การจำลองข้อมูลเพื่อประเมินประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบชนิดผสม

น.ส.นภสร รัตนวุฒิขจร

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A SIMULATION-BASED EVALUATION OF MIXED SYSTEMATIC SAMPLING DESIGN

Miss Napasorn Rattanawuttikajohn

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Statistics

Department of Statistics

FACULTY OF COMMERCE AND ACCOUNTANCY

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การจำลองข้อมูลเพื่อประเมินประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบชนิดผสม
โดย	น.ส.นภสร รัตนวุฒิขจร
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.นัท กุลวานิช

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร.วิเลิศ ภูริวัชร)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.อัครินทร์ ไพบูลย์พานิช)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ ดร.นัท กุลวานิช)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ณัตติฤดี เจริญรักษ์)

..... กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

นภสร รัตนวุฒิขจร : การจำลองข้อมูลเพื่อประเมินประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบชนิดผสม. (A SIMULATION-BASED EVALUATION OF MIXED SYSTEMATIC SAMPLING DESIGN) อ.ที่ปรึกษาหลัก : ผศ. ดร.นัท กุลวานิช

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS) กับการเลือกตัวอย่างแบบมีระบบชนิดวงกลม (Circular Systematic Sampling : CSS) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval) สำหรับกรณีช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม เมื่อประชากรมีแนวโน้มเชิงเส้น ด้วยค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) และเปรียบเทียบประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบทั้ง 3 วิธีด้วยค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency : RE) โดยการจำลองข้อมูลของประชากรเป็น 3 ขนาด แบ่งเป็น ขนาดเล็กหลักร้อย ได้แก่ 300, 500 และ 700 ขนาดกลางหลักพัน ได้แก่ 3,000, 5,000 และ 7,000 ขนาดใหญ่หลักหมื่น ได้แก่ 30,000, 50,000 และ 70,000 ด้วยโปรแกรม R กำหนดขนาดตัวอย่างที่ทำให้สัดส่วนระหว่างขนาดประชากรและขนาดตัวอย่างไม่เป็นจำนวนเต็ม ทำซ้ำทั้งหมด 1,000 ครั้ง พบว่าการเลือกตัวอย่างแบบมีระบบด้วยวิธี MRSS มีค่า MSE สูงกว่าการเลือกตัวอย่างอีกทั้ง 2 วิธี แต่เมื่อค่า g = 2 จะทำให้ค่าของ MSE ของการเลือกตัวอย่างทั้ง 3 วิธีมีค่ามากขึ้น โดยที่ค่า MSE ของการเลือกตัวอย่างแบบมีระบบชนิดผสมมีค่าต่ำกว่าการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและวิธีใช้ช่วงเศษส่วน ทั้งนี้เป็นผลเนื่องมาจากค่า g เป็นค่าที่กำหนดความเป็นเชิงเส้น เมื่อค่า g เพิ่มมากขึ้น ความเป็นเชิงเส้นของประชากรจะลดลง ทำให้ตัวประมาณค่าเฉลี่ยตัวอย่างที่คำนวณได้มีค่าต่างจากค่าเฉลี่ยประชากรมากตามไปด้วย

จึงสามารถสรุปได้ว่าตัวประมาณค่าเฉลี่ยที่ได้จากวิธีการเลือกตัวอย่างแบบมีระบบชนิดผสม มีแนวโน้มที่จะให้ค่า MSE สูงที่สุด เมื่อเทียบกับการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา สถิติ ลายมือชื่อนิสิต .....

ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6280155526 : MAJOR STATISTICS
KEYWORD: Mixed Systematic Random Sampling, Simulation

Napasorn Rattanawuttikajohn : A SIMULATION-BASED EVALUATION OF MIXED SYSTEMATIC SAMPLING DESIGN. Advisor: Assoc. Prof. NAT KULVANICH, Ph.D.

The objective of this study was to examine the effectiveness of mean estimators produced using Mixed Systematic Sampling (MRSS), Circular Systematics Sampling (CSS), and Fractional Interval. In the case of non-integer sample selection intervals, the efficiency of the three systematic sampling methods was compared with Relative Efficiency (RE) by simulating the population data into three sizes: small in hundred digits; 300, 500, and 700, medium in thousands of digits; 3,000, 5,000, and 7,000, and large in tens digits; 30,000, 50,000, and 70,000 from R program. A sample size was set so that the percentage between the population size and the sample size is non-integer, and this was repeated 1,000 times. The systematic sample selection by MRSS approach was found to have a larger MSE value than the other two ways, However, when g = 2, the MSE of the three sampling techniques was greater, with the mixed systematic sampling having a lower MSE than the circular systematic sampling and the fractional interval. This is because the linearity is defined by the g value. As the g value grows, the population's linearity declines, leading the derived sample mean estimate to diverge considerably from the population mean.

Therefore, it can be determined that mean estimators derived using the Mixed Systematic Random Sampling approach have the highest MSE values when

Field of Study: Statistics Student's Signature .....
Academic Year: 2022 Advisor's Signature .....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลุล่วงไปด้วยดีเพราะได้รับความอนุเคราะห์และเอาใจใส่ให้คำปรึกษาจาก อาจารย์นัท กุลวานิช ผู้เป็นอาจารย์ที่ปรึกษาที่ได้ให้คำแนะนำ แก้ไขปรับปรุงข้อบกพร่อง รวมทั้งให้องค์ความรู้แนวทางในการศึกษาค้นคว้ามาโดยตลอด ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ ประธานกรรมการสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.อัครินทร์ ไพบูลย์พานิช กรรมการสอบวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ณัตติฤดี เจริญรักษ์ และรองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ที่กรุณาสละเวลาเวลามาเป็นกรรมการในการสอบ ตลอดจนให้คำแนะนำที่มีคุณค่าต่อการนำไปปรับปรุงแก้ไขงานวิจัยฉบับนี้ให้สมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณ คณาจารย์ภาควิชาสถิติ รวมทั้งคณาจารย์ในภาควิชาสาขาอื่นๆ ของคณะพาณิชยศาสตร์และการบัญชีที่เมตตากรุณาถ่ายทอดองค์ความรู้แก้ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

นภสร รัตนวุฒิขจร

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญ

หน้า

..... ค

บทคัดย่อภาษาไทย ..... ค

..... ง

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง 1

สารบัญภาพ 1

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์การวิจัย 2

1.3 สมมติฐานการวิจัย 2

1.4 ขอบเขตการวิจัย 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4

2.1 ทฤษฎีที่เกี่ยวข้อง 4

2.1.1 การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น (Linear Systematic Sampling : LSY) ... 4

2.1.2 การเลือกตัวอย่างมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS) 4

2.1.3 การเลือกตัวอย่างมีระบบแบบวงกลม (Circular Systematic Sampling : CSS) 5

2.1.4 การเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval) 6

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2 งานวิจัยที่เกี่ยวข้อง 6
บทที่ 3 วิธีการดำเนินงานวิจัย 10
3.1 วิธีการดำเนินงาน 10
3.1.1 แผนผังการดำเนินงาน 10
3.1.2 ขั้นตอนในการดำเนินงานวิจัย 11
บทที่ 4 ผลการดำเนินงานวิจัย 14
บทที่ 5 สรุปผลการศึกษา อภิปรายผล และข้อเสนอแนะ 21
บรรณานุกรม 23
ภาคผนวก 24
ประวัติผู้เขียน 34

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าประสิทธิภาพสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบแบบผสมและวงกลม 14

ตารางที่ 2 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าประสิทธิภาพสัมพัทธ์ได้จากการเลือกตัวอย่างแบบมีระบบแบบผสมและใช้ช่วงเศษส่วน..... 17

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญภาพ

รูปที่ **หน้า**

รูปที่ 1 แสดงค่าประสิทธิภาพเชิงสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสมและวงกลม 17

รูปที่ 2 แสดงค่าประสิทธิภาพเชิงสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสมและใช้ช่วงเศษส่วน 20

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการดำเนินการวิจัย การเก็บรวมรวมข้อมูลจากประชากรเพื่อตอบคำถามหรือปัญหา มีความสำคัญอย่างยิ่ง เพราะข้อมูลที่ถูกเลือกจากประชากรแล้วนำมาวิเคราะห์นั้นจะต้องนำมาตอบคำถามเพื่อให้ได้ผลสรุปที่ดีที่สุดสำหรับงานวิจัย ดังนั้นผู้วิจัยจึงต้องมีความรู้ความเข้าใจที่ถูกต้องเกี่ยวกับประชากร กลุ่มตัวอย่าง รวมทั้งวิธีการเลือกกลุ่มตัวอย่างเพื่อให้ได้หน่วยตัวอย่างที่เป็นตัวแทนที่ดี มีความครอบคลุมลักษณะของประชากร การเลือกตัวอย่างแบบมีระบบ (Systematic Sampling : SYS) เป็นวิธีการเลือกตัวอย่างแบบหนึ่งที่ได้รับความนิยมกันอย่างแพร่หลาย วิธีการเลือกตัวอย่างชนิดนี้ถูกพัฒนาขึ้นจากการเลือกตัวอย่างแบบง่าย (Simple Random Sampling : SRS) เพื่อให้แน่ใจว่าหน่วยตัวอย่างที่ถูกเลือกมานั้นกระจายอย่างทั่วถึงจากทั้งประชากร

