่าความคลาดเคลื่อนที่ยอมให้เกิดขึ้นได้ในการหาค่าพารามิเตอร์ด้วยตัวประมาณค่า ($\hat{\theta}$) โดยที่ P คือ ความน่าจะเป็นหรือโอกาสที่จะเกิดเหตุการณ์นั้นขึ้น เราจะเรียก $\hat{\theta}$ ว่า เป็นตัวประมาณค่าที่มีความคงเส้นคงวาของ θ ซึ่งหมายความว่าความน่าจะเป็นที่ค่า $\hat{\theta}$ จะประมาณเข้าใกล้ θ และมีค่าสูงขึ้น เมื่อขนาดกลุ่มตัวอย่าง (n) มีค่าเพิ่มมากขึ้น ๆ คือ มีค่าความน่าจะเป็นเข้าใกล้ 1

3. **ความมีประสิทธิภาพ** (Efficiency)

ความมีประสิทธิภาพ หมายถึง การพิจารณาถึงความถูกต้องแม่นยำในการประมาณค่าของตัวประมาณค่าหนึ่ง ๆ เกณฑ์ที่ใช้พิจารณาความมีประสิทธิภาพของตัวประมาณค่า คือ ความแปรปรวนที่เปรียบเทียบกับกลุ่มของตัวประมาณค่าที่ไม่ลำเอียงด้วยกัน

โดยทั่วไปจะนิยามประสิทธิภาพของตัวประมาณค่าใด ๆ ว่าเป็นอัตราส่วนระหว่างค่าความแปรปรวนของตัวประมาณค่าที่มีประสิทธิภาพนั้น กับ ค่าความแปรปรวนของตัวประมาณค่าตัวอื่น ๆ กล่าวคือ ถ้าค่าความแปรปรวนของ $\hat{\theta}_i$ หรือ Var ($\hat{\theta}_i$)น้อยกว่าค่าความแปรปรวนของ $\hat{\theta}_j$ หรือ Var($\hat{\theta}_j$) เมื่อทั้ง $\hat{\theta}_i$ และ $\hat{\theta}_j$ เป็นตัวประมาณค่าที่ไม่เอนเอียงแล้วจะได้ว่า

เมื่อ

$$E_r = \frac{Var(\hat{\theta}_i)}{Var(\hat{\theta}_j)} \qquad ;0\langle E_r \langle 1$$

ดังนั้นเมื่อค่า E_r ที่ได้จากการคำนวณน้อยกว่า 1 แสดงว่า $\hat{\theta}_i$ เมื่อเทียบกับ $\hat{\theta}_j$ แล้ว $\hat{\theta}_i$ มีประสิทธิภาพมากกว่า (Absolutely efficient) เพราะว่าความแปรปรวนของ $\hat{\theta}_i$ น้อยกว่าความแปรปรวนของ $\hat{\theta}_j$

ในการพิจารณาคุณภาพของตัวประมาณค่าพารามิเตอร์$\hat{\theta}$ ลักษณะที่น่าจะพิจารณาอีกลักษณะหนึ่งก็คือ ค่าต่าง ๆ ของ $\hat{\theta}$ มีความแตกต่างจากค่าที่ต้องการประมาณเพียงใด ทั้งนี้เพราะถ้าค่าส่วนใหญ่ของ$\hat{\theta}$ มีความแตกต่างจากค่า θ น้อย ก็แสดงว่าโอกาสที่ค่าประมาณจะตกอยู่ใกล้ค่าของ θ มาก ตัวประมาณค่าที่มีลักษณะเช่นนี้จะดีกว่าตัวประมาณค่าที่มีค่าเป็นไปได้กระจายออกจากค่าที่ต้องการประมาณมาก ดังนั้นเกณฑ์ที่จะใช้เปรียบเทียบคุณภาพลักษณะนี้คือ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) แต่ในทางปฏิบัติเราไม่สามารถที่จะหาตัวประมาณค่าที่มีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองสำหรับทุก ๆ ค่าของ θ ได้ จากการพิจารณาตัวประมาณค่าที่ไม่เอนเอียง พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจะเท่ากับค่าความแปรปรวน จึงสามารถหาตัวประมาณที่ดีที่สุดในกลุ่มตัวประมาณที่ไม่เอนเอียง คือ ตัวประมาณที่มีค่าความแปรปรวนต่ำสุด

