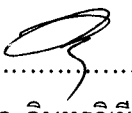
แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด: กรณีศึกษาของการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548

นาวาตรีพงศ์พัฒน์ อิศรกุล

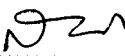
ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาวิจัย วัดผลและสถิติการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา
กรกฎาคม 2560
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

คณะกรรมการควบคุมดุษฎีนิพนธ์และคณะกรรมการสอบดุษฎีนิพนธ์ ได้พิจารณาดุษฎีนิพนธ์ของ นาวาตรีพงศ์พัฒน์ อิศรกุล ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาวิจัย วัดผลและสถิติการศึกษา ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมดุษฎีนิพนธ์

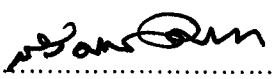


.....อาจารย์ที่ปรึกษาหลัก
(ดร.นฤมล อินทรวิเชียร)



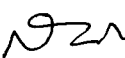
..... อาจารย์ที่ปรึกษาร่วม
(ดร.สมศักดิ์ ลิลา)

คณะกรรมการสอบดุษฎีนิพนธ์



.....ประธาน
(ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก)

.....กรรมการ
(ดร.นฤมล อินทรวิเชียร)



..... กรรมการ
(ดร.สมศักดิ์ ลิลา)

..... กรรมการ
(ดร.สมพงษ์ ปั้นหุ่น)

คณะศึกษาศาสตร์อนุมัติให้รับดุษฎีนิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาวิจัย วัดผลและสถิติการศึกษา ของมหาวิทยาลัยบูรพา

..... คณบดีคณะศึกษาศาสตร์
(รองศาสตราจารย์ ดร.วิชิต สุรัตน์เรืองชัย)
วันที่ 16 เดือน มิถุนายน พ.ศ. 2560

กิตติกรรมประกาศ

ดุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี เนื่องจากผู้วิจัยได้รับความช่วยเหลือ ดูแล เอาใจใส่เป็นอย่างดีจากหลาย ๆ ฝ่าย คือ อาจารย์ ผู้ช่วยศาสตราจารย์ ดร.สังวรณ์ งัดกระโทก (ประธานกรรมการสอบ) อาจารย์ ดร.สมพงษ์ ปั้นหุ่น (กรรมการสอบ) อาจารย์ ดร.นฤมล อินทรวิเชียร (อาจารย์ที่ปรึกษา) และอาจารย์ ดร.สมศักดิ์ ลิลา (อาจารย์ที่ปรึกษาร่วม) ในการแนะนำ ตรวจแก้ไข ให้ข้อเสนอแนะ ติดตามความก้าวหน้าในการดำเนินการวิจัย รวมทั้งเจ้าหน้าที่งานบริการการศึกษาทุกท่าน โดยเฉพาะ นางสาวจิรนันท์ พุ่มเหรียญ ทำให้ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาของอาจารย์และเจ้าหน้าที่ทุกท่านเป็นอย่างยิ่ง และขอขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้ ขอขอบพระคุณในความเอื้อเฟื้อของนิสิตหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาวิจัย วัดผล และสถิติการศึกษา ทุกท่านที่ได้ร่วมทุกข์สุขตลอดระยะเวลาการศึกษา ตลอดจนพี่วาสนา กลมอ่อน ซึ่งเป็นพี่สาวที่แสนดีและได้คอยให้การช่วยเหลือ และกราบขอบพระคุณบุพการีผู้ให้ทุกสิ่งทุกอย่างกับผู้วิจัย รวมถึงผู้มีพระคุณทุกท่านที่มิได้เอ่ยนามไว้ ณ ที่นี้

นาวาตรีพงศ์พัฒน์ อิศรกุล

52810149: สาขาวิชา: วิจัย วัดผล และสถิติการศึกษา; ปร.ด. (วิจัย วัดผล และสถิติการศึกษา)
คำสำคัญ: การสุ่มแบบหลายขั้นตอน/ การสุ่มแบบกริด/ ระบบสารสนเทศภูมิศาสตร์ และการทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน

นาวาตรีพงศ์พัฒน์ อิศรกุล: แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด: กรณีศึกษาของการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548 (SPATIAL INTERPOLATION MODEL OF GRID SAMPLING: A CASE STUDY OF O-NET IN 2005) คณะกรรมการควบคุมดุษฎีนิพนธ์: นฤมล อินทรวิเชียร, วท.ด., สมศักดิ์ ลิลา, กศ.ด., 239 หน้า. ปี พ.ศ. 2560.

วิธีการสุ่มแบบกริดเป็นวิธีการสุ่มตัวอย่างที่แพร่หลายสำหรับงานวิจัยทางด้านวิทยาศาสตร์กายภาพ ในกรณีที่ประชากรมีจำนวนมากและอาศัยในพื้นที่ขนาดใหญ่ ซึ่งการวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อทดสอบว่าวิธีการสุ่มแบบกริดมีความเหมาะสมกับงานวิจัยทางพฤติกรรมและสังคมศาสตร์หรือไม่อย่างไร โดยมีขั้นตอนการวิจัย 2 ขั้นตอน ได้แก่

ขั้นตอนที่ 1 การทดสอบวิธีการสุ่มแบบกริด โดยอาศัยแบบจำลองที่สร้างขึ้นโดยการเปรียบเทียบกับวิธีการสุ่มแบบอื่น ๆ จากผลการทดลองพบว่า เมื่อใช้วิธีการสุ่มแบบกริดกับแบบจำลอง S0 และ S00 ซึ่งมีค่าสหสัมพันธ์เชิงพื้นที่เป็นศูนย์ ทำให้ค่าเฉลี่ยของความคลาดเคลื่อนและค่าความแตกต่างความแปรปรวนมีค่ามาก ซึ่งแสดงให้เห็นว่าการสุ่มตัวอย่างแบบกริดไม่เหมาะสมกับประชากรที่มีสหสัมพันธ์เชิงพื้นที่เป็นศูนย์ ในส่วนแบบจำลอง S1-S4 ซึ่งเป็นแบบจำลองที่มีสหสัมพันธ์เชิงพื้นที่ทางบวก เมื่อนำมาทดสอบกับวิธีการสุ่มต่าง ๆ ได้ผลการทดลอง คือ วิธีการสุ่มแบบแบ่งชั้นภูมิ วิธีการสุ่มแบบมีระบบ และวิธีการสุ่มแบบกริด เป็นวิธีการที่มีความแม่นยำสูง แต่วิธีการสุ่มอย่างง่ายและวิธีการสุ่มแบบแบ่งกลุ่มมีความแม่นยำต่ำในทุกแบบจำลอง นอกจากนั้นในการทดสอบความแม่นยำการประมาณค่าเชิงพื้นที่ของวิธีการสุ่มแบบกริดด้วยแบบจำลอง SI1- SI3 พบว่า เมื่อสหสัมพันธ์เชิงพื้นที่มีค่าเป็นบวกจะสามารถทำการประมาณค่าเชิงพื้นที่ได้อย่างถูกต้องและแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดขึ้นมีความน่าเชื่อถือและสามารถนำไปใช้งานได้

ขั้นตอนที่ 2 การทดสอบโดยอาศัยฐานข้อมูลคะแนน O-NET เพื่อตรวจสอบความแม่นยำในการประมาณค่าของวิธีการสุ่มแบบกริด วิธีการสุ่มแบบหลายขั้นตอน และวิธีการสุ่มแบบกริด-ชั้นภูมิ จากผลการทดลอง พบว่า วิธีการสุ่มแบบกริดมีความแม่ยำต่ำที่สุด โดยมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 3.18 คะแนน วิธีการสุ่มแบบหลายขั้นตอนมีความแม่นยำเป็นอันดับ 2 โดยมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.52 คะแนน และวิธีการสุ่มแบบกริด-ชั้นภูมิมีความแม่นยำมากที่สุด โดยค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.05 คะแนน นอกจากนี้แผนที่ระบบสารสนเทศภูมิศาสตร์ที่ได้จากวิธีการสุ่มตัวอย่างแบบกริด หรือแบบกริด-ชั้นภูมิ ได้ให้สารสนเทศทั้งในระดับประเทศและระดับท้องถิ่น ซึ่งแตกต่างจากวิธีการสุ่มแบบหลายขั้นตอนที่ให้สารสนเทศเฉพาะระดับประเทศเท่านั้น

52810149: MAJOR: MEASUREMENT AND STATISTICS; PH.D. (RESEARCH, MEASUREMENT AND STATISTICS)

KEYWORDS: MULTISTAGE RANDOM SAMPLING/ GRID SAMPLING/ GEOGRAPHIC INFORMATION SYSTEM AND ORDINARY NATIONAL EDUCATIONAL TEST.

LCdr. PONGPAT ISSARAKUL: SPATIAL CORRELATION MODEL OF GRID SAMPLING: A CASE STUDY OF O-NET IN 2005. DISSERTATION ADVISORS: NARUMOL INTRAWICHIEN, Ph.D., SOMSAK LILA, Ed.D., 239 P. 2017.

Grid sampling is a popular sampling method for research in the physical sciences, in the case of a large population and large areas. The purpose of this research is to test whether grid-based random sampling is appropriate for behavioral and social science research, two steps are required.

Step 1: Testing the Grid Sampling method by testing a built model and by comparison with other sampling methods. The results of the experiment are: when using grid sampling with S0 and S00 models, which has zero spatial correlation, the mean of the discrepancies and the difference of the variance is very large. This shows that grid-based sampling is not appropriate for population with zero spatial correlation. In the S1-S4 model with positive spatial correlation when tested with various random methods, the results are: A Stratified Random Sampling method, Systematic Random Sampling method and the Grid Sampling method are all highly accurate methods. However, the accuracy of Simple Random Sampling method and Cluster Random Sampling methods are all low. In addition, in the spatial accuracy estimation of grid-based method with the SI1-SI3 model, it is found that: When the spatial correlation is positive, the spatial correlation can be performed correctly and the map of Geographical Information System (GIS) is credible and can be used.

Step 2: Using the database of O-NET scoring test to verify the accuracy of the grid sampling method, Multistage Random Sampling and Stratified Random Sampling method, the results are: The grid method has the lowest accuracy. The average error rate is 3.18 points. The multistage random sampling method is the second most accurate with 2.52 mean error scores and the most accurate method is Grid-Stratified Random Sampling method. The average error rate is 2.05 points. In addition, the GIS maps derived from the Grid Sampling method or Grid-Stratified Random Sampling method provided information at national and local level. This is different from Multistage Random Sampling that provides information only at national level.

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง ..... ฌ

สารบัญภาพ ..... ญ

บทที่

1 บทนำ 1

ความเป็นมาและความสำคัญของปัญหา 1

คำถามวิจัย 4

วัตถุประสงค์ของการวิจัย 5

กรอบแนวคิดในการวิจัย 5

ประโยชน์ที่ได้รับจากการวิจัย 7

ขอบเขตของการวิจัย 7

นิยามศัพท์เฉพาะ 8

2 เอกสารและงานวิจัยที่เกี่ยวข้อง 12

ตอนที่ 1 ความหมาย ประเภท และวิธีการของการสุ่มตัวอย่างจากประชากร 12

ตอนที่ 2 แนวคิดและทฤษฎีของการสุ่มแบบกริด 17

ตอนที่ 3 การสุ่มตัวอย่างจำนวนมาก ๆ ที่ใช้ในการทดสอบทางการศึกษาของประเทศไทย 49

ตอนที่ 4 การสุ่มตัวอย่างมีผลกระทบต่อความแม่นยำของโพลล์ 55

ตอนที่ 5 แนวคิดในการสร้างแบบจำลอง 58

ตอนที่ 6 งานวิจัยในประเทศและต่างประเทศที่เกี่ยวข้อง 66

3 วิธีการดำเนินการวิจัย 82

ประชากรและกลุ่มตัวอย่าง 82

ตัวแปรที่ใช้ในการวิจัย 85

ขั้นตอนการวิจัย 86

เครื่องมือที่ใช้ในการวิจัย 93

แบบจำลองการสุ่ม 105

สารบัญ (ต่อ)

บทที่		หน้า
	แบบจำลองการประมาณค่าเชิงพื้นที่	117
	การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน	122
	การสุ่มตัวอย่างแบบหลายขั้นตอน	129
	การสุ่มตัวอย่างแบบกริด	133
	การสุ่มตัวอย่างแบบกริด-ชั้นภูมิ	141
	การวิเคราะห์ข้อมูล	145
4	ผลการวิเคราะห์ข้อมูล	146
	สัญลักษณ์ที่ใช้ในการวิเคราะห์	147
	ตอนที่ 1 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยแบบจำลองที่สร้างขึ้น	147
	ตอนที่ 2 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยฐานข้อมูลคะแนน O-NET	163
5	สรุป อภิปรายผล และข้อเสนอแนะ	181
	วัตถุประสงค์การวิจัย	181
	ประชากรและกลุ่มตัวอย่าง	181
	ตัวแปรที่ใช้ในการวิจัย	182
	เครื่องมือที่ใช้ในการทดลอง	183
	แบบจำลองการสุ่ม	183
	วิธีการสุ่มตัวอย่างแบบหลายขั้นตอน	184
	วิธีการสุ่มตัวอย่างแบบกริด	185
	วิธีการสุ่มตัวอย่างแบบกริด-ชั้นภูมิ	185
	การวิเคราะห์ข้อมูล	186
	การสรุปผล	187
	การอภิปรายผล	189
	ข้อเสนอแนะ	192
บรรณานุกรม		195
ภาคผนวก		202

สารบัญ (ต่อ)

บทที่ หน้า

ภาคผนวก ก..... 203

ภาคผนวก ข 215

ภาคผนวก ค 227

ประวัติย่อของผู้วิจัย..... 239

สารบัญตาราง

ตารางที่ หน้า

2-1 ตารางของทาโร ยามาเน่ 29
2-3 จำนวนโรงเรียนและนักเรียนกลุ่มตัวอย่างชั้นประถมศึกษาปีที่ 4 50
2-4 จำนวนโรงเรียนและนักเรียนกลุ่มตัวอย่างชั้นมัธยมศึกษาปีที่ 2 51
3-1 ความหมายของสีตามช่วงเกรดเฉลี่ยสะสม 123
3-2 ความหมายของสีตามช่วงคะแนน O-NET 125
3-3 ความหมายของสีตามช่วงของค่าสัมประสิทธิ์สหสัมพันธ์ 128
3-4 ร้อยละของจำนวนโรงเรียนตามภูมิภาค 130
3-5 จำนวนโรงเรียนเรียนที่ต้องถูกสุ่มออกมา 131
3-6 จำนวนจังหวัดที่ต้องถูกสุ่มออกมา 131
3-7 สรุปการสุ่มโรงเรียนแยกตามรายภาค 132
3-8 จำนวนกลุ่มตัวอย่างในแต่ละชั้นภูมิ 144
4-1 เปรียบเทียบความแม่นยำในแต่ละวิธีการสุ่ม 159
4-2 รายละเอียดการประมาณค่าเชิงพื้นที่และค่าคลาดเคลื่อนร้อยละ 163
4-3 ผลลัพธ์ที่ได้จากวิธีการสุ่มแบบหลายขั้นตอน 164
4-4 ผลลัพธ์ที่ได้จากวิธีการสุ่มแบบกริด 168
4-5 ผลการทดสอบความสัมพันธ์ระหว่างขนาดโรงเรียนและคะแนน O-NET 170
4-6 ค่าสถิติและจำนวนกลุ่มตัวอย่างที่ได้จากการจัดสรรแบบสัดส่วน 171
4-7 ค่าเฉลี่ยความคลาดเคลื่อนของแต่ละวิธีการสุ่ม 175

สารบัญภาพ

ภาพที่ หน้า

1-1 กรอบแนวคิดในการวิจัย 7

2-1 การเก็บตัวอย่างดินแบบกริด 18

2-2 พื้นที่ของ N_{ij} 19

2-3 พื้นที่ของ N_{ij} 20

2-4 การเพิ่มชั้นข้อมูลเชิงเส้น 24

2-5 การแสดงผลลัพธ์แผนที่และตำแหน่งที่ตั้ง 25

2-6 การใช้งานฟังก์ชั่น Regular points 25

2-7 การใช้โปรแกรม QGIS สุ่มจุดกริด 26

2-8 ตำแหน่งจุดกริดทั้งหมด 27

2-9 ตำแหน่งจุดกริดที่ใช้ในการเก็บรวบรวมข้อมูล 27

2-10 การขยายรัศมี เพื่อหาค่า N'_{ij} 32

2-11 การหากลุ่มตัวอย่างด้วยวงกลมการสุ่ม 34

2-12 รูปแบบของสหสัมพันธ์เชิงพื้นที่ กรณีที่เป็นบวก ศูนย์ และลบ 35

2-13 แผนที่แสดงลัทธิความเชื่อทางการเมืองในยุคสงครามเย็น 36

2-14 แผนที่แสดงการนับถือศาสนาตามเขตพื้นที่ 37

2-15 ปัญหาของการสุ่มกลุ่มตัวอย่างในพื้นที่ที่ประชากรเป็นวิวิธพันธ์ 39

2-16 การสุ่มกลุ่มตัวอย่าง 2 ขั้นตอนด้วยการขยายขนาดรัศมีวงกลมของการสุ่ม 40

2-17 การนำวิธีสุ่มแบบกริดมาประยุกต์ใช้ในระบบการสุ่มหลายขั้นตอน 41

2-18 ความแตกต่างระหว่างวิธีการประมาณค่าในช่วงและวิธีการถดถอย 45

2-19 การประมาณค่าเชิงพื้นที่แบบ TIN 46

2-20 การประมาณค่าเชิงพื้นที่แบบ IDW 47

2-21 แผนที่ประเทศไทยแสดงผลการสำรวจระดับสติปัญญาเด็กนักเรียนไทย
พ.ศ. 2554 54

2-22 แบบจำลองแผนที่ขนาด 3 x 3 ช่อง 58

2-23 แบบจำลองแผนที่แสดงค่าตัวแปร z 59

2-24 การแบ่งชั้นภูมิของแบบจำลอง 60

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

2-25 การแบ่งกลุ่มของแบบจำลอง 61
2-26 ตำแหน่งกริดเซลล์ที่ถูกสุ่มด้วยวิธีการสุ่มแบบกริด 61
2-27 แบบจำลองแผนที่แสดงค่าตัวแปร z 63
2-28 ตำแหน่งกริดเซลล์ที่ถูกสุ่มด้วยวิธีการสุ่มแบบกริด 64
2-29 การประมาณค่าเชิงพื้นที่ 64
2-30 แบบจำลองแสดงค่าตัวแปรเป็นเฉดสี 65
2-32 ผลการทดลองที่ได้จากรูปแบบการสุ่ม 6 วิธี 67
2-33 แผนที่แสดงปริมาณตัวด้วงน้ำในเขตแหลมไอบีเรีย 68
2-34 ตำแหน่งที่ใช้ในการสุ่มตัวอย่างในเขต The cerco de almadenejos 70
2-35 ผลการทดลองที่ได้จากรูปแบบการสุ่ม 11 วิธี 71
2-36 ตำแหน่งที่ใช้เก็บข้อมูลเห็บ/ ไรในพื้นที่ Reserve ducke ในป่าอะเมซอน 72
2-37 ตำแหน่งที่ใช้เก็บตัวอย่างดินในสวนยางเมือง Yangjiang ประเทศจีน 73
2-38 ตำแหน่งที่ใช้ในการเก็บข้อมูลในเมือง Kyivska, Khmelnytska และ Zhytomyrska 75
3-1 แบบจำลอง Sampling 0 (S0) 82
3-2 การสุ่มอย่างง่ายของแบบจำลอง S0 83
3-3 การประมาณค่าเชิงพื้นที่ 87
3-4 การสร้างตารางคำนวณ 87
3-5 การหาค่าคลาดเคลื่อนร้อยละ 88
3-6 ผังงานการวิจัยขั้นตอนที่ 1 89
3-10 แบบจำลอง Sampling 0 (S0) 105
3-11 การสุ่มอย่างง่ายของแบบจำลอง S0 106
3-12 การสุ่มแบบมีระบบของแบบจำลอง S0 106
3-13 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S0 107
3-14 การสุ่มแบบจุดกริดของแบบจำลอง S0 107
3-15 แบบจำลอง Sampling 00 (S00) 108
3-16 แบบจำลอง Sampling 1 (S1) 109
3-17 การสุ่มแบบชั้นภูมิของแบบจำลอง S1 109

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

3-18 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S1 110
3-19 แบบจำลอง Sampling 2 (S2) 110
3-20 การสุ่มแบบชั้นภูมิของแบบจำลอง S2 111
3-21 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S2 112
3-22 แบบจำลอง Sampling 3 (S3) 112
3-23 การสุ่มแบบมีระบบของแบบจำลอง S3 113
3-24 การสุ่มแบบชั้นภูมิของแบบจำลอง S3 114
3-25 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S3 114
3-26 การสุ่มแบบจุดกริดของแบบจำลอง S3 115
3-27 แบบจำลอง Sampling 4 (S4) 115
3-28 การสุ่มแบบชั้นภูมิของแบบจำลอง S4 116
3-29 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S4 116
3-30 แบบจำลอง Spatial Interpolation 1 (SI1) 117
3-31 ตำแหน่งจุดกริดของแบบจำลอง SI1 118
3-32 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI1 118
3-33 แบบจำลอง Spatial Interpolation 2 (SI2) 119
3-34 ตำแหน่งจุดกริดของแบบจำลอง SI2 119
3-35 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI2 120
3-36 แบบจำลอง Spatial interpolation 3 (SI3) 120
3-37 ตำแหน่งจุดกริดของแบบจำลอง SI3 121
3-38 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI3 121
3-39 ผลการเรียนเฉลี่ยสะสมของโรงเรียนทั้งหมด 123
3-40 ผลการเรียนเฉลี่ยสะสมของโรงเรียนในเขตตอนบนของประเทศ 124
3-41 ผลการเรียนเฉลี่ยสะสมของโรงเรียนในเขตตอนล่างของประเทศ 124
3-42 คะแนน O-NET ของโรงเรียนทั้งหมด 126
3-43 คะแนน O-NETของโรงเรียนในเขตตอนบนของประเทศ 127
3-44 คะแนน O-NET ของโรงเรียนในเมืองหลวงและปริมณฑล 127

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

3-45 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง เกรดเฉลี่ย Gpax และ คะแนน O-NET 128

3-46 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง เกรดเฉลี่ย Gpaxและ คะแนน O-NET ในเขตตอนบนของประเทศ 129

3-47 ตำแหน่งจุดกริดที่สุ่มด้วยโปรแกรม QGIS 134

3-48 ตำแหน่งจุดกริดที่อยู่ในภาคเหนือ 135

3-49 ตำแหน่งจุดกริด i = 2, j = 10 อยู่ในจังหวัดเชียงใหม่ 136

3-50 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันออกเฉียงเหนือ 136

3-51 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันออก 137

3-52 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันตก 137

3-53 ตำแหน่งจุดกริดที่อยู่ในภาคใต้ 139

3-54 ตำแหน่งจุดกริดที่อยู่ในภาคกลาง 139

3-55 ตำแหน่งจุดกริด i = 3, j = 6 140

3-56 (ก) และ (ข) พื้นที่บริเวณจุดกริด i=2, j=3 ที่ประชากรมีลักษณะเป็นวิวิธพันธ์ 141

3-57 (ก) และ (ข) การแบ่งชั้นภูมิด้วยขนาดโรงเรียน 143

3-58 กลุ่มตัวอย่างที่ได้จากวิธีการสุ่มสองขั้นตอน 144

4-1 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S0 148

4-2 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S00 148

4-3 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S1 149

4-4 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S2 150

4-5 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S3 151

4-6 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S4 152

4-7 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S0 153

4-8 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S00 154

4-9 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S1 155

4-10 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S2 156

4-11 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S3 157

4-12 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S4 158

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

4-13 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงระดับทัศนคติ 160
4-14 แผนที่ค่าคลาดเคลื่อนร้อยละ 161
4-15 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงคะแนนการทดสอบ..... 161
4-16 แผนที่ค่าคลาดเคลื่อนร้อยละ 162
4-17 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงระดับทัศนคติ 162
4-18 แผนที่ค่าคลาดเคลื่อนร้อยละ 163
4-19 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบหลายขั้นตอน 166
4-20 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบกริด 169
4-21 การแบ่งชั้นภูมิของโรงเรียนทั่วประเทศ 171
4-22 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบกริด-ชั้นภูมิ 174
4-23 คะแนนเฉลี่ย O-NET ของประเทศไทยที่ได้จากวิธีการสุ่มแบบหลายขั้นตอน 177
4-24 คะแนนเฉลี่ย O-NET ของประเทศไทยที่ได้จากวิธีการสุ่มแบบกริด 179

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

ข้อมูลสถิติเป็นสิ่งสำคัญอย่างหนึ่งที่ใช้ประกอบ การตัดสินใจ จึงจำเป็นต้องมีการรวบรวมรายละเอียดและข้อเท็จจริงเกี่ยวกับเรื่องที่ทำการศึกษานั้น ถ้าข้อมูลหรือหลักฐานที่รวบรวมได้ไม่ถูกต้องหรือเชื่อถือไม่ได้ ผลสรุปที่ได้จะผิดไปจากความเป็นจริง เช่น กรณีการเลือกตั้งผู้ดำรงตำแหน่งประธานาธิบดีของประเทศอเมริกาใน ปี ค.ศ. 1936 ซึ่งมี แลนดอน (Landon) เป็นตัวแทนของพรรครีพับลิกัน และรูสเวลท์ (Roosvelt) ตัวแทนของพรรคเดโมแครต วารสารไดเจสท์ได้สำรวจความนิยมของประชาชน พบว่า แลนดอน จะได้ชัยชนะด้วยคะแนนเสียงอย่างท่วมท้น แต่ผลที่เกิดขึ้นกลับตรงกันข้าม โดยรูสเวลท์ได้คะแนนเสียงมากอย่างท่วมท้น ซึ่งวารสารลิทเทอรารี่ ไดเจสท์ได้ทำนายผิดพลาดไปถึง 19% ส่งผลให้วารสารลิทเทอรารี่ ไดเจสท์ต้องปิดตัวลงในปีถัดมา (ธัญลักษณ์ เหลืองวิสุทธิ์, 2540, หน้า 5-8)

งานวิจัยหลายเรื่องได้ทำการทดลองกับประชากรโดยตรง โดยเฉพาะเมื่อประชากรมีขนาดเล็ก และมีจำนวนหน่วยตัวอย่างน้อยกว่า 50 หน่วย ควรศึกษาทั้งประชากรจะได้ข้อมูลที่ถูกต้องเที่ยงตรงที่สุด (Henry, 1991, p. 14) ซึ่งการเก็บรวบรวมข้อมูลจากทุกหน่วยของประชากร หรือเรียกอีกอย่างว่าการทำสำมะโน (Census) ซึ่งการทำสำมะโนมีข้อดีคือการใช้ข้อมูลของทั้งประชากร ทำให้สารสนเทศที่ได้มีความถูกต้องสูง แต่ข้อเสียของการทำสำมะโนต้องใช้งบประมาณ เวลา และกำลังคนเป็นจำนวนมาก และปัญหาที่สำคัญที่สุด คือ ไม่สามารถทำได้จริงสำหรับโครงการที่มีขนาดใหญ่ เนื่องจากประชากรมีจำนวนเป็นอนันต์ นอกจากนั้นการทำสำมะโนกับประชากรบางอย่างไม่สามารถปฏิบัติได้จริง เนื่องจากไม่สามารถหาหน่วยตัวอย่างได้ครบทั้งหมด หรือหน่วยตัวอย่างบางหน่วยนำมาทดสอบแล้วอาจเสื่อมสภาพไปหรือสร้างความเสียหายต่อกลุ่มตัวอย่าง (สำนักสถิติแห่งชาติ, ม.ป.ป., หน้า 3) ด้วยปัญหาดังกล่าว จึงต้องนำบางส่วนของประชากรมาใช้ในการทดลอง/ วิจัย แล้วนำสรุปผลการวิเคราะห์จากกลุ่มตัวอย่างที่ได้อ้างอิงไปยังประชากร หรือเรียกอีกอย่างว่า สถิติอนุมาน (Inferential statistics) ซึ่งการสำรวจด้วยกลุ่มตัวอย่างมีข้อได้เปรียบการทำสำมะโน เนื่องจากในการทำวิจัยจะถูกจำกัดด้านทรัพยากรด้านต่าง ๆ เช่น ค่าใช้จ่าย จำนวนเจ้าหน้าหน้าที่ภาคสนาม และเวลา ดังนั้น การใช้กลุ่มตัวอย่างขนาดเล็กกว่าประชากร ทำให้ประหยัดงบประมาณ เจ้าหน้าหน้าที่ภาคสนาม และใช้เวลาในการรวบรวมข้อมูลน้อยกว่า จึงสามารถปฏิบัติงานได้รวดเร็วและนำเสนอผลการสำรวจได้ทันต่อสถานการณ์

จากที่กล่าวมาแล้ว ทำให้ทราบถึงเหตุผลความจำเป็นที่ผู้วิจัยต้องใช้กลุ่มตัวอย่างสำหรับการวิจัย ซึ่งการเลือกกลุ่มตัวอย่างที่เหมาะสมสำหรับการทำวิจัยเป็นสิ่งสำคัญอย่างมาก เพราะการวิจัยกลุ่มตัวอย่างเพียงส่วนน้อยจะสะท้อนภาพของประชากรทั้งหมด ถ้ากลุ่มตัวอย่างที่สุ่มออกมาไม่เหมาะสมจะทำให้ผลจากการศึกษาที่ได้สรุปอ้างไปยังประชากรไม่ถูกต้องตามความเป็นจริง ดังนั้น กลุ่มตัวอย่างที่จะทำการศึกษาจึงต้องเป็นตัวแทนประชากรที่ดี ซึ่งมีค่าสถิติใกล้เคียงกับพารามิเตอร์ เช่น ค่าเฉลี่ยกลุ่มตัวอย่างใกล้เคียงกับค่าเฉลี่ยประชากร เป็นต้น เมื่อนำมาจัดกระทำแล้วมีโอกาสผิดพลาดน้อย