การเลือกตัวอย่างแบบมีระบบสามารถแบ่งได้เป็น 2 ชนิดใหญ่ๆ ได้แก่ การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น (Linear Systematic Sampling : LSY) และ การเลือกตัวอย่างแบบมีระบบชนิดวงกลม (Circular Systematic Sampling : CSS) โดยการเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นนั้น (David G. Hankin et al., 2019) จำแนกได้ 2 กรณี คือ ช่วงของการเลือกตัวอย่างเป็นจำนวนเต็ม จะนิยมใช้การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น การเลือกตัวอย่างด้วยวิธีนี้เป็นการที่หน่วยตัวอย่างในประชากรมีลักษณะเรียงเป็นลำดับ ทำให้ได้กลุ่มตัวอย่างที่ได้มีลักษณะใกล้เคียงกับการเลือกตัวอย่างแบบง่าย และช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม จะนิยมใช้การเลือกตัวอย่างแบบมีระบบชนิดวงกลม และมีการนำวิธีการใช้ช่วงเศษส่วน (Fractional Interval Method) มาปรับใช้ด้วย จากงานจัยของ (Jammaree Thuksoon, 2000) ระบุว่า การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นและชนิดวงกลม มีประสิทธิภาพพอๆกันเมื่อมีขนาดตัวอย่างที่เท่ากัน เนื่องจากการเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นจะไม่เหมาะสมกับข้อมูลชุดนั้นๆ เมื่อช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม จะทำให้ได้จำนวนหน่วยตัวอย่างไม่เท่ากับขนาดตัวอย่างที่ต้องการ และจะทำให้ความแปรปรวนของค่าเฉลี่ยเป็นตัวประมาณที่เอนเอียง (Bias Estimator) จึงควรใช้วิธีการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและใช้ช่วงเศษส่วนจะเหมาะสมกว่า (Huang. C.K., 2004) จึงได้พัฒนาวิธีการเลือกตัวอย่างแบบมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS) ซึ่งวิธีดังกล่าวนี้เกิดขึ้นจากผสมผสานระหว่างการเลือกตัวอย่างแบบง่าย และการเลือกตัวอย่างแบบมีระบบไว้ด้วยกัน โดยการเลือกตัวอย่างชนิดนี้สามารถแก้ไขปัญหาดังที่กล่าวมาข้างต้นได้ แต่วิธีการนี้ก็ยังไม่ได้รับความนิยมเมื่อเทียบกับวิธีการเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นและชนิดวงกลม ต่อมา (Nakharin Kaewsrasean & Ampai Thongteeraparp, 2012) ได้ทำการศึกษาเปรียบเทียบการ

เลือกตัวอย่างแบบมีระบบชนิดผสม กับการเลือกตัวอย่างแบบมีระบบอีก 3 วิธี ได้แก่ การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น(LSY) การเลือกตัวอย่างแบบมีระบบเชิงเส้นคู่(BSY) และการเลือกตัวอย่างแบบมีระบบชนิด Modified Balanced Circular(MBC) เมื่อข้อมูลประชากรมีแนวโน้มเชิงเส้น (Linear Trend) ศึกษาเฉพาะกรณีช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม ผลการศึกษาพบว่าค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการเลือกตัวอย่างแบบมีระบบชนิดผสมมีแนวโน้มสูงกว่าการเลือกตัวอย่างทั้ง 3 ชนิด อาจเป็นผลมาจากผู้วิจัยได้กำหนดค่าความแปรปรวนของ Error Term ไว้ค่อนข้างสูงจึงทำให้ได้ค่า MSE ที่สูงมาก

ดังนั้น ในงานวิจัยครั้งนี้จะต่อยอดงานวิจัยของ (Nakharin Kaewsrasean & Ampai Thongteeraparp, 2012) ที่ศึกษาในกรณีที่ช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม และช่วงของการเลือกตัวอย่างเป็นหลักหน่วย โดยงานวิจัยฉบับนี้จะแบ่งขนาดของตัวอย่างเป็น ขนาดเล็ก ขนาดกลางและขนาดใหญ่ และกำหนดช่วงของการเลือกตัวอย่างเป็นหลักหน่วย หลักสิบและหลักร้อย เพื่อศึกษาว่าเมื่อขนาดประชากรมีลักษณะคงที่ แต่ขนาดตัวอย่าง ช่วงของการเลือกตัวอย่างมีลักษณะที่ต่างกัน มีผลต่อประสิทธิภาพของการเลือกตัวอย่างหรือไม่ และ Error Term ที่มีการแจกแจงแบบปกติด้วยค่าเฉลี่ยคงที่ แต่มีการเปลี่ยนแปลงของค่าความแปรแปรวน จะทำให้ MSE มีแนวโน้มเปลี่ยนแปลงไปอย่างไร นอกจากการนำวิธีการเลือกตัวอย่างแบบมีระบบชนิดผสมมาแก้ไขปัญหาในกรณีที่ช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็มแล้ว จะทำการเปรียบเทียบวิธีการเลือกตัวอย่าง 3 วิธี ได้แก่ การเลือกตัวอย่างแบบมีระบบชนิดผสม (MRSS) การเลือกตัวอย่างแบบมีระบบชนิดวงกลม (CSS) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน ว่าวิธีการใดมีประสิทธิภาพดีที่สุดสำหรับกรณีที่ช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 วัตถุประสงค์การวิจัย

เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสม (MRSS) กับการเลือกตัวอย่างแบบมีระบบชนิดวงกลม (CSS) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval) สำหรับกรณีช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม เมื่อประชากรมีแนวโน้มเชิงเส้น

1.3 สมมติฐานการวิจัย

การเลือกตัวอย่างแบบมีระบบชนิดผสม (MRSS) มีประสิทธิภาพดีที่สุด สำหรับช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม เมื่อประชากรมีแนวโน้มเชิงเส้น

1.4 ขอบเขตการวิจัย

ประชากรที่ศึกษามีแนวโน้มเชิงเส้น ดังสมการ

$$Y_i = \alpha + \beta i + e_i$$

โดยที่ i = ลำดับที่ของประชากร = 1, 2, 3, ..., N

e_i มีการแจกแจงแบบปกติ $E(e_i) = 0$ และ $Var(e_i) = i^g$

เมื่อ g คือ ค่าคงที่กำหนด (Predetermined constant) โดยกำหนด g = 0, 1, 2

กำหนด i และ e_i เป็นอิสระจากกัน และ i มีการแจกแจงแบบ discrete uniform

$i \in \{1, 2, 3, ..., N\}$

$$Var(Y) = E(Var(Y|i)) + Var(E(Y|i))$$

จะได้ $$Var(Y) = \begin{cases} \beta^2\left(\frac{N^2 - 1}{12}\right) + 1 ; g = 0 \\ \beta^2\left(\frac{N^2 - 1}{12}\right) + \frac{N + 1}{2} ; g = 1 \\ \beta^2\left(\frac{N^2 - 1}{12}\right) + \frac{N^3 - 1}{3N - 3} ; g = 2 \end{cases}$$

จำลองข้อมูลประชากร 3 ขนาด แบ่งเป็น ขนาดเล็กหลักร้อย ได้แก่ 300, 500 และ 700 ขนาดกลาง หลักพัน ได้แก่ 3,000, 5,000 และ 7,000 ขนาดใหญ่หลักหมื่น ได้แก่ 30,000, 50,000 และ 70,000

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.5 ประโยชน์ที่คาดว่าจะได้รับ

งานวิจัยนี้สามารถเปรียบเทียบประสิทธิภาพของการเลือกตัวอย่างของวิธีการเลือกตัวอย่างแบบมีระบบชนิดผสม, การเลือกตัวอย่างแบบมีระบบชนิดวงกลม และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน จากการจำลองข้อมูล เพื่อเป็นแนวทางสำหรับผู้วิจัยในการเลือกวิธีการเลือกตัวอย่างให้เหมาะกับลักษณะของข้อมูลต่อไปในอนาคต

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น (Linear Systematic Sampling : LSY)

(Prachum Suwattee, 2009) กล่าวถึงการเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นว่าเป็นวิธีการเลือกตัวอย่างที่คล้ายกับการเลือกตัวอย่างแบบง่าย เหมาะสำหรับประชากรที่มีการจัดเรียงลำดับตามลักษณะที่ต้องการ เป็นการเลือกตัวอย่าง n หน่วย จากประชากร N หน่วย จะได้จำนวนชุดตัวอย่างที่เป็นไปได้ทั้งหมด k ชุด โดยมีขั้นตอนดังนี้

1) เรียงลำดับข้อมูลจากน้อยไปมาก
2) เลือกตัวอย่างเริ่มต้นมีค่าเป็นจำนวนเต็ม rs จาก $1 \le rs \le k$ เมื่อ $k = \frac{N}{n}$
3) ทำการเลือกหน่วยตัวอย่างไปทุกๆ k จะได้หน่วยตัวอย่างเช่น rs, rs + k, rs + 2k, rs + 3k, ... จนได้ครบจำนวนขนาดตัวอย่างที่ต้องการ

2.1.2 การเลือกตัวอย่างมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS)