4. **ความพอเพียง** (Sufficiency)

ตัวประมาณค่า $\hat{\theta}$ จะเป็นตัวประมาณค่าที่มีความพอเพียง ถ้าตัวประมาณค่าให้สารสนเทศที่ก่อเกิดประโยชน์ได้ทั้งหมดที่ต้องการเกี่ยวกับพารามิเตอร์ที่ต้องการประมาณ เช่น $\bar{x}$ เป็นตัวประมาณที่พอเพียงของ μ ก็หมายความว่าไม่มีตัวประมาณค่าของ μ ตัวอื่น เช่นฐานนิยม (Mode) มัธยฐาน (Median) ที่จะสามารถให้สารสนเทศเกี่ยวกับ μ เพิ่มขึ้นได้อีกดีไปกว่า $\bar{x}$

สำหรับการวิจัยครั้งนี้ จะเปรียบเทียบคุณสมบัติตัวประมาณค่าพารามิเตอร์ใน สามด้านคือ ด้านความไม่เอนเอียง ด้านความคงเส้นคงวา และความมีประสิทธิภาพ สำหรับ ด้านความพอเพียงนั้นมิได้นำมาศึกษา

ตอนที่ 6 การแจกแจงการสุ่มตัวอย่าง (Sampling Distribution)

งานวิจัยส่วนใหญ่โดยเฉพาะทางสังคมศาสตร์ โดยปกติแล้วไม่สามารถเก็บรวบรวม ข้อมูลจากทุกหน่วยของประชากร เนื่องจากมีข้อจำกัดหลายด้าน เช่น เวลา งบประมาณ กำลังคน ดังนั้นจึงจำเป็นต้องศึกษากับลักษณะของกลุ่มตัวอย่างมากกว่าที่จะศึกษากับประชากร ทั้งหมดหรือศึกษาความสัมพันธ์ระหว่างค่าที่ได้จากกลุ่มตัวอย่างกับค่าพารามิเตอร์ของประชากร เพื่อการสรุปอ้างอิงอันจะนำไปสู่การตัดสินอย่างใดอย่างหนึ่ง

ในการศึกษาข้อมูลจากกลุ่มตัวอย่าง เมื่อสุ่มตัวอย่างขนาด n จากประชากรหนึ่ง ซึ่งมีขนาด N จะมีกลุ่มตัวอย่างที่เป็นไปได้ทั้งหมด ${}_{N}C_{n}$ กลุ่มตัวอย่าง (All Possible Samples of size n) ในแต่ละกลุ่มตัวอย่าง เราสามารถคำนวณค่าสถิติต่าง ๆ เช่น ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน มัธยฐาน ฐานนิยม ส่วนเบี่ยงเบนควอไทล์ ฯลฯ ค่าสถิติที่คำนวณได้ จากกลุ่มตัวอย่างประชากรขนาดเดียวกัน เมื่อนำมาแจกแจงความถี่ จะให้การแจกแจงการสุ่ม ของค่าสถิตินั้น หรือจะได้รูปของการแจกแจงที่เรียกว่า การแจกแจงการสุ่มตัวอย่าง (ศิริชัย กาญจนวาสี, 2533)

ลักษณะสำคัญของการแจงแจงการสุ่มตัวอย่าง คือ (อุทุมพร จามรมาน, 2537)

1. ค่าตรงกลาง(Central Tendency Measure) เรียกว่า ค่าที่คาดหวัง (Expected Value)