ข้อมูลสถิติที่มีประโยชน์จริง ๆ จะต้องเป็นข้อมูลที่มีความถูกต้อง เป็นที่น่าเชื่อถือได้ การเก็บรวบรวมข้อมูลสถิติที่ดีจำเป็นต้องอาศัยวิธีการเก็บรวบรวมที่เหมาะสม (สุรินทร์ นิยมมางกูร, 2546, หน้า 1) ซึ่งวิธีการสุ่มตัวอย่างมีหลายวิธี แต่ในงานวิจัยครั้งนี้จะมุ่งเน้นการสุ่มตัวอย่างโดยใช้ความน่าจะเป็นในการเลือก ซึ่งการเลือกตัวอย่างวิธีนี้ ให้ผลน่าเชื่อถือ มักทำให้ได้ตัวอย่างที่ไม่เอนเอียง เป็นตัวแทนที่ดีของประชากร (สุชาดา บวรกิติวงศ์, 2548, หน้า 117-120) การเลือกตัวอย่างแบบใช้ความน่าจะเป็น ได้แก่ การสุ่มอย่างง่าย (Simple random sampling) การสุ่มแบบมีระบบ (Systematic random sampling) การสุ่มแบบชั้นภูมิ (Stratified random sampling) การสุ่มแบบแบ่งกลุ่ม (Cluster random sampling) การสุ่มแบบหลายขั้นตอน (Multistage random sampling) ดังนั้นการเลือกใช้วิธีการสุ่มต้องพิจารณาถึงความเหมาะสมและสอดคล้องกับงานวิจัย ภายใต้ลักษณะของประชากรและทรัพยากรในการวิจัย วิธีการสุ่มบางอย่างให้ค่าสถิติที่มีความคลาดเคลื่อนต่ำ แต่วิธีการนั้นอาจไม่สามารถนำมาใช้ปฏิบัติกับงานวิจัยได้ ดังนั้นผู้วิจัยต้องตัดสินใจเลือกวิธีการสุ่ม

วิธีการสุ่มตัวอย่างแต่ละวิธีมีจุดเด่นจุดด้อย คือ วิธีการสุ่มตัวอย่างแบบแบ่งชั้นภูมิเป็นวิธีการที่มีประสิทธิภาพสูง เพราะสามารถควบคุมตัวแปรแทรกซ้อนได้ ทำให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนของประชากร อย่างไรก็ตามวิธีการนี้ไม่สามารถนำมาใช้กับโครงการขนาดใหญ่ได้ ยกตัวอย่าง เช่น โครงการประเมินผลนักเรียนนานาชาติ PISA ในปี ค.ศ. 2012 ซึ่งกลุ่มตัวอย่าง คือ นักเรียน 510,000 คน จากประชากรนักเรียน 65 ประเทศ เป็นต้น (โครงการ PISA ประเทศไทย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554) ซึ่งในการวิจัยแบบนี้ไม่สามารถที่จะแบ่งนักเรียนทั้งหมดออกเป็นชั้นภูมิได้ เพราะประชากรมีจำนวนมากและกระจายอยู่ทั่วโลก ซึ่งปัญหาเรื่องการเตรียมกรอบตัวอย่างเป็นปัญหาที่ยุ่งยาก เพราะต้องใช้เวลาอย่างมากในการสำรวจและมีค่าใช้จ่ายที่สูง นอกจากนี้ยังมีปัญหาในด้านการปรับปรุงให้กรอบตัวอย่างให้ทันสมัยอยู่เสมอ เพราะประชากรเป็นสิ่งมีชีวิตสามารถเคลื่อนย้ายที่ได้หรือเสียชีวิตได้

ในส่วนวิธีการสุ่มแบบง่ายและวิธีการสุ่มแบบมีระบบก็มีปัญหาในการเตรียมกรอบตัวอย่างเหมือนกันกับวิธีการสุ่มแบบชั้นภูมิ ถึงแม้ผู้วิจัยจะสามารถเตรียมกรอบตัวอย่างได้ แต่กลุ่มตัวอย่าง

ที่ได้จากวิธีการสุ่มแบบง่ายและแบบมีระบบ จะมีลักษณะกระจัดกระจาย เนื่องจากสมาชิกทุก ๆ ตัวในประชากรมีโอกาสถูกเลือกมาเป็นกลุ่มตัวอย่าง ดังนั้นกลุ่มตัวอย่างที่ได้จะกระจัดกระจายไปทั่วประชากร ซึ่งการเก็บรวมรวบข้อมูลจะกระทำได้ยากมากและสิ้นเปลืองค่าใช้จ่ายอย่างสูง นอกจากนี้ยังมีอุปสรรคในการวางแผนควบคุมงานสนาม เนื่องจากหน่วยสำรวจอยู่กระจายไปทั่วพื้นที่ ทำให้ห่างไกลกัน ย่อมสร้างความยุ่งยากในการติดต่อสื่อสารและการควบคุมงานสนามของผู้วิจัยหรือเจ้าของโครงการ (มนตรี พิริยะกุล, 2530, หน้า 252)

การวิจัยโครงการขนาดใหญ่มีปัญหาที่สำคัญ ได้แก่ ปัญหาการเตรียมกรอบตัวอย่าง ปัญหาค่าใช้จ่ายในการสำรวจ และปัญหาการควบคุมงานสนาม แต่มีวิธีการสุ่มแบบหนึ่งที่เหมาะสมกับการสุ่มกรณีพื้นที่ขนาดใหญ่และประชากรมีจำนวนมากนั้นคือวิธีการสุ่มแบบแบ่งกลุ่ม โดยใช้พื้นที่เป็นหน่วยของการสุ่ม ทำให้สมาชิกในกลุ่มตัวอย่างอยู่บริเวณเดียวกัน ดังนั้นจะเกิดความสะดวกในการเก็บรวบรวมข้อมูลและลดค่าใช้จ่ายได้อย่างมาก แต่ปัญหาของการสุ่มแบบกลุ่ม คือ มีประสิทธิภาพต่ำกว่าวิธีสุ่มอย่างแบบอื่น ๆ ในยุคปัจจุบันการสุ่มตัวอย่างสำหรับโครงการขนาดใหญ่จะใช้วิธีการสุ่มแบบหลายขั้นตอน ซึ่งวิธีการนี้มีข้อดีคล้ายกับการสุ่มแบบกลุ่มกล่าวคือ ช่วยลดเวลา ค่าใช้จ่าย และแรงงานในการสุ่มตัวอย่าง ซึ่งเหมาะสำหรับงานวิจัยที่ประชากรมีขนาดใหญ่ที่แบ่งเป็นลำดับชั้นลดหลั่น แต่วิธีการนี้มีปัญหาในเรื่องความซับซ้อนยุ่งยากในการประมาณค่าพารามิเตอร์ เนื่องจากต้องมีการคำนวณค่าสถิติในทุกขั้นตอนที่ทำการสุ่มตัวอย่างออกมา (สุชาดา บวรกิติวงศ์, 2548, หน้า 120; สุรินทร์ นิยมมางกูร, 2541, หน้า 198) และการประมาณค่าพารามิเตอร์จะสลับซับซ้อนถ้ามีจำนวนชั้นของการสุ่มมาก (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

จุดด้อยประการหนึ่งของวิธีการสุ่มแบบหลายขั้นตอน คือ ความแปรปรวนของตัวประมาณค่าของพารามิเตอร์จะมีค่าสูงขึ้นตามจำนวนชั้นที่สุ่มตัวอย่าง เพราะการสุ่มตัวอย่างครั้งหนึ่งก่อให้เกิดความแปรปรวนขึ้นระหว่างหน่วยที่เกี่ยวข้องเสมอ เช่น เมื่อสุ่มระดับหมู่บ้านทำให้เกิดความแปรปรวนระหว่างหมู่บ้าน เมื่อสุ่มระดับตำบล ทำให้เกิดความแปรปรวนในระหว่างตำบล เมื่อสุ่มระดับอำเภอเกิดความแปรปรวนระหว่างอำเภอ เป็นต้น ดังนั้นถ้าโครงการใดมีการสุ่มตัวอย่างมากชั้น ค่าความแปรปรวนของตัวประมาณของค่าพารามิเตอร์จะมากขึ้นด้วย (มนตรี พิริยะกุล, 2530, หน้า 254-255) ดังนั้น ความแปรปรวนของตัวประมาณค่าของพารามิเตอร์ของวิธีการสุ่มแบบกริดจะมีค่าต่ำกว่าวิธีการสุ่มแบบหลายขั้นตอนเสมอ

ด้วยเหตุผลที่กล่าวมาทั้งหมด ผู้วิจัยขอนำเสนอวิธีการสุ่มแบบกริด ซึ่งเป็นวิธีการสุ่มตัวอย่างที่แพร่หลายสำหรับงานวิจัยทางด้านวิทยาศาสตร์กายภาพที่ศึกษาวิจัย ในกรณีที่ประชากรมีจำนวนมากและอาศัยในพื้นที่ขนาดใหญ่ โดยใช้พิกัดโลกเป็นตำแหน่งในการอ้างอิง ซึ่งจุดตัดที่เกิดจากเส้นละติจูดและเส้นลองจิจูดตัดกันคือตำแหน่งที่ใช้ในการสุ่มตัวอย่าง วิธีการสุ่มแบบนี้

มีจุดเด่นเหมือนกับการสุ่มแบบแบ่งกลุ่ม กล่าวคือ กลุ่มตัวอย่างที่ถูกสุ่มออกมาจะอยู่ใกล้บริเวณจุดกริดโดยรอบ ดังนั้น กลุ่มตัวอย่างที่สุ่มออกมาจะไม่กระจัดกระจาย ทำให้สะดวกและลดค่าใช้จ่าย รวมทั้งประหยัดเวลาและแรงงานในการเก็บรวบรวมข้อมูล นอกจากนั้นวิธีการสุ่มแบบกริดนี้สามารถแก้ปัญหาความยุ่งยากในการประมาณค่าพารามิเตอร์หลาย ๆ ครั้งได้ เนื่องจากวิธีการสุ่มแบบกริดมีการสุ่มเพียงแค่ชั้นเดียวเท่านั้น ในส่วนจุดด้อย คือ ในกรณีที่พื้นที่มีสหสัมพันธ์เชิงพื้นที่เป็นศูนย์ หรือถ้าโปรแกรมสุ่มจุดกริดได้ตำแหน่งที่ประชากรย่อย (Subpopulation) จะได้ค่าสถิติที่มีความคลาดเคลื่อน อาจทำให้ผลการวิจัยผิดพลาดไปจากความเป็นจริง

ในการวิจัยทางการศึกษาเป็นสาขาหนึ่งที่สำคัญ เนื่องจากประเทศจะพัฒนาหรือล้าหลังย่อมเกิดจากคุณภาพประชากรเป็นหลัก ถ้าประชาชนมีความรู้และคุณธรรมจะนำพาให้ประเทศชาติก้าวหน้า ยกตัวอย่าง เช่น ประเทศสิงคโปร์ เป็นต้น แต่สำหรับประเทศไทยนั้นปัญหาทางการศึกษาเป็นปัญหาระดับประเทศที่มีความซับซ้อนสูง ซึ่งในการวิจัยครั้งนี้ให้ความสนใจในเรื่องการสอบ O-NET เนื่องจากการสอบ O-NET นั้นเป็นโครงการขนาดใหญ่ที่ค่าใช้จ่ายในการจัดสอบประมาณปีละ 600 ล้านบาท โดยนักเรียน ม.6 เกือบทั้งหมดของประเทศต้องเข้าสอบ ทำให้ฐานข้อมูลคะแนน O-NET มีความครบถ้วนสมบูรณ์ นอกจากนี้ตำแหน่งที่ตั้งของโรงเรียนมีถิ่นที่อยู่กระจัดกระจายอยู่ทุกพื้นที่ประเทศไทย ทำให้ผู้วิจัยมีแนวความคิดในการนำวิธีการสุ่มแบบกริดมาประยุกต์ใช้กับฐานข้อมูลคะแนน O-NET เนื่องจากต้องการทดสอบความแม่นยำของวิธีการสุ่มแบบกริดที่ใช้ในเชิงวิทยาศาสตร์กายภาพว่ามีความเหมาะสมกับงานวิจัยเชิงพฤติกรรมศาสตร์และสังคมศาสตร์หรือไม่ อย่างไร ถ้าผลการทดลอง พบว่า ค่าสถิติใกล้เคียงกับค่าพารามิเตอร์ ในอนาคตการสอบ O-NET หรือ NT สามารถนำวิธีการสุ่มแบบกริดไปประยุกต์ใช้ได้ ซึ่งจะช่วยลดงบประมาณ เวลา และจำนวนเจ้าหน้าที่ภาคสนาม เพราะวิธีการสุ่มดังกล่าวมีความเหมาะสำหรับโครงการขนาดใหญ่ที่ประชากรอาศัยอยู่อย่างกระจัดกระจายในพื้นที่กว้าง

ในส่วนขั้นตอนการนำเสนอข้อมูลที่ได้หลังจากวิเคราะห์ข้อมูล จะนำเสนอแบบระบบสารสนเทศภูมิศาสตร์ โดยใช้โปรแกรม QGIS เป็นเครื่องมือในการแสดงผลลัพธ์เป็นแผนที่ภาพ 2 มิติ เพื่อให้เข้าใจความหมายและมองเห็นภาพของผลลัพธ์ได้ดีมากขึ้น ซึ่งจะเป็นประโยชน์ต่อการวิจัยในด้านการจัดการศึกษาของประเทศได้รวดเร็ว ลึกซึ้ง มีรายละเอียดครบถ้วนและเกิดประสิทธิภาพ

คำถามวิจัย

1. ค่าสถิติที่ได้จากวิธีการสุ่มแบบกริดและค่าสถิติที่ได้จากวิธีการสุ่มแบบอื่น ๆ (วิธีการสุ่มอย่างง่าย วิธีการสุ่มแบบมีระบบ วิธีการสุ่มแบบแบ่งชั้นภูมิ วิธีการสุ่มแบบกลุ่ม และวิธีการสุ่ม

แบบหลายขั้นตอน) มีความคลาดเคลื่อนแตกต่างกันหรือไม่ วิธีการใดช่วยให้การประมาณค่าสถิติใกล้เคียงกับค่าพารามิเตอร์ที่สุด

2. แผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดจากวิธีการสุ่มแบบกริด (แบบกริด-ชั้นภูมิ) กับแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดจากวิธีการสุ่มแบบหลายขั้นตอนมีประสิทธิภาพแตกต่างกันหรือไม่ วิธีการสุ่มแบบใดให้สารสนเทศที่เป็นประโยชน์ต่อการตัดสินใจมากที่สุด

วัตถุประสงค์ของการวิจัย

1. เพื่อตรวจสอบความแม่นยำในการประมาณค่าของวิธีการสุ่มแบบต่าง ๆ
2. เพื่อเปรียบเทียบประสิทธิภาพระหว่างวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ
3. เพื่อทดสอบวิธีการสุ่มแบบกริดกับฐานข้อมูลคะแนน O-NET
4. เพื่อเปรียบเทียบประสิทธิภาพของแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดจากวิธีการสุ่มแบบกริด กับวิธีการสุ่มแบบหลายขั้นตอน

กรอบแนวคิดในการวิจัย

วิธีการสุ่มแบบกริดเป็นวิธีการสุ่มตัวอย่างที่แพร่หลายสำหรับงานวิจัยทางด้านวิทยาศาสตร์กายภาพ ในกรณีที่ประชากรมีจำนวนมากและอาศัยในพื้นที่ขนาดใหญ่ ซึ่งการวิจัยในครั้งนี้มีจุดมุ่งหมาย เพื่อทดสอบว่าวิธีการสุ่มแบบกริดมีความเหมาะสมกับงานวิจัยทางพฤติกรรมและสังคมศาสตร์หรือไม่ อย่างไร โดยมีขั้นตอนการวิจัยแบ่งออกเป็น 2 ขั้นตอนใหญ่ ๆ ดังนี้

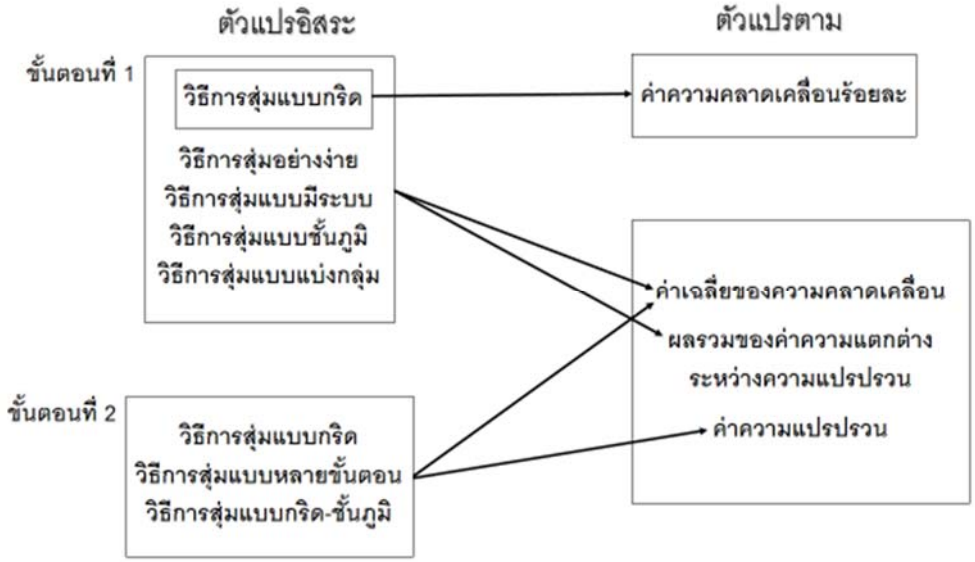
ขั้นตอนที่ 1 การทดสอบวิธีการสุ่มแบบกริด โดยอาศัยแบบจำลองที่สร้างขึ้น (S0, S00, S1, S2, S3 และ S4) เพื่อตรวจสอบความแม่นยำในการประมาณค่าของการสุ่มแบบกริดเทียบกับวิธีการสุ่มแบบอื่น ๆ ได้แก่ 1) วิธีการสุ่มอย่างง่าย 2) วิธีการสุ่มแบบมีระบบ 3) วิธีการสุ่มแบบชั้นภูมิ และ 4) วิธีการสุ่มแบบแบ่งกลุ่ม โดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนและผลรวมของความแตกต่างระหว่างความแปรปรวนในการทดสอบ นอกจากนี้ยังมีการสร้างแบบจำลอง SI1 SI2 และ SI3 เพื่อทดสอบความแม่นยำในการประมาณค่าเชิงพื้นที่ของวิธีการสุ่มแบบกริด โดยใช้ค่าความคลาดเคลื่อนร้อยละในการทดสอบ

ขั้นตอนที่ 2 การทดสอบวิธีการสุ่มแบบกริด โดยอาศัยฐานข้อมูลคะแนน O-NET เพื่อตรวจสอบความแม่นยำในการประมาณค่าจากวิธีการสุ่มแบบกริด วิธีการสุ่มแบบหลายขั้นตอน และวิธีการสุ่มแบบกริด-ชั้นภูมิ โดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนและค่าความแปรปรวนในการทดสอบ จากนั้นทำการเปรียบเทียบประสิทธิภาพของแผนที่ระบบสารสนเทศภูมิศาสตร์ที่ได้จากวิธีการสุ่มทั้ง 3 วิธี

เหตุผลความจำเป็นและแนวคิดในแต่ละขั้นตอนการวิจัย คือ ขั้นตอนที่ 1 เป็นการทดสอบวิธีการสุ่มแบบกริดด้วยการเขียนคำสั่งโปรแกรมคอมพิวเตอร์ เพื่อจำลองข้อมูลและสร้างแผนที่ทำให้สามารถศึกษาตัวแปรเชิงพื้นที่ที่มีการจัดเรียงตัว/ กระจายตัวได้หลากหลายรูปแบบ เนื่องจากเป็นแผนที่สมมติ ตลอดจนมีความสะดวกและรวดเร็วในการสุ่มตัวอย่างได้เป็นพันหรือหมื่นครั้ง เพื่อให้ผลลัพธ์ที่ได้มีความเสถียรและน่าเชื่อถือ สำหรับวิธีการสุ่มที่นำมาใช้ในขั้นตอนที่ 1 เป็นวิธีการการสุ่มที่มีมาตรฐานตามแนวทางการวิจัยด้านพฤติกรรมและสังคมศาสตร์ (ได้แก่ วิธีการสุ่มอย่างง่าย วิธีการสุ่มแบบมีระบบ วิธีการสุ่มแบบชั้นภูมิ และวิธีการสุ่มแบบแบ่งกลุ่ม) ซึ่งวิธีการสุ่มเหล่านี้จะนำมาใช้เปรียบเทียบกับวิธีการสุ่มแบบกริด หากวิธีการสุ่มแบบกริดมีความแม่นยำต่ำกว่าแสดงว่าวิธีการสุ่มแบบกริดไม่มีความเหมาะสมที่จะนำมาประยุกต์ใช้งาน

ในส่วนขั้นตอนที่ 2 เป็นการทดสอบวิธีการสุ่มแบบกริดด้วยฐานข้อมูลคะแนน O-NET เนื่องจากในขั้นตอนที่ 1 เป็นแบบจำลองที่ผู้วิจัยสร้างขึ้นเองจึงทำให้ความน่าเชื่อถือลดลง ซึ่งการทดสอบโดยการใช้ข้อมูลจริงเป็นการทดสอบความตรงภายนอกของวิธีการสุ่มแบบกริด สำหรับวิธีการสุ่มที่นำมาใช้ในขั้นตอนที่ 2 คือ วิธีการสุ่มแบบหลายขั้นตอน (วิธีการสุ่มแบบแบ่งกลุ่ม 2 ขั้นตอน) เนื่องจากวิธีการสุ่มแบบกริดและวิธีการสุ่มแบบหลายขั้นตอนเป็นวิธีการที่เหมาะสำหรับการวิจัยโครงการขนาดใหญ่ ซึ่งมีปัญหาที่สำคัญ ได้แก่ ปัญหาการเตรียมกรอบตัวอย่าง ปัญหาค่าใช้จ่ายในการสำรวจ และปัญหาการควบคุมงานสนาม ถึงแม้วิธีการสุ่มแบบแบ่งกลุ่มมีจุดด้อยในด้านประสิทธิภาพที่ต่ำกว่าวิธีสุ่มอย่างแบบอื่น ๆ แต่วิธีการนี้สามารถใช้งานได้จริงในกรณีที่ประชากรเป็นแบบไม่จำกัด (Infinite population) ซึ่งในปัจจุบันวิธีการสุ่มตัวอย่างจากประชากรขนาดใหญ่ คือ วิธีการสุ่มแบบหลายขั้นตอน ดังนั้น ในการวิจัยขั้นตอนที่ 2 ผู้วิจัยจะไม่ใช้วิธีการสุ่มอย่างง่าย การสุ่มแบบมีระบบ และการสุ่มแบบชั้นภูมิ ในการเปรียบเทียบกับวิธีการสุ่มแบบกริดโดยตรง แต่จะนำวิธีการสุ่มแบบชั้นภูมิมาประยุกต์ใช้กับวิธีการสุ่มแบบกริด เพราะวิธีการสุ่มแบบชั้นภูมิเป็นวิธีการที่มีประสิทธิภาพสูงและสามารถควบคุมตัวแปรแทรกซ้อนได้ ทำให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนของประชากร และเรียกวิธีการสุ่มนี้ว่า “การสุ่มแบบกริด-ชั้นภูมิ” (ซึ่งมีรายละเอียดตามหัวข้อการสุ่มตามหัวข้อการสุ่มแบบกริดสองขั้นตอน)

จากที่กล่าวมาทั้งหมดสามารถสรุปเป็นกรอบแนวคิดในการวิจัยได้ตามภาพที่ 1-1



ภาพที่ 1-1 กรอบแนวคิดในการวิจัย

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้วิธีการสุ่มแบบกริดที่ใช้ในเชิงวิทยาศาสตร์กายภาพมาประยุกต์ใช้ในเชิงพฤติกรรมศาสตร์และสังคมศาสตร์ ซึ่งจะช่วยลดงบประมาณ เวลา และจำนวนเจ้าหน้าที่ภาคสนาม

2. ผลการศึกษาครั้งนี้จะเป็นแนวปฏิบัติให้ผู้วิจัยทางด้านการศึกษาพฤติกรรมศาสตร์และสังคมศาสตร์ นำเสนอข้อมูลโดยใช้แผนที่ระบบสารสนเทศภูมิศาสตร์ ซึ่งจะทำให้งานวิจัยที่เกิดขึ้นมีความชัดเจนและง่าย ต่อการศึกษาทำความเข้าใจมากกว่าการนำเสนอข้อมูลด้วยตัวเลขและสัญลักษณ์ทางคณิตศาสตร์ รวมทั้งช่วยให้ผู้บริหาร/ผู้บังคับบัญชาสามารถประเมินสถานการณ์และตัดสินใจได้อย่างถูกต้องและรวดเร็ว

3. ผลการวิเคราะห์วิธีการสุ่มแบบกริดสามารถใช้เป็นแนวทางเบื้องต้นในการวางแผนพัฒนาการศึกษาของไทยเชิงพื้นที่ ทั้งในด้านกายภาพ เศรษฐกิจ และสังคมต่อไป

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง โดยมีการดำเนินการวิจัย 2 ขั้นตอน ซึ่งขั้นตอนที่ 1 เป็นการทดลองการสุ่มแบบกริดโดยอาศัยแบบจำลองที่สร้างขึ้น และขั้นตอนที่ 2 เป็นการทดลองการสุ่มแบบกริดโดยอาศัยข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา (สพฐ.) จัดทำขึ้น ซึ่งประชากรที่ใช้ในการศึกษามีดังนี้

1.1 ประชากรสำหรับกรณีแบบจำลองการสุ่ม คือ ค่าตัวแปรทุก ๆ ค่าที่อยู่ในตารางแบบจำลองนั้น ๆ เช่น แบบจำลอง S0 (ตามที่ผู้วิจัยกำหนด) เป็นตารางแบบ 9x9 ตารางหน่วย ดังนั้นมีประชากรทั้งหมด 81 หน่วย เป็นต้น

1.2 ประชากรสำหรับกรณีศึกษา คือ โรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 ทั้งหมด 2,584 โรงเรียน

2. กลุ่มตัวอย่างสำหรับกรณีแบบจำลองการสุ่ม คือ 9 และ 36 หน่วย สำหรับกลุ่มตัวอย่างจำนวน 9 หน่วย ใช้กับแบบจำลอง S00, S0, S1, S2 และ S3 (ตามที่ผู้วิจัยกำหนด) และสำหรับกลุ่มตัวอย่าง จำนวน 36 หน่วย ใช้กับแบบจำลอง S4, SI1, SI2 และ SI3 (ตามที่ผู้วิจัยกำหนด) ในส่วนกรณีศึกษานั้น กลุ่มตัวอย่าง คือ โรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 จำนวน 335 โรงเรียน

3. ตัวแปรในการวิจัย แบ่งเป็นตัวแปรอิสระและตัวแปรตามมีรายละเอียด ดังนี้

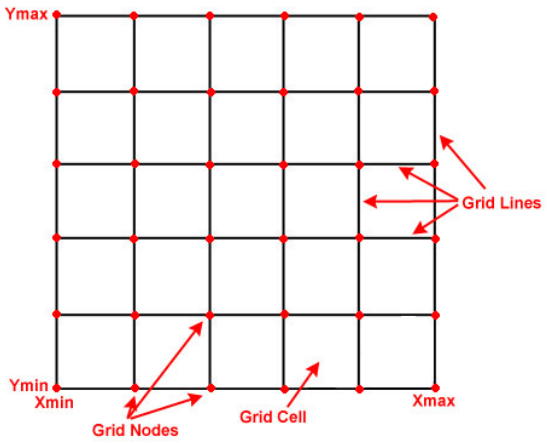
3.1 ตัวแปรอิสระ คือ วิธีการสุ่มตัวอย่าง โดยทำการสุ่มตัวอย่างจากแบบจำลองหรือจากฐานข้อมูล O-NET ซึ่งวิธีการสุ่มในการวิจัยครั้งนี้มี 6 วิธี ได้แก่ 1) การสุ่มตัวอย่างแบบง่าย 2) การสุ่มตัวอย่างแบบมีระบบ 3) การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ 4) การสุ่มตัวอย่างแบบกลุ่ม 5) การสุ่มแบบหลายขั้นตอน และ 6) การสุ่มตัวอย่างแบบกริด

3.2 ตัวแปรตาม คือ ค่าเฉลี่ยของความคลาดเคลื่อน ผลรวมของค่าความแตกต่าง ความแปรปรวน ค่าคลาดเคลื่อนร้อยละ และความแปรปรวน

นิยามศัพท์เฉพาะ

1. กริดเซลล์ (Grid cell) หมายถึง หน่วยย่อยที่เล็กที่สุดของกริด ซึ่งมีรูปร่างเป็นตารางสี่เหลี่ยมมุมฉากขนาดเท่า ๆ กัน เกิดจากเส้นแนวตั้ง (2 คู่ที่ติดกัน) กับเส้นแนวนอน (2 คู่ที่ติดกัน) ตัดกัน

2. จุดกริด (Grid node) หมายถึง ตำแหน่งที่เกิดการตัดกันระหว่างเส้นแนวตั้งและเส้นแนวนอน ตามรูปภาพที่ 1-2