(Huang. C.K., 2004) ได้เสนอวิธีการเลือกตัวอย่างแบบ Mixed Systematic Random Sampling หรือ วิธีการเลือกตัวอย่างชนิดผสม ซึ่งวิธีดังกล่าวได้พัฒนามาจากการเลือกตัวอย่างแบบง่าย และการเลือกตัวอย่างแบบมีระบบ โดยมีขั้นตอนดังนี้

สมมติว่าประชากรมีลักษณะจำกัด (Finite Population) $U = (u_1, u_2, u_3, ..., u_N)$ ประกอบด้วย N หน่วย ที่มีหมายเลขตั้งแต่ 1 ถึง N ขนาดของประชากรสามารถเขียนได้ในรูป $N = nk + r = (n - r)k + r(k + 1)$ โดยที่ $1 \le r \le n$ เมื่อ r คือ เศษที่เหลือจากการหาร N ด้วย n

1) จัดเรียงประชากรในลักษณะเป็นวงกลม
2) เลือกตัวอย่างเริ่มต้น rs จาก $1 \le rs \le N$
3) แบ่งประชากรออกเป็น 2 กลุ่มย่อย โดยประชากรย่อยกลุ่มที่ 1 ประกอบด้วย

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

(n – r) k หน่วย ที่มีตัวอย่าง {rs, rs + 1, rs + 2, ..., rs + (n - r) k - 1} และประชากรย่อยกลุ่มที่ 2 ประกอบด้วยหน่วยตัวอย่างที่เหลือจากประชากรย่อยกลุ่มแรก

4) จากนั้นใช้วิธีการเลือกตัวอย่างแบบง่าย (SRS) s_{t1} จาก (n – r)k หน่วยของประชากรย่อยกลุ่มแรก และใช้วิธีเลือกตัวอย่างแบบมีระบบ (SYS) ชนิดเชิงเส้นกับประชากรที่ถูก จัดเรียงในลักษณะวงกลม s_{t2} จาก r หน่วย ได้ดังนี้

rs + (n – r) k + l(k + 1) – 1 โดยที่ l = 1, 2, 3, ..., r

5) ผลลัพธ์ของตัวอย่าง คือ s_t ขนาด n ที่เกิดจากการรวมกันของ s_{t1} และ s_{t2}

การเลือกตัวอย่างมีระบบชนิดวงกลม (Circular Systematic Sampling : CSS) ตัวประมาณความแปรปรวนของค่าเฉลี่ยตัวอย่าง คือ

$$\hat{V}(\bar{y}_{MRSS}) = \frac{1}{N^2}\sum_{u_i,u_j \in s_t}^{n}\sum_{j>i}^{n}\left(\frac{1}{\pi_{ij}} - \frac{N^2}{n^2}\right)(y_i - y_j)^2$$

โดยที่ y_i = หน่วยตัวอย่างที่ i เมื่อ i = 1, 2, 3, ..., n

N = ขนาดของประชากร

n = ขนาดของตัวอย่าง

π_{ij} = ความน่าจะเป็นร่วม (Joint inclusion probabilities)

s_t = หน่วยตัวอย่างทั้งหมดมีขนาดเท่ากับ n

u_i = หน่วยตัวอย่างจากกลุ่มที่ 1

u_j = หน่วยตัวอย่างจากกลุ่มที่ 2

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.3 การเลือกตัวอย่างมีระบบแบบวงกลม (Circular Systematic Sampling : CSS)

(Subramani J, 2018) จัดเรียงหน่วยประชากร N หน่วย $U_1, U_2, U_3, ..., U_N$ เป็นวงกลม นำเสนอโดย Lahiri

1) เลือกตัวอย่างเริ่มต้น rs จาก $1 \le rs \le N$

2) สำหรับการเลือกตัวอย่างแบบวงกลมขนาด n ให้เลือกตัวอย่างทุกๆช่วงสัดส่วนที่คำนวณไว้

จากตัวอย่างเริ่มต้น r จนกว่าจะได้จำนวนตัวอย่างครบ n หน่วย

ให้หน่วยตัวอย่างที่เลือกได้ เป็นดังนี้ U_{rs}, $U_{rs + k}$, $U_{rs + 2k}$, ..., $U_{rs + (n - 1)k}$ เป็นตัวอย่างระบบวงกลมของตัวอย่างขนาด n หน่วย สำหรับตัวอย่างเริ่มต้น rs ถ้า rs + jk > N ให้เลือกตัวอย่างที่ตรงกับ {rs + jk}(mod N) ถ้า {rs + jk}(mod N) = 0 แสดงว่า N หน่วยถูกเลือก

2.1.4 การเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval)

(Robert M. Groves et al., 2004) ได้เสนอให้จัดเรียงข้อมูลในลักษณะวงกลม คำนวณช่วงของการเลือกตัวอย่าง k แต่จะปัดเศษหลังจากการเลือกตัวอย่างเสร็จเรียบร้อย

ให้จำนวนประชากร (N) = 12,500 ขนาดตัวอย่าง (n) = 1,000

ช่วงของการสุ่ม $k = \frac{N}{n} = \frac{12,500}{1,000} = 12.5$

(กำหนดให้ค่า k ใช้ทศนิยม 1 ตำแหน่งเท่านั้น)

1) สมมติ k = 12.5 เพิ่มทศนิยมขึ้น 1 ตำแหน่งให้เป็นจำนวนเต็ม 125
2) เลือกตัวอย่างเริ่มต้น rs จาก 1 ≤ rs ≤ 125 (กำหนด rs = 28)
3) หน่วยถัดไปให้บวกทีละ 125 จะได้หน่วยตัวอย่างเป็น 28, 153, 278, 403, 528, ...
4) วางทศนิยมลง 1 ตำแหน่ง จะได้ 2.8, 15.3, 27.8, 40.3, 52.8, ...
5) เลือกตัวอย่างที่มีหมายเลขตรงกับจำนวนเต็ม (ตัดทศนิยมทิ้งทั้งหมด) จะได้ 5 หน่วยตัวอย่างแรกคือ 2, 15, 27, 40, 52, ... โดยที่ช่วงของการเลือกตัวอย่างจะแตกต่างกันไป และมีค่าเฉลี่ย 12.5 ซึ่งมีค่าเท่ากับ k

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.2 งานวิจัยที่เกี่ยวข้อง

(Charin Kukusamit & Amnuay Maneesriwongul, 1998) ได้ศึกษาเปรียบเทียบการเลือกตัวอย่างแบบมีระบบ 3 ชนิด ได้แก่ การเลือกตัวอย่างแบบมีระบบชนิดวงกลม(CSY) การเลือกตัวอย่างแบบมีระบบชนิดใหม่(NSY) และการเลือกตัวอย่างแบบมีระบบชนิดสองเส้น

(BSY) งานวิจัยฉบับนี้จึงได้ทำการศึกษาเปรียบเทียบวิธีการเลือกตัวอย่างทั้ง 3 ชนิดได้แก่ CSY, NSY และ BSY โดยใช้ประชากร 2 กลุ่ม คือ กลุ่ม A ผลผลิตข้าวรายจังหวัดปีเพาะปลูก 2535/36 และกลุ่ม B คือ จำนวนสหกรณ์การเกษตร จำแนกตามจังหวัดปี 2535/36 จำนวนของประชากรทั้งสองกลุ่มนี้ เท่ากับ 69 และ 73 ตามลำดับ และกำหนดขนาดตัวอย่างต่างๆกัน ตั้งแต่ n = 12 จนถึง 30 ผู้วิจัยทำการศึกษาจากการคำนวณค่าความคลาดเคลื่อนมาตรฐาน พบว่า ประชากรทั้งสองกลุ่มนี้มีค่าความคลาดเคลื่อนมาตรฐานคล้ายคลึงกัน ความคลาดเคลื่อนมาตรฐานของการเลือกตัวอย่างแบบมีระบบชนิดใหม่จะมีค่ามากที่สุดสำหรับทุกขนาดตัวอย่าง สำหรับกรณีที่ตัวอย่างมีขนาดเล็ก ค่าความคลาดเคลื่อนมาตรฐานจะมีค่ามากในทุกกรณี แต่เมื่อเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานของวิธีการเลือกตัวอย่างแบบมีระบบชนิดใหม่ กับ การเลือกตัวอย่างแบบมีระบบชนิดวงกลม และ การเลือกตัวอย่างแบบมีระบบชนิดสองเส้น เมื่อตัวอย่างมีขนาดใหญ่ ทำให้ทราบว่าค่าความคลาดเคลื่อนมาตรฐานของวิธีการเลือกตัวอย่างแบบมีระบบชนิดสองเส้น มีค่าน้อยที่สุด และจะยิ่งมีค่าน้อยลงเรื่อยๆ เมื่อ เศษส่วนการสุ่มตัวอย่างมีค่าตั้งแต่ 39% ขึ้นไป ในขณะที่ค่าความคลาดเคลื่อนมาตรฐานของการเลือกตัวอย่างแบบมีระบบชนิดวงกลม กับ การเลือกตัวอย่างแบบมีระบบชนิดใหม่ มีค่าพอๆกัน