ค่าคาดหวัง คือ ค่าเฉลี่ยของค่าสถิติของทุกกลุ่มเท่าที่จะเป็นไปได้ หรือค่าตรงกลาง ของการแจกแจงการสุ่ม หากค่าคาดหวังมีค่าเท่ากับค่าพารามิเตอร์ เช่น ค่ามัชฌิมเลขคณิต จากการสุ่มตัวอย่าง มีค่าเท่ากับ ค่ามัชฌิมเลขคณิตของประชากร (μ) ในกรณีนี้เรียกว่า $\bar{x}$ เป็นตัวประมาณค่าพารามิเตอร์ที่ไม่เอนเอียง (Unbiased Estimator)

2. ค่าการกระจาย (Dispersion หรือ Variability) เรียกว่า ความคลาดเคลื่อน มาตรฐาน (Standard Error)

ความคลาดเคลื่อนมาตรฐาน คือ ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงการสุ่ม ของค่าสถิติ เช่น ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงการสุ่มของค่ามัชฌิมเลขคณิต

3. ลักษณะการแจกแจง (Form)

ลักษณะการแจกแจงการสุ่มขึ้นอยู่กับการนำค่าสถิติใดมาแจกแจง ถ้าใช้ค่ามัชฌิม เลขคณิตของกลุ่มตัวอย่างเป็นตัวประมาณค่าพารามิเตอร์ของประชากร โดยที่ลักษณะของ

I17281143

ประชากรจะมีลักษณะใดก็ตาม การแจกแจงการสุ่มของค่ามัชฌิมเลขคณิตจะมีลักษณะเข้าใกล้หรือเป็นโค้งปกติ ไม่ว่ากลุ่มตัวอย่างจะมีขนาดเท่าใดก็ตาม ซึ่งเป็นผลสืบเนื่องมาจาก ทฤษฎีเซ็นทรัล-ลิมิต (Central - limit Theorem)

ทฤษฎีเซ็นทรัล-ลิมิต (Central - limit Theorem)

ทฤษฎีเซ็นทรัล-ลิมิต เป็นทฤษฎีที่มีประโยชน์และสำคัญอันหนึ่งในทฤษฎีทั้งหลายของความน่าจะเป็นและสถิติ ทำให้สามารถคาดการณ์เกี่ยวกับการแจกแจงโดยประมาณของค่าเฉลี่ยของกลุ่มตัวอย่างได้ ตามทฤษฎีเซ็นทรัล-ลิมิต (Burington and May, 1970 ; Glass and Hopkins, 1984) กล่าวไว้ว่า ถ้าประชากรหนึ่งมีค่าเฉลี่ย μ และมีความแปรปรวนจำกัด (Finite variance) เป็น σ^2 เมื่อสุ่มตัวอย่างที่เป็นอิสระจากกันขนาด n แล้ว การแจกแจงการสุ่มตัวอย่างของค่า $\bar{x}$ จะมีลักษณะดังนี้

1. การแจกแจงการสุ่มตัวอย่างของ $\bar{x}$ (Sampling Distribution of Mean) จะมีการแจกแจงใกล้เคียง หรือเข้าใกล้การแจกแจงปกติที่มีค่าเฉลี่ย μ และ ความแปรปรวนσ^2 / n เมื่อ n มีขนาดเพิ่มขึ้น ไม่ว่าประชากรมีการแจกแจงแบบใด

2. เมื่อมีลักษณะการแจกแจงใกล้เคียงหรือเป็นโค้งปกติแล้ว

- ค่าเฉลี่ยของการแจกแจงการสุ่มของค่าสถิติ เท่ากับค่าเฉลี่ยประชากร หรือ

$$E(\bar{x}) = \mu$$

- ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงการสุ่มของค่าสถิติ $\bar{x}$ เรียกว่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (Standard error of the mean) หรือ$\sigma_{\bar{x}}$ มีค่าเท่ากับส่วนเบี่ยงเบนมาตรฐานของประชากร หารด้วย รากที่สองของขนาดกลุ่มตัวอย่าง นั่นคือ