ภาพที่ 1-2 ภาพอธิบายระบบกริด (Dickenson, 2009)

3. วงกลมการสุ่ม หมายถึง วงกลมที่สร้างขึ้น เพื่อใช้ในการหากลุ่มตัวอย่าง โดยมีจุดศูนย์กลางเป็นตำแหน่งจุดกริด (i, j) ดังนั้นสมาชิกหรือหน่วยข้อมูล (Element/ unit) ที่อยู่ภายในเส้นรอบวง คือ กลุ่มตัวอย่างที่ใช้ในการศึกษา ซึ่งรัศมีวงกลมการสุ่มมีระยะไกลสุดเท่ากับครึ่งหนึ่งของระยะทางระหว่างจุดกริด 2 จุด ที่อยู่ติดกัน

4. ตัวแปรเชิงพื้นที่ (Spatial variable) หมายถึง คุณลักษณะหรือคุณสมบัติของสิ่งที่ศึกษาวิจัยที่เปลี่ยนแปลงไปตามพิกัดภูมิศาสตร์ เช่น ตัวแปรคะแนนเฉลี่ย O-NET ของโรงเรียนในกรุงเทพและปริมณฑลมีค่ามากกว่าตัวแปรคะแนนเฉลี่ย O-NET ของโรงเรียนในภาคตะวันออกเฉียงเหนือ

5. ค่าตัวแปร Z หมายถึง ค่าสังเกตของกลุ่มตัวอย่างหรือสิ่งที่ศึกษาวิจัย โดยตำแหน่งที่อยู่ของกลุ่มตัวอย่างมีค่าพิกัดตามแนวเส้นละติจูดและแนวเส้นลองจิจูดเปรียบได้กับแกน X และ Y ส่วนค่าตัวแปรที่ใช้ในการศึกษาวิจัยคือค่า Z เสมือนเป็นมิติที่ 3 ของข้อมูลเชิงพื้นที่ ยกตัวอย่างเช่น ในการศึกษาระดับความลึกของท้องทะเล แกน X คือละติจูด แกน Y คือลองติจูด และ แกน Z คือ ระดับความลึกพื้นท้องทะเล เป็นต้น

6. สหสัมพันธ์เชิงพื้นที่ (Spatial autocorrelation) หมายถึง คือความสัมพันธ์ระหว่างค่าของตัวแปรกับระยะห่างของ 2 สิ่งใด ๆ โดยสิ่งใด (Object) ที่อยู่ใกล้กันจะมีค่าตัวแปรเชิงพื้นที่ใกล้เคียงกัน แต่สิ่งใด (Object) ที่อยู่ห่างไกลกันค่าตัวแปรเชิงพื้นที่แตกต่างกัน

7. การประมาณค่าเชิงพื้นที่ (Spatial interpolation) หมายถึง กระบวนการของการใช้ข้อมูลจุดกริดที่ทราบค่า เพื่อประมาณค่าพื้นที่หรือจุดกริดอื่น ๆ ที่ไม่ทราบค่า เนื่องจากการหาค่าตัวแปร Z ที่ตำแหน่งใด ๆ สามารถหาได้จากการเก็บข้อมูลจากกลุ่มตัวอย่างในบริเวณนั้น แต่การนำเสนอข้อมูลในรูปแบบแผนที่จำเป็นต้องทราบข้อมูลค่าตัวแปร Z ทุกตำแหน่งบนแผนที่

ซึ่งการเก็บข้อมูลของประชากรทั่วทั้งแผนที่เป็นสิ่งที่ทำได้ยากมากและมีค่าใช้จ่ายที่สูงมาก ดั้งนั้นต้องเก็บข้อมูลจากจุดตัวอย่างเพียงบางจุด เพื่อใช้ประมาณค่าจุดที่เหลือทั้งหมด การประมาณค่าเชิงพื้นที่มีหลักการที่สำคัญ 2 ข้อ คือ ค่าตัวแปร Z ต้องมีการเปลี่ยนแปลงแบบต่อเนื่องอย่างค่อยเป็นค่อยไป และค่าตัวแปร Z ต้องมีสหสัมพันธ์เชิงพื้นที่

8. ค่าเฉลี่ยของความคลาดเคลื่อนหมายถึง ค่าเฉลี่ยของผลต่างระหว่างค่าเฉลี่ยของประชากรกับค่าค่าเฉลี่ยของกลุ่มตัวอย่าง เนื่องจากค่าเฉลี่ยของประชากรมีเพียงค่าเดียวเท่านั้น แต่ค่าความเฉลี่ยของกลุ่มตัวอย่างจะเปลี่ยนแปลงไปตามค่าตัวแปรของสมาชิกในกลุ่ม ทำให้ค่าเฉลี่ยของกลุ่มตัวอย่างมีหลายค่า ดังนั้นผลต่างระหว่างค่าเฉลี่ยของประชากรกับกลุ่มตัวอย่างจึงมีหลายค่าด้วย การนำผลต่างที่เกิดขึ้นในทุก ๆ ครั้งมารวมกันแล้วหารด้วยจำวนครั้งเรียกว่า “ค่าเฉลี่ยของความคลาดเคลื่อน”

9. ผลรวมของค่าความแตกต่างความแปรปรวน หมายถึง ผลรวมของผลต่างระหว่างค่าความแปรปรวนของประชากรกับค่าความแปรปรวนของกลุ่มตัวอย่าง เนื่องจากค่าความแปรปรวนของประชากรมีเพียงค่าเดียวเท่านั้น แต่ค่าความแปรปรวนของกลุ่มตัวอย่างจะเปลี่ยนแปลงไปตามค่าตัวแปรของสมาชิกในกลุ่ม ทำให้ค่าความแปรปรวนของกลุ่มตัวอย่างมีหลายค่า ดังนั้นผลต่างระหว่างค่าความแปรปรวนของประชากรกับกลุ่มตัวอย่างจึงมีหลายค่าด้วย การนำผลต่างที่เกิดขึ้นในทุก ๆ ครั้งมารวมกันเรียกว่า “ผลรวมของค่าความแตกต่างความแปรปรวน”

10. ค่าจริง (True value) หมายถึง ค่าที่แสดงคุณลักษณะหรือคุณสมบัติของสิ่งที่ต้องการวัด โดยค่าจริงนี้มีความถูกต้องและไม่มีความผิดพลาดคลาดเคลื่อน ในทางสถิติ “ค่าจริง” มีความหมายเหมือนกับคำว่า “ค่าพารามิเตอร์” ยกตัวอย่าง เช่น แบบจำลองแผนที่แสดงค่าตัวแปร Z ตามภาพที่ 2-27 ค่าจริงบริเวณตรงกลางของแผนที่ คือ 4

11. ค่าประมาณ (Approximate value) หมายถึง ค่าที่หาได้จากวิธีการประมาณค่า เพื่อใช้ทดแทนค่าจริง ในทางสถิติ “ค่าประมาณ” มีความหมายเหมือนกับคำว่า “ค่าสถิติ” ยกตัวอย่างเช่น แบบจำลองแผนที่แสดงค่าตัวแปร Z ตามภาพที่ 2-29 ค่าประมาณบริเวณตรงกลางของแผนที่ คือ 3.5

12. ค่าคลาดเคลื่อนร้อยละ (Percent error) หมายถึง ค่าปริมาณความแตกต่างระหว่างค่าความจริงกับค่าประมาณ เมื่อเปรียบเทียบกับค่าจริง ในรูปแบบอัตราส่วนร้อยละ โดยมีสูตรการคำนวน คือ $E = \frac{\left|e_{i'j'}\right|}{\mu_{i'j'}} \text{x} 100\%$ ยกตัวอย่างเช่น แบบจำลองแผนที่แสดงค่าตัวแปร Z ตามภาพที่ 2-31 ค่าคลาดเคลื่อนร้อยละบริเวณตรงกลางของแผนที่ คือ 12.5%

13. พื้นที่ตามภูมิภาค หมายถึง พื้นที่มีอาณาเขตชัดเจนและแน่นอน โดยมีเส้นแบ่งเขตแดนตามธรรมชาติ เช่น ภูเขา ทิวเขา แม่น้ำ ฯลฯ หรือเส้นแบ่งเขตแดนที่สร้างขึ้น เช่น ถนน รั้ว กำแพง ฯลฯ ซึ่งเป็นที่รับรู้ของประชาชนโดยทั่วไปหรือมีกฎหมายในการแบ่งอาณาเขต

14. พื้นที่ของ $N_{i,j}$ หมายถึง พื้นที่ที่ถูกแบ่งเขตแดนด้วยทฤษฎีการสุ่มแบบกริด ซึ่งพื้นที่จะถูกแบ่งออกเป็นส่วน ๆ ตามจำนวนของจุดกริดตามภาพที่ 2-11 และภาพที่ 2-12 โดยการแบ่งพื้นที่ตามจุดกริดนั้นไม่ให้ความสำคัญกับแนวเขต/ เขตแดนตามภูมิภาค

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยทางพฤติกรรมและสังคมศาสตร์ ซึ่งในบางกรณีประชากรมีขนาดใหญ่มาก ทำให้ไม่สามารถที่จะศึกษาจากประชากรทั้งหมดได้ เนื่องจากติดขัดด้วยปัจจัยหลายอย่าง เช่น งบประมาณ พนักงาน และระยะเวลา จึงจำเป็นต้องศึกษาเฉพาะกลุ่มตัวอย่าง เพื่อทำการอนุมานย้อนกลับไปยังประชากร ดังนั้น การเลือกใช้วิธีการในการสุ่มจึงถือเป็นเรื่องสำคัญ ถ้ากลุ่มตัวอย่างที่สุ่มได้ไม่เป็นตัวแทนที่ดีของประชากร จะทำให้การสรุปลักษณะของประชากรเกิดความผิดพลาด เพื่อให้งานวิจัยที่ทำมีความถูกต้อง ความเชื่อมั่นสูง และมีความคลาดเคลื่อนต่ำ ผู้วิจัยจึงต้องศึกษารายละเอียดของแต่ละวิธีการสุ่ม เพื่อให้เหมาะสมกับลักษณะประชากร ตลอดจนขนาดของกลุ่มตัวอย่าง ทฤษฎีต่าง ๆ ที่เกี่ยวข้องกับการสุ่มแบบกริด และงานวิจัยที่ใช้วิธีการสุ่มแบบกริดในการทดลอง ในการศึกษาครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องในแต่ละหัวข้อและนำเสนอตามลำดับ ดังนี้

1. ตอนที่ 1 ความหมาย ประเภท และวิธีการของการสุ่มตัวอย่างจากประชากร
2. ตอนที่ 2 แนวคิดและทฤษฎีของการสุ่มแบบกริด
3. ตอนที่ 3 การสุ่มตัวอย่างจำนวนมาก ๆ ที่ใช้ในการทดสอบทางการศึกษาของประเทศไทย
4. ตอนที่ 4 การสุ่มตัวอย่างมีผลกระทบต่อความแม่นยำของโพลล์
5. ตอนที่ 5 แนวคิดในการสร้างแบบจำลอง
6. ตอนที่ 6 งานวิจัยในประเทศและต่างประเทศที่เกี่ยวข้อง

ตอนที่ 1 ความหมาย ประเภท และวิธีการสุ่มตัวอย่างจากประชากร

1. การสุ่มตัวอย่างจากประชากร (Random selection)

วัตถุประสงค์ของ การสุ่มตัวอย่าง คือ เพื่อให้กลุ่มตัวอย่างที่สุ่มได้เป็นตัวแทนที่ดีของประชากร ผลการวิจัยที่ศึกษาจากกลุ่มตัวอย่างและอ้างอิงไปยังประชากรของงานวิจัย ซึ่งวิธีการสุ่มตัวอย่างจำแนกเป็น 2 ประเภทใหญ่ ๆ คือ การสุ่มตัวอย่างที่ใช้ทฤษฎีความน่าจะเป็น (Probability sampling) และการสุ่มตัวอย่างที่ไม่ใช้ทฤษฎีความน่าจะเป็น (Nonprobability sampling) ซึ่งความแตกต่างระหว่างประเภทของวิธีการสุ่มมีรายละเอียด ดังนี้

การเลือกตัวอย่างประชากรโดยไม่ใช้ทฤษฎีความน่าจะเป็น (Non-probability sampling)

การเลือกตัวอย่างโดยไม่คำนึงถึงทฤษฎีความน่าจะเป็น อาจทำให้บางส่วนของประชากรไม่มีโอกาสได้รับการเลือกหรือมีความน่าจะเป็นในการถูกสุ่มน้อยมาก ซึ่งการเลือกตัวอย่างด้วยวิธีนี้ต้องขึ้นอยู่กับความรู้และประสบการณ์ของผู้วิจัยหรือพนักงานสำรวจรวบรวมข้อมูล ดังนั้น การเลือกตัวอย่างจากประชากรจึงมีความเสี่ยงสูงที่ทำให้เกิดความลำเอียงในการสุ่มตัวอย่าง นอกจากนั้นการเลือกตัวอย่างโดยไม่อาศัยหลักความน่าจะเป็นไม่มีความเหมาะสมกับการทดสอบสมมติฐานโดยใช้สถิติอนุมาน ดังนั้นการเลือกตัวอย่างแบบนี้จึงไม่มีหลักประกันใด ๆ ว่ากลุ่มตัวอย่างที่ได้ จะเป็นตัวแทนที่ดีของประชากร

การสุ่มตัวอย่างโดยใช้ทฤษฎีความน่าจะเป็น (Probability sampling)

การสุ่มตัวอย่างโดยใช้ทฤษฎีความน่าจะเป็นคือการสุ่มที่สมาชิกแต่ละหน่วยของประชากรมีโอกาสที่จะถูกเลือก ซึ่งเป็นวิธีการสุ่มที่มีระเบียบและกฎเกณฑ์ที่น่าเชื่อถือได้ ทำให้ปราศจากอคติหรือความลำเอียงในการสุ่มตัวอย่าง จึงทำให้เชื่อถือได้ว่ากลุ่มตัวอย่างที่ถูกสุ่มออกมาเป็นตัวแทนที่ดีของประชากร และผลสรุปที่ได้สามารถอ้างไปถึงประชากรได้ โดยอาศัยวิธีการทางสถิติอ้างอิง นอกจากนั้นวิธีการสุ่มแบบนี้สามารถรองรับวิธีการทางสถิติชั้นสูงได้หลายวิธี เนื่องจากวิธีการทางสถิติชั้นสูงสร้างขึ้นโดยใช้ทฤษฎีความน่าจะเป็น การสุ่มตัวอย่างประเภทนี้ ได้แก่ การสุ่มตัวอย่างอย่างง่าย, การสุ่มแบบมีระบบ, การสุ่มแบบชั้นภูมิ, การสุ่มแบบแบ่งกลุ่ม และการสุ่มแบบหลายขั้นตอน ซึ่งมีรายละเอียดดังนี้

2. การสุ่มตัวอย่างอย่างง่าย (Simple random sampling)

การสุ่มตัวอย่างอย่างง่าย เป็นวิธีการสุ่มตัวอย่างที่เปิดโอกาสให้สมาชิกทุกหน่วยมีโอกาสถูกเลือกมาเท่าเทียมกัน ทำให้วิธีการนี้ปราศจากอคติและความลำเอียง โดยผู้วิจัยจะต้องมีบัญชีรายชื่อของประชากรทั้งหมด และกำหนดหมายเลขให้กับรายชื่อสมาชิกหรือค่าสังเกตแต่ละหน่วย จากนั้นผู้วิจัยทำการสุ่มตัวอย่างตามจำนวนที่ต้องการจากบัญชีรายชื่อการสุ่มตัวอย่างแบบนี้มีความเหมาะสมกับประชากรที่มีลักษณะคล้ายคลึงกันหรือเป็นเอกพันธ์ นอกจากนั้นวิธีการนี้เป็นที่นิยม และเป็นพื้นฐานของการสุ่มตัวอย่างวิธีอื่น ๆ การสุ่มตัวอย่างอย่างง่ายสามารถทำได้ 3 วิธี คือ

การจับฉลาก

ในการใช้วิธีการนี้ผู้วิจัยต้องจัดทำฉลาก โดยกำหนดชื่อหรือหมายเลขขึ้น เพื่อเป็นรหัสแทนสมาชิกแต่ละหน่วยของประชากร จากนั้นทำการจับฉลากจนได้จำนวนสมาชิกครบตามจำนวนที่ต้องการ (สำนักสถิติแห่งชาติ, ม.ป.ป.)

ตารางการสุ่ม

ในการใช้วิธีการนี้ผู้วิจัยต้องเลขที่หรือหมายเลขขึ้น เพื่อเป็นรหัสแทนสมาชิกแต่ละหน่วยของประชากร จากนั้นสุ่มหลัก (Column) และแถว (Row) ของตัวเลขเริ่มต้น แล้วอ่านจากซ้ายไปขวา เมื่อจบแถวให้ขึ้นแถวใหม่ต่อไปตามลำดับ ถ้าได้หมายเลขซ้ำหรือหมายเลขเกินจากประชากรก็ตัดออก จนได้จำนวนตัวอย่างครบตามที่ต้องการ (สำนักสถิติแห่งชาติ, ม.ป.ป.)

การใช้คอมพิวเตอร์ในการสุ่ม

วิธีการนี้ใช้คอมพิวเตอร์เป็นเครื่องมือในการสร้างรหัสแทนสมาชิกทั้งหมดของประชากร เพื่อให้คอมพิวเตอร์สุ่มรหัสตามจำนวนที่ผู้วิจัยต้องการ การสุ่มโดยใช้คอมพิวเตอร์มีข้อดี คือ วิธีการนี้เหมาะกับการเลือกหน่วยตัวอย่างที่ประชากรมีขนาดใหญ่หรือต้องสุ่มตัวอย่างหลายสิบหลายร้อยครั้ง ซึ่งเป็นการทดแทนแรงงานมนุษย์ โดยขั้นแรกผู้วิจัยต้องเขียนคำสั่งในโปรแกรม (Source code) เพื่อรับข้อมูลที่เป็นตัวเลขและอักขระทางคีย์บอร์ด จากนั้นคอมพิวเตอร์จะประมวลผล และแสดงผลกลุ่มตัวอย่างที่ถูกสุ่มออกมาทางหน้าจอคอมพิวเตอร์ ในกรณีที่ผลลัพธ์มีขนาดใหญ่และมีข้อมูลมาก ต้องแสดงผลข้อมูลในรูปแบบไฟล์ข้อมูล (สำนักสถิติแห่งชาติ, ม.ป.ป.)

จุดเด่นการสุ่มตัวอย่างอย่างง่ายเป็นวิธีการไม่สลับซับซ้อนและสะดวกต่อการนำไปใช้ (สุชาดา บวรกิติวงศ์, 2548, หน้า 117; สุรินทร์ นิยมมางกูร, 2546, หน้า 64) ทำการวิเคราะห์และอธิบายผลการทดลองได้ง่ายแต่จุดด้อยการสุ่มตัวอย่างอย่างง่าย ถ้าประชากรขนาดใหญ่ใช้เวลาดำเนินการมาก และมีค่าใช้จ่ายสูง เพราะต้องจัดทำบัญชีรายชื่อสมาชิกทุกหน่วยของประชากรและเป็นวิธีการที่ก่อให้เกิดความคลาดเคลื่อนได้มาก (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

3. การสุ่มตัวอย่างแบบมีระบบ (Systematic random sampling)

การสุ่มตัวอย่างแบบนี้เหมาะสำหรับประชากรที่ถูกจัดไว้เป็นระบบอยู่แล้ว เช่น เรียงตามเลขบัตรประจำตัวประชาชน หรือเรียงตามบ้านเลขที่ จากนั้นสุ่มตัวอย่างจากประชากรเป็นช่วง ๆ ซึ่งการสุ่มหน่วยตัวอย่างแรกอาจจะใช้วิธีจับฉลากหรือตารางสุ่มก็ได้ และเมื่อสุ่มได้หน่วยแรกแล้ว หน่วยที่ 2 และหน่วยอื่น ๆ ถัดไป จะห่างเท่า ๆ กัน ซึ่งวิธีการคำนวณทำได้ โดยเอาประชากรทั้งหมด (N) หารด้วยขนาดตัวอย่างที่ต้องการ (n) เท่ากับช่วงห่างของชุดตัวอย่าง (I) ยกตัวอย่าง เช่น ห้อง ม.3/ 1 มีนักเรียน (N) 30 คน ซึ่งเลขที่ของนักเรียนเรียงตามอักษรจากเลขที่ 1 ถึง 30 ถ้าต้องการสุ่มตัวอย่าง (n) 6 คน ดังนั้น ช่วงห่างของชุดตัวอย่าง (I) คือ 5 จากหน่วยประชากรที่ 1 ถึง 5 นั้นต้องถูกเลือกมา 1 หน่วย เพื่อเป็นตัวเริ่มต้นของการสุ่ม สมมุติว่าตัวแรกสุ่มได้เลขที่ 2 หน่วยถัดไป คือ 2 + 5, 2 + 2 (5), 2 + 3 (5), 2 + 4 (5) และ 2 + 5 (5) สรุปนักเรียนห้อง ม.3/ 1ที่ถูกสุ่มออกมา 6 คน คือ เลขที่ 2, 7, 12, 17, 22 และ 27 (สำนักสถิติแห่งชาติ, ม.ป.ป.)

จุดเด่นการสุ่มตัวอย่างแบบมีระบบคือวิธีการที่สะดวกง่ายต่อการปฏิบัติ ทำการวิเคราะห์และอธิบายผลการทดลองได้ง่าย (สมบัติ ท้ายเรือคำ, 2551, หน้า 54) แต่จุดด้อยการสุ่มตัวอย่างแบบมีระบบคือ ในกรณีที่ประชากรขนาดใหญ่ใช้เวลาดำเนินการมาก และมีค่าใช้จ่ายสูง เพราะต้องจัดทำบัญชีรายชื่อสมาชิกทุกหน่วยของประชากร ถ้าบัญชีรายชื่อของประชากรจัดเรียงอย่างเป็นระบบหรือสมาชิกของประชากรมีการเปลี่ยนแปลงขึ้นลงเป็นวงจรหรือช่วง ทำให้เกิดความลำเอียงในการสุ่ม (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

4. การสุ่มตัวอย่างแบบชั้นภูมิ (Stratified random sampling)

เป็นการสุ่มกลุ่มตัวอย่างที่แบ่งกลุ่มประชากรเป็นชั้น ๆ (Stratum) ตามตัวแปรคุณลักษณะที่มีผลต่อตัวแปรตาม เช่น เพศ ช่วงอายุ ระดับการศึกษา รายได้ อาชีพ ฯลฯ โดยพยายามแบ่งให้แต่ละชั้นมีความเป็นเอกพันธ์หรือลักษณะคล้ายคลึงกัน แต่มีความแตกต่างระหว่างชั้นมากที่สุด ดังนั้น ทุกภาคส่วนของประชากรมีโอกาสเป็นตัวแทนหรือได้รับการสุ่มเลือก ยกตัวอย่าง เช่น การศึกษาปัญหาการมั่วสุมตามสถานบันเทิงวัยรุ่น ผู้วิจัยสามารถแบ่งวัยรุ่นออกเป็นระดับการศึกษาได้ เนื่องจากในแต่ละระดับการศึกษามีนักเรียนนักศึกษาที่มั่วสุมตามสถานบันเทิงแตกต่างกันหรือเป็นวิวิธพันธ์ เช่น นักศึกษาระดับอุดมศึกษามีการมั่วสุมกันมากที่สุด เนื่องจากมีกฎหมายรองรับและมีเงินทุนมากกว่านักเรียนระดับ ม.ต้น แต่ถ้าพิจารณาในระดับการศึกษาเท่ากันพบว่ามีความเป็นเอกพันธ์ เช่น นักเรียนทั้งสายอาชีพและสายสามัญมีการมั่วสุมตามสถานบันเทิง

จุดเด่นการสุ่มตัวอย่างแบบชั้นภูมิ เป็นวิธีการควบคุมตัวแปรแทรกซ้อนได้ดี, กลุ่มตัวอย่างที่ได้มีความเป็นตัวแทนประชากรย่อย, ความคลาดเคลื่อนจากการเลือกตัวอย่าง (Sampling error) น้อยกว่าการสุ่มตัวอย่างอย่างง่ายและการสุ่มตัวอย่างแบบมีระบบ เมื่อขนาดกลุ่มตัวอย่างเท่ากัน, และสามารถเลือกใช้วิธีการสุ่มที่แตกต่างกันในแต่ละชั้น ซึ่งเป็นประโยชน์อย่างมากในทางปฏิบัติ เพราะในบางครั้งชั้นภูมิแต่ละชั้นภูมิมีลักษณะที่แตกต่างกันมาก ผู้วิจัยสามารถที่จะเลือกใช้วิธีสุ่มตัวอย่างที่ตามความเหมาะสม แต่จุดด้อยการสุ่มตัวอย่างแบบชั้นภูมิคือการแบ่งประชากรเป็นประชากรย่อยปฏิบัติได้ยากหรือไม่สามารถปฏิบัติได้ และการประมาณค่าพารามิเตอร์มีความสลับซับซ้อน ถ้าแต่ละชั้นใช้วิธีสุ่มแตกต่างกัน (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

5. การสุ่มตัวอย่างแบบแบ่งกลุ่ม (Cluster sampling)

เป็นการสุ่มตัวอย่างด้วยการแบ่งประชากรออกเป็นกลุ่ม ๆ โดยแต่ละกลุ่มมีลักษณะภายในกลุ่มมีความหลากหลายหรือเป็นวิวิธพันธ์ แต่ระหว่างกลุ่มมีความคล้ายคลึงกันหรือเป็นเอกพันธ์ ซึ่งการสุ่มตัวอย่างแบบแบ่งกลุ่มเหมาะสำหรับโครงการวิจัยที่มีขนาดใหญ่ที่มีประชากรจำนวนมาก ยกตัวอย่าง เช่น การศึกษาปัญหาการมั่วสุมตามสถานบันเทิงวัยรุ่น ผู้วิจัยสามารถแบ่งวัยรุ่นออกเป็นตามภาคได้ เนื่องจากในแต่ละภาคของประเทศมีนักเรียนนักศึกษาที่มั่วสุมตามสถานบันเทิงเหมือนกัน

หรือเป็นเอกพันธ์ แต่ถ้าพิจารณาในระดับเมืองพบว่าแต่ละเมืองมีระดับความรุนแรงแตกต่างกัน เป็นวิวิธพันธ์ เช่น ในเมืองสำคัญ ๆ ตามแต่ละภาคมีระดับปัญหาที่รุนแรงมากกว่าชนบท เป็นต้น

จุดเด่นการสุ่มตัวอย่างแบบแบ่งกลุ่ม คือ ปฏิบัติได้ง่ายและสะดวก ไม่จำเป็นต้องจัดทำบัญชีรายชื่อสมาชิกทุกหน่วยของประชากร และสามารถสุ่มโดยใช้พื้นที่เป็นหน่วยของการสุ่ม ทำให้ลดเวลา ค่าใช้จ่าย และแรงงานในการศึกษาจากกลุ่มตัวอย่าง เนื่องจากกลุ่มตัวอย่างอาศัยอยู่รวมกัน แต่จุดด้อยการสุ่มตัวอย่างแบบแบ่งกลุ่ม คือ การแบ่งประชากรเป็นประชากรย่อยปฏิบัติได้ยากหรือไม่สามารถปฏิบัติได้ และถ้าแบ่งประชากรไม่ดีจนกระทั่งประชากรย่อยมีความแปรปรวนระหว่างกลุ่มสูงจะทำให้ประสิทธิภาพของการสุ่มตัวอย่างต่ำ (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131) การคำนวณค่าความแปรปรวนของข้อมูลยุ่งยากกว่าการสุ่มตัวอย่างแบบง่าย และโดยภาพรวมวิธีนี้มีประสิทธิภาพต่ำกว่าวิธีเลือกกลุ่มตัวอย่างแบบใช้ความน่าจะเป็นวิธีอื่น ๆ (สุชาดา บวรกิติวงศ์, 2548, หน้า 119; สมบัติ ท้ายเรือคำ, 2551, หน้า 55)

6. การสุ่มตัวอย่างแบบหลายขั้นตอน (Multistage sampling)

เป็นการสุ่มตัวอย่างที่มีขั้นตอนการสุ่มตั้งแต่ 2 ขั้นตอนขึ้นไป โดยในแต่ละขั้นตอนต้องเลือกใช้วิธีการสุ่มอย่างเหมาะสม ทำให้วิธีการสุ่มมีหลายรูปแบบมาผสมผสานกัน ในการสุ่มครั้งแรกคือการสุ่มประชากรจากหน่วยที่ใหญ่ที่สุดจนถึงการสุ่มครั้งสุดท้ายคือหน่วยย่อยที่ผู้วิจัยต้องการ การสุ่มวิธีนี้นิยมใช้ในกรณีเป็นโครงการสำรวจที่ประชากรขนาดใหญ่และอาศัยอยู่อย่างกระจัดกระจายกว้างขวาง เช่น โครงการศึกษาความสุขของประชากรโลกสามารถดำเนินการได้ดังนี้

1. สุ่มทวีปต่าง ๆ เช่น ใช้การสุ่มแบบแบ่งกลุ่มได้แก่ ทวีปแอฟริกา ทวีปอเมริกา ทวีปโอเชียเนีย ทวีปยุโรป และทวีปเอเชีย

2. สุ่มประเทศในแต่ละทวีป เช่น ทวีปยุโรป ใช้การสุ่มตัวอย่างอย่าง่ายได้ ฝรั่งเศส เยอรมนี นอร์เวย์ สวีเดน กรีซ สโลวีเนียโรมาเนีย และ ยูเครน

3. สุ่มเมืองในแต่ละประเทศ เช่น ประเทศสวีเดน ใช้การสุ่มแบบมีระบบได้แก่ มณฑล Uppsala มณฑล Norrbotten มณฑล Dalarna มณฑล Jamtland และ มณฑล Blekinge