(Jammaree Thuksoon, 2000) ได้ศึกษาเปรียบเทียบประสิทธิภาพของตัวสถิติจากการเลือกตัวอย่างแบบมีระบบ 4 ชนิด ได้แก่ การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น (LSY), การเลือกตัวอย่างแบบมีระบบชนิดวงกลม (CSY), การเลือกตัวอย่างแบบมีระบบชนิดสองเชิงเส้น (BSY) และการเลือกตัวอย่างแบบมีระบบชนิดใหม่ (NSY) โดยทำการศึกษาผลของค่าความคลาดเคลื่อนมาตรฐานที่มีต่อวิธีการเลือกตัวอย่างแบบมีระบบ โดยจำลองข้อมูลที่มีการแจกแจงแบบปกติจากโปรแกรมคอมพิวเตอร์ SPSS โดยกำหนดประชากรเท่ากับ 20, 30, 50, 70 และ 100 ผู้วิจัยทำการศึกษาด้วยการคำนวณค่าความแปรปรวน ความคลาดเคลื่อนมาตรฐาน และค่าสัมประสิทธิ์ความผันแปรของตัวอย่างในแต่ละวิธี แล้วนำมาเปรียบเทียบกัน ผลการศึกษาพบว่า ค่าความแปรปรวน ความคลาดเคลื่อนมาตรฐาน และค่าสัมประสิทธิ์ความผันแปรมีการเปลี่ยนแปลงอย่างไม่มีรูปแบบที่ชัดเจนแต่ค่าสถิติทั้ง 3 มีแนวโน้มที่ลดลง เมื่อขนาดตัวอย่างมีขนาดมากขึ้น การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นและชนิดวงกลม สำหรับกรณีที่ขนาดตัวอย่างเท่ากัน พบว่าค่าสถิติทั้ง 3 เท่ากัน หมายความ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ว่า การเลือกตัวอย่างทั้งสองวิธีดังกล่าวมีประสิทธิภาพพอๆกัน ในส่วนการเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นและชนิดเชิงเส้นคู่ ทั้งสองวิธีนี้ไม่สามารถนำมาเปรียบเทียบกันได้ เนื่องจากทั้งสองวิธีไม่มีขนาดตัวอย่างที่ไม่เท่ากัน การเลือกตัวอย่างแบบมีระบบชนิดวงกลมและชนิดเชิงเส้นคู่ไม่สามารถสรุปได้ว่าวิธีใดมีประสิทธิภาพมากกว่ากัน และการเลือกตัวอย่างแบบมีระบบชนิดใหม่ เมื่อแจกแจงขนาดตัวอย่างที่เป็นไปได้ทั้งหมด จะได้ขนาดตัวอย่างที่มีค่าซ้ำกันหลายค่า เมื่อขนาดตัวอย่างมีค่าเท่ากัน แต่กลับมีค่าความแปรปรวนที่แตกต่างกัน จึงไม่สามารถเปรียบเทียบความแปรปรวนของการเลือกตัวอย่างแบบมีระบบชนิดใหม่กับการเลือกตัวอย่างแบบอื่นๆได้

(Nakharin Kaewsrasean & Ampai Thongteeraparp, 2012) ได้ทำการศึกษาการเปรียบเทียบประสิทธิภาพการเลือกตัวอย่างแบบมีระบบ 4 วิธี คือ การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้น(LSY) การเลือกตัวอย่างแบบมีระบบชนิดเชิงเส้นคู่(BSY) การเลือกตัวอย่างแบบมีระบบชนิด Modified Balanced Circular (MBC) และการเลือกตัวอย่างแบบมีระบบชนิด Mixed Random Systematic Sampling (MRSS) เมื่อประชากรมีแนวโน้มเชิงเส้น (Linear Trend) ด้วยข้อมูลจำลอง โดยโปรแกรมสำเร็จรูปทางสถิติ SAS 9.1 ทำซ้ำจำนวน 2,000 ครั้ง กำหนดขนาดประชากรเป็น 1500 1000 700 500 100 และ 25 และกำหนดขนาดของตัวอย่าง 2 ขนาด คือ ขนาดใหญ่และขนาดเล็ก ทุกขนาดของประชากร เปรียบเทียบประสิทธิภาพการเลือกตัวอย่างแบบมีระบบ 4 วิธี ด้วยค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (MSE) และค่าประสิทธิภาพเชิงสัมพัทธ์ (RE) ผลการศึกษาพบว่าค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการเลือกตัวอย่างแบบมีระบบชนิดผสมมีแนวโน้มสูงกว่าการเลือกตัวอย่างอีก 3 วิธี โดยค่าเฉลี่ยตัวอย่างที่ได้จากการเลือกตัวอย่างด้วยวิธี MBC ให้ค่า MSE และ RE ต่ำที่สุด

(Huang. C.K., 2004) ได้ทำการศึกษาเปรียบเทียบวิธีการเลือกตัวอย่าง 4 วิธี ได้แก่ การเลือกตัวอย่างแบบระบบชนิด Mixed Random Systematic Sampling (MRSS) ,การเลือกตัวอย่างแบบง่าย (SRS), การเลือกตัวอย่างแบบระบบชนิดวงกลม (CSS) และการเลือกตัวอย่างแบบระบบชนิด New Partially Systematic Sampling เนื่องจากประสิทธิภาพสัมพัทธ์ขึ้นอยู่กับโครงสร้างของประชากร จึงเลือกพิจารณาประชากรบางประเภท ได้แก่ ประชากรที่มีแนวโน้มเชิงเส้น (Linear) , เอ็กซ์โพเนนเชียล (Exponential) และ ไฮเพอร์โบลิก (Hyperbolic) ผลการศึกษาพบว่า การเลือกตัวอย่างแบบระบบชนิด Mixed Random

Systematic Sampling (MRSS) ในประชากรที่มีแนวโน้มเชิงเส้น ให้ผลลัพธ์ที่ดีกว่าการเลือกตัวอย่างแบบง่าย (SRS) สำหรับประชากรทั้ง 3 ประเภท เมื่อเทียบกับการเลือกตัวอย่างแบบระบบวงกลม (CSS) และการเลือกตัวอย่างแบบระบบชนิด New Partially Systematic Sampling โดย Huang อ้างว่าวิธีการเลือกตัวอย่างแบบระบบชนิด Mixed Random Systematic Sampling (MRSS) ที่เสนอนั้นมีประสิทธิภาพมากกว่าในบางสถานการณ์ ดังนั้น วิธีการเลือกตัวอย่างแบบระบบชนิด Mixed Random Systematic Sampling (MRSS) จึงมีประสิทธิภาพดีพอๆกับ การเลือกตัวอย่างแบบระบบวงกลม (CSS) และการเลือกตัวอย่างแบบระบบชนิด New Partially Systematic Sampling นอกจากนี้การเลือกตัวอย่างแบบมีระบบชนิดผสมจะมีประสิทธิภาพสูงขึ้นเมื่อกลุ่มตัวอย่างที่ต้องการมีขนาดเล็กลง จึงควรพิจารณาในลักษณะที่ขนาดของการเลือกตัวอย่างแบบง่ายมีขนาดเล็ก

(Subramani J, 2018) ได้ทำการศึกษาการเลือกตัวอย่างแบบวงกลม เมื่อข้อมูลมีแนวโน้มเชิงเส้น โดยคำนวณค่าเฉลี่ยตัวอย่างจากการเลือกระบบชนิดวงกลมและความแปรปรวน สำหรับขนาดตัวอย่าง และขนาดประชากรคงที่ ประสิทธิภาพของการเลือกแบบระบบชนิดวงกลมกับการเลือกตัวอย่างแบบง่ายโดยไม่ใส่คืนที่ พบว่าการเลือกตัวอย่างแบบระบบชนิดวงกลมมีประสิทธิภาพดีกว่าการเลือกแบบง่ายในกรณีที่ไม่ใส่คืนที่

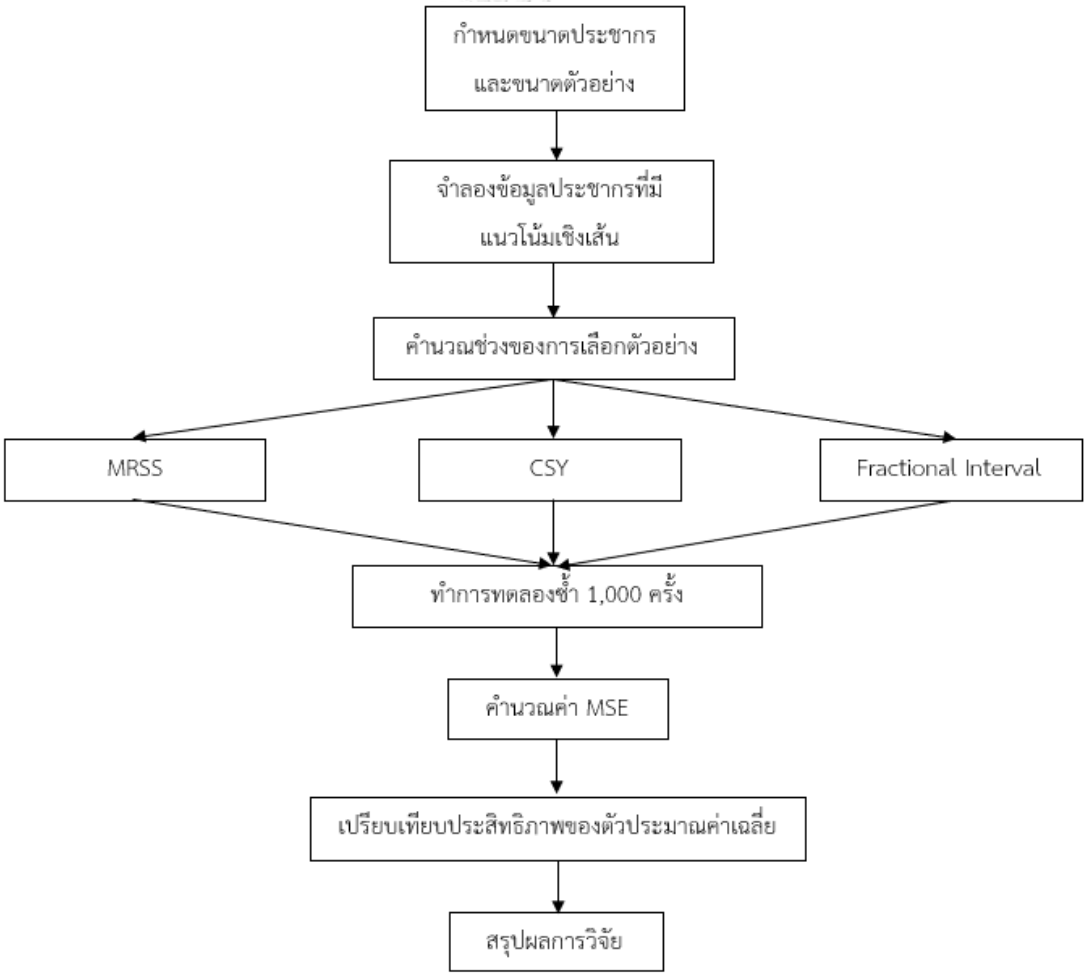
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 3
วิธีการดำเนินงานวิจัย