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

ดังนั้น ในการสุ่มตัวอย่างจากประชากรแล้วใช้ข้อมูลจากกลุ่มตัวอย่าง เพื่อประมาณค่าพารามิเตอร์นั้น ในกรณีที่ประชากรมีการแจกแจงปกติ การแจกแจงของค่าประมาณมัชฌิมเลขคณิตจากการสุ่มตัวอย่างจะเป็นการแจกแจงปกติเสมอ ไม่ว่ากลุ่มตัวอย่างจะมีขนาดเท่าใด สำหรับกรณีที่ประชากรไม่ได้มีการแจกแจงแบบปกติ การแจกแจงของค่าประมาณมัชฌิมเลขคณิตจากการสุ่มตัวอย่างจะยังคงเข้าสู่การแจกแจงปกติ ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ นั่นคือ ถ้าไม่ทราบการแจกแจงของประชากร หากสุ่มตัวอย่างให้มีขนาดใหญ่ การแจกแจงของค่าเฉลี่ยของกลุ่มตัวอย่างจะเป็นการแจกแจงปกติ ซึ่งลักษณะดังกล่าวจะก่อให้เกิดตัวประมาณค่าที่ดีดังได้กล่าวมาแล้วได้

ตอนที่ 7 การชิมูเลชั่นและวิธีการ มอนติ คาร์โล (Simulation and The Monte Carlo Mothods)

เนย์เลอร์ (Naylor) (อ้างถึงใน Rubinstein,1981) ได้ให้คำจำกัดความไว้ดังนี้

> การชิมูเลชั่น (Simulation) เป็นเทคนิคทางจำนวนสำหรับการดำเนินการทดลองในเครื่องคอมพิวเตอร์ ซึ่งนำไปสู่ชนิดของโมเดลทางคณิตศาสตร์และโมเดลเชิงเหตุผลที่อธิบาย การปฏิบัติของระบบธุรกิจและเศรษฐกิจ (หรือบางองค์ประกอบของสิ่งนั้น) ที่อยู่นอกเหนือข้อจำกัดของเวลาในสภาพความเป็นจริง

นอกจากนั้นยังให้ คำจำกัดความของการชิมูเลชั่น ว่าเป็นเทคนิคที่มีการปฏิบัติการทดลองสุ่มของแบบจำลองระบบ คำจำกัดความดังกล่าวข้างต้น ค่อนข้างกว้าง อย่างไรก็ตาม การชิมูเลชั่น ในความหมายแนวแคบ หรือ การชิมูเลชั่นแบบกำหนด (Stochastic Simulation) ว่าเป็นการทดลองสุ่มทางสถิติด้วยรูปแบบ (model) และการสุ่มนี้นำไปสู่ปัญหาทั้งหมดของการวิเคราะห์การออกแบบทางสถิติ

เพราะเหตุว่า การสุ่มตัวอย่างจากการแจกแจงแบบพาทิคูลาร์ (Particular Distribution) นำไปสู่การใช้เลขสุ่ม (Random Number) ดังนั้น การชิมูเลชั่นแบบกำหนด ในบางครั้งจึงเรียกว่า มอนติ คาร์โล ชิมูเลชั่น (Monte Carlo Simulation)

วิธีการมอนติ คาร์โล เป็นเทคนิคหนึ่งที่ใช้เลขสุ่ม (Random Number) และเลขสุ่มแบบเทียม (Psudorandom Number) สำหรับหาคำตอบของรูปแบบที่ต้องการศึกษา (Rubinstein,1981) เลขสุ่มที่ได้จากเครื่องคอมพิวเตอร์ จะมีความเป็นอิสระในการสุ่มซึ่งเป็นสิ่งสำคัญ และจะมีลักษณะการแจกแจงแบบยูนิฟอร์ม ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1 [0 , 1] คอมพิวเตอร์จะสร้างเลขสุ่มแต่ละหลัก ตั้งแต่ 0 ถึง 9 ที่ครอบคลุมโอกาสที่จะเกิดขึ้นได้ทั้งหมด