จุดเด่นของการสุ่มแบบหลายขั้นตอนคือการสร้างกรอบตัวอย่างที่สมบูรณ์และทันสมัย สามารถทำได้อย่างสะดวกและรวดเร็ว (สุรินทร์ นิยมมางกูร, 2546, หน้า 198) และเป็นวิธีการที่เหมาะสมกับประชากรขนาดใหญ่ที่แบ่งเป็นลำดับชั้นลดหลั่น แต่จุดด้อยการสุ่มแบบหลายขั้นตอน คือ ความยุ่งยากในการประมวลผลและการวิเคราะห์ข้อมูล ซึ่งจะต้องมีการคำนวณทุกขั้นตอนที่ทำการสุ่มตัวอย่างออกมา (สุชาดา บวรกิติวงศ์, 2548, หน้า 120; สุรินทร์ นิยมมางกูร, 2546, หน้า 198) และการประมาณค่าพารามิเตอร์จะสลับซับซ้อนถ้ามีจำนวนชั้นของการสุ่มมาก (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

สรุปได้ว่าการสุ่มมีหลายวิธี ซึ่งแต่ละมีกระบวนการ ขั้นตอน จุดเด่นจุดด้อยแตกต่างกัน การเลือกวิธีการสุ่มจึงเป็นสิ่งที่สำคัญ ดังนั้นก่อนทำวิจัยหรือเก็บรวบรวมข้อมูลต้องพิจารณารายละเอียดและองค์ประกอบต่าง ๆ ได้แก่ ขนาดของประชากรและกลุ่มตัวอย่าง ลักษณะความเป็นเอกพันธ์/ วิวิธพันธ์ของประชากร ทรัพยากร/ งบประมาณในการวิจัย เป็นต้น เพื่อเลือกวิธีการสุ่มที่เหมาะสมกับงานวิจัยมากที่สุด โดยที่งานวิจัยมีความถูกต้องและยอมรับได้

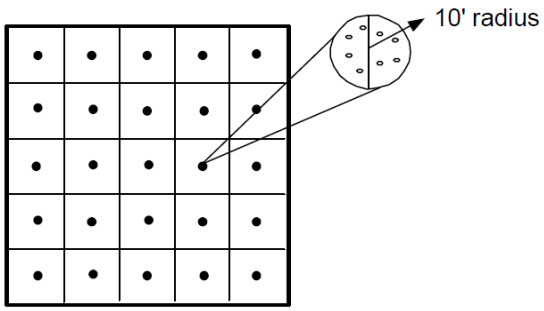
ตอนที่ 2 แนวคิดและทฤษฎีของการสุ่มแบบกริด

1. ที่มาและความสำคัญของการสุ่มแบบกริด

การเก็บตัวอย่างดินเพื่อทำการทดสอบคุณสมบัติได้รับความนิยมในช่วง ปี ค.ศ. 1960 ถึง ค.ศ. 1970 การเก็บตัวอย่างในยุคนั้นอาศัยปัจจัย 3 ประการ คือ การใช้ประโยชน์ของที่ดิน หน่วยแผนที่ดิน (Soil mapping unit) และภูมิประเทศเป็นตัวแบ่งพื้นที่ของการสุ่ม ปัจจัยเหล่านี้ทำให้ทราบถึงปริมาณแร่ธาตุในดิน ซึ่งสารเคมีและคุณสมบัติทางกายภาพของตัวอย่างดินที่สุ่มได้เป็นตัวบ่งชี้ถึงปริมาณของแร่ธาตุในดิน การเจริญเติบโตของพืช และการแพร่กระจายในดิน (Mallarino & Wittry, 2001, p. 159) วิธีการเก็บตัวอย่างดินให้ได้ความถูกต้องแม่นยำต้องพิจารณาในเรื่องการไถเตรียมดิน, การให้ปุ๋ยในอดีต, การชลประทานและการระบายน้ำ, ประเภทของดินและภูมิประเทศ ซึ่งปัจจัยเหล่านี้จะเป็นตัวบ่งชี้ในการเตรียมดินสำหรับเพาะปลูกและการใส่ปุ๋ย ถ้าเก็บตัวอย่างดินได้ความครอบคลุมและเป็นตัวแทนที่ดีของประชากร จะช่วยลดความสิ้นเปลืองในการทำเกษตรกรรมได้

2. หลักการของการสุ่มแบบกริด

การสุ่มตัวอย่างแบบกริดสามารถแบ่งได้ 2 ประเภท คือ การสุ่มตัวอย่างแบบกริดเซลล์และการสุ่มตัวอย่างแบบจุดกริด (Dinkins & Jones, 2008, p. 1) การสุ่มตัวอย่างแบบกริดเซลล์ คือ การแบ่งพื้นที่ออกเป็นสี่เหลี่ยมอย่างละเท่า ๆ กัน และทำการเก็บตัวอย่างที่กริดเซลล์ที่ได้วางแผนไว้ แต่การสุ่มตัวอย่างแบบจุดกริดเป็นการเก็บตัวอย่างตรงจุดอ้างอิงหรือเก็บตัวอย่างบริเวณใกล้ ๆ กับจุดอ้างอิง สำหรับขนาดของกริดเซลล์มีไม่ควรเกิน 40 เอเคอร์ แต่ในกรณีพื้นที่นั้นมีตัวแปรมากหรือเป็นพื้นที่มีปัญหาต้องใช้ขนาดกริดเซลล์ที่เล็กกว่านั้น การเก็บตัวอย่างดินใช้กลุ่มตัวอย่างประมาณ 10-20 จุด (Midwest Labolatories, 2009) และเพื่อให้การสุ่มตัวอย่างมีความถูกต้องเหมาะสม โดยใช้ความหนาแน่นในการสุ่มเท่ากับพื้นที่ 2.5 เอเคอร์ ต่อ 1 จุดกริด และภายในวงกลมของแต่ละจุดกริดให้สุ่มตัวอย่างออกมา 8-10 ตำแหน่ง โดยรัศมีวงกลมมีขนาด 10 ฟุต ดังภาพที่ 2-1



ภาพที่ 2-1 การเก็บตัวอย่างดินแบบกริด (Midwest Labolatories, 2009, p. 2)

3. การสุ่มแบบกริดด้วย GPS

การเก็บตัวอย่างดินโดยการใช้ข้อมูลจาก GPS มีประโยชน์ในการตัดสินใจด้านการจัดการด้านแร่ธาตุ ซึ่งการใส่ปุ๋ยที่ถูกต้องนั้นต้องพิจารณาถึงปริมาณแร่ธาตุของเดิมที่มีอยู่ ซึ่งการทดสอบดินจะทำให้ทราบถึงปริมาณฟอสฟอรัสหรือโพแทสเซียม การใส่ปุ๋ยให้ตรงตามความต้องการของพืชย่อมให้ผลผลิตทางการเกษตรมากกว่าการใส่ปุ๋ยแบบผสม

ขั้นตอนในการการเก็บตัวอย่างดินด้วย GPS

1. การสร้างแผนที่ต้องการกำหนดขอบเขตของแผนที่ โดยการขับรถ atv ที่ติดอุปกรณ์บันทึกข้อมูลการเคลื่อนที่ของยานพาหนะ (GPS data logger) วิ่งรอบพื้นที่ และทำการบันทึกข้อมูลทุก ๆ 5 วินาที

2. ขนาดกริดเซลล์ปกติเท่ากับ 1 เอเคอร์ แต่สามารถปรับแต่งได้ตามความเหมาะสม เช่น 2 เอเคอร์ หรือ 0.5 เอเคอร์

3. ในพื้นที่กริดเซลล์เดียวกันให้สุ่มตัวอย่างดินออกมา 5 ตำแหน่ง (หรือตามความเหมาะสม) นำดินที่เก็บได้มาผสมรวมกัน เพื่อเตรียมนำไปวิเคราะห์

4. บันทึกข้อมูลตำแหน่งของจุดกริดลงเครื่อง GPS

5. ตัวอย่างดินจะถูกนำไปวิเคราะห์ในห้องทดลอง เพื่อวิเคราะห์ ค่าความเป็นกรดเป็นด่าง ฟอสฟอรัส โพแทสเซียม และแมกนีเซียม

6. ผลที่ได้จากการทดลองจะถูกนำไปสร้างเป็นแผนที่แร่ธาตุ ซึ่งแผนที่นี้จะนำไปใช้วางแผนการเพาะปลูก

สรุปได้ว่าวิธีการสุ่มแบบกริดถูกพัฒนาขึ้นจากการเก็บตัวอย่างดิน ในช่วงปี ค.ศ.1960 แต่ในปัจจุบันได้ใช้เครื่องมือ GPS เพื่อความแม่นยำและความสะดวกในการหาตำแหน่งพิกัดในการเก็บตัวอย่าง ซึ่งหลักการของการสุ่มแบบกริดคือแบ่งพื้นที่ออกเป็นกริดเซลล์อย่างละเท่า ๆ กัน

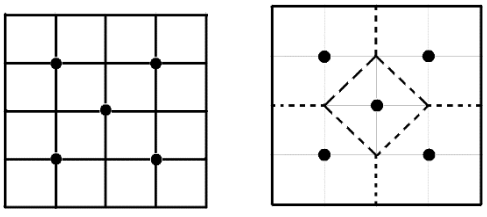
โดยขนาดกริดเซลล์ที่เหมาะสม คือ 1 เอเคอร์ และทำการเก็บสุ่มตัวอย่างดินตรงบริเวณโดยรอบจุดกริด (รัศมี 10 ฟุต) จำนวน 5 ตำแหน่ง จากนั้นตัวอย่างดินจะถูกนำไปวิเคราะห์คุณสมบัติและแร่ธาตุในดิน เพื่อสร้างแผนที่แร่ธาตุและวางแผนการเพาะปลูก ทำให้ลดความสิ้นเปลืองในการทำเกษตรกรรม จากการศึกษาหลักการและรายละเอียดของวิธีการสุ่มแบบกริดพบว่าเป็นวิธีการสุ่มที่ใช้งานได้จริงและมีประโยชน์ ซึ่งเหมาะสมที่จะนำวิธีดังกล่าวมาทำการวิจัย

4. ทฤษฎีการสุ่มแบบกริด

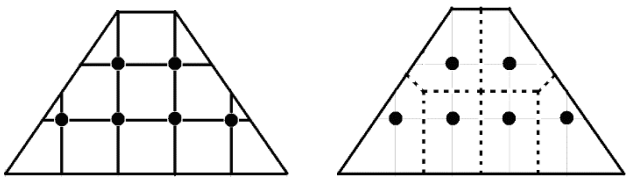
การสุ่มตัวอย่างในงานวิจัยเชิงสังคมและพฤติกรรมศาสตร์มีวิธีการสุ่มพื้นฐาน 4 วิธี ได้แก่ การสุ่มแบบอย่างง่าย การสุ่มแบบมีระบบ การสุ่มแบบชั้นภูมิ การสุ่มแบบแบ่งกลุ่ม ถ้านำวิธีการสุ่มหลาย ๆ วิธีมาใช้ร่วมกันจะเรียกว่าวิธีการสุ่มแบบหลายขั้นตอน หรืออาจเรียกชื่อตามการผสมผสาน ยกตัวอย่างเช่น Stratified-cluster sampling, Three-stage cluster sampling และ Multi-stage cluster sampling เป็นต้น ซึ่งนักคณิตศาสตร์/ นักสถิติได้คิดค้นวิธีการและทฤษฎีการสุ่มเพื่อใช้อธิบายหลักการและแนวคิดในเชิงสัญลักษณ์ เช่น การประมาณค่าเฉลี่ยและยอดรวมของประชากร การกำหนดขนาดตัวอย่าง การจัดสรรขนาดตัวอย่าง การประมาณค่าอัตราส่วน การประมาณค่าสัดส่วน เป็นต้น แต่วิธีการสุ่มแบบกริดเป็นการสุ่มตัวอย่างในงานวิจัยประเภทวิทยาศาสตร์กายภาพ โดยการสุ่มแบบจุดกริดถูกพัฒนาขึ้นจากการการเก็บตัวอย่างดิน เพื่อทำการทดสอบปริมาณแร่ธาตุ, ความเป็นกรดเป็นด่าง หรือการระบายน้ำและอากาศของดิน ดังนั้นทฤษฎีการสุ่มแบบกริดที่มีอยู่เดิมควรปรับปรุงใหม่ให้มีความเหมาะสมกับตัวแปรเชิงพื้นที่ที่เป็นความเชื่อ ความพึงพอใจ ความคิดเห็น เจตคติ ฯลฯ

นิยามและสัญลักษณ์

เมื่อกำหนดจำนวนจุดกริดและตำแหน่งแล้วพื้นที่ทั้งหมดจะถูกแบ่งออกเป็นส่วน ๆ ตามจำนวนของจุดกริดนั้น ยกตัวอย่าง เช่น



ภาพที่ 2-2 พื้นที่ของ N_{ij}



ภาพที่ 2-3 พื้นที่ของ N_{ij}

จากภาพที่ 2-2 และภาพที่ 2-3 แสดงถึงการแบ่งพื้นที่ทั้งหมดออกเป็นส่วน ๆ เท่ากับจำนวน จุดกริดที่ได้กำหนดไว้ โดยมีเส้นแบ่งอยู่ระหว่างกึ่งกลางของจุดกริดทั้งสอง ซึ่งการแบ่งพื้นที่ไม่ต้องคำนึงถึงความแตกต่างของขนาดพื้นที่

เมื่อ x คือ จำนวนจุดบนละติจูด

y คือ จำนวนจุดบนลองติจูด

โดยที่ **i** = 1, 2, 3, ..., **x** และ **j** = 1, 2, 3, ..., **y**

Δx คือ ระยะห่างระหว่างจุดกริดตามแนวแกน x

Δy คือ ระยะห่างระหว่างจุดกริดตามแนวแกน y

N คือ ขนาดประชากร

N_{ij} คือ ขนาดประชากรที่อยู่ในพื้นที่ของจุดกริด (i, j) ซึ่งพื้นที่ได้ถูกแบ่งออกเป็นส่วน ๆ ตามจำนวนของจุดกริด ดังภาพที่ 2-2 และภาพที่ 2-3

$$N = \sum_{i}^{x} \sum_{j}^{y} N_{ij}$$

ถ้า n คือ ขนาดกลุ่มตัวอย่าง

n_{ij} คือ ขนาดกลุ่มตัวอย่างที่อยู่เฉพาะภายในวงกลมของการสุ่ม

ดังนั้น $n = \sum_{i}^{x} \sum_{j}^{y} n_{ij}$

a คือ ตัวแปรในการวิจัย เช่น ตัวแปรคะแนนเฉลี่ย O-NET ของโรงเรียนต่าง ๆ ในปีการศึกษา 2548 เป็นต้น

b คือ ตัวที่/ ลำดับที่ของสมาชิก เช่น โรงเรียนที่ 1, โรงเรียนที่ 2 เป็นต้น

$a^{(b)}$ คือ ค่าตัวแปรสุ่มในการวิจัย a ลำดับที่ b โดยที่ $a=\left\{a^{(1)},a^{(2)},a^{(3)},...,a^{(b)}\right\}$

$a_{ij}^{(b)}$ คือ ค่าของตัวแปร a ตัวที่ b ซึ่งสุ่มมาจากจุดกริด (i,j) โดยที่

$$b=1,2,3,...,n_{ij},n_{ij}+1,n_{ij}+2,...,N_{ij}$$

$z_{ij}=\sum_{b}^{n_{ij}} a_{ij}^{(b)}$ คือ ผลรวมของค่าตัวแปรของกลุ่มตัวอย่างที่จุดกริด (i,j)

$\bar{z}_{ij}=\frac{1}{n_{ij}}\sum_{b}^{n_{ij}} a_{i,j}^{(b)}=\frac{z_{ij}}{n_{ij}}$ คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่จุดกริด (i,j)

$\bar{z}=\frac{1}{N}\sum_{i}^{x}\sum_{j}^{y} N_{ij}\ \bar{z}_{ij}$ คือ ค่าเฉลี่ยของกลุ่มตัวอย่าง

$s_{ij}^2=\frac{1}{n_{ij}-1}\sum_{b}^{n_{ij}}(a_{ij}^{(b)}-\bar{z}_{ij})^2$ คือ ความแปรปรวนของกลุ่มตัวอย่างที่จุดกริด (i,j)

$s^2=\frac{\sum_{i}^{x}\sum_{j}^{y}\sum_{b}^{n_{ij}}(a_{ij}^{(b)}-\bar{z})^2}{n-1}$ คือ ความแปรปรวนของกลุ่มตัวอย่าง

$\mu_{ij}=\frac{1}{N_{ij}}\sum_{b}^{N_{ij}} a_{ij}^{(b)}$ คือ ค่าเฉลี่ยของประชากรที่จุดกริด (i,j)

$\mu=\frac{1}{N}\sum_{i}^{x}\sum_{j}^{y}\sum_{b}^{N_{ij}} a_{ij}^{(b)}$ คือ ค่าเฉลี่ยของประชากร

$\sigma_{ij}^2=\frac{1}{N_{ij}-1}\sum_{b}^{N_{ij}}(a_{ij}^{(b)}-\mu_{ij})^2$ คือ ความแปรปรวนของประชากรที่จุดกริด (i,j)

$\sigma^2=\frac{\sum_{i}^{x}\sum_{j}^{y}\sum_{b}^{N_{ij}}(a_{ij}^{(b)}-\mu)^2}{N}$ คือ ความแปรปรวนของประชากร

ในทางปฏิบัติการหาค่า $\bar{z}$ และ s^2 นั้น ทำได้ยาก เนื่องจากไม่ทราบขนาดตัวอย่างของ N_{ij} เพราะการแบ่งพื้นที่ตามภูมิภาคจะถูกแบ่งตามภาค จังหวัด อำเภอ ตำบล หมู่บ้าน แต่การแบ่งพื้นที่ตามจุดกริดนั้นไม่ให้ความสำคัญกับแนวเขต/เขตแดนตามภูมิภาค จึงไม่สามารถหาขนาดตัวอย่างของ N_{ij} ได้ เพื่อแก้ปัญหาดังกล่าวสามารถใช้ตัวประมาณค่าของ $\bar{z}$ ทดแทนได้ ดังนี้

$\hat{\bar{z}} = \frac{1}{n}\sum_{i}^{x}\sum_{j}^{y} n_{ij}\ \bar{z}_{ij}$ คือ ตัวประมาณค่าของค่าเฉลี่ยของกลุ่มตัวอย่าง

$\hat{s}^2 = \frac{\sum_{i}^{x}\sum_{j}^{y}(n_{ij}-1)s_{ij}^2}{\sum_{i}^{x}\sum_{j}^{y} n_{ij} - xy}$ คือ ตัวประมาณค่าของความแปรปรวนของกลุ่มตัวอย่าง

ในกรณีที่มีการจัดสรรตัวอย่างตามสัดส่วน ถ้าสัดส่วนของ $\frac{n_{ij}}{n}=\frac{N_{ij}}{N}$ แล้วตัวประมาณค่า $\hat{\bar{z}}$ จะเท่ากับ $\bar{z}$

5. การสุ่มตำแหน่งจุดกริดและค้นหากลุ่มตัวอย่าง

โปรแกรม QGIS

การสุ่มตัวอย่างแบบจุดกริดจำเป็นต้องใช้แผนที่เป็นเครื่องมือในการวางแผนเพื่อเปรียบเทียบตำแหน่งของจุดกริดกับค่าพิกัดภูมิศาสตร์ ทำให้จุดกริดที่สร้างขึ้นสามารถอ้างอิงไปยังตำแหน่งที่มีอยู่จริงบนพื้นผิวโลก เช่น ตำแหน่งอาคาร ตำแหน่งบ้าน ถนน แม่น้ำ เพื่อใช้เป็นพื้นที่อ้างอิงในการเก็บข้อมูล ทำให้เจ้าหน้าที่สำรวจข้อมูลทำงานได้อย่างสะดวกและไม่สับสน ในสมัยก่อนการกำหนดจุดกริดลงบนแผนที่ อาจต้องใช้ไม้บรรทัดขนานและปากคีบ (Divider) เป็นอุปกรณ์ช่วย แต่การวางจุดพิกัดบนแผนที่ด้วยมือเปล่าเป็นสิ่งที่ล้าสมัยและยุ่งยากในการทำวิจัย ในปัจจุบันความก้าวหน้าของวิทยาศาสตร์และเทคโนโลยีทำให้มีการพัฒนาซอฟต์แวร์ทางภูมิสารสนเทศ เพื่อความสะดวกรวดเร็วในการใช้งาน ซึ่งโปรแกรมของระบบภูมิสารสนเทศที่มีใช้อยู่ในปัจจุบันมีหลายโปรแกรม เช่น ArcGIS, ArcView, QGIS ฯลฯ แต่โปรแกรมที่ผู้วิจัยจะนำมาใช้ในการศึกษาครั้งนี้คือโปรแกรม QGIS เนื่องจากเป็นโปรแกรมรหัสเปิด (Free open source program) ซึ่งโปรแกรมนี้มี มีความสามารถต่าง ๆ ในการใช้ด้านภูมิสารสนเทศ เช่น การเรียกดูและแสดงผลข้อมูลของแผนที่ภาพเชิงเส้น การเรียกดูและแสดงผลข้อมูลของแผนที่ภาพเชิงตัวเลข การสร้างข้อมูลแผนที่ด้วยวิธีการ Digitize ฯลฯ และความสามารถต่าง ๆ เหล่านี้ได้ถูกนำมาประยุกต์ใช้งานหลากหลายประเภท เช่น การวิเคราะห์การแพร่ของโรคระบาดติดตาม ตรวจสอบผลการจัดเก็บภาษี การคำนวณปริมาตรการกักเก็บน้ำของพื้นที่ ฯลฯ

การกำหนดจำนวนจุดกริดของการสุ่ม

การกำหนดจำนวนจุดกริดของของการสุ่มนั้นไม่มีกฎเกณฑ์หรือรูปแบบที่แน่นอน ผู้วิจัยจะต้องเป็นผู้กำหนดจำนวนจุดกริดขึ้นเองโดยคำนึงถึงปัจจัยต่าง ๆ ดังนี้

1. ความถูกต้องของงานวิจัย

สำหรับงานวิจัยที่ต้องการความถูกต้องแม่นยำมากควรใช้กลุ่มตัวอย่างให้ครอบคลุมประชากรทั่วทั้งแผนที่ โดยการเพิ่มจำนวนจุดกริดให้มากขึ้น เพื่อให้ความหนาแน่นของการสุ่ม (Sampling density) มีค่าเพิ่มขึ้นตาม เนื่องจากความถูกต้องแม่นยำมีค่าแปรผันตรงตามความหนาแน่นของการสุ่ม

2. งบประมาณ

ในการเก็บรวบรวมข้อมูล พนักงานสำรวจจะต้องลงพื้นที่เพื่อสัมภาษณ์ สังเกตหรือสอบถามกับกลุ่มตัวอย่าง ทำให้เกิดค่าใช้จ่ายต่าง ๆ เช่น ค่าเดินทาง ค่าอาหาร ค่าที่พัก ฯลฯ ถ้าใช้จำนวนจุดกริดมากค่าใช้จ่ายต่าง ๆ ก็จะเพิ่มสูงขึ้นด้วย

3. ระยะเวลา

ในส่วนของระยะเวลาในการทำวิจัยมีความสัมพันธ์กับจำนวนจุดกริดคล้ายคลึงกับงบประมาณ กล่าวคือ ถ้าใช้จำนวนจุดกริดคลายคลึงกับงบประมาณ กล่าวคือ ถ้าใช้จำนวนจุดกริดจะทำให้ใช้ระยะเวลาเวลาในการวิจัยยาวนาน แต่ในกรณีงานวิจัยที่มีงบประมาณมาก อาจแก้ไขโดยการจ้างพนักงานสำรวจเพิ่มเติม

4. พนักงานสำรวจ

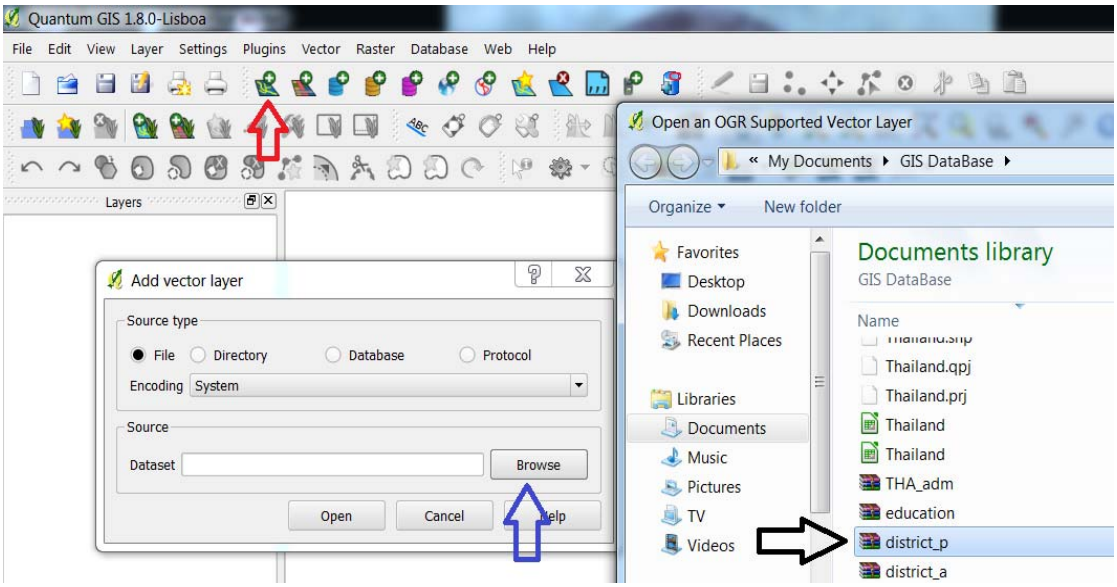
ถ้ามีพนักงานสำรวจหลายคน ผู้วิจัยสามารถกำหนดจำนวนจุดกริดให้มีจำนวนมากได้ เนื่องจากผู้วิจัยสามารถสั่งการให้พนักงานสำรวจแยกย้ายกันไปเก็บข้อมูล ตามแต่ละพื้นที่ต่าง ๆ ได้ แต่ผู้วิจัยต้องคำนึงถึงความรู้ความสามารถ ประสบการณ์ รวมทั้งคุณธรรมของพนักงานสำรวจ

วิธีการสุ่มแบบกริดมีแนวคิดคล้ายกับการสุ่มตัวอย่างแบบแบ่งกลุ่มประเภทการสุ่มตัวอย่างแบบพื้นที่ (Area random sampling) ซึ่งวิธีการสุ่มตัวอย่างแบบพื้นที่ทำการแบ่งประชากรออกเป็นกลุ่มตามสภาพทางภูมิศาสตร์ ได้แก่ จังหวัด อำเภอ เขตพื้นที่การศึกษา เป็นต้น จากนั้นทำการสุ่มตัวอย่างบางพื้นที่ เพื่อเก็บรวบรวมข้อมูล ซึ่งหลักการนี้เหมือนกันกับวิธีการสุ่มแบบกริด โดยวิธีการสุ่มแบบกริดจะสุ่มตำแหน่งจุดกริดและแบ่งแผนที่ออกเป็นพื้นที่ N_{ij} ดังนั้นการหาความเหมาะสมของจำนวนจุดกริดทั้งหมดต่อแผนที่จึงต้องพิจารณาขนาดของพื้นที่ N_{ij} ให้มีขนาดใกล้เคียงกับพื้นที่ที่ได้จากวิธีการสุ่มตัวอย่างแบบพื้นที่ ซึ่งเป็นไปตามเนื้อหาหรือเรื่องที่จะวิจัย ยกตัวอย่างเช่น ถ้าสมมติให้วิธีการสุ่มตัวอย่างแบบพื้นที่ใช้เขตพื้นที่การศึกษาระดับมัธยมศึกษาจำนวน 42 เขต เป็นหน่วยการสุ่ม ดังนั้นต้องใช้วิธีการสุ่มแบบกริดจำนวน 42 จุดกริด

หมายเหตุ การกำหนดจำนวนจุดกริด ไม่มีผลต่อขนาดของกลุ่มตัวอย่าง แต่จะส่งผลกระทบต่อการจัดสรรกลุ่มตัวอย่าง