3.1 วิธีการดำเนินงาน

การวิจัยเรื่อง “การจำลองข้อมูลเพื่อประเมินประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบชนิดผสม” สามารถเขียนแผนผังและขั้นตอนการดำเนินงานได้ดังนี้

3.1.1 แผนผังการดำเนินงาน



3.1.2 ขั้นตอนในการดำเนินงานวิจัย

ขั้นตอนการทำวิจัย มีดังนี้

1) กำหนดขนาดของประชากร(N) ขนาดตัวอย่าง(n) ดังนี้

- ขนาดประชากรหลักร้อย

จำนวนประชากร	ขนาดของกลุ่มตัวอย่าง
300	48
300	11
500	63
500	16
700	77
700	22

- ขนาดประชากรหลักพัน

จำนวนประชากร	ขนาดของกลุ่มตัวอย่าง
3,000	429
3,000	64
3,000	9
5,000	555
5,000	73
5,000	14
7,000	764
7,000	89
7,000	17

- ขนาดประชากรหลักหมื่น

จำนวนประชากร	ขนาดของกลุ่มตัวอย่าง
30,000	978
30,000	71
50,000	1,213
50,000	81

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จำนวนประชากร	ขนาดของกลุ่มตัวอย่าง
70,000	814
70,000	99

2) จำลองข้อมูลประชากรที่มีแนวโน้มเชิงเส้น (Nakharin Kaewsrasean & Ampai Thongteeraparp, 2012) คือ

$$Y_i = 5 + 10i + e_i$$

โดยที่ i = ลำดับที่ของประชากร = 1, 2, 3, ..., N

e_i มีการแจกแจงแบบปกติ $E(e_i) = 0$ และ $Var(e_i) = i^g$

เมื่อ g คือ ค่าคงที่กำหนด (Predetermined constant)

โดยกำหนด g = 0,1,2

3) คำนวณช่วงของการเลือกตัวอย่าง $k = \frac{N}{n}$

4) ดำเนินการเลือกหน่วยตัวอย่างจากประชากรด้วยวิธีการเลือกตัวอย่างแบบมีระบบทั้ง 3 ชนิด ได้แก่การเลือกตัวอย่างแบบมีระบบชนิดผสม (MRSS), การเลือกตัวอย่างแบบมีระบบชนิดวงกลม (CSY) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval)

5) ทำการเลือกตัวอย่างแต่ละวิธี โดยทำซ้ำทั้งหมด 1,000 ครั้ง

6) คำนวณค่าเฉลี่ยและความคลาดเคลื่อนกำลังสอง (MSE)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

$$MSE(\bar{y}) = \frac{\sum_{h=1}^{1,000} (\bar{y}_h - \bar{Y})^2}{1,000}$$

โดยที่ $\bar{y}_h$ คือ ค่าเฉลี่ยตัวอย่างในการทำซ้ำครั้งที่ h

$\bar{Y}$ คือ ค่าเฉลี่ยประชากร

7) เปรียบเทียบประสิทธิภาพของการเลือกตัวอย่างแบบมีระบบทั้ง 2 กรณี ด้วยค่าประสิทธิภาพเชิงสัมพัทธ์ (Relative Efficiency : RE) ดังนี้

$$RE = \frac{MSE(\bar{y}_A)}{MSE(\bar{y}_B)}$$

ใช้เกณฑ์ในการเปรียบเทียบดังนี้

- ถ้า RE = 1 แสดงว่า $MSE(\bar{y}_A)$ มีค่าเท่ากับ $MSE(\bar{y}_B)$

  สรุปได้ว่า $\bar{y}_A$ และ $\bar{y}_B$ มีประสิทธิภาพเท่ากัน

- ถ้า RE < 1 แสดงว่า $MSE(\bar{y}_A)$ มีค่าต่ำกว่า $MSE(\bar{y}_B)$

  สรุปได้ว่า $\bar{y}_A$ มีประสิทธิภาพสูงกว่า $\bar{y}_B$

- ถ้า RE > 1 แสดงว่า $MSE(\bar{y}_A)$ มีค่าสูงกว่า $MSE(\bar{y}_B)$

  สรุปได้ว่า $\bar{y}_A$ มีประสิทธิภาพต่ำกว่า $\bar{y}_B$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
ผลการดำเนินงานวิจัย

งานวิจัยนี้มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS) กับการเลือกตัวอย่างแบบมีระบบชนิดวงกลม (Circular Systematic Sampling : CSS) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval) สำหรับกรณีช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม เมื่อประชากรมีแนวโน้มเชิงเส้น โดยจำลองข้อมูลด้วยโปรแกรม R ซึ่งกำหนดให้ประชากรมีลักษณะเชิงเส้น $Y_i = 5 + 10i + e_i$ โดยที่ $E(e_i) = 0$ และ $Var(e_i) = i^g$ เมื่อ $g = 0, 1, 2$ ผลที่ได้จากการศึกษามีดังต่อไปนี้

ตารางที่ 1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าประสิทธิภาพสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบแบบผสมและวงกลม

N	n	r	g	$MSE(\bar{y}_{MRSS})$	$MSE(\bar{y}_{CSS})$	$RE = \frac{MSE(\bar{y}_{MRSS})}{MSE(\bar{y}_{CSS})}$
300	48	12	0	7582.224711	7127.714618	1.063766595
			1	7399.353465	7597.31128	0.973943701
			2	6647051.44	43574237.15	0.152545446
	11	3	0	64112.34623	7140.104754	8.979188462
			1	54308.32668	11057.95885	4.911243335
			2	91138813.81	209197680.7	0.435658816
500	63	59	0	239.1226759	672.1057598	0.355781322
			1	8.01638478	1817.982648	0.004409495
			2	6177383.648	144648784.1	0.042706088
500	16	4	0	114798.3113	9548.769143	12.02231508
			1	154371.1541	12719.42177	12.13664873
			2	944354280.1	639291398.8	1.477189091
700	77	7	0	37779.5325	831.4065004	45.44050651
			1	42549.45877	4388.649955	9.695341211
			2	562024132.8	717050941.8	0.783799449
	22	18	0	25960.4148	8500.367721	3.054034325