ศตวรรษแห่งการเริ่มต้นของวิธีการมอนติ คาร์โล ถูกใช้ในการตรวจสอบสมการบอลท์ซแมน (Boltzmann Equation) ในปี ค.ศ. 1908 นักศึกษาวิชาสถิติที่มีชื่อเสียงได้ใช้วิธีการมอนติ คาร์โล สำหรับการประมาณค่าสัมประสิทธิ์สหสัมพันธ์ในการแจกแจงที (t-distribution)

" มอนติ คาร์โล " ถูกนำมาใช้โดย Neumann และ Ulam ในช่วงสงครามโลกครั้งที่สอง เป็นรหัสของคำในการทำงานลับที่ ลอส แอลมอส (Los Alamos) และถูกแนะนำโดยบ่อนการพนันในเมืองมอนติ คาร์โล ประเทศโมนาโค วิธีมอนติ คาร์โล ยังถูกนำไปใช้ในการแก้ปัญหาที่เกี่ยวข้องกับลูกระเบิดปรมาณู (Atomic Bomb) งานในครั้งนี้ นำไปสู่การชิมูเลชั่นของการปฏิบัติที่เกี่ยวข้องกับการกระจายอย่างสุ่มของนิวตรอนของวัตถุที่มีการแตกตัว

ชิมูเลชั่นของการปฏิบัติที่เกี่ยวข้องกับการกระจายอย่างสุ่มของนิวตรอนของวัตถุที่มีการแตกตัว ด้วยเหตุนี้ ในเวลาไม่นานนัก วิธีมอนติ คาร์โล ถูกใช้หาค่าจำนวนพหุมิติที่ซับซ้อน การแก้สมการ ดิฟเฟอร์เรนเชียล (Differential Equation) สมการเชิงจำนวน (Integral Equation) ที่ยาก แก่การวิเคราะห์ เป็นต้น

วิธีการมอนติ คาร์โล ถูกนำไปใช้หลายทาง ไม่เพียงแต่ใช้ในการแก้ปัญหาที่กำหนด ขึ้น (Stochastic Problems) เท่านั้น แต่ยังใช้สำหรับการแก้ปัญหาที่เกิดขึ้นในแบบ Determistic Problems อีกด้วย นอกจากนั้น มีการนำไปใช้ประโยชน์ในการสุ่มค่าต่าง ๆ จากการแจกแจง ความน่าจะเป็น (Rubinstein,1981)

ในปัจจุบัน วิธีการมอนติ คาร์โล เป็นเทคนิคที่มีความสามารถสูง และใช้ทั่วไปในการ แก้ปัญหาการวิเคราะห์ข้อมูลให้สมบูรณ์

สำหรับภาษาที่ใช้ในการเขียนโปรแกรมคอมพิวเตอร์นั้น สามารถใช้ได้มากมายหลาย ภาษา เช่น ภาษาฟอร์แทรน (FORTRAN), ALGOL, PASCAL และ PL/1 เป็นตัวอย่างของ ภาษาที่ใช้ในจุดประสงค์ทั่วไป ขณะที่ GPSS, SIMCRIP และ SIMULA เป็นตัวอย่างของภาษาที่ ใช้สำหรับการสุ่มที่พิเศษ (Rubinstein, 1981) ปัจจุบัน มีภาษาอื่น ๆ มากมายที่ถูกพัฒนาและมี ประสิทธิภาพมากขึ้น สามารถนำมาใช้ในการเขียนโปรแกรมคอมพิวเตอร์สำหรับการทำงานโดยใช้ เทคนิคมอนติ คาร์โล ชิมูเลชั่น ได้