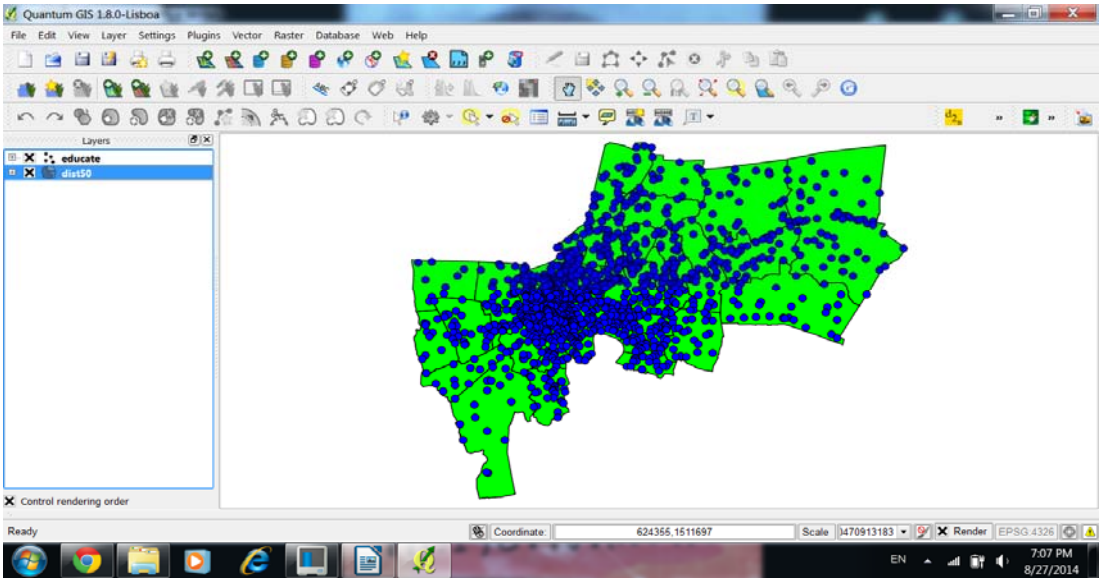
วิธีการสร้างจุดกริดบนแผนที่

การกำหนดระบบพิกัดในวิธีการสุ่มแบบกริดเหมือนกันกับการกำหนดระบบพิกัดภูมิศาสตร์ โดยใช้เส้นละติจูดเป็นเส้นแนวนอนและใช้เส้นลองติจูดเป็นเส้นแนวตั้ง ซึ่งจุดตัดระหว่างสองแนวเส้นนั้นคือจุดกริด ซึ่งการกำหนดตำแหน่งจุดกริดต้องอาศัยทั้งศาสตร์และศิลป์ เนื่องจากไม่มีวิธีการที่แน่นอน เป็นเพียงแต่แนวคิด ด้วยเหตุนี้ทำให้วิธีการสุ่มแบบจุดกริดเป็นวิธีการที่มีความยืดหยุ่นสูง สามารถปรับเปลี่ยนให้เข้ากับรูปร่างของแผนที่ได้ นอกจากนี้ ผู้วิจัยสามารถเลือกจำนวนจุดกริดได้ตามความต้องการ ดังนั้นวิธีการสุ่มแบบนี้จึงเหมาะกับการใช้งานจริง ซึ่งการสร้างระบบพิกัดมีวิธีการ ดังนี้



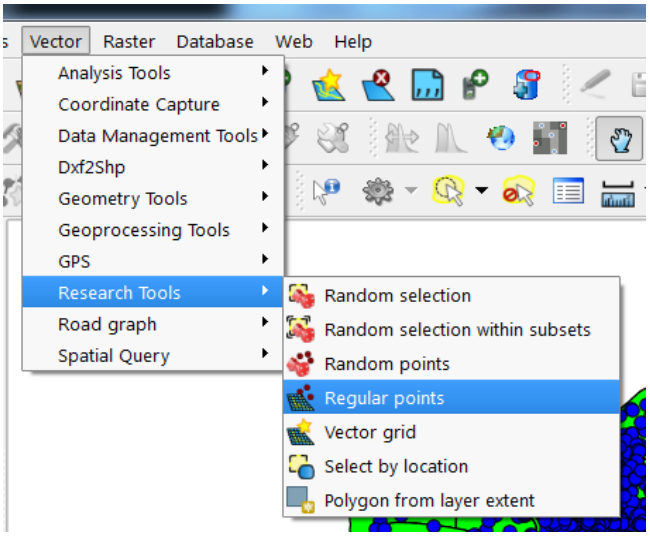
ภาพที่ 2-4 การเพิ่มชั้นข้อมูลเชิงเส้น

1. เมื่อเปิดโปรแกรม Quantum GIS 1.8.0-Lisboa (ซึ่งสามารถดาวน์โหลดได้ฟรีในเว็บไซต์ QGIS) จะปรากฏหน้าต่างขึ้นมาดังภาพที่ 2-4 ให้เลือกเครื่องมือเพิ่มชั้นข้อมูลเชิงเส้น (ตามลูกศรสีแดง) จากนั้นจะมีหน้าต่างขึ้นมาอีกครั้ง ให้เลือกแสดง (ตามลูกศรสีน้ำเงิน) แล้วทำการเลือกรายการชั้นข้อมูล (ตามลูกศรสีดำ) เมื่อเรียกข้อมูลทั้งหมดที่ต้องใช้เรียบร้อยแล้ว โปรแกรมจะแสดงแผนที่และตำแหน่งที่ตั้ง ตามภาพที่ 2-5 เป็นแผนที่กรุงเทพมหานครและตำแหน่งที่ตั้งสถานศึกษา



ภาพที่ 2-5 การแสดงผลลัพธ์แผนที่และตำแหน่งที่ตั้ง

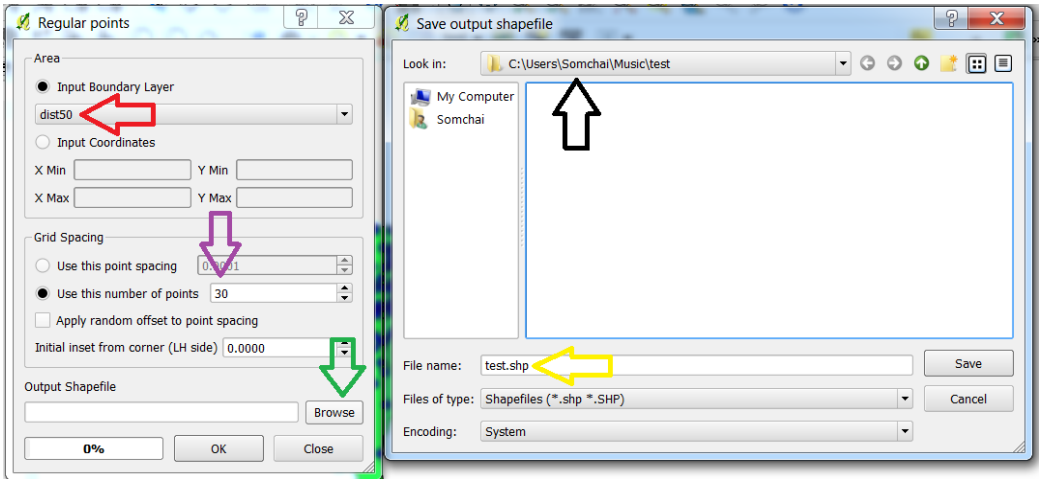
2. เลือกเครื่องมือ Vector > Research tools > Regular points เช่นเดียวกันกับภาพที่ 2-6



ภาพที่ 2-6 การใช้งานฟังก์ชั่น Regular points

3. เมื่อเลือกฟังก์ชั่น Regular points แล้วจะมีหน้าต่างชื่อ Regular points ปรากฏขึ้นมา ซึ่งการกำหนดขอบเขตของจุดกริดในส่วนของ Area นั้น ให้เลือก Input boundary layer และเลือก

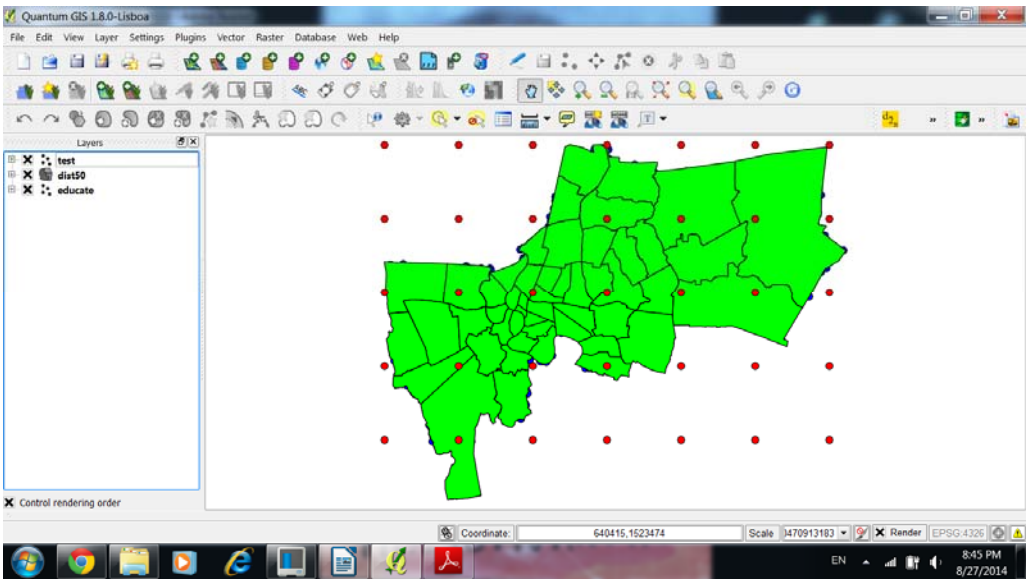
รายการแผนที่ ซึ่งกรณีตัวอย่างนี้คือแผนที่กรุงเทพมหานคร (dist50) (ตามลูกศรสีแดง) สำหรับ Grid spacing เป็นการกำหนดขนาดของกริดเซลล์ ให้เลือกจำนวนของจุดกริดและป้อนจำนวนจุดกริดตามต้องการ(ตามลูกศรสีม่วง) ดังภาพที่ 2-7



ภาพที่ 2-7 การใช้โปรแกรม QGIS สุ่มจุดกริด

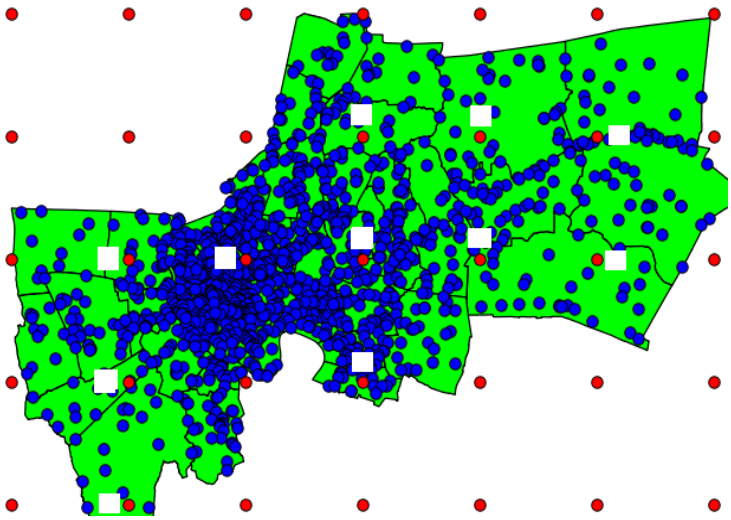
หมายเหตุ จำนวนจุดกริดที่ป้อนนั้น ต้องมีจำนวนมากกว่าจำนวนจุดกริดที่หามาจากปัจจัยในการกำหนดจุดกริด เพราะจุดกริดบางตำแหน่งจะอยู่นอกแผนที่ จึงจำเป็นต้องมีการเผื่อไว้ เช่น ถ้าผู้วิจัยวางแผนไว้ว่าจะใช้จุดกริด 10-15 จุดกริด อาจลองใช้ 20, 25, 30 และ 35 จุดกริด

4. ตามภาพที่ 2-7 ให้เลือก Browse เพื่อเลือกที่อยู่ที่จะบันทึกโครงการ (ตามลูกศรสีดำ) และระบุชื่อโครงการ (ตามลูกศรสีเหลือง) หลังจากที่เลือก Save และ OK แล้ว โปรแกรมจะแสดงตำแหน่งจุดกริด ดังภาพที่ 2-8



ภาพที่ 2-8 ตำแหน่งจุดกริดทั้งหมด

5. ในขั้นตอนสุดท้ายคือการเลือกจุดกริดที่อยู่บนพื้นที่เป้าหมาย เพื่อใช้เป็นจุดอ้างอิงในการสุ่มกลุ่มตัวอย่าง ซึ่งในการเก็บรวบรวมข้อมูลในกรณีนี้ได้สมมุติไว้ว่าผู้วิจัยกำหนดจุดกริดไว้ 10-15 ตำแหน่ง ดังนั้นจุดกริดที่อยู่ในเขตกรุงเทพ ซึ่งถูกเลือกมา 11 จุดกริด (นั้นคือจุดกริดที่อยู่ใกล้กับสี่เหลี่ยมสีขาว) มีตำแหน่งดังภาพที่ 2-9



ภาพที่ 2-9 ตำแหน่งจุดกริดที่ใช้ในการเก็บรวบรวมข้อมูล

การสุ่มตัวอย่างในแต่ละจุดกริด

หลังจากกำหนดจำนวนจุดกริดและสร้างจุดกริดบนแผนที่เรียบร้อยแล้ว ขั้นต่อไป คือ การหาขนาดกลุ่มตัวอย่างทั้งหมดและจัดสรรตัวอย่างในแต่ละจุดกริด ในขั้นตอนสุดท้าย คือ การเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างที่อยู่ตรงจุดกริดและบริเวณโดยรอบ ซึ่งมีรายละเอียดขั้นตอนการปฏิบัติ ดังนี้

1. การหาขนาดของกลุ่มตัวอย่าง

ผลการวิจัยที่ได้นั้นจะมีความน่าเชื่อถือหรือไม่ขึ้นอยู่กับองค์ประกอบหลายอย่าง การกำหนดขนาดตัวอย่าง (Sample size) ให้มีความเหมาะสมเป็นสิ่งสำคัญประการหนึ่งในการตัดสินความถูกต้องของงานวิจัย ถ้าใช้ขนาดกลุ่มตัวอย่างที่มากเกินไปจะทำให้สิ้นเปลืองทรัพยากรในด้านต่างเช่น งบประมาณ กำลังคน ระยะเวลาในการเก็บข้อมูล ฯลฯ แต่ถ้าใช้ขนาดกลุ่มตัวอย่างที่น้อยเกินไปจะทำให้ข้อมูลที่เก็บรวบรวมมาได้ไม่มีความเป็นตัวแทนที่ดีของประชากร เพื่อให้ได้จำนวนหน่วยตัวอย่างที่มีความเหมาะสมและน่าเชื่อถือจึงจำเป็นต้องนำปัจจัยต่าง ๆ ที่เกี่ยวข้อง เช่น ความแปรปรวนของประชากร ขนาดประชากร ระดับความเชื่อมั่น และค่าความคลาดเคลื่อน เป็นต้น มาพิจารณาร่วมกันในการกำหนดขนาดกลุ่มตัวอย่าง การกำหนดขนาดตัวอย่างของวิธีการสุ่มแบบกริดมีรูปแบบเหมือนกับการหาขนาดตัวอย่างด้วยวิธีการสุ่มแบบอื่น ๆ ซึ่งการหาขนาดตัวอย่างมีหลายวิธี เช่น การกำหนดเกณฑ์ ตารางสำเร็จรูปของทาโร ยามาเน่ ตารางสำเร็จรูปของเครจซี่และมอร์แกน การประมาณค่าประชากรด้วยสูตร และการวิเคราะห์อำนาจการทดสอบ โดยมีรายละเอียดดังนี้

1.1 การกำหนดเกณฑ์ ในกรณีที่ทราบจำนวนประชากรที่แน่นอนสามารถใช้เกณฑ์ในการกำหนดร้อยละของประชากร ถ้าขนาดประชากรเป็นหลักร้อย ต้องใช้กลุ่มตัวอย่างอย่างน้อย 25% ถ้าขนาดประชากรเป็นหลักพัน ต้องใช้กลุ่มตัวอย่างอย่างน้อย 10% ถ้าขนาดประชากรเป็นหลักหมื่น ต้องใช้กลุ่มตัวอย่างอย่างน้อย 5% ถ้าขนาดประชากรเป็นหลักแสน ต้องใช้กลุ่มตัวอย่างอย่างน้อย 1% (ธีรวุฒิ เอกะกุล, 2543)

1.2 ตารางสำเร็จรูปของทาโร ยามาเน่ เป็นตารางที่ใช้หาขนาดของกลุ่มตัวอย่างเพื่อประมาณค่าสัดส่วนของประชากร โดยคาดว่าสัดส่วนของลักษณะที่สนใจในประชากร เท่ากับ 0.5 และระดับความเชื่อมั่น 95% (ธีรวุฒิ เอกะกุล, 2543)

ตารางที่ 2-1 ตารางของทาโร ยามาเน่ (มารยาท โยทองยศ และปราณี สวัสดิสรรพ์, 2551)

ขนาดประชากร	ขนาดของกลุ่มตัวอย่างที่ขนาดความคลาดเคลื่อน (e)					
	± 1%	± 2%	± 3%	± 4%	± 5%	± 10%
500					222	83
1,000				385	286	91
1,500			638	441	316	94
2,000			714	476	333	95
2,500		1250	769	500	345	96
3,000		1364	811	517	353	97
3,500		1458	843	530	359	97
4,000		1538	870	541	364	98
4,500		1607	891	549	367	98
5,000		1667	909	556	370	98
6,000		1765	938	566	375	98
7,000		1842	959	574	378	99
8,000		1905	976	580	381	99
9,000		1957	989	584	383	99
10,000	5000	2000	1000	588	385	99
15,000	6000	2143	1034	600	390	99
20,000	6667	2222	1053	606	392	100
25,000	7143	2273	1064	610	394	100
50,000	8333	2381	1087	617	397	100
100,000	9091	2439	1099	621	398	100
∞	10000	2500	1111	625	400	100

* หมายถึง ขนาดตัวอย่างไม่เหมาะสมที่จะ Assume ให้เป็นการกระจายแบบปกติ จึงไม่สามารถใช้สูตรคำนวณขนาดของกลุ่มตัวอย่างได้

1.3 ตารางสำเร็จรูปของ เครจซี่และมอร์แกน เป็นตารางนี้ใช้ในการประมาณค่าสัดส่วนของประชากร และกำหนดให้สัดส่วนของลักษณะที่สนใจในประชากร เท่ากับ 0.5 ระดับความคลาดเคลื่อนที่ยอมรับได้ 5% และระดับความเชื่อมั่น 95% (ธีรวุฒิ เอกะกุล, 2543)

ตารางที่ 2-2 ตารางของเครจซี่และมอร์แกน (มารยาท โยทองยศ และปราณี สวัสดิสรรพ์, 2551)

ขนาดประชากร	ขนาดตัวอย่าง	ขนาดประชากร	ขนาดตัวอย่าง	ขนาดประชากร	ขนาดตัวอย่าง	ขนาดประชากร	ขนาดตัวอย่าง	ขนาดประชากร	ขนาดตัวอย่าง
10	10	100	80	280	162	800	260	2,800	338
15	14	110	86	290	165	850	265	3,000	341
20	19	120	92	300	169	900	269	3,500	346
25	24	130	97	320	175	950	274	4,000	351
30	28	140	103	340	181	1,000	278	4,500	354
35	32	150	108	360	186	1,100	285	5,000	357
40	36	160	113	380	191	1,200	291	6,000	361
45	40	170	118	400	196	1,300	297	7,000	364
50	44	180	123	420	201	1,400	302	8,000	367
55	48	190	127	440	205	1,500	306	9,000	368
60	52	200	132	460	210	1,600	310	10,000	370
65	56	210	136	480	214	1,700	313	15,000	375
70	59	220	140	500	217	1,800	317	20,000	377
75	63	230	144	550	226	1,900	320	30,000	379
80	66	240	148	600	234	2,000	322	40,000	380
85	70	250	152	650	242	2,200	327	50,000	381
90	73	260	155	700	248	2,400	331	75,000	382
95	76	270	159	750	254	2,600	333	100,000	384

1.4 การประมาณค่าประชากรด้วยสูตร

สูตรการหาขนาดกลุ่มตัวอย่าง เพื่อประมาณค่าประชากร ในกรณีที่ทราบจำนวนประชากร (มารยาท โยทองยศ และปราณี สวัสดิสรรพ์, 2551)

$$n = \frac{N\sigma^2}{\frac{Ne^2}{Z^2} + \sigma^2}$$

n คือ จำนวนตัวอย่าง

N คือ จำนวนประชากร

σ^2 คือ ความแปรปรวนของประชากร

e คือ ความคลาดเคลื่อน

Z คือ คะแนนมาตรฐานที่เกี่ยวข้องกับระดับความเชื่อมั่น (จากตารางโค้งปกติ)

1.5 การวิเคราะห์อำนาจการทดสอบ

การกำหนดขนาดตัวอย่างสามารถหาได้จากการวิเคราะห์อำนาจการทดสอบ ในปี ค.ศ. 1977 Cohen ได้พัฒนาชุดตารางสำเร็จที่ใช้ในการวิเคราะห์หาอำนาจการทดสอบ แต่เนื่องจากการใช้งานตารางสำเร็จรูปของ Cohen มีความยุ่งยากและไม่สะดวกในการใช้งาน ดังนั้น ในปัจจุบันมหาวิทยาลัย Heinrich-Heine-Universität ได้สร้างโปรแกรมโปรแกรม G*Power 3.1.9.2 for Windows ซึ่งเป็นโปรแกรมที่พัฒนามาจากชุดตารางสำเร็จที่ใช้สำหรับการทดสอบค่าสถิติหลายวิธีของ Cohen เพื่อให้นักศึกษาและประชาชนได้นำไปใช้กับงานวิจัยของตนเอง

2. การจัดสรรตัวอย่างตามสัดส่วน

การจัดสรรตัวอย่างตามสัดส่วนสำหรับการสุ่มแบบกริดคือการแบ่งจำนวนของกลุ่มอย่าง (ที่หาได้จากการหาขนาดกลุ่มตัวอย่าง) ออกเป็นส่วน ๆ ตามจำนวนจุดกริด เพื่อให้นักวิจัยใช้ในการวางแผนการเก็บรวบรวมข้อมูล และเป็นแนวทางการปฏิบัติงานสนามของผู้ช่วยนักวิจัย ทำให้ผู้ช่วยนักวิจัยรับรู้ได้ว่าจะต้องเก็บข้อมูลจากกลุ่มตัวอย่างที่จุดกริดนั้น ๆ จำนวนเท่าใด ซึ่งการจัดสรรจำนวนตัวอย่างให้กับจุดกริดต้องพิจารณาถึงขนาดของประชากรในพื้นที่เป็นสำคัญ ถ้าพื้นที่ใดที่มีจำนวนประชากรมาก ในพื้นที่นั้นต้องใช้กลุ่มตัวอย่างขนาดใหญ่ ถ้าพื้นที่ใดที่มีจำนวนประชากรเบาบาง ในพื้นที่นั้นต้องใช้กลุ่มตัวอย่างขนาดเล็ก ดังนั้นจุดกริดใดที่ขนาดประชากรในพื้นที่ (N_{ij}) มากก็ควรจัดสรรจำนวนตัวอย่าง (n_{ij}) ให้มาก แต่จุดกริดที่มีขนาดประชากรในพื้นที่ (N_{ij}) น้อยก็ควรจัดสรรจำนวนตัวอย่าง (n_{ij}) ให้น้อย โดยคิดเทียบตามหลักอัตราส่วนดังนี้การจัดสรรตามสัดส่วน

$$\frac{n_{ij}}{N_{ij}} = \frac{n}{N}$$

$$n_{ij} = \frac{n}{N} N_{ij}$$

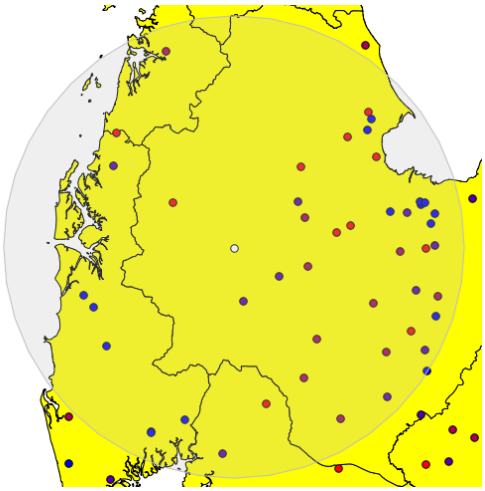
ในการหาค่า N_{ij} นั้นกระทำได้ยาก เนื่องจากพื้นที่จะถูกตัดออกเป็นส่วน ๆ อย่างละไม่เท่ากัน นอกจากนี้การกระจายตัวของประชากรเป็นลักษณะไม่เป็นระเบียบแบบแผน จึงทำให้ไม่สามารถคาดเดาได้ว่า พื้นที่ ๆ ถูกแบ่งนั้นมีประชากรเหลือเท่าใด เพื่อเป็นการลดความซับซ้อนและยุ่งยากในการวิจัยจึงจำเป็นต้องหาขนาดที่จุดกริด (i, j) แบบกะประมาณ โดยใช้วงกลมการสุ่ม ซึ่งมีหลักการ ดังนี้

การกะประมาณขนาดกลุ่มตัวอย่างที่จุดกริด (i, j)

$$n_{i,j} \approx n'_{i,j} = \frac{n}{N} N'_{i,j}$$

N'_{ij} คือ ขนาดของประชากรในวงกลมการสุ่ม

n'_{ij} คือ ขนาดของกลุ่มตัวอย่างที่ได้จาก N'_{ij}



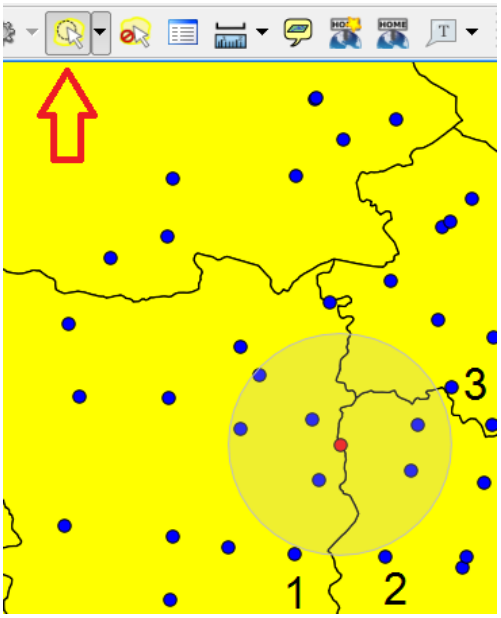
ภาพที่ 2-10 การขยายรัศมี เพื่อหาค่า N'_{ij}

จากภาพที่ 2-10 แสดงตำแหน่งจุดกริด i = 2, j = 3 (จุดกริดสีขาว) ซึ่งอยู่ในจังหวัดสุราษฎร์ธานี ถ้ากำหนดให้ประชากร คือ โรงเรียน จำนวน 2584 โรงเรียนทั่วประเทศไทย และต้องสุ่มตัวอย่างเท่ากับ 335 โรงเรียน เมื่อใช้จุดกริดนี้เป็นตำแหน่งศูนย์กลางวงกลมและขยายรัศมีวงกลมเท่ากับระยะทางครึ่งหนึ่งของระยะห่างระหว่างจุดกริด 2 จุดกริดที่ติดกัน จะได้กลุ่มตัวอย่างเท่ากับ 47 โรงเรียน ดังนั้น สามารถกะประมาณขนาดกลุ่มตัวอย่างที่จุดกริดนี้ คือ (335/ 2584) x 47 เท่ากับ 6 โรงเรียน แต่วิธีการจัดสรรแบบนี้ เมื่อนำขนาดกลุ่มตัวอย่างในแต่ละจุดกริดมารวมกัน อาจมากกว่า

หรือน้อยกว่าขนาดกลุ่มตัวอย่างที่วางแผนไว้ (ซึ่งคำนวณได้จากหัวข้อการหาขนาดของกลุ่มตัวอย่าง) จึงต้องทำการปรับเกลี่ยขนาดกลุ่มตัวอย่างในภาพรวมให้เท่ากับ 335 โรงเรียน แต่ในการวิจัยครั้งนี้ (ขั้นตอนที่ 2 กรณีศึกษาคะแนน O-NET ปี พ.ศ. 2548) เป็นการเปรียบเทียบประสิทธิภาพการสุ่มแบบกริด การสุ่มแบบกริดชั้นภูมิและการสุ่มแบบหลายขั้นตอน และเพื่อควบคุมตัวแปรแทรกซ้อนในการวิจัยจึงต้องตั้งค่าเริ่มต้นต่าง ๆ ให้เท่ากัน ดังนั้นขนาดตัวอย่างในแต่ละภูมิภาคของวิธีการสุ่มทั้งสามวิธีต้องเท่ากัน ทำให้ต้องจัดสรรตามสัดส่วนในลักษณะภูมิภาค ยกตัวอย่างเช่น เมื่อคำนวณการสุ่มแบบหลายขั้นตอนในภาคใต้ได้ขนาดกลุ่มตัวอย่าง 41 โรงเรียน ดังนั้นขนาดกลุ่มตัวอย่างของวิธีการสุ่มแบบกริด (หรือการสุ่มแบบกริด-ชั้นภูมิ) ในภาคใต้ต้องเท่ากับ 41 โรงเรียนด้วย และถ้าจุดกริดในภาคใต้มี 3 ตำแหน่ง ดังนั้นที่จุดกริด i = 2, j = 3 ต้องสุ่มโรงเรียนเท่ากับ 41/ 3 เท่ากับ 13 โรงเรียน แต่สำหรับงานวิจัยบางเรื่องที่ประชากรที่มีจำนวนไม่จำกัด (Infinite population) ทำให้ไม่สามารถจัดทำบัญชีรายชื่อ/ จัดเตรียมกรอบตัวอย่างหรือจัดทำข้อมูลเชิงคุณลักษณะ (Attribute data) เพื่อใช้สำหรับโปรแกรม อาจต้องใช้การจัดสรรตัวอย่างอย่างเท่าเทียมกัน (Equal allocation)