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

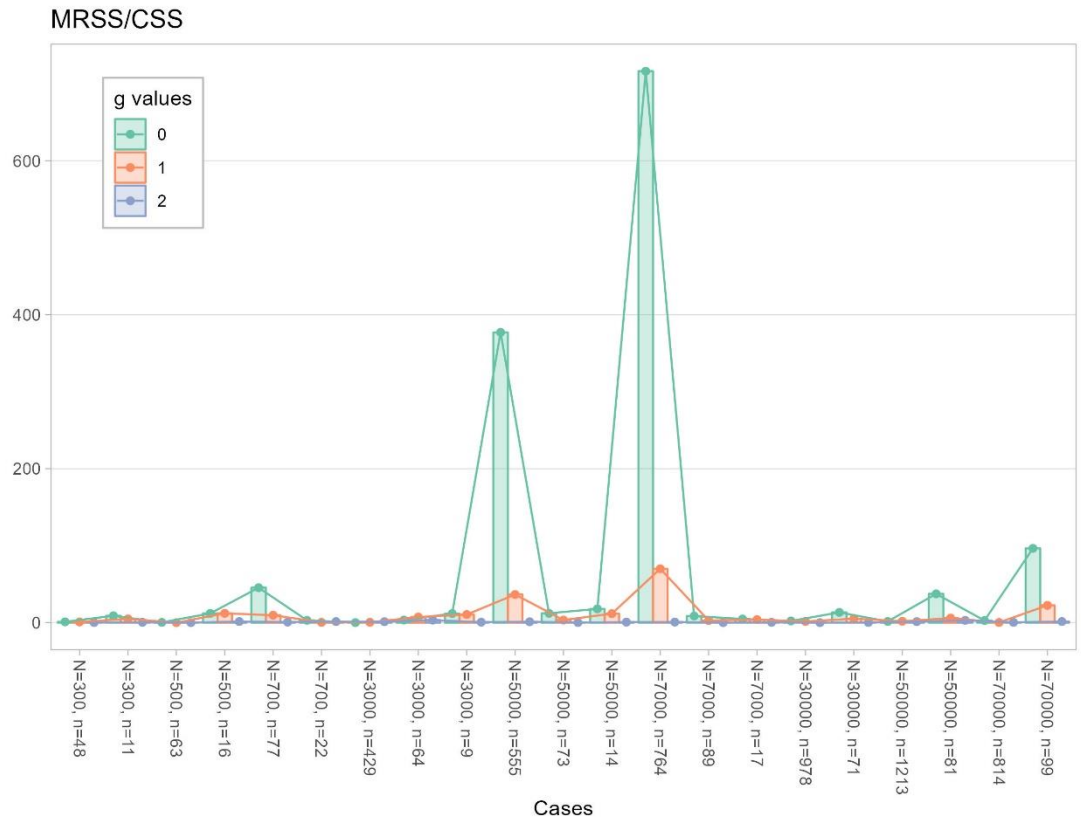
			1	5350.758157	13378.84424	0.399941733
			2	2391178157	1697673693	1.408502804
3000	429	426	0	25.51373382	471.1379834	0.054153422
			1	2297.905796	5001.252207	0.45946609
			2	50614056441	40139866251	1.260942329
	64	56	0	65627.91884	19062.07128	3.442853502
			1	685222.0298	93103.10146	7.359819588
			2	1.39605E+12	4.18953E+11	3.33224908
	9	3	0	10583773.4	894997.2791	11.82548109
			1	13195187.33	1251724.325	10.54160813
			2	1.30681E+12	1.9884E+12	0.657214063
5000	555	5	0	310933.7001	824.30892	377.205308
			1	362032.1409	9902.208538	36.56074698
			2	2.05539E+11	1.77376E+11	1.158775769
	73	36	0	667192.3262	54758.13235	12.18435139
			1	518114.3019	151577.2269	3.418153984
			2	2.64262E+11	1.56357E+12	0.169012077
	14	2	0	19519352.02	1091068.336	17.89012784
			1	18471567.5	1573954.319	11.73577103
			2	6.60433E+12	9.11409E+12	0.724628239
7000	764	124	0	581944.8143	812.1968391	716.5071155
			1	599177.6745	8556.113382	70.02918823
			2	1.91742E+11	2.34584E+11	0.817370251
	89	58	0	521347.229	60398.80205	8.631747838
			1	648937.2707	231168.9135	2.807199553
			2	38927243218	4.10123E+12	0.009491597
	17	13	0	6718004.506	1399850.79	4.799086129
			1	9569516.644	2408996.191	3.972408375
			2	1.18532E+13	2.68795E+13	0.440977458
30000	978	660	0	1815296.779	811988.2146	2.235619614

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

			1	1613452.955	1024665.921	1.574613659
			2	9.81494E+11	1.13172E+14	0.008672575
	71	38	0	20417542.96	1528424.393	13.35855608
			1	28316922.12	5491181.839	5.156799201
			2	4.73283E+14	2.0607E+15	0.229670815
50000	1213	267	0	10943530.39	6978605.701	1.568154279
			1	13206201.79	7315687.351	1.805189473
			2	1.50745E+15	1.11453E+15	1.352538906
	81	23	0	122469558	3272080.288	37.42865309
			1	92948264.28	15585259.82	5.963857219
			2	4.88901E+16	1.60786E+16	3.040701293
70000	814	810	0	165615.7063	59595.79562	2.778983057
			1	466579.4508	2056461.966	0.226884551
			2	2.57666E+15	6.33354E+15	0.406827587
70000	99	7	0	400492793	4147565.219	96.56093922
			1	423788412	18816967.89	22.52160999
			2	7.25536E+16	4.9466E+16	1.466737336

จากตารางที่ 1 การเลือกตัวอย่างด้วยวิธี CSS มีแนวโน้มที่จะมีค่า MSE น้อยกว่าการเลือกตัวอย่างด้วยวิธี MRSS แต่เมื่อ g = 2 ค่าของ MSE ของการเลือกตัวอย่างด้วยวิธี CSS จะเพิ่มขึ้น

CHULALONGKORN UNIVERSITY



รูปที่ 1 แสดงค่าประสิทธิภาพเชิงสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสมและวงกลม

ตารางที่ 2 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าประสิทธิภาพสัมพัทธ์ได้จากการเลือกตัวอย่างแบบมีระบบแบบผสมและใช้ช่วงเศษส่วน

N	n	r	g	$MSE(\bar{y}_{MRSS})$	$MSE(\bar{y}_{FI})$	$RE = \frac{MSE(\bar{y}_{MRSS})}{MSE(\bar{y}_{FI})}$
300	48	12	0	7582.224711	169.2406983	44.80142652
			1	7399.353465	419.7662807	17.6273174
			2	6647051.44	10672866.4	0.622799086
	11	3	0	64112.34623	5643.478904	11.36042985
			1	54308.32668	8517.903453	6.375785659
			2	91138813.81	145733381	0.625380494
500	63	59	0	239.1226759	281.1284595	0.850581532
			1	6177383.648	106378016.6	0.058070115

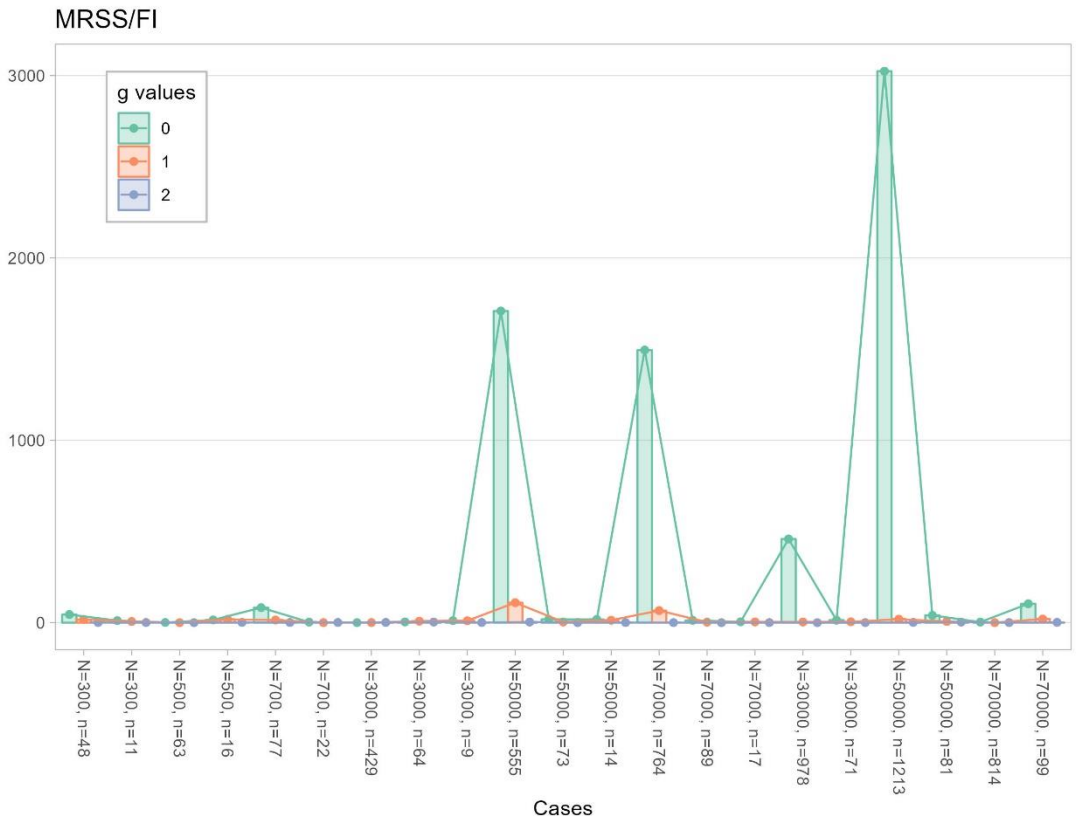
			2	114798.3113	7544.212403	15.21673902
500	16	4	0	154371.1541	8472.850012	18.21950747
			1	944354280.1	753370736.8	1.253505391
			2	37779.5325	458.869502	82.33175735
700	77	7	0	42549.45877	2908.627356	14.62870748
			1	562024132.8	573731946.5	0.979593582
			2	25960.4148	8162.497659	3.180449892
	22	18	0	5350.758157	15239.89826	0.35110196
			1	2391178157	2480167120	0.964119772
			2	25.51373382	127.4168148	0.200238358
3000	429	426	0	2297.905796	4421.252406	0.519740921
			1	50614056441	40566822161	1.247671219
			2	65627.91884	16007.98879	4.099697951
	64	56	0	685222.0298	82172.84727	8.338788939
			1	1.39605E+12	3.79781E+11	3.67594101
			2	10583773.4	927648.7781	11.40924631
	9	3	0	13195187.33	1244979.056	10.59872234
			1	1.30681E+12	1.81217E+12	0.721129265
			2	310933.7001	181.9319568	1709.065881
5000	555	5	0	362032.1409	3302.400929	109.6269498
			1	2.05539E+11	58028484841	3.542033941
			2	667192.3262	36243.35129	18.40868194
	73	36	0	518114.3019	131032.7763	3.954081693
			1	2.64262E+11	1.67353E+12	0.157907064
			2	19519352.02	1090126.48	17.9055847
	14	2	0	18471567.5	1373063.783	13.45281095
			1	6.60433E+12	8.84481E+12	0.746689802
		2	2	581944.8143	389.2377123	1495.088466
7000	764	124	0	599177.6745	9034.504908	66.32103039
			1	1.91742E+11	4.87056E+11	0.39367525