ตอนที่ 8 งานวิจัยที่เกี่ยวข้อง

ในการศึกษางานวิจัยที่เกี่ยวกับการพัฒนารูปแบบและวิธีการสุ่มตัวอย่างนั้น พบว่า มี ผู้ที่สนใจและทำการศึกษาเกี่ยวกับรูปแบบและวิธีการสุ่มตัวอย่างต่าง ๆ กัน ดังนี้

ผกามาศ สิงห์สง่า (2523) ได้ศึกษาแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้น 3 แบบ คือ แผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นทางเดียวโดยสุ่มตัวอย่างมาชั้นละ 1 หน่วย แผนแบบ การสุ่มตัวอย่างแบบแบ่งชั้นสองทางโดยใช้ตัวอย่างจำนวนน้อย และแผนแบบการสุ่มตัวอย่างแบบ แบ่งชั้นสองทาง โดยใช้จำนวนแบบมีระบบโดยให้ความน่าจะเป็นของการสุ่มหน่วยตัวอย่างเป็น ปฏิภาคกับขนาดของหน่วยตัวอย่างที่ต้องสำรวจ (ในแต่ละชั้นภูมิสุ่มตัวอย่างแบบมีระบบตาม จำนวนที่ต้องการจากการคำนวณหาขนาดของกลุ่มตัวอย่างในแต่ละชั้น) ศึกษาเปรียบเทียบ ประสิทธิภาพว่าแผนแบบใด จะให้ค่าประมาณที่มีความเชื่อถือได้มากกว่ากัน โดยใช้ข้อมูลจำนวน ครัวเรือนส่วนบุคคล ของกองสำรวจประชากร จำแนกตามขนาดของครัวเรือนและอำเภอของ กรุงเทพมหานคร จากรายงานสำมะโนประชากรและเคหะ พ.ศ. 2513 ของหน่วยสำมะโน ประชากรและเคหะ กองสำรวจประชากร สำนักงานสถิติแห่งชาติ พบว่า การใช้แผนการสุ่ม

ความแปรปรวนน้อยที่สุด ส่วนแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นทางเดียวโดยสุ่มตัวอย่างมาชั้นละ 1 หน่วย ให้ค่าประมาณที่มีความแปรปรวนมากที่สุด

สุรพล ปธานวนิช (2529) ได้ทำการทดสอบสมรรถภาพของการสุ่มตัวอย่างในงานวิจัยทางสังคมศาสตร์ โดยศึกษาเปรียบเทียบการสุ่มตัวอย่างแบบไม่อาศัยทฤษฎีความน่าจะเป็น กับการสุ่มตัวอย่างแบบอาศัยหลักความน่าจะเป็น ซึ่งวิธีการสุ่มตัวอย่างที่ใช้ทดสอบ คือ การสุ่มตัวอย่างแบบเจาะจง การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีระบบ การสุ่มตัวอย่างแบบแบ่งชั้น การสุ่มตัวอย่างแบบยกกลุ่ม และการสุ่มตัวอย่างแบบแบ่งชั้นตามคุณลักษณะของตัวแปรอิสระ 2 ตัว ประชากรที่ใช้เป็นข้าราชการพลเรือนระดับกรมหน่วยงานหนึ่ง จำนวน 7,057 คน ผลการศึกษา พบว่า การสุ่มตัวอย่างแบบเจาะจง ซึ่งเป็นตัวแทนของการสุ่มตัวอย่างที่ไม่อาศัยทฤษฎีความน่าจะเป็น มีประสิทธิภาพต่ำที่สุด สำหรับการสุ่มตัวอย่างที่อาศัยความน่าจะเป็นด้วยกัน การสุ่มตัวอย่างแบบยกกลุ่มจะมีประสิทธิภาพต่ำที่สุด และการสุ่มตัวอย่างแบบแบ่งชั้นให้ค่าประมาณใกล้เคียงกับพารามิเตอร์สม่ำเสมอที่สุด