3. การใช้วงกลมการสุ่มในการสุ่มตัวอย่าง

ขั้นตอนสุดท้ายของการสุ่มตัวอย่างในแต่ละจุดกริด คือ การการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างที่อยู่ตรงจุดกริดและบริเวณโดยรอบจุดกริด ซึ่งเป็นงานภาคสนามของผู้ช่วยนักวิจัยหรือเจ้าหน้าที่รวบรวมข้อมูล แต่ต้องกำหนด เปิดโปรแกรม QGIS และเลือกเมนูแสดงรายละเอียดแผนที่ดังภาพที่ 2-11 (ตามลูกศรสีแดง) นำเคอร์เซอร์เมาส์ไปยังตำแหน่งจุดกริด เพื่อให้จุดกริดนั้นกลายเป็นจุดศูนย์กลางของวงกลม จากนั้นขยายรัศมีวงกลมออกไปให้ครอบคลุมกลุ่มตัวอย่างจนได้ขนาดกลุ่มตัวตามที่ต้องการ (ซึ่งเป็นไปผลลัพธ์ที่ได้จากหัวข้อการจัดสรรตัวอย่างตามสัดส่วน) แต่ในบางครั้งอาจมีสมาชิกหรือหน่วยตัวอย่างบางตัวที่ก้ำกึ่งว่าจะอยู่ภายในรัศมีที่กำหนดหรือไม่ยกตัวอย่าง เช่น จุดกริดตามภาพที่ 2-11 ถ้าต้องการสุ่มตัวอย่างที่จุดกริดนี้ 7 โรงเรียน ซึ่งภายในรัศมีวงกลมมีกลุ่มตัวอย่างเท่ากับ 6 โรงเรียน แต่มีโรงเรียนที่มีระยะห่างจากจุดกริดใกล้เคียงกัน 3 โรงเรียน คือ โรงเรียนหมายเลข 1 2 และ 3 เป็นต้น ในส่วนการแก้ปัญหาสมาชิกที่มีความก้ำกึ่งแบบกรณีนี้ให้ใช้ปุ่มฟังก์ชั่น Feature by radius (ตามลูกศรสีแดง) ค่อย ๆ ขยายรัศมีวงกลมการสุ่มจนกระทั่งแตะขอบโรงเรียนเพียงแค่โรงเรียนเดียว ซึ่งตามรูปภาพโรงเรียนหมายเลข 2 คือ โรงเรียนที่อยู่



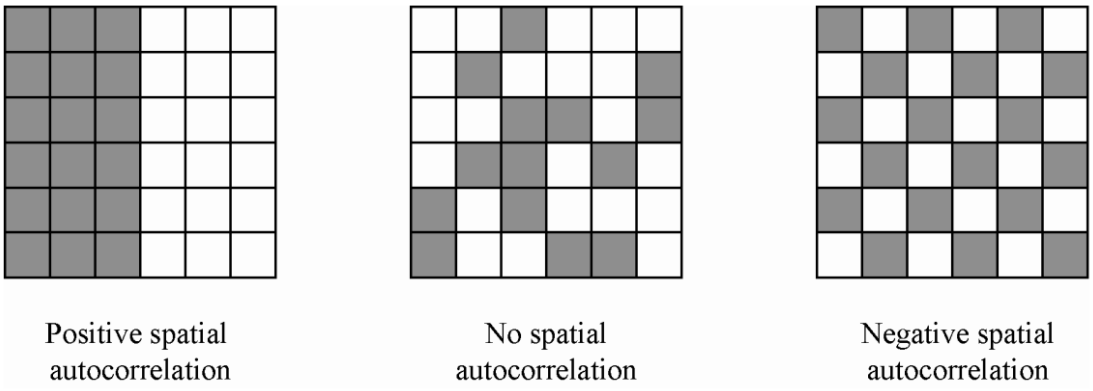
ภาพที่ 2-11 การหากลุ่มตัวอย่างด้วยวงกลมการสุ่ม

6. ข้อตกลงเบื้องต้นของวิธีการสุ่มแบบกริด

ข้อตกลงเบื้องต้นเป็นกรอบบังคับหรือปัจจัยพื้นฐานบางประการในการทำวิจัย ยกตัวอย่าง เช่น สถิติทดสอบ **Z**-test หรือ การวิเคราะห์ความแปรปรวน มีข้อตกลงเบื้องต้น ได้แก่ 1) กลุ่มตัวอย่างได้มาโดยการสุ่ม 2) การแจกแจงของประชากรเป็นโค้งปกติ (Normal distribution) 3) ข้อมูลอยู่ในมาตราอันตรภาค (Interval scale) ขึ้นไป เป็นต้น ซึ่งเป็นการกำหนดเงื่อนไขของการทำวิจัย ในส่วนการสุ่มแบบกริดได้นำระบบพิกัดภูมิศาสตร์มาประยุกต์ใช้ในการสุ่มตัวอย่าง ดังนั้น พื้นที่ในการสุ่มควรต้องมีขนาดใหญ่ เช่น แผนที่จังหวัดหรือแผนที่ประเทศ เป็นต้น วิธีการสุ่มแบบกริดจึงไม่มีความเหมาะสมกับพื้นที่ขนาดเล็ก เช่น การสุ่มตัวอย่างนักเรียนภายในห้องเรียน เป็นต้น นอกจากนั้นธรรมชาติของประชากรบนแผนที่เป็นอีกสิ่งหนึ่งที่มีสำคัญในการสุ่มแบบกริด เนื่องจากถ้าโปรแกรม QGIS สุ่มตำแหน่งของจุดกริดได้ในพื้นที่ที่เป็นประชากรกลุ่มย่อย (Subpopulation) จะทำให้กลุ่มตัวอย่างที่ได้จากจุดกริดนั้นไม่เป็นตัวแทนของ N_{ij} และเมื่อนำข้อมูลจากกลุ่มตัวอย่างนี้มาวิเคราะห์ผลจะทำให้เกิดความผิดพลาด ตลอดจนการนำเสนอข้อมูลที่ไม่ถูกต้อง ดังนั้นวิธีการสุ่มแบบกริดมีจึงเหมาะกับธรรมชาติของประชากรที่มีค่าสังเกตที่มีการเปลี่ยนแปลงแบบต่อเนื่องอย่างค่อยเป็นค่อยไป ซึ่งการกระจายตัวของค่าตัวแปรในแผนที่แบบนี้ ในสาขาวิชาภูมิสารสนเทศศาสตร์เรียก “สหสัมพันธ์เชิงพื้นที่บวก”

สหสัมพันธ์เชิงพื้นที่ คือ ความสัมพันธ์ระหว่างค่าของตัวแปรที่ตำแหน่งหนึ่งกับค่าตัวแปรที่ตำแหน่งใกล้เคียงในระบบพื้นผิวสองมิติ (2-D) สหสัมพันธ์เชิงพื้นที่เป็นปรากฏการณ์ทางธรรมชาติ ซึ่งในโลกความเป็นจริงปรากฏการณ์เหล่านี้มีรูปแบบ (แผนที่) ที่แน่นอนและเป็นระบบมากกว่าการเกิดขึ้นแบบสุ่ม (Griffith, 2009, p. 1) ในปี ค.ศ. 1950 Pat Moran ได้ค้นพบทฤษฎี Moran's I ซึ่งทฤษฎีนี้เป็นการอธิบายถึงลักษณะความสัมพันธ์ระหว่างค่าตัวแปรและพิกัดของตำแหน่ง โดยค่า I หมายถึง สัมประสิทธ์ของสหสัมพันธ์เชิงพื้นที่ (Spatial autocorrelation) และค่า I ที่เป็นไปได้มีค่าตั้งแต่ -1 จนถึง 1 Moran ได้แบ่งรูปแบบของสหสัมพันธ์เชิงพื้นที่ออกเป็น 3 รูปแบบ ได้แก่ สหสัมพันธ์เชิงพื้นที่ทางบวก, สหสัมพันธ์เชิงพื้นที่ทางลบ และไม่มีสหสัมพันธ์เชิงพื้นที่ (Moran, 1950)

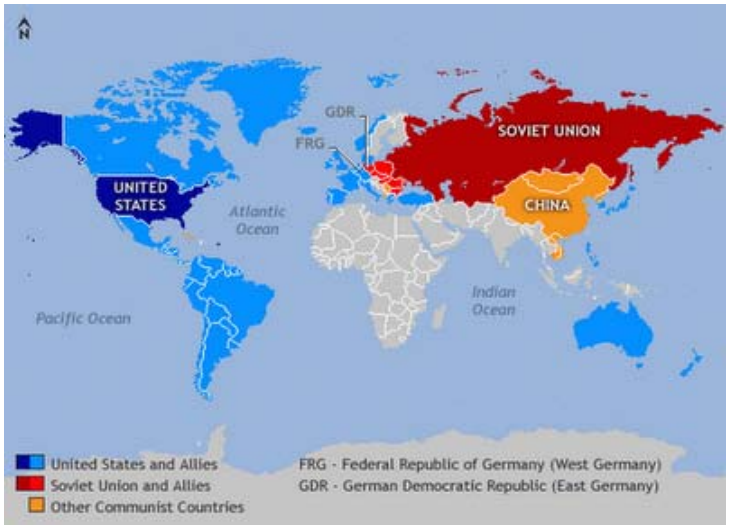
สหสัมพันธ์เชิงพื้นที่ทางบวก หมายถึง ค่าตัวแปรมีการกระจายตัวแบบแบ่งกลุ่ม พื้นที่ใดที่มีค่าตัวแปรมากต้องรวมกลุ่มกับพื้นที่ที่มีค่าตัวแปรมากเหมือนกัน ถ้าพื้นที่ใดที่มีค่าตัวแปรน้อยบริเวณใกล้เคียงจะมีค่าตัวแปรน้อยตามกัน ในส่วนสหสัมพันธ์เชิงพื้นที่ทางลบหมายถึงค่าตัวแปรมีการกระจายตัวแบบมีระบบ ซึ่งพื้นที่ที่มีค่าตัวแปรมากต้องอยู่ใกล้กับพื้นที่ที่มีค่าตัวแปรน้อย ในทางกลับกันพื้นที่ใดที่มีค่าตัวแปรน้อยต้องอยู่ใกล้กับพื้นที่ที่มีค่าตัวแปรมาก ดังนั้น แผนที่ของสหสัมพันธ์เชิงพื้นที่ทางลบจึงมีการกระจายตัวแบบสลับกันไปมาเหมือนตารางหมากฮอส สำหรับสหสัมพันธ์เชิงพื้นที่เป็นศูนย์หมายถึงค่าตัวแปรมีการกระจายตัวแบบสุ่ม ซึ่งไม่มีรูปแบบและไม่มีความแน่นอน ดังภาพที่ 2-12



ภาพที่ 2-12 รูปแบบของสหสัมพันธ์เชิงพื้นที่ กรณีที่เป็นบวก ศูนย์ และลบ (Radil, 2011)

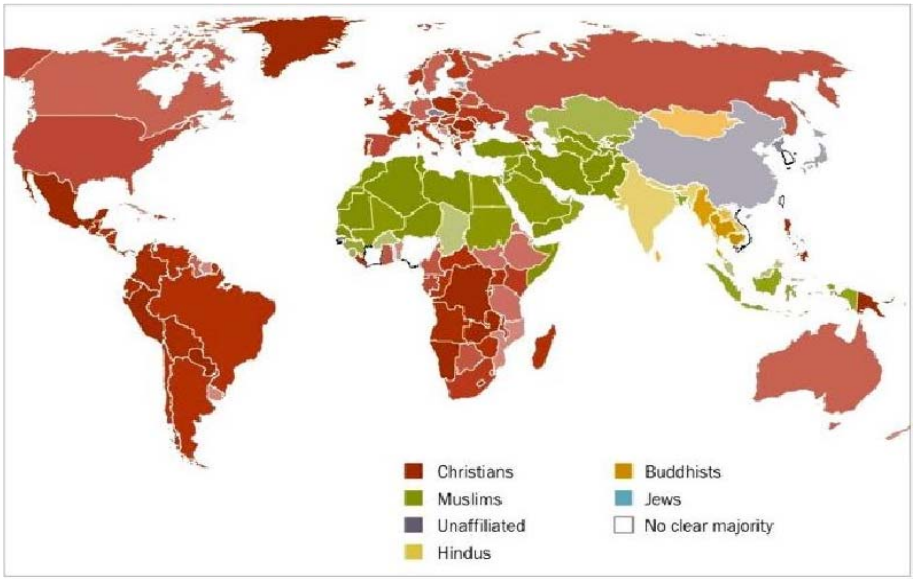
แนวคิดสหสัมพันธ์เชิงพื้นที่ทางบวกมีความคล้ายคลึงกับกฎข้อที่ 1 ของภูมิศาสตร์ Waldo Tobler ได้คิดค้นกฎข้อที่ 1 ของภูมิศาสตร์ [Tobler's first law (TFL)] ในปี ค.ศ. 1970 และมีใจความสำคัญว่า "ทุก ๆ สิ่งที่อยู่ใกล้กันย่อมมีความเกี่ยวข้องกัน มากกว่าสิ่งที่อยู่ไกลออกไป"

(“Everything is related to everything else, but near things are more related than distant things”) (Tobler, 1970) กฎข้อนี้เป็นกฎแห่งธรรมชาติและมีความเป็นพื้นฐานทางระบบสารสนเทศภูมิศาสตร์ ทำให้กฎข้อนี้ได้ถูกพัฒนาและต่อยอดในการประมาณค่าเชิงพื้นที่และการสร้างแบบจำลอง เพื่อก่อให้เกิดความเข้าใจได้ง่ายและชัดเจน ขอยกตัวอย่างแผนที่ดังนี้



ภาพที่ 2-13 แผนที่แสดงลัทธิความเชื่อทางการเมืองในยุคสงครามเย็น (Allison, 2008)

ภาพที่ 2-13 แสดงถึงการเผยแพร่ของลัทธิความเชื่อและอุดมการณ์ทางการเมือง รวมทั้งระบอบการเมือง ซึ่งเป็นการต่อสู้กันระหว่างกลุ่มประเทศ 2 กลุ่ม ฝ่ายหนึ่งคือ สหรัฐอเมริกาและกลุ่มพันธมิตรเรียกว่าค่ายตะวันตก ซึ่งปกครองด้วยระบอบเสรีประชาธิปไตย เนื่องจากเยอรมันตะวันตกอยู่ใกล้สหรัฐอเมริกา จึงมีแนวคิดระบอบเสรีประชาธิปไตย แต่อีกฝ่ายหนึ่ง คือ สหภาพโซเวียตเรียกว่าค่ายตะวันออก ซึ่งปกครองด้วยระบอบคอมมิวนิสต์ โดยประเทศจีนและเยอรมันตะวันออกที่อยู่ใกล้สหภาพโซเวียต ได้รับอิทธิพลความเชื่อจากสหภาพโซเวียต ทำให้จีนและเยอรมันตะวันออก ปกครองด้วยระบอบคอมมิวนิสต์ ในส่วนประเทศที่เป็นสีขาวคือกลุ่มประเทศไม่ฝักใฝ่ฝ่ายใด (Non-aligned movement) ซึ่งกลุ่มประเทศไม่ฝักใฝ่ฝ่ายใดนี้ได้ก่อตั้งขึ้นจากประเทศสมาชิกส่วนใหญ่ที่เป็นอาณานิคมของประเทศยุโรป (ยงยุทธ มัยลาภ, 2549) ซึ่งต้องการรวมกลุ่มกันเพื่อแสดงจุดยืนที่ไม่เอนเอียงฝักใฝ่ กับทั้งขั้วของสหรัฐฯหรือขั้วของสหภาพโซเวียต ประเทศไม่ฝักใฝ่ฝ่ายใดมุ่งสนับสนุนให้ประเทศสมาชิกสามารถกำหนดอนาคตของตนเองได้ สนับสนุนความเป็นอิสระของแต่ละประเทศ การรักษาไว้ซึ่งอธิปไตยแห่งรัฐ การไม่แทรกแซงในกิจการภายในของสมาชิก และการอยู่ร่วมกันโดยสันติของประเทศสมาชิก



ภาพที่ 2-14 แผนที่แสดงการนับถือศาสนาตามเขตพื้นที่ (Pew Research Religion & Public Life Project, 2012)

ภาพที่ 2-14 แสดงถึงการแพร่กระจายความเชื่อทางศาสนา การเผยแพร่ของศาสนาคริสต์มีความเจริญรุ่งเรืองอย่างมาก ในยุคล่าอาณานิคมของพวกจักรวรรดิ์นิยมชาวยุโรปประมาณคริสต์ศตวรรษที่ 16-19 ชาวยุโรปมีความศรัทธาต่อศาสนาคริสต์ และเชื่อมั่นว่า อารยธรรมของพวกตนดีกว่าชนชาติอื่น สำหรับชนชาติอื่นที่แตกต่างจากตนทั้งเชื้อชาติ ศาสนาและวัฒนาธรรม เป็นพวกล้าหลัง มิชชันนารีชาวยุโรปจำนวนนึงจึงยินดีในการช่วยเหลือชาวพื้นเมืองพร้อมกับการเผยแผ่ศาสนาในดินแดนที่ห่างไกลความเจริญ เพราะบาทหลวงและนักบุญมีความเชื่อว่าการช่วยเหลือให้คนเหล่านั้นได้เข้าถึงศรัทธาในพระเจ้าของตนและเรียนรู้อารยธรรมของชาวตะวันตก เป็นการสร้างกุศลที่ยิ่งใหญ่ให้แก่ผู้เผยแผ่ศาสนา ด้วยเหตุนี้ทำให้คริสต์ศาสนิกชนมีปริมาณเพิ่มมากขึ้นทั้งในทวีปอเมริกา แอฟริกา และออสเตรเลีย (มูลนิธิคริสต์ศาสนิกชนออร์โธดอกซ์ในประเทศไทย, ม.ป.ป.) ในทวีปอเมริกาใต้ อเมริกากลาง และบางส่วนของทวีปอเมริกาเหนือเป็นอาณานิคมของสเปน ทวีปอเมริกาเหนือเป็นอาณานิคมของ อังกฤษและฝรั่งเศส ทวีปแอฟริกาเป็นอาณานิคมของ อังกฤษ ฝรั่งเศส เยอรมนี และอิตาลี และทวีปออสเตรเลียเป็นอาณานิคม อังกฤษ ในส่วนศาสนาอิสลามมีถิ่นกำเนิดในคาบสมุทรอาหรับ ตั้งแต่คริสต์ศตวรรษที่ 7 จากนั้น ศาสนาอิสลามก็แพร่หลายออกไปสู่ดินแดนอื่น ๆ เช่น อินเดีย และเอเชียตะวันออกเฉียงใต้ ในประเทศอินเดียมีการนับถือศาสนาอิสลาม ตั้งแต่คริสต์ศตวรรษที่ 13 เป็นต้นมา (มหาวิทยาลัยเชียงใหม่, 2554) หลักฐานเกี่ยวกับการเผยแพร่ของศาสนาอิสลามในเอเชียตะวันออกเฉียงใต้คือ บันทึกของ มาร์โค โปโล (Marco

Polo) ในปี ค.ศ. 1292 มีการบันทึกว่าเมือง Perlak (เมือง Peurleuak) เป็นเมืองท่าที่มีการค้าขาย อยู่ตรงปากทางเข้าช่องแคบมะละกา จึงมีพ่อค้าชาวอินเดียซึ่งนับถือศาสนาอิสลามเดินทางเข้ามาค้าขาย มาร์โค โปโล กล่าวว่า พ่อค้ามุสลิมชาวอินเดียได้เข้ามาตั้งหลักแหล่งในเมืองนี้ และได้เผยแพร่ศาสนาอิสลามให้แก่ชาวเมือง ทำให้ชาวเมืองเปลี่ยนมานับถือศาสนาอิสลาม จนกระทั่งศาสนาอิสลามเริ่มเป็นที่ยอมรับในหมู่เจ้าผู้ครองเมืองต่าง ๆ ในหมู่เกาะอินโดนีเซีย ตั้งแต่สมัยคริสต์ศตวรรษที่ 13 นักโบราณคดีได้ค้นพบจารึกที่หลุมฝังศพของสุลต่านของ Samudra ชื่อ Sultan Malik al Saleh ซึ่งเป็นกษัตริย์ที่นับถือศาสนาอิสลามองค์แรกของเมือง Samudra (มหาวิทยาลัยเชียงใหม่, 2554)

ข้อตกลงเบื้องต้นของวิธีการสุ่มแบบกริดมี 2 ข้อ ได้แก่

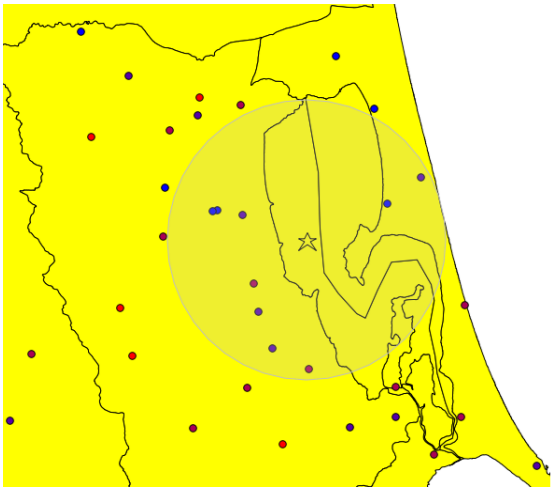
1. เทคนิคการสุ่มแบบจุดกริดเหมาะสำหรับการสุ่มตัวอย่างจากประชากรที่มีจำนวนมาก ๆ ในพื้นที่ขนาดใหญ่ โดยการกระจายตัวของประชากรเป็นแบบสม่ำเสมอหรือไม่สม่ำเสมอก็ได้ บางพื้นที่อาจมีความหนาแน่นของประชากรมากกว่าอีกพื้นที่หนึ่ง ดังนั้นวิธีการสุ่มแบบนี้จึงเหมาะสำหรับโครงการขนาดใหญ่

2. การกระจายตัวของค่าตัวแปรบนแผนที่ต้องเป็นไปตามกฎข้อที่ 1 ของภูมิศาสตร์หรือสหสัมพันธ์เชิงพื้นที่ทางบวก ถ้าค่าตัวแปรบนแผนที่เป็นแบบสหสัมพันธ์เชิงพื้นที่ทางลบหรือไม่มีสหสัมพันธ์เชิงพื้นที่จะทำให้การประมาณค่าเชิงพื้นที่เกิดความคลาดเคลื่อนอย่างมาก

7. การสุ่มแบบกริดสองขั้นตอน

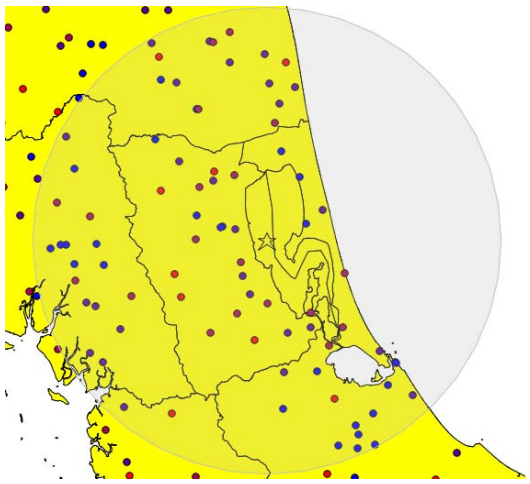
ในกรณีที่จุดกริดใด ๆ อยู่ในพื้นที่ ๆ มีสมาชิกที่มีค่าสังเกตแตกต่างกันมาก (มีความแปรปรวนสูง) หรือเป็นพื้นที่ ๆ มีประชากรหนาแน่นมาก ๆ เมื่อเทียบกับพื้นที่อื่น เนื่องจากมีการอพยพของประชาชนจากท้องถิ่นอื่นเข้ามาอาศัยอยู่พื้นที่นั้น ทำให้เกิดความแตกต่างกันทางวัฒนธรรม ขนบธรรมเนียม ประเพณี ความเชื่อ ภาษา เศรษฐกิจ สังคม ฯลฯ เช่น เมืองหลวง เมืองอุตสาหกรรม เมืองท่า ฯลฯ ดังนั้นพื้นที่นี้จะมีสมาชิกที่มีค่าสังเกตแตกต่างกันมาก ทำให้ธรรมชาติของประชากรไม่เป็นไปตามข้อตกลงเบื้องต้น เมื่อสุ่มตำแหน่งจุดกริดแล้วจุดกริดใดตกลงในพื้นที่ดังกล่าว การเลือกตัวอย่างที่อยู่ใกล้หรือติดกับจุดกริดจึงเป็นสิ่งที่ไม่ถูกต้อง เพราะค่าเฉลี่ยที่วัดได้กลุ่มตัวอย่างที่อยู่ใกล้ ๆ นั้น อาจมีค่าแตกต่างจากค่าเฉลี่ยของกลุ่มตัวอย่างทุก ๆ ตัวในพื้นที่นั้น วิธีแก้ปัญหาในกรณีพื้นที่มีลักษณะประชากรเป็นวิวิธพันธ์ คือ การสุ่มให้ได้กลุ่มตัวอย่างที่มาจากทุกกลุ่มก้อนทุกภาคส่วนของประชากร ซึ่งหลักการนี้เป็นหลักการของการสุ่มแบบชั้นภูมิ โดยการแบ่งประชากรเป็นประชากรกลุ่มย่อย (Subpopulation) ซึ่งสมาชิกภายในกลุ่มประชากรย่อยต้องมีค่าตัวแปรใกล้เคียงกัน และเมื่อสุ่มกลุ่มตัวอย่างในแต่ละกลุ่มประชากรย่อยด้วยวิธีจัดสรรตัวอย่างแบบสัดส่วน ทำให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนที่ดีของประชากร ซึ่งการสุ่ม

แบบนี้ คือ การสุ่มสองขั้นตอน ในขั้นตอนแรกใช้วิธีการสุ่มแบบกริดเพื่อเลือกตำแหน่งในแผนที่ และในขั้นตอนที่สองใช้วิธีการสุ่มแบบชั้นภูมิเพื่อทำให้ประชากรกลุ่มย่อยต่าง ๆ มีโอกาสที่จะถูกเลือก



ภาพที่ 2-15 ปัญหาของการสุ่มกลุ่มตัวอย่างในพื้นที่ที่ประชากรเป็นวิวิธพันธ์

จากภาพที่ 2-15 กำหนดให้จุดกริดที่ได้จากการสุ่มมีรูปร่างเป็นดวงดาว จุดเล็ก คือ ตำแหน่งที่ตั้งของโรงเรียน สีของจุดวงกลมเล็กประกอบด้วย สีแดง สีแดงปนม่วง สีน้ำเงินปนม่วง และสีน้ำเงิน ซึ่งแต่ละสีแทนความหมายค่าเฉลี่ยคะแนน ONET ของโรงเรียน โดยสีแดงแทนความหมายระดับคะแนน ONET ต่ำที่สุด และสีน้ำเงินแทนความหมายระดับคะแนน ONET สูงที่สุด (เรียงกันแบบน้อยไปหามาก) วงกลมใหญ่ คือ วงกลมของการสุ่ม และจุดเล็กที่อยู่ภายในวงกลมใหญ่ คือ กลุ่มตัวอย่างที่สุ่มได้มีทั้ง 9 จุด ซึ่งเป็นไปตามแผนการสุ่มที่วางไว้ เมื่อพิจารณากลุ่มตัวอย่างที่ได้พบว่า จุดเล็กเกือบทั้งหมดมีสีน้ำเงิน หรือสีน้ำเงินปนม่วง ดังนั้น ค่าเฉลี่ยกลุ่มตัวอย่างทั้ง 9 โรงเรียน จะมีค่าเฉลี่ยค่อนข้างสูง แต่เมื่อพิจารณาจุดเล็กอื่น ๆ ที่อยู่นอกรัศมีของการสุ่มพบว่ามีจุดเล็กสีแดงและสีแดงปนม่วง ดังนั้น การเลือกกลุ่มตัวอย่างที่อยู่ใกล้ชิดกับจุดกริดจึงเป็นสิ่งที่ไม่ถูกต้อง เพราะพื้นที่นี้มีลักษณะเป็นแบบวิวิธพันธ์



ภาพที่ 2-16 การสุ่มกลุ่มตัวอย่าง 2 ขั้นตอนด้วยการขยายขนาดรัศมีวงกลมของการสุ่ม

จากภาพที่ 2-16 เมื่อขยายขนาดวงกลมไปจนสุด (รัศมีเท่ากับครึ่งหนึ่งของระยะทางระหว่างจุดกริด) ทำให้ขนาดกลุ่มตัวอย่างเท่ากับ 81 โรงเรียน แต่แผนการสุ่มกำหนดให้สุ่มกลุ่มตัวอย่างเพียงแค่ 9 โรงเรียน เท่านั้น ดังนั้นต้องนำโรงเรียนทั้ง 81 โรง มาแบ่งเป็นชั้นภูมิ 4 ชั้นภูมิ คือ กลุ่มโรงเรียนสีแดง กลุ่มโรงเรียนสีแดงปนม่วง กลุ่มโรงเรียนสีน้ำเงินปนม่วง และกลุ่มโรงเรียนสีน้ำเงิน จากนั้นทำการจัดสรรแบบสัดส่วน โดยมีหลักการว่ากลุ่มโรงเรียนใดเป็นกลุ่มตัวอย่างขนาดใหญ่ต้องสุ่มออกมามาก แต่ถ้ากลุ่มโรงเรียนใดเป็นกลุ่มตัวอย่างขนาดเล็กต้องสุ่มออกมาน้อย

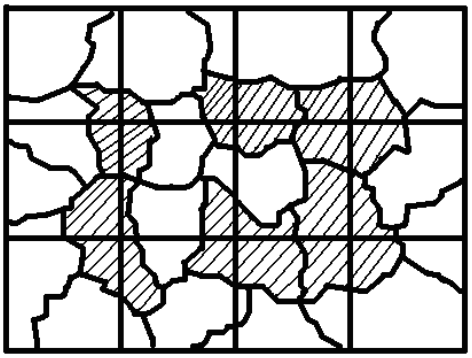
การสุ่มตัวอย่างสองขั้นตอนจะทำให้ค่าเฉลี่ยคะแนน ONET ในแต่ละจุดกริดใกล้เคียงกับค่าพารามิเตอร์ของประชากรในแต่ละจุดกริด และเมื่อพิจารณาในภาพรวมทั้งแผนที่ ถ้าใช้วิธีการสุ่มแบบสองขั้นตอนครบทุกจุดจะทำให้ค่าเฉลี่ยคะแนน ONET ของกลุ่มตัวอย่างใกล้เคียงกับค่าพารามิเตอร์