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

			2	521347.229	40790.50099	12.78109404
	89	58	0	648937.2707	223297.3008	2.906158151
			1	6177383.648	106378016.6	0.058070115
			2	38927243218	5.6938E+12	0.006836781
	17	13	0	6718004.506	1453833.784	4.620888977
			1	9569516.644	2329042.163	4.108777761
			2	1.18532E+13	2.85836E+13	0.414685937
30000	978	660	0	1815296.779	3955.321136	458.950542
			1	1613452.955	420421.4183	3.837703992
			2	9.81494E+11	2.06496E+14	0.004753098
	71	38	0	20417542.96	1477100.255	13.82271981
			1	28316922.12	5990789.689	4.726742816
			2	4.73283E+14	2.26986E+15	0.208507663
50000	1213	267	0	10943530.39	3618.463026	3024.358771
			1	13206201.79	691237.0481	19.10517069
			2	1.50745E+15	1.00161E+15	1.505022609
	81	23	0	122469558	3061177.569	40.00733548
			1	92948264.28	14640039.24	6.348908139
			2	4.88901E+16	1.47454E+16	3.315618728
70000	814	810	0	165615.7063	50782.54814	3.261272078
			1	466579.4508	1970670.21	0.236761812
			2	2.57666E+15	6.01417E+15	0.428431131
70000	99	7	0	400492793	3861509.579	103.7140488
			1	423788412	21261370.41	19.93231874
			2	7.25536E+16	4.66181E+16	1.556339161

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จากตารางที่ 2 การเลือกตัวอย่างด้วยวิธี FI มีแนวโน้มที่จะมีค่า MSE น้อยกว่าการเลือกตัวอย่างด้วยวิธี MRSS แต่เมื่อ $g = 2$ ค่าของ MSE ของการเลือกตัวอย่างด้วยวิธี FI จะเพิ่มขึ้น



รูปที่ 2 แสดงค่าประสิทธิภาพเชิงสัมพัทธ์ที่ได้จากการเลือกตัวอย่างแบบมีระบบชนิดผสมและใช้ช่วงเศษส่วน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 5
สรุปผลการศึกษา อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย

จากการศึกษาและเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยด้วยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าประสิทธิภาพสัมพัทธ์ (RE) ของ การเลือกตัวอย่างแบบมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS) กับการเลือกตัวอย่างแบบมีระบบชนิดวงกลม (Circular Systematic Sampling : CSS) และการเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval) สำหรับกรณีช่วงของการเลือกตัวอย่างไม่เป็นจำนวนเต็ม เมื่อประชากรมีแนวโน้มเชิงเส้น โดยการจำลองข้อมูลประชากรด้วยโปรแกรม R พบว่า ขนาดประชากร ขนาดตัวอย่างและช่วงของการเลือกตัวอย่างไม่มีผลกับประสิทธิภาพในการเลือกตัวอย่าง การเลือกตัวอย่างแบบมีระบบด้วยวิธี MRSS มีค่า MSE สูงกว่าการเลือกตัวอย่างอีกทั้ง 2 วิธี แต่เมื่อค่า g = 2 จะทำให้ค่าของ MSE ของการเลือกตัวอย่างทั้ง 3 วิธีมีค่ามากขึ้น โดยที่ค่า MSE ของการเลือกตัวอย่างแบบมีระบบชนิดผสมมีค่าต่ำกว่าการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและวิธีใช้ช่วงเศษส่วน ทั้งนี้เป็นผลเนื่องมาจากค่า g เป็นค่าที่กำหนดความเป็นเชิงเส้น เมื่อค่า g เพิ่มมากขึ้น ความเป็นเชิงเส้นของประชากรจะลดลง ทำให้ตัวประมาณค่าเฉลี่ยตัวอย่างที่คำนวณได้มีค่าต่างจากค่าเฉลี่ยประชากรมากตามไปด้วย

อภิปรายผลการวิจัยและข้อเสนอแนะ

จากการศึกษาวิจัยครั้งนี้สอดกับงานวิจัยก่อนหน้า (Nakharin Kaewsrasean & Ampai Thongteerapaprp, 2012) การเลือกตัวอย่างแบบมีระบบชนิดผสม มีค่า MSE สูงกว่าการเลือกตัวอย่างอีกทั้ง 2 วิธีในทุกกรณี แต่เมื่อ g = 2 จะทำให้ค่า MSE ของการเลือกตัวอย่างแบบมีระบบชนิดผสมมีค่าลดลงเมื่อเทียบกับการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและวิธีการใช้ช่วงเศษส่วน ขณะที่เมื่อ g = 0 และ g = 1 ค่า MSE ของการเลือกตัวอย่างด้วยวิธี MRSS สูงกว่าการเลือกตัวอย่างอีก 2 วิธีค่อนข้างมาก และการเลือกตัวอย่างแบบมีระบบชนิดผสม เมื่อ g = 2 นั่นคือเมื่อประชากรมีลักษณะสูญเสียความเป็นเชิงเส้นจะทำให้มีประสิทธิภาพที่สูงขึ้น ซึ่งสอดคล้องกับงานวิจัยของ (Huang. C.K., 2004) ที่บอกว่าประสิทธิภาพสัมพัทธ์ขึ้นอยู่กับโครงสร้างประชากร และยังกล่าวอีกว่าการเลือกตัวอย่างแบบมีระบบชนิดผสมจะมีประสิทธิภาพที่ดีกว่าในบางกรณี ทั้งนี้อาจจะเป็นผลมาจากค่า r หรือเศษเหลือจากการหาร N ด้วย n ที่มีค่าเข้าใกล้มีค่าเข้าใกล้ n หรือขนาดตัวอย่างมากขึ้น อันส่งให้ต้องทำการเลือกตัวอย่างด้วยวิธี Simple random sampling ในสัดส่วนที่มากขึ้นจะทำให้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ประสิทธิภาพที่ได้จากการเลือกตัวอย่างด้วยวิธีนี้มีค่าน้อยลง ซึ่งหมายความว่า การเลือกตัวอย่างแบบมีระบบชนิดผสมจะยิ่งได้ประสิทธิภาพที่ดีขึ้นถ้าค่า r ค่ามีต่างจากค่า n

ข้อเสนอแนะ

ในกรณีที่ข้อมูลประชากรมีลักษณะเป็นเชิงเส้นแนะนำวิธีการเลือกตัวอย่างแบบมีระบบชนิดวงกลมและใช้ช่วงเศษส่วนเพราะทั้ง 2 วิธีดังกล่าวมีประสิทธิภาพใกล้เคียงกันและยังมีประสิทธิภาพในการประมาณค่าเฉลี่ยที่ดีกว่าวิธีการเลือกตัวอย่างแบบมีระบบชนิดผสม และเมื่อค่า i มีค่ามากขึ้นจะส่งผลให้ความแปรปรวนมีค่ามากขึ้นตามไปด้วยแสดงว่าการเลือกตัวอย่างแบบ Mixed Systematic Sampling เหมาะกับข้อมูลที่มีแนวโน้มการกระจายที่ไม่คงที่เช่น ข้อมูลเศรษฐศาสตร์ สังคมศาสตร์ เป็นต้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บรรณานุกรม

Charin Kukusamit, & Amnuay Maneesriwongul. (1998). Comparison of Three Random Sampling Method : Circular Systematic Sampling, New Systematic Sampling, and Bilinear Systematic Sampling. *Nida Development Journal*, *38*(1), 13-18.

David G. Hankin, Michael S. Mohr, & Ken B. Newman. (2019). Systematic Sampling. In *Sampling Theory: For the Ecological and Natural Resource Sciences* (pp. 48-67). Oxford University Press.

Huang. C.K. (2004). Mixed random systematic sampling designs. *Metrika*, *59*, 1-11.

Jammaree Thuksoon. (2000). Comparative Study of Efficiency of Statistics from Various Types of Systematic Sampling [Unpublished master's thesis]. *Chiang Mai University*.

Nakharin Kaewsrasean, & Ampai Thongteeraparp. (2012). Comparative Efficiency of Four Systematic Sampling Methods. *Veridian E-Journal SU*, *5*(3), 420-428.

Prachum Suwattee. (2009). *Sample surveys: sampling designs and analysis* (Vol. 1). National Institute of Development Administration.

Robert M. Groves, Floy J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, & Roger Tourangeau. (2004). *Servey Methedology* (Vol. 1). Wiley-Interscience.