นิเวศน์ คำรัตน์ (2534) ศึกษาการเปรียบเทียบค่าประมาณมัชฌิมเลขคณิตของประชากรจากกลุ่มตัวอย่างที่ได้จากการสุ่มแบบต่าง ๆ ได้แก่ การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีระบบ การสุ่มตัวอย่างแบบแบ่งชั้น การสุ่มตัวอย่างแบบสองชั้น และการสุ่มตัวอย่างแบบสามชั้น โดยใช้ขนาดของกลุ่มตัวอย่างที่แตกต่างกัน 3 ขนาด คือ จำนวน 1,095 , 1,555 และ 2,636 คน ประชากรที่ใช้ในการวิจัย คือ ข้าราชการครู ในเขตการศึกษา 1 จนถึงเขตการศึกษา 12 กรมสามัญศึกษา กระทรวงศึกษาธิการ ปีการศึกษา 2531 จำนวน 78,035 คน ดำเนินการสุ่มตัวอย่างในแต่ละวิธีและแต่ละขนาดของกลุ่มตัวอย่าง โดยใช้คอมพิวเตอร์กระทำการสุ่มซ้ำ 1,000 ครั้ง เปรียบเทียบค่าประมาณมัชฌิมเลขคณิตของอายุของประชากรตามคุณสมบัติของตัวประมาณค่าที่ดี 3 ด้าน คือความไม่เอนเอียง ความคงเส้นคงวา และความมีประสิทธิภาพ พบว่า วิธีการสุ่มตัวอย่างแบบสองชั้นให้ค่าประมาณมัชฌิมเลขคณิตของประชากรที่มีความเอนเอียง ที่ขนาดของกลุ่มตัวอย่าง จำนวน 1,095 คน วิธีการสุ่มตัวอย่างทั้ง 5 วิธีมีความคงเส้นคงวา และวิธีการสุ่มแบบแบ่งชั้น ให้ค่าประมาณมัชฌิมเลขคณิตของประชากรมีประสิทธิภาพมากที่สุด

สมชัย วงษ์นายะ (2534) ได้ศึกษาเปรียบเทียบค่าประมาณพารามิเตอร์จากแบบแผนการสุ่มตัวอย่างต่างแบบ โดยใช้ข้อมูลจากคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ ปีการศึกษา 2531 จำนวน 7,335 คน ซึ่งใช้การสุ่มตัวอย่างทั้งหมด 7 วิธี ได้แก่ การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีระบบ การสุ่มตัวอย่างแบบแบ่งชั้นที่ใช้ตัวแปรจำแนกชั้นต่างกัน คือ ขนาดโรงเรียน คุณภาพของโรงเรียน อำเภอ และการสุ่มตัวอย่างแบบแบ่งชั้น 2 ระยะโดยใช้ตัวแปร

จำแนกชั้นอำเภอขนาดโรงเรียน และตัวแปรจำแนกชั้นอำเภอ-คุณภาพของโรงเรียน ด้วยขนาดกลุ่มตัวอย่าง 3 ขนาด คือ ขนาดกลุ่มตัวอย่างที่มีระดับความเชื่อมั่น 90 %, 95% และ 99% (424 คน, 588 คนและ 962 คน) ดำเนินการสุ่มตัวอย่างแต่ละวิธีโดยใช้คอมพิวเตอร์ในการสุ่มซ้ำ 1,000 ครั้ง เปรียบเทียบค่าประมาณมัชฌิมเลขคณิตและค่าประมาณความแปรปรวน ผลการศึกษา พบว่า วิธีการสุ่มตัวอย่างแบบแบ่งชั้นที่ใช้ขนาดโรงเรียนเป็นตัวจำแนกชั้น ด้วยขนาดกลุ่มตัวอย่างที่มีความเชื่อมั่น 95% ให้ค่าเฉลี่ยของค่าประมาณมัชฌิมเลขคณิต ใกล้เคียงกับค่าพารามิเตอร์มากที่สุดและวิธีการสุ่มแบบแบ่งชั้นที่ใช้คุณภาพของโรงเรียนเป็นตัวจำแนกชั้นด้วยขนาดกลุ่มตัวอย่างที่มีความเชื่อมั่น 90% ให้ค่าเฉลี่ยของค่าประมาณความแปรปรวน ใกล้เคียงกับค่าพารามิเตอร์มากที่สุด นอกจากนี้พบว่าวิธีการสุ่มตัวอย่างแบบเป็นระบบ ด้วยขนาดกลุ่มตัวอย่างที่มีความเชื่อมั่น 99% ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองน้อยที่สุด