การสุ่มกลุ่มตัวอย่างแบบจุดกริดชนิดสองขั้นตอนหรือหลายขั้นตอน แตกต่างจากการสุ่มกลุ่มตัวอย่างสองขั้นตอนชนิดอื่น ๆ เพราะการสุ่มกลุ่มตัวอย่างในกรณีที่ประชากรมีจำนวนมาก ทำให้ต้องสุ่มกลุ่มตัวอย่างสองขั้นตอนหรือหลายขั้นตอน เพื่อให้กลุ่มตัวอย่างมีขนาดเล็กและเหมาะสมในการศึกษา/ ติดตามข้อมูล ถ้าโครงการใดที่มีการสุ่มหลาย ๆ ครั้ง แสดงว่าโครงการนั้นเป็นโครงการขนาดใหญ่ที่ศึกษาวิจัยกับประชากรจำนวนมาก ๆ ในส่วนของการสุ่มกลุ่มตัวอย่างแบบกริด มีการกระจายตำแหน่งจุดกริดให้ครอบคลุมพื้นที่ทั้งหมด โดยขนาดกลุ่มตัวอย่างในแต่ละจุดกริดขึ้นอยู่กับการจัดสรรแบบสัดส่วน ดังนั้นขนาดกลุ่มตัวอย่างจึงมีขนาดเล็ก ตั้งแต่ครั้งแรกของการสุ่ม แต่วัตถุประสงค์ของการสุ่มกลุ่มตัวอย่างแบบจุดกริดชนิดสองขั้นตอนหรือ

หลายขั้นตอน เพื่อใช้สำหรับแก้ปัญหาธรรมชาติของประชากรที่ไม่มีสหสัมพันธ์เชิงพื้นที่ เนื่องจากจุดกริดอยู่ในพื้นที่ ๆ ประชากรมีลักษณะวิวิธพันธ์ (Heterogeneity) จึงต้องนำวิธีการสุ่มแบบแบ่งชั้นภูมิเข้ามาช่วยแก้ปัญหา เพื่อให้ทุก ๆ ภาคส่วนของประชากรในพื้นที่นั้น จะมีโอกาสที่จะถูกเลือกมากยิ่งขึ้น ผลลัพธ์ที่ได้คือค่าเฉลี่ยของกลุ่มตัวอย่างใกล้เคียงกับค่าพารามิเตอร์มากขึ้น (ค่าความคลาดเคลื่อนลดลง) ทำให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนที่ดีของประชากร

8. การประยุกต์ระหว่างการสุ่มแบบกริดกับการสุ่มด้วยวิธีการอื่น ในระบบการสุ่มหลายขั้นตอน

เมื่อเลือกจำนวนจุดกริดและกำหนดตำแหน่งของจุดกริดแล้ว จุดกริดเหล่านี้จะอยู่ภายในขอบเขตของพื้นที่สมมติ ซึ่งสมาชิกทั้งหมดภายในพื้นที่นี้คือ $N_{i,j}$ ในขณะเดียวกันจุดกริดเหล่านี้ก็จะอยู่ภายในขอบเขตตามภูมิภาคและมีจำนวนสมาชิกเท่ากับ $N'_{i,j}$ อีกด้วย จึงเหมือนกับว่าการใช้โปรแกรมสุ่มตำแหน่งจุดกริดเป็นการสุ่มพื้นที่ตามภูมิภาคจากแผนที่



ภาพที่ 2-17 การนำวิธีสุ่มแบบกริดมาประยุกต์ใช้ในระบบการสุ่มหลายขั้นตอน

สมมติให้ประเทศหนึ่งมีเมืองทั้งหมด 22 เมือง (ตามภาพที่ 2-17) ถ้าต้องการสุ่มให้ได้ 6 เมือง จากทั้งหมด เพื่อทำการเก็บข้อมูลมาใช้ศึกษาวิจัย ผู้วิจัยสามารถใช้เทคนิคการสุ่มแบบจุดกริดโดยการใช้เส้นละติจูด 2 เส้น กับเส้นลองติจูด 3 เส้น ทำให้เกิดจุดกริด 6 จุด ซึ่งจุดกริดทั้ง 6 จุดนี้อยู่ในเมืองต่าง ๆ ซึ่งเป็นพื้นที่แรเงา จากนั้นสามารถใช้เทคนิคการสุ่มแบบอื่น ๆ เพื่อสุ่มเลือก เขต, แขวง ฯลฯ ต่อไป

การประยุกต์ใช้เทคนิคการสุ่มแบบจุดกริดกับเทคนิคการสุ่มอื่น ๆ นั้น ควรใช้การสุ่มแบบจุดกริดเริ่มต้นก่อนวิธีอื่น ๆ เนื่องจากวิธีการสุ่มนี้เหมาะที่จะใช้งานกับพื้นที่ขนาดใหญ่มาก ๆ ดังนั้นกลุ่มตัวอย่างที่ได้จะมาจากพื้นที่ห่างไกลกัน ซึ่งมีวิถีชีวิตไม่เหมือนกัน ทำให้พื้นฐานทางความคิดแตกต่างกันด้วย จึงเชื่อได้ว่ากลุ่มตัวอย่างนี้จะเป็นตัวแทนที่ดีของประชากร เพราะสมาชิกในกลุ่มตัวอย่างมีความคิดเห็นที่หลากหลาย ซึ่งได้มาจากทุกกลุ่มก้อนของประชากร

9. การวิเคราะห์ค่าสถิติ

Relative efficiency เป็นวิธีการเปรียบเทียบประสิทธิภาพตัวประมาณค่าไม่ลำเอียง (Unbiased estimator) 2 ตัว จากพารามิเตอร์เดียวกัน โดยการเปรียบเทียบอัตราส่วนของความแปรปรวน ซึ่งตัวประมาณค่ามีหลายชนิด เช่น ค่าสัดส่วน ค่าเฉลี่ย ค่ายอดรวม เป็นต้น ในกรณีตัวประมาณค่า คือ ค่าเฉลี่ยของกลุ่มตัวอย่าง ดังนั้น ค่าความแปรปรวน คือ เครื่องมือในการวัดการกระจาย ถ้าความแปรปรวนต่ำแสดงว่าโดยถัวเฉลี่ยแล้วค่าเฉลี่ยนั้นเบนห่างไปจากพารามิเตอร์น้อย ถ้าความแปรปรวนสูงแสดงว่าโดยถัวเฉลี่ยแล้วค่าเฉลี่ยนั้นเบนห่างไปจากพารามิเตอร์มาก แต่ในการเปรียบเทียบประสิทธิภาพระหว่างวิธีประมาณค่า 2 วิธี ตามหลักการระเบียบวิธีเชิงตัวเลขนั้นพิจารณาจากความแตกต่างระหว่างผลเฉลยเชิงตัวเลขกับผลเฉลยเชิงแม่นตรง ซึ่งวิธีการสุ่มแบบกริดเป็นวิธีการที่พัฒนาจากการเก็บตัวอย่างดิน (Grid soil sampling) ของวิทยาศาสตร์กายภาพ ดังนั้นในการวิจัยครั้งนี้ได้นำแนวคิดวิธีการเชิงตัวเลขมาประยุกต์ใช้ โดยใช้ค่าเฉลี่ยและความแปรปรวนของกลุ่มตัวอย่างในแต่ละวิธีการสุ่มเป็นตัวชี้วัดความเป็นตัวแทนที่ดีของประชากร โดยนำค่าเฉลี่ยและค่าความแปรปรวนเทียบกับค่าพารามิเตอร์ ซึ่งผลต่างระหว่างค่าเฉลี่ยที่ได้จากประชากรกับค่าเฉลี่ยที่ได้จากกลุ่มตัวอย่างเรียกว่า “ค่าความคลาดเคลื่อน”

$$e = \left|\sum e_{i,j}\right| = |\mu - \bar{z}| \approx \left|\mu - \hat{\bar{z}}\right|$$

ค่าความคลาดเคลื่อนคือผลต่างระหว่างค่าพารามิเตอร์กับค่าเฉลี่ยกลุ่มตัวอย่าง และมีค่าใกล้เคียงกับผลต่างระหว่างค่าพารามิเตอร์กับตัวประมาณค่าของค่าเฉลี่ยกลุ่มตัวอย่าง

$$= \left| \frac{1}{N} \sum_{i}^{x} \sum_{j}^{y} N_{ij} \mu_{ij} - \frac{1}{N} \sum_{i}^{x} \sum_{j}^{y} N_{ij} \ \bar{z}_{ij} \right|$$

$$\approx \left| \frac{1}{N} \sum_{i}^{x} \sum_{j}^{y} N_{ij} \mu_{ij} - \frac{1}{n} \sum_{i}^{x} \sum_{j}^{y} n_{ij} \ \bar{z}_{ij} \right|$$

เพื่อให้การวิเคราะห์มีความน่าเชื่อและได้รับการยอมรับ ดังนั้นจำนวนครั้งของการสุ่มตัวอย่างต้องมีขนาดมากเกินพอหรืออาจเข้าใกล้ค่าอนันต์ (Infinity) ก็ได้ ซึ่งผลการวิเคราะห์ที่ได้จะปราศจากข้อสงสัยท้วงติงในเรื่องความพอเพียงของจำนวนครั้งการทดลอง เมื่อสุ่มกลุ่มตัวอย่างจากประชากรในแต่ละครั้งจะได้สมาชิกในกลุ่มตัวอย่างไม่เหมือนกัน ทำให้ความคลาดเคลื่อนในแต่ละครั้งมีค่าแตกต่างกัน ดังนั้นต้องใช้ค่าเฉลี่ยของความคลาดเคลื่อนในการเปรียบเทียบประสิทธิภาพของวิธีการสุ่มแบบต่าง ๆ ซึ่งมีสูตร ดังนี้

$$\bar{e} = \frac{\sum_{k=1}^{m} e_k}{m}$$

$\bar{e}$ คือ ค่าเฉลี่ยของความคลาดเคลื่อน

e_k คือ ค่าความคลาดเคลื่อน เมื่อทำการการสุ่มตัวอย่างครั้งที่ **k**

m คือ จำนวนครั้งของการสุ่มทั้งหมด

ในการเปรียบเทียบค่าความแปรปรวนของกลุ่มตัวอย่างต้องใช้ความแปรปรวนของประชากรเป็นเกณฑ์ ซึ่งผลต่างระหว่างค่าความแปรปรวนที่ได้จากประชากรกับค่าความแปรปรวนที่ได้จากกลุ่มตัวอย่างเรียกว่า “ค่าคลาดเคลื่อนสมบูรณ์ของความแปรปรวน” หรือ “ค่าความแตกต่างความแปรปรวน”

$$|\Delta v| = \left|\sigma^2 - s^2\right|$$

เนื่องจากการสุ่มกลุ่มตัวอย่างจากประชากรในแต่ละครั้งจะได้สมาชิกในกลุ่มตัวอย่างไม่เหมือนกัน ทำให้ค่าความแตกต่างความแปรปรวนในแต่ละครั้งมีค่าแตกต่างกันด้วย ดังนั้นต้องใช้ผลรวมของความแตกต่างระหว่างความแปรปรวนในการเปรียบเทียบประสิทธิภาพของวิธีการสุ่มแบบต่าง ๆ ซึ่งมีสูตรดังนี้

$$\Sigma|\Delta v| = |\Delta v|_1 + |\Delta v|_2 + |\Delta v|_3 + ... + |\Delta v|_m$$

$\Sigma|\Delta v|$ คือ ผลรวมของค่าความแตกต่างความแปรปรวน

m คือ จำนวนครั้งของการสุ่มทั้งหมด

10. เทคนิคการใช้งาน

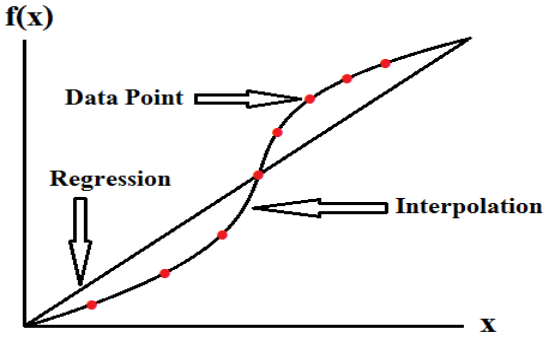
เนื่องจากธรรมชาติของกลุ่มประชากรในการสุ่มแบบจุดกริด ต้องมีสหสัมพันธ์เชิงพื้นที่ทางบวก ดังนั้นผู้วิจัยต้องศึกษางานวิจัยที่เกี่ยวข้องหรือมั่นใจว่าประชากรที่ต้องการศึกษาวิจัยเป็นไปตามข้อตกลงเบื้องต้น ในกรณีที่ไม่สามารถหางานวิจัยที่เกี่ยวข้องได้หรือไม่มั่นใจในเรื่องธรรมชาติของประชากรให้ทดลองสุ่มตัวอย่างจุดกริดบางตำแหน่งก่อน ซึ่งการสุ่มตัวอย่างที่จุดกริดนั้นไม่ต้องสุ่มให้ครบตามจำนวนที่วางแผน เพื่อทดสอบก่อนว่าค่าสถิติต่าง ๆ ที่ได้จากกลุ่มตัวอย่างนั้น เป็นไปตามที่คาดการณ์ไว้หรือไม่ ซึ่งมีวิธีการตรวจสอบคือสมาชิกของกลุ่มตัวอย่างที่อยู่บริเวณใกล้กับจุดกริด (ภายในวงกลมการสุ่ม) ต้องมีค่าสถิติใกล้เคียงกัน เพราะประชาชนที่อาศัยอยู่ในท้องถิ่นเดียวกันต้องมีวิถีชีวิต วัฒนธรรม กฎหมาย และความเชื่อคล้ายกัน ดังนั้นค่าความแปรปรวนต้องต่ำ

(Homogeneity) แต่เมื่อเก็บรวบรวมข้อมูลที่จุดกริดที่ห่างออกไปค่าสถิติที่ได้ต้องมีค่าต่างจากเดิม เนื่องจากท้องถิ่นที่ห่างออกไปนั้นอาจมีปัจจัยต่าง ๆ ที่ส่งผลต่อความคิดเห็น ความเชื่อ และทัศนคติแตกต่างกัน ยกตัวอย่างเช่น สภาพอากาศ สังคม เศรษฐกิจ การเมืองและการปกครอง ศาสนา เป็นต้น ดังนั้นค่าสถิติทั้งสองจุดกริดต้องแตกต่างกัน

หลังจากเก็บข้อมูลและหาค่าสถิติแล้ว ให้พิจารณาการเปลี่ยนแปลงค่าสถิติที่จุดกริดต่าง ๆ ในภาพรวมว่าเป็นไปตามกฎข้อที่ 1 ของระบบภูมิศาสตร์หรือไม่ ถ้าข้อมูลที่เก็บรวบรวมมาได้มีสหสัมพันธ์เชิงพื้นที่แสดงว่าการคาดการณ์ถูกต้องและสามารถดำเนินการในขั้นตอนต่อไปได้ แต่ถ้าข้อมูลที่เก็บรวบรวมมาได้ไม่มีสหสัมพันธ์เชิงพื้นที่ต้องเปลี่ยนเทคนิคการสุ่มกลุ่มตัวอย่างทันที นอกจากนี้ในการปฏิบัติงานสนามผู้ช่วยนักวิจัยต้องมั่นใจว่ากลุ่มตัวอย่างที่ทำการสำรวจข้อมูลนั้นต้องเป็นคนในท้องถิ่นหรือทำงานในพื้นที่นั้นจริง ๆ จนกระทั่งได้ซึมซับวัฒนธรรมและขนบธรรมเนียมประเพณี โดยมีวิถีชีวิตสอดคล้องกับสังคมชุมชน ถ้ากลุ่มตัวอย่างที่ถูกเก็บข้อมูลเป็นเพียงนักท่องเที่ยวหรือผู้ที่เดินทางผ่านมา ข้อมูลที่ได้อาจไม่สอดคล้องกับแนวคิดของคนในชุมชนและส่งผลให้งานวิจัยผิดเพี้ยน ซึ่งผู้ช่วยผู้วิจัยอาจใช้การสัมภาษณ์อย่างคร่าว ๆ ก่อนทำการสำรวจข้อมูลก็ได้

11. การนำวิธีการประมาณค่าในช่วงมาประยุกต์ใช้กับทฤษฎีการสุ่มแบบกริด

ในทางสถิติวิธีการวิเคราะห์การถดถอย (Regression analysis) เป็นวิธีที่ใช้หาสมการความสัมพันธ์ระหว่างตัวแปรเชิงปริมาณตั้งแต่สองตัวขึ้นไป เพื่อใช้อธิบายถึงลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งขนาดและทิศทางความสัมพันธ์ ซึ่งสมการความสัมพันธ์สามารถใช้ในการพยากรณ์ค่าของตัวแปรตาม (Dependent variable) จากตัวแปรอิสระ (Independent variable) ได้ ซึ่งหลักการของวิธี Regression คือ การสร้างเส้นกราฟหรือเส้นโค้งที่ให้ทำให้เกิดผลต่างกำลังสองระหว่างค่าจริงกับค่าประมาณที่น้อยที่สุด วิธีการดังกล่าวเรียกว่า วิธีกำลังสองน้อยที่สุด (Method of least squares) วิธีการ Least square regression สามารถแบ่งออกเป็นหลายประเภท เช่น Linear regression, Nonlinear regression, Polynomial regression และ Multiple regression เป็นต้น แต่ในระเบียบวิธีเชิงตัวเลข (Numerical methods) มีวิธีการอีกอย่างหนึ่งในการประมาณค่าตัวแปรตาม [f(x)] จากตัวแปรอิสระ [x] วิธีการนั้น คือ การประมาณค่าในช่วง (Interpolation) ซึ่งหลักการของวิธีการ Interpolation คือ การสร้างเส้นกราฟให้ผ่านจุดข้อมูล (Data points) ที่ถูกสุ่มวัดค่าทั้งหมด ดังภาพที่ 2-18 วิธีการประมาณค่าในช่วงมีหลากหลายวิธี เช่น Lagrange interpolation, Newton interpolation, Interpolation using direct method และSpline interpolation เป็นต้น



ภาพที่ 2-18 ความแตกต่างระหว่างวิธีการประมาณค่าในช่วงและวิธีการถดถอย

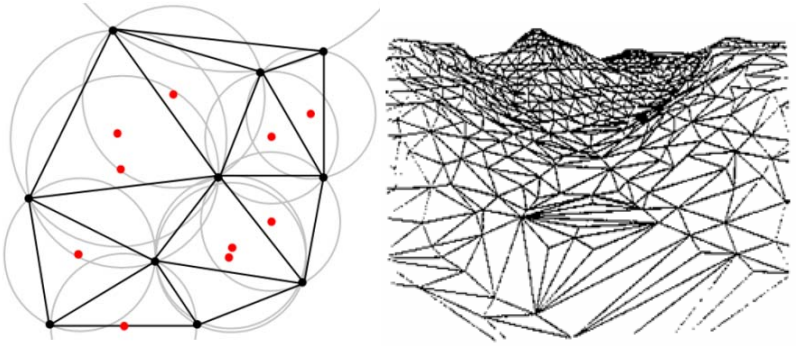
1. การประมาณค่าในช่วง (Interpolation)

การวิเคราะห์พื้นผิว (Surface analysis) เป็นการวิเคราะห์การกระจายตัวของค่าตัวแปรเชิงพื้นที่บนแผนที่ โดยตำแหน่งของจุดกริดเกิดจากการตัดกันของแนวเส้นละติจูดและแนวเส้นลองจิจูดเปรียบได้กับแกน X และ Y และค่าตัวแปรที่ตำแหน่งจุดกริดเปรียบได้กับแกน Z ซึ่งเสมือนเป็นมิติที่ 3 ของข้อมูลเชิงพื้นที่ (สุระ พัฒนเกียรติ, 2554) ตัวอย่างของค่าตัวแปร Z เช่น ความพึงพอใจ ความคิดเห็น ความเชื่อ และเจตคติของคนในพื้นที่ เป็นต้น ในการหาค่าตัวแปร Z ที่ตำแหน่งใด ๆ สามารถหาได้จากการเก็บข้อมูลจากกลุ่มตัวอย่างในบริเวณนั้น แต่การนำเสนอข้อมูลในรูปแบบแผนที่ จำเป็นต้องทราบข้อมูลค่าตัวแปร Z ทุกตำแหน่งบนแผนที่ ซึ่งการเก็บข้อมูลของประชากรทั่วทั้งแผนที่ เป็นสิ่งที่ทำได้ยากมากและมีค่าใช้จ่ายที่สูงมาก ดั้งนั้นต้องเก็บข้อมูลจากจุดตัวอย่างเพียงบางจุด เพื่อใช้ประมาณค่าจุดที่เหลือทั้งหมด ทำให้ผลที่ได้จากการประมาณค่าเชิงพื้นที่สามารถแสดงเป็นภาพ 3 มิติ พร้อมทั้งแสดงสหสัมพันธ์เชิงพื้นที่ (Spatial relationship) ระหว่างค่า Z กับค่าพิกัดทางภูมิศาสตร์

การประมาณค่าเชิงพื้นที่ (Spatial interpolation) มีหลักการที่สำคัญ 2 ข้อ คือ ค่าตัวแปร Z ต้องมีการเปลี่ยนแปลงแบบต่อเนื่องอย่างค่อยเป็นค่อยไป และค่าตัวแปร Z ต้องมีสหสัมพันธ์เชิงพื้นที่ โดยค่าตัวแปร Z ของจุดที่ไม่ทราบค่าจะมีค่าใกล้เคียงกับจุดที่ทราบค่าที่อยู่ใกล้ที่สุด ซึ่งวิธีการประมาณค่า (Interpolation methods) มีหลายวิธี เช่น Inverse distance weighted (IDW), Natural neighbors, Spline และ Kriging แต่การประมาณค่าโดยโปรแกรม QGIS มี 2 วิธี ได้แก่

1) Triangulated irregular network (TIN) คือ การแสดงลักษณะของพื้นผิวโดยการใช้รูปสามเหลี่ยมหลายรูป ซึ่งมีด้านประชิดกันและใช้จุดยอดร่วมกันเรียงต่อเนื่องกันไป โดยค่า Z จัดเก็บอยู่ที่จุดยอดของสามเหลี่ยม จุดเหล่านี้จะกระจายตัวไม่สม่ำเสมอ โดยพื้นที่ที่มีความแตกต่างของค่า Z มาก ตำแหน่งของจุดต้องอยู่ใกล้ ๆ กัน แต่พื้นที่ที่มีค่า Z ไม่แตกต่างกันนัก ตำแหน่งของ

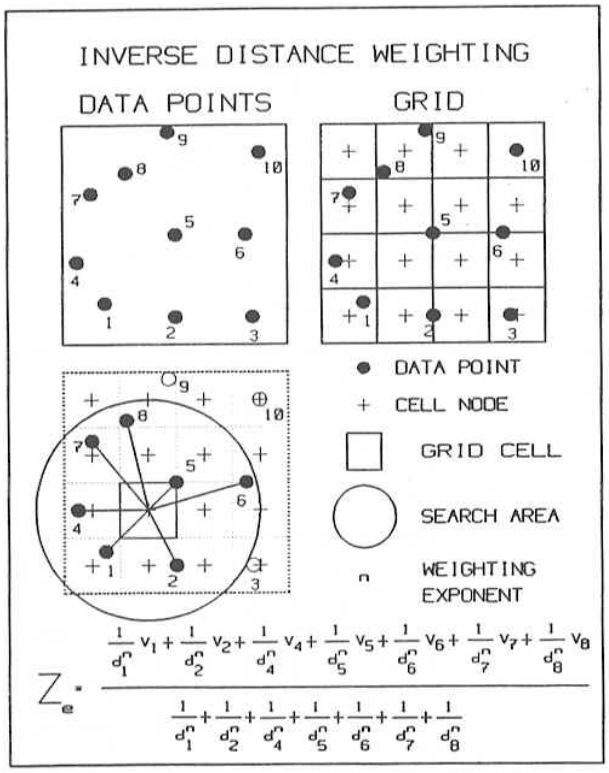
จุดสามารถอยู่ห่างกัน ดังภาพที่ 2-19 (ก) (Longley, Goodchild, Maguire, & Rhind, 1999). และ (ข) (สุระ พัฒนเกียรติ, 2554)



(ก) (ข)

ภาพที่ 2-19 การประมาณค่าเชิงพื้นที่แบบ TIN

2) Inverse distance weighted (IDW) เป็นวิธีการประมาณค่าเฉลี่ยถ่วงน้ำหนัก โดยปัจจัยระยะทางแปรผกผันกับค่าถ่วงน้ำหนักกล่าวคือ จุดกริดที่อยู่ใกล้ (กับกริดเซลล์ที่ต้องการหาค่า) มีผลกระทบหรืออิทธิพลต่อกริดเซลล์มากกว่าจุดกริดที่อยู่ไกลออกไป โดยสามารถกำหนดจำนวนจุดหรืออาจใช้ทุกจุดที่อยู่ภายในรัศมีที่กำหนดมาทำการประมาณค่ากริดเซลล์นั้น ๆ ได้ ดังนั้นผลกระทบหรืออิทธิพลจะน้อยลงเรื่อย ๆ ตามระยะทางที่ไกลออกไปจากเซลล์ที่ต้องการประมาณค่า โดยมีรายละเอียดตามภาพที่ 2-20



ภาพที่ 2-20 การประมาณค่าเชิงพื้นที่แบบ IDW (Krajewski & Gibbs, 1994)

2. การวิเคราะห์ความคลาดเคลื่อน

ในปี ค.ศ. 1968 Donald Shepard ค้นพบวิธีการประมาณค่าเชิงพื้นที่แบบค่าเฉลี่ยถ่วงน้ำหนัก (Shepard, 1968) โดยใช้จุดข้อมูลที่ทราบตัวแปร $\left(f_k \right)$ จากการวัดมาประมาณค่าจุดตำแหน่งที่ไม่ทราบค่า $\left(f_{uk} \right)$ ซึ่งหลักการในการคำนวนคือจุดที่ต้องการทราบค่า $\left(f_{uk} \right)$ ได้รับอิทธิพลหรือผลกระทบจากจุดข้อมูลที่ทราบค่า $\left(f_k \right)$ ที่ใกล้ที่สุด ดังนั้นจุดที่อยู่ใกล้กันต้องมีค่าตัวแปรที่ใกล้เคียงกัน เมื่อจุดที่ไม่ทราบค่าอยู่ระหว่างจุดข้อมูลจำนวนหลายจุดต้องใช้การเฉลี่ยถ่วงน้ำหนักจากจุดข้อมูลเหล่านั้น โดยยึดถือระยะทางเป็นสำคัญกล่าวคือค่าถ่วงน้ำหนัก $\left(w_k \right)$ เป็นค่าผกผันของระยะทาง $\left(d_k \right)$ ระหว่างจุดที่ทราบค่ากับจุดที่ไม่ทราบค่า ซึ่งมีสูตรในการคำนวนดังนี้ (Shepard, 1968)

$$f_{uk} = \frac{\sum_{k=1}^{q} w_k f_k}{\sum_{k=1}^{q} w_k} \qquad (1)$$

โดยที่ (f_{uk}) คือ ค่าประมาณของจุดใด ๆ บนแผนที่

(f_k) คือ ค่าตัวแปรของจุดข้อมูลตัวที่ k

(w_k) คือ ค่าถ่วงน้ำหนักของจุดข้อมูลตัวที่ k

q คือ จำนวนจุดที่อยู่บริเวณใกล้เคียงทั้งหมดที่ใช้ในการประมาณค่า

ซึ่งค่าถ่วงน้ำหนัก (w_k) เป็นค่าผกผันของระยะทาง (d_k) ระหว่างจุดที่ทราบค่ากับจุดที่ไม่ทราบค่า ซึ่งมีความสัมพันธ์ดังนี้ (Shepard, 1968)

$$w_k = \frac{1}{d_k^p} \qquad (2)$$

โดยที่ (d_k) คือ ระยะทางระหว่างจุดข้อมูลที่ถูกวัดค่ากับจุดใด ๆ ที่ต้องการทราบค่า

p คือ ตัวเลขชี้กำลัง ซึ่งมีค่า (0,2) ในกรณีพื้นผิว 2 มิติ

จากสมการที่ (1) และ (2) สามารถยุบรวมกันได้ดังนี้ (Shepard, 1968)

$$f_{uk} = \frac{\sum_{k=1}^{q}\left(\frac{1}{d_k^p}\right) f_k}{\sum_{k=1}^{q}\left(\frac{1}{d_k^p}\right)}$$

เมื่อนำแนวคิดของ Shepard's method กับแนวคิด Grid sampling มาบูรณาการณ์ร่วมกัน สามารถกระทำได้ดังนี้

ถ้าจุดข้อมูล f_k คือ จุดกริดที่ทราบค่าตัวแปร z โดยการสุ่มตัวอย่างและหาค่าเฉลี่ย ได้เท่ากับ $\overline{z}_{i,j}$ แต่การประมาณค่า f ที่ตำแหน่งใด ๆ บนแผนที่ไม่ได้ใช้จุดกริดทั้งหมด ซึ่งการคำนวนค่าตัวแปร z จะใช้จุดกริดที่อยู่บริเวณใกล้เคียง ดังนั้นให้ $\overline{z}_k$ แทนค่า f_k โดยที่ $\{\overline{z}_k\} \subset \{\overline{z}_{i,j}\}$ และมีความสัมพันธ์ในรูปสมการ ดังนี้

$$\overline{z}_{uk} = \frac{\sum_{k=1}^{q}\left(\frac{1}{d_k^p}\right)\overline{z}_k}{\sum_{k=1}^{q}\left(\frac{1}{d_k^p}\right)}$$