Subramani J. (2018). On Circular systematic sampling in the presence of linear trend. *Biometrics & Biostatistics International Journal*, *7*(4), 286-292.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาคผนวก

การเลือกตัวอย่างมีระบบชนิดผสม (Mixed Systematic Random Sampling : MRSS)

สำหรับ N = 300, n = 48 และ g = 0

```
# mrss ----


NN = 300 # pop size
n = 48 # sample size
g = 0

# generate population  ----
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

k = floor(NN/n)
r = NN %% n
pop = seq(1:NN)
rand =  sample(1:NN,1)
t_last = rand+((n-r)*k)-1
t_last

# define group1,2
if (t_last < NN){
  group1 = pop[rand:t_last]

  group2_1 = (t_last+1):NN
  group2_2 = 1:(rand-1)
  group2 = c(group2_1,group2_2)
} else if (t_last > NN) {
  group1_1 = pop[rand: NN]
  group1_2 = pop[1:(t_last-NN)]
  group1 = c(group1_1,group1_2)

  group2 = setdiff(pop,group1)
} else if (t_last==NN){
  group1 = pop[rand:t_last]
  group2 = NULL
}

group1
group2

# SRS for group1
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
select1 = sample(group1, n-r)


# SYS for group2
select2 = group2[(1:r)*(k+1)]
select2


# combine 2 sampling groups
c_group <- c(select1,select2)

c_group
```

สำหรับ N = 300, n = 48 และ g = 1

```
# mrss ----


NN = 300 # pop size
n = 48 # sample size
g = 1

# generate population  ----
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

k = floor(NN/n)
r = NN %% n
pop = seq(1:NN)
rand =  sample(1:NN,1)
t_last = rand+((n-r)*k)-1
t_last

# define group1,2
if (t_last < NN){
  group1 = pop[rand:t_last]

  group2_1 = (t_last+1):NN
  group2_2 = 1:(rand-1)
  group2 = c(group2_1,group2_2)
} else if (t_last > NN) {
  group1_1 = pop[rand: NN]
  group1_2 = pop[1:(t_last-NN)]
  group1 = c(group1_1,group1_2)
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
  group2 = setdiff(pop,group1)
} else if (t_last==NN){
  group1 = pop[rand:t_last]
  group2 = NULL
}

group1
group2

# SRS for group1
select1 = sample(group1, n-r)


# SYS for group2
select2 = group2[(1:r)*(k+1)]
select2


# combine 2 sampling groups
c_group <- c(select1,select2)

c_group
```

สำหรับ N = 300, n = 48 และ g = 2

```
# mrss ----


NN = 300 # pop size
n = 48 # sample size
g = 2

# generate population  ----
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

k = floor(NN/n)
r = NN %% n
pop = seq(1:NN)
rand =  sample(1:NN,1)
t_last = rand+((n-r)*k)-1
t_last

# define group1,2
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
if (t_last < NN){
  group1 = pop[rand:t_last]

  group2_1 = (t_last+1):NN
  group2_2 = 1:(rand-1)
  group2 = c(group2_1,group2_2)
} else if (t_last > NN) {
  group1_1 = pop[rand: NN]
  group1_2 = pop[1:(t_last-NN)]
  group1 = c(group1_1,group1_2)

  group2 = setdiff(pop,group1)
} else if (t_last==NN){
  group1 = pop[rand:t_last]
  group2 = NULL
}

group1
group2

# SRS for group1
select1 = sample(group1, n-r)


# SYS for group2
select2 = group2[(1:r)*(k+1)]
select2


# combine 2 sampling groups
c_group <- c(select1,select2)

c_group
```

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

การเลือกตัวอย่างมีระบบแบบวงกลม (Circular Systematic Sampling : CSS)

สำหรับ N = 300, n = 48 และ ε = 0

```
# css ----
# create sys function
sys = function(N,n){
  k = ceiling(N/n)
  r = sample(1:k, 1)
  seq(r, r + k*(n-1), k)
}
```

```
NN = 300 # pop size
n = 48 # sample size
g = 0

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

# do css

k = ceiling(NN/n) # interval
rand = sample(1:NN,1) # random start point


# there are 2 possibilities,
#1. the last one(r+k*n) dont exceed NN, only one range is
found(rand to NN)
#2. the last one is larger than NN, 2 ranges(r to NN; 1
to rand)

if ((rand+k*n) <= NN) {
  u = seq(from = rand, to = NN, by = k)
  u = u[1:n]

} else if ((rand+k*n) > NN) {
  # create vector of rand to NN and 1 to rand-1
  range1 = seq(from=rand, to = NN, by=1)
  range2 = seq(from=1, to = rand-1, by=1)

  # combine 2 range
  range = c(range1,range2,range1,range2)

  # access u vector from index
  u_index = seq(from = 1, to = length(range), by = k)

  u = range[u_index]

  u = u[1:n]

}


u <- unique(u)
u
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
# project on linear pop, then find mean
mean_y = mean(y[u])
# mean_y is css

u # the selected positions of y pop
```

สำหรับ N = 300, n = 48 และ g = 1

```
# css ----
# create sys function
sys = function(N,n){
  k = ceiling(N/n)
  r = sample(1:k, 1)
  seq(r, r + k*(n-1), k)
}

NN = 300 # pop size
n = 48 # sample size
g = 1

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

# do css

k = ceiling(NN/n) # interval
rand = sample(1:NN,1) # random start point


# there are 2 possibilities,
#1. the last one(r+k*n) dont exceed NN, only one range is
found(rand to NN)
#2. the last one is larger than NN, 2 ranges(r to NN; 1
to rand)

if ((rand+k*n) <= NN) {
  u = seq(from = rand, to = NN, by = k)
  u = u[1:n]

} else if ((rand+k*n) > NN) {
  # create vector of rand to NN and 1 to rand-1
  range1 = seq(from=rand, to = NN, by=1)
  range2 = seq(from=1, to = rand-1, by=1)
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
  # combine 2 range
  range = c(range1,range2,range1,range2)

  # access u vector from index
  u_index = seq(from = 1, to = length(range), by = k)

  u = range[u_index]

  u = u[1:n]

}


u <- unique(u)
u

# project on linear pop, then find mean
mean_y = mean(y[u])
# mean_y is css

u # the selected positions of y pop
```

สำหรับ N = 300, n = 48 และ g = 2

```
# css ----
# create sys function
sys = function(N,n){
  k = ceiling(N/n)
  r = sample(1:k, 1)
  seq(r, r + k*(n-1), k)
}

NN = 300 # pop size
n = 48 # sample size
g = 2

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}

# do css

k = ceiling(NN/n) # interval
rand = sample(1:NN,1) # random start point
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
# there are 2 possibilities,
#1. the last one(r+k*n) dont exceed NN, only one range is
found(rand to NN)
#2. the last one is larger than NN, 2 ranges(r to NN; 1
to rand)

if ((rand+k*n) <= NN) {
  u = seq(from = rand, to = NN, by = k)
  u = u[1:n]

} else if ((rand+k*n) > NN) {
  # create vector of rand to NN and 1 to rand-1
  range1 = seq(from=rand, to = NN, by=1)
  range2 = seq(from=1, to = rand-1, by=1)

  # combine 2 range
  range = c(range1,range2,range1,range2)

  # access u vector from index
  u_index = seq(from = 1, to = length(range), by = k)

  u = range[u_index]

  u = u[1:n]

}


u <- unique(u)
u

# project on linear pop, then find mean
mean_y = mean(y[u])
# mean_y is css

u # the selected positions of y pop
```

การเลือกตัวอย่างแบบมีระบบโดยใช้ช่วงเศษส่วน (Fractional Interval)

สำหรับ N = 300, n = 48 และ g = 0

```
# fi ----

NN = 300 # pop size
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
n = 48 # sample size
g = 0

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}


# do factorial

k = round(NN/n, digits = 1)*10
rand = sample(10:k,1)

u = seq(from = rand, to = NN*10, by = k)
u1 = floor(u/10)

u1
u2<-unique(u1)

# project on linear pop, then find mean
mean_y = mean(y[u1])
# mean_y is factorial
u1 # the selected positions of y pop
```

สำหรับ N = 300, n = 48 และ g = 1

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

```
# fi ----

NN = 300 # pop size
n = 48 # sample size
g = 1

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}


# do factorial

k = round(NN/n, digits = 1)*10
rand = sample(10:k,1)
```

```
u = seq(from = rand, to = NN*10, by = k)
u1 = floor(u/10)

u1
u2<-unique(u1)

# project on linear pop, then find mean
mean_y = mean(y[u1])
# mean_y is factorial
u1 # the selected positions of y pop
```

สำหรับ N = 300, n = 48 และ g = 2

```
# fi ----

NN = 300 # pop size
n = 48 # sample size
g = 2

# population generate
y=NULL
for (i in 1:NN) {
  y[i]= 5 + 10*i+rnorm(1,0,i^g)
}


# do factorial

k = round(NN/n, digits = 1)*10
rand = sample(10:k,1)

u = seq(from = rand, to = NN*10, by = k)
u1 = floor(u/10)

u1
u2<-unique(u1)

# project on linear pop, then find mean
mean_y = mean(y[u1])
# mean_y is factorial
u1 # the selected positions of y pop
```

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน

ชื่อ-สกุล	นภสร รัตนวุฒิขจร
วัน เดือน ปี เกิด	5 พฤษภาคม 2540
สถานที่เกิด	กรุงเทพฯ
วุฒิการศึกษา	ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาสถิติ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ในปีการศึกษา 2561
ที่อยู่ปัจจุบัน	141/33 ถ.ตลิ่งชัน-สุพรรณบุรี ต.ละหาร อ.บางบัวทอง จ.นนทบุรี 11110

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University