ดวงใจ ปวีณอภิชาต (2535) ได้ศึกษาเปรียบเทียบค่าประมาณมัชฌิมเลขคณิตและค่าประมาณความแปรปรวนที่ได้จากการสุ่มตัวอย่างแบบแบ่งชั้น ที่มีตัวแปรจำแนกชั้นภูมิแตกต่างกัน 3 ลักษณะ คือ ประเภทโรงเรียน ขนาดโรงเรียน และเขตอำเภอ โดยใช้วิธีการกำหนดขนาดของกลุ่มตัวอย่างย่อยที่แตกต่างกัน 2 ลักษณะ คือ วิธีกำหนดแบบนีย์แมน และวิธีกำหนดแบบสัดส่วน รวมแผนการสุ่มทั้งหมด 6 วิธี แต่ละแผนการสุ่มใช้ขนาดตัวอย่างที่ระดับความเชื่อมั่นที่แตกต่างกัน 3 ระดับ คือ 90% 95% และ 99% โดยใช้คอมพิวเตอร์กระทำการสุ่มซ้ำ 1,000 ครั้ง เปรียบเทียบค่าประมาณของคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของประชากร โดยใช้เกณฑ์ในการเปรียบเทียบความมีประสิทธิภาพ 4 เกณฑ์ คือ ความใกล้เคียงความแปรปรวน ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง และค่าประสิทธิภาพสัมพันธ์ ของค่าประมาณมัชฌิมเลขคณิต และค่าประมาณความแปรปรวนของประชากร ผลการวิจัยพบว่า ขนาดโรงเรียน เป็นตัวแปรจำแนกชั้นภูมิที่มีประสิทธิภาพมากที่สุดในการประมาณค่ามัชฌิมเลขคณิตทุกขนาดกลุ่มตัวอย่าง และประมาณค่าความแปรปรวนเมื่อใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่น 95% ในส่วนของวิธีกำหนดขนาดกลุ่มตัวอย่างย่อยแบบนีย์แบบมีประสิทธิภาพในการประมาณค่ามัชฌิมเลขคณิต และการประมาณค่าความแปรปรวนมากกว่าวิธีกำหนดแบบสัดส่วน เมื่อใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่น 99%

จากงานวิจัยแผนการสุ่มตัวอย่างต่างแบบดังกล่าวข้างต้นในรายละเอียดจะพบว่าแผนการสุ่มตัวอย่างหลายขั้นตอนที่ศึกษา เมื่อกำหนดวิธีการสุ่มตัวอย่างจากประชากรเป็นวิธีการต่าง ๆแล้ว ในการสุ่มตัวอย่างจากกลุ่มย่อยจะใช้วิธีการสุ่มแบบง่าย

ดังนั้น ผู้วิจัยจึงเห็นว่า ยังมีวิธีการสุ่มกลุ่มตัวอย่างย่อยอีกวิธีการหนึ่ง คือ วิธีการสุ่มแบบมีระบบ ว่าจะให้ผลที่แตกต่างกันหรือไม่ หรือวิธีสุ่มใดจะมีคุณภาพในการประมาณค่าได้ดีกว่ากัน เมื่อประชากรมีลักษณะที่สามารถเลือกวิธีการสุ่มได้หลากหลาย