ค่า $\overline{z}_{uk}$ เป็นค่าประมาณที่ได้จากจุดกริด $\left(\overline{z}_k\right)$ ที่ตำแหน่ง (i,j) ดังนั้นค่าคลาดเคลื่อนหาได้ดังนี้

$$e_{i'j'} = \mu_{i'j'} - \overline{z}_{uk}$$

โดยที่ $\left(i', j'\right)$ คือพิกัดหรือตำแหน่งของจุด $\overline{z}_{uk}$ ซึ่งเป็นจุดที่ต้องการทราบค่าและไม่ใช่จุดกริด $\left(i, j\right)$

$e_{i'j'}$ คือ ค่าความคลาดเคลื่อนที่ตำแหน่ง $\left(i', j'\right)$

$\mu_{i'j'}$ คือ ค่าพารามิเตอร์ที่หาได้จากประชากรที่ตำแหน่ง $\left(i', j'\right)$

ค่าคลาดเคลื่อนร้อยละ (Percent error)

$$E = \frac{\left|e_{i'j'}\right|}{\mu_{i'j'}} \text{x} 100\%$$

ค่าคลาดเคลื่อนร้อยละคือตัวบ่งชี้ประสิทธิภาพของการประมาณค่าเชิงพื้นที่ของแบบจำลองที่ถูกสุ่มด้วยวิธีการสุ่มแบบกริด ณ ตำแหน่งจุดกริดนั้น

ตอนที่ 3 การสุ่มตัวอย่างจำนวนมาก ๆ ที่ใช้ในการทดสอบทางการศึกษาของประเทศไทย

1. โครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ (TIMSS) Trends in International Mathematics and Science Study (TIMSS) เป็นโครงการที่สมาคมนานาชาติ (International Association for the Evaluation of Educational Achievement หรือ IEA) จัดทำขึ้นร่วมกับประเทศสมาชิก เพื่อประเมินผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์และคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 4 (Grade 4) และชั้นมัธยมศึกษาปีที่ 2 (Grade 8) โดยเน้นการประเมินในด้านเนื้อหา (Content domain) และด้านพฤติกรรมการเรียนรู้ (Cognitive domain) ของวิชาคณิตศาสตร์และวิทยาศาสตร์ ในส่วนโครงการวิจัยนานาชาติ TIMSS 2011 ได้นำแบบทดสอบมาใช้ประเมินผล ซึ่งมีขอบเขตของหัวข้อเรื่อง ดังนี้

1. เนื้อหาวิชาคณิตศาสตร์ ได้แก่ จำนวน พีชคณิต เรขาคณิตและการวัด ข้อมูลและโอกาส

2. พฤติกรรมการเรียนรู้วิชาคณิตศาสตร์ ได้แก่ ความรู้ การประยุกต์ใช้ความรู้ การใช้เหตุผล

3. เนื้อหาวิชาวิทยาศาสตร์ ได้แก่ ชีววิทยา เคมี ฟิสิกส์ โลก ดาราศาสตร์และอวกาศ

4. พฤติกรรมการเรียนรู้วิชาวิทยาศาสตร์ ได้แก่ ความรู้ การประยุกต์ใช้ความรู้ การใช้เหตุผล (โครงการ TIMSS 2011 THAILAND) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ข; 2556 ค)

นอกจากนี้ยังได้นำแบบสอบถามมาใช้สำรวจข้อมูลจาก นักเรียน ครู และผู้บริหาร สำหรับแบบสอบถามนักเรียนเป็นเรื่องเกี่ยวกับข้อมูลทั่วไปของนักเรียนและโรงเรียน การเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ในโรงเรียน สำหรับแบบสอบถามครูวิชาวิทยาศาสตร์และคณิตศาสตร์เป็นเรื่องเกี่ยวกับข้อมูลทั่วไปของครูและโรงเรียน แหล่งการเรียนรู้ หัวข้อที่สอนและความครอบคลุมของเนื้อหา การมอบหมายการบ้าน การประเมินผล และการเตรียมการสอน และสำหรับผู้บริหารโรงเรียนเป็นเรื่องเกี่ยวกับจำนวนนักเรียน ลักษณะของโรงเรียน เวลาที่ใช้ในการสอน แหล่งการเรียนรู้และเทคโนโลยี การมีส่วนร่วมของผู้ปกครองในกิจกรรมของโรงเรียน บรรยากาศการเรียนรู้ของนักเรียน การพัฒนาตนเองของครู และการปฏิบัติงานในฐานะผู้บริหาร (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ข; 2556 ค) โครงการ TIMSS 2011 มี 63 ประเทศ และ 14 รัฐเข้าร่วมการวิจัยครั้งนี้ ในส่วนประเทศไทยในชั้นประถมศึกษาปีที่ 4 มีนักเรียนเข้าร่วมประเมิน 4,448 คน จาก 168 โรงเรียน และชั้นมัธยมศึกษาปีที่ 2 มีนักเรียนเข้าร่วมประเมิน 6,124 คน จาก 172 โรงเรียน โดยเก็บรวบรวมข้อมูลเมื่อเดือนกุมภาพันธ์ พ.ศ. 2554 จำนวนนักเรียนและจำนวนโรงเรียนกลุ่มตัวอย่างชั้นประถมศึกษาปีที่ 4 จำแนกตามสังกัด แสดงดังตาราง 2-3 และมัธยมศึกษาปีที่ 2 จำแนกตามสังกัดแสดงดังตาราง 2-4 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ข; 2556 ค)

ตารางที่ 2-3 จำนวนโรงเรียนและนักเรียนกลุ่มตัวอย่างชั้นประถมศึกษาปีที่ 4 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ข)

สังกัด	จำนวน	
	นักเรียน (ร้อยละ)	โรงเรียน (ร้อยละ)
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)	2,340 (52)	108 (64)
สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)	1,025 (23)	28 (17)
สำนักการศึกษา กรุงเทพมหานคร (กทม.)	299 (7)	10 (6)
กรมการปกครองส่วนท้องถิ่น (เทศบาล/ ท้องถิ่น)	339 (8)	10 (6)
สำนักงานคณะกรรมการการอุดมศึกษา (สาธิต)	445 (10)	12 (7)
รวม	4,448	168

ตารางที่ 2-4 จำนวนโรงเรียนและนักเรียนกลุ่มตัวอย่างชั้นมัธยมศึกษาปีที่ 2 (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ข)

สังกัด	จำนวน	
	นักเรียน (ร้อยละ)	โรงเรียน (ร้อยละ)
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (โรงเรียน สพฐ. ที่มาจากโรงเรียนขยายโอกาสเดิม, สพฐ.1)	1,225 (20)	38 (21)
สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (โรงเรียนในสังกัดกรมสามัญศึกษาเดิม, สพฐ.2)	3,172 (52)	82 (48)
สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.)	572 (9)	20 (12)
สำนักการศึกษา กรุงเทพมหานคร (กทม.)	349 (5)	10 (6)
กรมการปกครองส่วนท้องถิ่น (เทศบาล/ ท้องถิ่น)	384 (6)	12 (7)
สำนักงานคณะกรรมการการอุดมศึกษา (สาธิต)	422 (8)	10 (6)
รวม	6,124	172

2. โครงการประเมินผลนักเรียนนานาชาติ PISA

โครงการประเมินผลนักเรียนนานาชาติ (Programme for International Student Assessment หรือ PISA) เป็นโครงการประเมินผลการศึกษาของประเทศในกลุ่มสมาชิก โดย องค์การเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ (Organisation for economic co-operation and development หรือ OECD) เป็นผู้ดำเนินการจัดการทดสอบและประเมินผลการเรียนรู้ เพื่อเปรียบเทียบศักยภาพในการแข่งขันของเยาวชนในระดับการศึกษาขั้นพื้นฐาน โดยการตรวจสอบว่าเยาวชนมีความรู้และทักษะที่จำเป็นที่ต้องใช้ในชีวิตจริงบนโลกยุคใหม่ที่มีการเปลี่ยนแปลงอย่างรวดเร็ว ซึ่ง PISA เน้นการประเมินสมรรถนะของนักเรียนวัย 15 ปี ที่จะใช้ความรู้และทักษะเพื่อเผชิญกับโลกในชีวิตจริงมากกว่าการเรียนรู้ในห้องเรียนตามหลักสูตรการศึกษา (โครงการ PISA ประเทศไทย) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554)

โครงการประเมินผลนักเรียนนานาชาติ PISA มีวัตถุประสงค์เพื่อต้องการหาตัวชี้วัดคุณภาพการศึกษาให้แก่ประเทศสมาชิกในโครงการ ใน PISA 2012 ใช้กลุ่มตัวอย่างนักเรียนประมาณ 510,000 คน จาก 65 ประเทศ มาเป็นตัวแทนประชากร ในส่วนประเทศไทยมีกลุ่มตัวอย่างนักเรียนจำนวน 6,606 คน จาก 239 โรงเรียนของทุกสังกัด โดยสุ่มตัวอย่างโรงเรียนแบบแบ่งชั้นภูมิตามสัดส่วน ซึ่งชั้นภูมิทั้งหมด 9 ชั้นภูมิ ดังนี้

1. โรงเรียน สพฐ. ที่มาจากโรงเรียนขยายโอกาสเดิม
2. โรงเรียน สพฐ. ในสังกัดกรมสามัญศึกษาเดิม
3. โรงเรียนในสังกัดสำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน
4. โรงเรียนในสังกัดสำนักประสานและพัฒนาการจัดการศึกษาท้องถิ่น
5. โรงเรียนสาธิตของมหาวิทยาลัย
6. โรงเรียน/ วิทยาลัยอาชีวศึกษาของเอกชน
7. วิทยาลัยอาชีวศึกษาของรัฐ
8. โรงเรียนในสังกัดสำนักการศึกษา กรุงเทพมหานคร
9. โรงเรียนจุฬาภรณราชวิทยาลัย (เฉพาะกลุ่มโรงเรียนจุฬาภรณราชวิทยาลัยเข้าร่วมการประเมินทุกโรงเรียน) (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2556 ก)

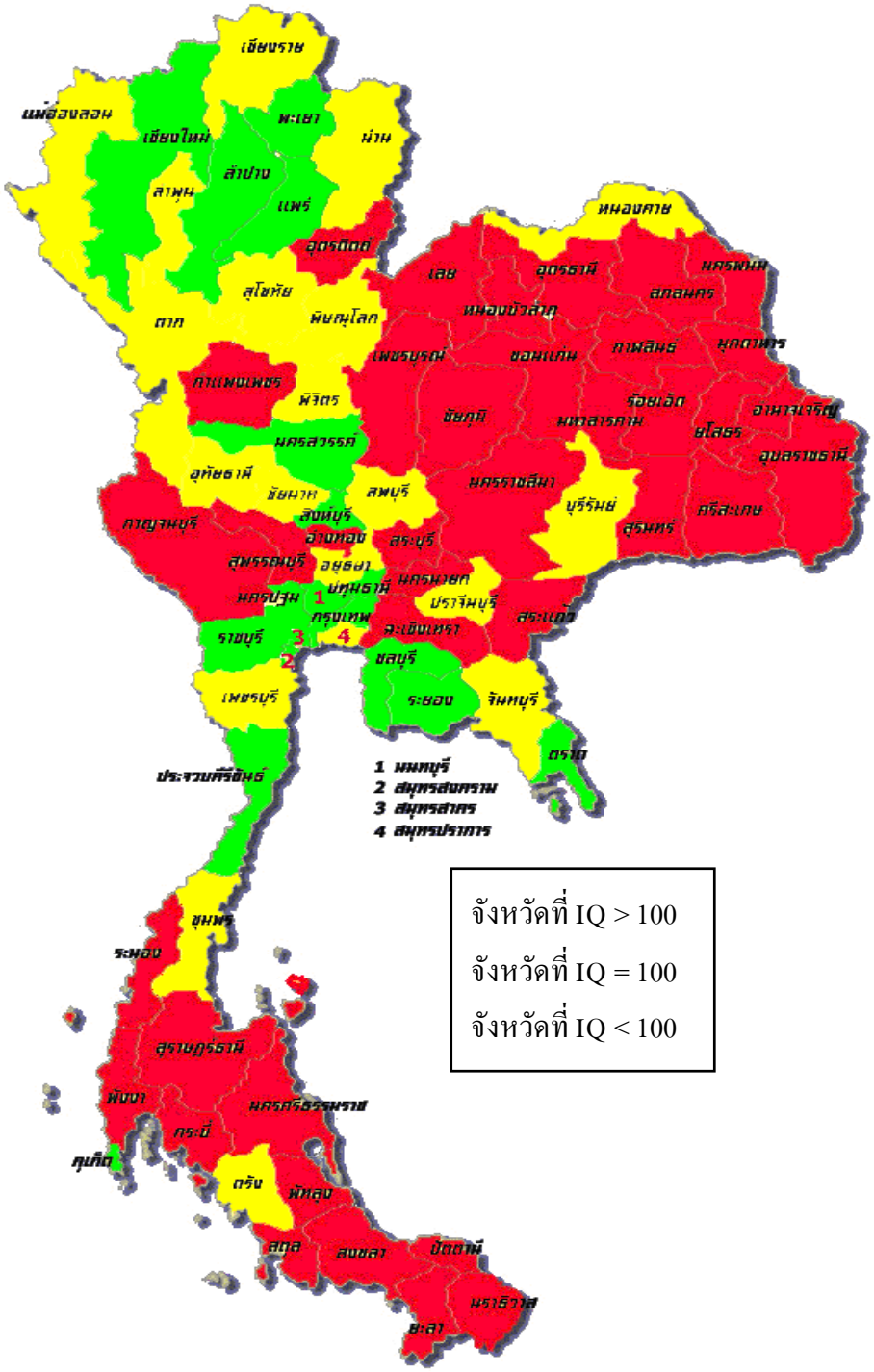
นอกจากการสุ่มตามประเภทโรงเรียนแล้ว ยังมีการสุ่มตัวอย่างตามภาคพื้นที่ภูมิศาสตร์ เพื่อให้กลุ่มตัวอย่างครอบคลุมทุกพื้นที่ของประเทศ โดยพิจารณาให้มีปริมาณของพื้นที่ และจำนวนโรงเรียนให้ใกล้เคียงกันมากที่สุด ได้แก่ กทม. และปริมณฑล, ภาคกลาง, ภาคเหนือตอนบน, ภาคเหนือตอนล่าง, ภาคอีสานตอนบน, ภาคอีสานตอนล่าง, ภาคใต้, ภาคตะวันตก และภาคตะวันออก (สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, 2554)

3. โครงการสำรวจสถานการณ์ระดับสติปัญญาเด็กนักเรียนไทย ปี พ.ศ. 2554 การส่งเสริมพัฒนาระดับสติปัญญาของเด็กและเยาวชนในประเทศ ให้พัฒนาได้อย่างเหมาะสม จึงถือเป็นเป้าหมายพื้นฐานที่จะพัฒนาคุณภาพของประชากรที่จะเติบโตไปเป็นกำลังสำคัญของชาติในอนาคต และส่งผลต่อความเจริญรุ่งเรืองของประเทศต่อไป ดังนั้นกรมสุขภาพจิต กระทรวงสาธารณสุข ได้ดำเนินโครงการสำรวจสถานการณ์ระดับสติปัญญาเด็กนักเรียนไทยทั่วประเทศ ในปีงบประมาณ 2554 เพื่อใช้เป็นฐานข้อมูลพื้นสติปัญญาเด็กนักเรียนไทย ทั้งในระดับประเทศ ระดับภาค และระดับจังหวัด ซึ่งการสำรวจนี้ได้ทำกรอบการสุ่มตัวอย่างระดับประเทศ ระดับภาค และครอบคลุมถึงระดับจังหวัด เพื่อให้ผู้บริหารและหน่วยงานที่มีส่วนเกี่ยวข้องได้รับข้อมูลและรายละเอียดที่ชัดเจนและเชื่อถือได้ สามารถวางแผนแก้ไขปัญหาได้ถูกต้องและตรงจุด สามารถประเมินผล ติดตามความก้าวหน้าในการดำเนินการในพื้นที่เป้าหมาย ในการสำรวจระดับ IQ ใช้กลุ่มตัวอย่าง จำนวน 72,780 คน อายุ 6-15 ปี ในระดับชั้นเรียนประถมศึกษาปีที่ 1-มัธยมศึกษาปีที่ 3 จากโรงเรียน 787 โรงเรียน (สถาบันราชานุกูล กรมสุขภาพจิต, 2554) ซึ่งมีรายละเอียดดังนี้

1. โรงเรียนสังกัดคณะกรรมการศึกษาขั้นพื้นฐาน (สพฐ.) จำนวน 612 โรงเรียน
2. สังกัดคณะกรรมการส่งเสริมการศึกษาเอกชน จำนวน 153 โรงเรียน
3. สังกัดคณะกรรมการอุดมศึกษา (สาธิตและราชภัฏ) จำนวน 16 โรงเรียน

4. สังกัดกรุงเทพมหานคร จำนวน 6 โรงเรียน

นอกจากการสุ่มตามประเภทโรงเรียนแล้ว ยังมีการสุ่มตัวอย่างตามจังหวัด โดยการสุ่มตัวอย่าง 7-19 โรงเรียนต่อจังหวัด คิดเป็นจำนวนประมาณ 851-1,163 คน ต่อจังหวัด ผลการสำรวจพบว่า ค่าเฉลี่ยระดับสติปัญญาเด็กนักเรียนไทยในภาพระดับประเทศ เท่ากับ 98.59 (ค่าเฉลี่ย IQ ในเกณฑ์ปกติ 90-109) ซึ่งถือเป็นค่าระดับสติปัญญาที่อยู่ในเกณฑ์ปกติแต่ค่อนไปทางต่ำ โดยเฉพาะเมื่อเทียบกับสถานการณ์ระดับสติปัญญาเด็กในประเทศอาเซียนหลาย ๆ ประเทศ เช่น ฮ่องกง สิงคโปร์ จีน ญี่ปุ่น ส่วนค่าเฉลี่ยระดับสติปัญญาเมื่อเทียบตามภาคพบว่า เรียงลำดับจากมากไปน้อยได้ดังนี้ คือ กรุงเทพมหานคร ภาคกลาง ภาคเหนือ ภาคใต้ และภาคตะวันออกเฉียงเหนือ (สถาบันราชานุกุล กรมสุขภาพจิต, 2554) ซึ่งมีรายละเอียดตามภาพที่ 2-21



ภาพที่ 2-21 แผนที่ประเทศไทยแสดงผลการสำรวจระดับสติปัญญาเด็กนักเรียนไทย พ.ศ. 2554
(สถาบันราชานุกูล กรมสุขภาพจิต, 2554)

ตอนที่ 4 การสุ่มตัวอย่างมีผลกระทบต่อความแม่นยำของโพล

โพล หมายถึง การสำรวจความคิดเห็น ความเชื่อ เจตคติ หรือพฤติกรรมทางสังคมของสาธารณชนต่อเรื่องใดเรื่องหนึ่ง เช่น ด้านสังคมและวัฒนธรรม ด้านเศรษฐกิจ หรือด้านการเมือง ณ ช่วงเวลาใดเวลาหนึ่ง ดังนั้น ผลการสำรวจดังกล่าวจะสะท้อนความคิดเห็นของประชาชนต่อเรื่องนั้น ๆ ณ ช่วงเวลาที่ทำการสำรวจเท่านั้น เมื่อเวลาเปลี่ยนแปลงไปความคิดเห็นดังกล่าวอาจเปลี่ยนแปลงตามไปด้วย (วัฒนา สุนทรชัย, 2549, หน้า 102) ถ้าผู้สำรวจมีการกำหนดประเด็นเรื่องที่ต้องการสำรวจอย่างชัดเจน มีการสุ่มตัวอย่างที่เป็นตัวแทนประชากร มีการซักถามที่มีความเป็นกลาง ตลอดจนมีการวิเคราะห์ข้อมูลและการรายงานผลอย่างมีประสิทธิภาพ การทำโพลนั้นจะกลายเป็นงานวิจัยเชิงสำรวจที่มีมาตรฐาน (ศิริชัย กาญจนวาสี, 2556, หน้า 1)

1. เหตุการณ์สุ่มตัวอย่างที่ผิดพลาดในประเทศสหรัฐอเมริกา

ใน ปี ค.ศ.1824 บริษัทหนังสือพิมพ์ฮาริสเบอรก์ แพนซิลวาเนียน (Harrisburg Pensylvanian) ได้ส่งผู้สื่อข่าวออกไปสัมภาษณ์ประชาชน เพื่อสอบถามความนิยมต่อผู้สมัครรับเลือกตั้งประธานาธิบดีสหรัฐอเมริกา ซึ่งขณะนั้นเรียกกันว่า สตรอโพล (Straw poll) ในเวลาต่อมาหนังสือพิมพ์ฉบับอื่น ๆ ก็ได้ทำการสำรวจความนิยมของประชาชนตามบ้าง จนทำให้การทำโพลเป็นที่นิยมอย่างมาก และมีการทำโพลในด้านอื่น ๆ นอกจากด้านการเมือง ได้แก่ เศรษฐกิจ สังคม การศึกษา การกีฬา วัฒนธรรม ฯลฯ เช่น สาขาวิชายอดนิยมของนักศึกษามหาวิทยาลัย ดารายอดนิยมของประเทศ เป็นต้น (พรพรรณ วีระปรียากูร, 2548) สตรอโพลที่มีชื่อเสียงที่สุดในอเมริกา ซึ่งจัดทำโดยวารสารลิทเทอรารี่ ไดเจสท์ โดยการส่งไปรษณียบัตรลงคะแนนไปสู่ประชาชนเป็นจำนวนหลายล้านคนทั่วประเทศ เพื่อสำรวจความนิยมของผู้สมัครชิงตำแหน่งประธานาธิบดี ซึ่งวารสารลิทเทอรารี่ ไดเจสท์ สามารถทำนายผลการเลือกตั้งได้อย่างถูกต้องในปี ค.ศ. 1920, ค.ศ. 1924, ค.ศ. 1928 และ ค.ศ. 1932 (ธัญลักษณ์ เหลืองวิสุทธิ์, 2540, หน้า 5-8)

ในปี ค.ศ. 1936 วารสารลิทเทอรารี่ ไดเจสท์ ได้ส่งบัตรลงคะแนนจำนวน 10 ล้านใบตามรายชื่อในสมุดโทรศัพท์และทะเบียนรถยนต์ เพื่อสอบถามความนิยม แลนดอน (Landon) เป็นตัวแทนของพรรครีพับลิกัน และรูสเวลท์ (Roosvelt) ตัวแทนของพรรคเดโมแครต จากผลการสำรวจของวารสารลิทเทอรารี่ ไดเจสท์ ทำนายว่า แลนดอนจะชนะ 57% แต่ในปีนั้นมีโพลอีก 3 สำนัก ได้แก่ ครอสลีย์ (Crossley) กัลล์ลัป (Gallup) โรเพอร์ (Roper) ทำนายว่ารูสเวลท์ได้รับชัยชนะ ซึ่งผลการเลือกตั้งปรากฏว่ารูสเวลท์เป็นผู้ชนะ และวารสารลิทเทอรารี่ ไดเจสท์ ทำนายผิดพลาดไปถึง 19% ส่งผลให้วารสารลิทเทอรารี่ ไดเจสท์ ได้ปิดตัวลงในปีถัดมา ผลจากความล้มเหลวของสตรอโพลที่ดำเนินการโดยวารสารลิทเทอรารี่ ไดเจสท์ได้รับการวิเคราะห์และสรุปว่าผู้ที่มีโทรศัพท์หรือรถยนต์นั้นเป็นครอบครัวที่มีฐานะเป็นชนชั้นกลางขึ้นไป ซึ่งในยุคนั้นประชาชน

มีโทรศัพท์เพียงแค่ 33% ซึ่งคนที่มีฐานะทางเศรษฐกิจดีมีความนิยมแลนดอน แต่คนยากจนจากความตกต่ำทางเศรษฐกิจมีความชื่นชอบรูสเวลท์มากกว่า หรือกล่าวได้อีกอย่างหนึ่งว่ากลุ่มตัวอย่างของสตรอโพลของไดเจสท์นั้นไม่เป็นตัวแทนประชากรที่ดี (ธัญลักษณ์ เหลืองวิสุทธิ์, 2540, หน้า 5-8)

ในการเลือกตั้งปี ค.ศ. 1948 เกิดประวัติศาสตร์ซ้ำรอยเหมือนกันกับสตรอโพลของไดเจสท์ เมื่อสำนักโพลระดับชาติทั้ง 3 สำนัก ได้แก่ ครอสลีย์ กัลล์ลัป โรเพอร์ ได้ทำนายผลการเลือกตั้งผิดพลาด กล่าวคือ สำนักโพลต่าง ๆ ได้ทำนายว่า ดิวอี้ (Dewey) จะได้รับชัยชนะเหนือทรูแมน (Truman) แต่ผลการเลือกตั้งปรากฏว่า ประธานาธิบดีทรูแมนได้รับชัยชนะ เมื่อผลของโพลตรงข้ามกับความเป็นจริง ทำให้นักหนังสือพิมพ์เกิดวิกฤติศรัทธาต่อโพลและผู้จัดทำโพล จากวิเคราะห์ความผิดพลาดของโพล พบว่า กลุ่มตัวอย่างที่ได้มาจากฐานคะแนนของพรรครีพับลิกันมีมากเกินไป และช่วงเวลาในการเก็บข้อมูลห่างจากวันเลือกตั้งเป็นระยะเวลายาวนาน ทำให้มีการแก้ไขปรับปรุงและพัฒนาเทคนิคการสุ่ม และให้ความสำคัญกับช่วงเวลามากขึ้น โดยมีการทำโพลในหลาย ๆ ช่วง เพื่อดูแนวโน้มการตัดสินใจของประชาชนว่ามีการเปลี่ยนแปลงอย่างไร จนกระทั่งเป็นการทำโพลในวันเลือกตั้งที่เรียกว่า เอ็กซิท โพล (Exit poll) ประกอบกับนำเทคโนโลยีมาใช้ในการประมวลผล เช่น โทรศัพท์ คอมพิวเตอร์ เป็นต้น (ธัญลักษณ์ เหลืองวิสุทธิ์, 2540, หน้า 5-8)

2. เหตุการณ์สุ่มตัวอย่างที่ผิดพลาดในประเทศไทย

จากการติดตามการทำโพลล์เลือกตั้งของไทยนั้น ในบางครั้งการทำโพลของแต่ละสำนักทำโพลเพื่อทำนายผลการเลือกตั้ง ในช่วงเวลาเดียวกัน แต่ผลการสำรวจมีความแตกต่างกันอย่างมาก ทั้งนี้เกิดจากมาตรฐานในการทำโพลที่แตกต่างกัน ทำให้ผลการสำรวจมีความแม่นยำแตกต่างกันไป (ศิริชัย กาญจนวาสี, 2556, หน้า 10) ยกตัวอย่าง เช่น ศูนย์วิจัยกรุงเทพโพล และสำนักวิจัยเอแบคโพล ได้ทำโพลในช่วงเวลาใกล้เคียงกัน (วัฒนา สุนทรชัย, 2549, หน้า 105) ผลโพลปรากฏดังนี้

ศูนย์วิจัยกรุงเทพโพล สำรวจประชาชนในเขตกรุงเทพ จำนวนตัวอย่างทั้งสิ้น 1,414 คน (ความคลาดเคลื่อนไม่เกิน $\pm$ 3% ณ ระดับความเชื่อมั่น 95%) ระหว่างวันที่ 9-15 ธันวาคม พ.ศ. 2548 เรื่อง บุคคลในสังคมไทยที่มีพฤติกรรมน่าชื่นชมยกย่องมากที่สุด (วัฒนา สุนทรชัย, 2549, หน้า 105) ผลเป็นดังนี้

อันดับที่ 1 พ.ต.ท.ทักษิณ ชินวัตร 31.1%

อันดับที่ 2 นายสนธิ ลิ้มทองกุล 15.7%

อันดับที่ 3 นายอภิรักษ์ โกษะโยธิน 5.9%

สำนักวิจัยเอแบคโพล สำรวจประชาชนใน 32 จังหวัดทั่วประเทศ จำนวนตัวอย่างทั้งสิ้น 14,538 คน (ความคลาดเคลื่อนไม่เกิน ± 1% ณ ระดับความเชื่อมั่น 95%) ระหว่างวันที่ 9-24 ธันวาคม พ.ศ. 2548 เรื่องบุคคลแห่งปี (วัฒนา สุนทรชัย, 2549, หน้า 105) ผลเป็นดังนี้

อันดับที่ 1 พ.ต.ท.ทักษิณ ชินวัตร 31.5%

อันดับที่ 2 นายสนธิ ลิ้มทองกุล 15.4%

อันดับที่ 3 นายอภิสิทธิ์ เวชชาชีวะ 10.5%

จากการสำรวจความนิยมของโพลทั้งสองพบว่าอันดับที่ 3 มีความแตกต่างกัน ซึ่งสามารถวิเคราะห์ได้ว่า นายอภิรักษ์ โกษะโยธิน เป็นผู้ว่าราชการกรุงเทพมหานครได้ทำประโยชน์ให้กับคนท้องถิ่นจึงทำให้ได้คะแนนมากกว่า แต่ในการสำรวจข้อมูลทั่วประเทศนั้น นายอภิสิทธิ์ เวชชาชีวะเป็นหัวหน้าพรรคการเมืองระดับประเทศ ทำให้ได้รับความนิยมมากกว่า ดังนั้นกลุ่มตัวอย่างที่ได้จากกรุงเทพไม่สามารถเป็นตัวแทนประชากรของประเทศไทย

สรุปได้ว่าการเลือกวิธีการสุ่มตัวมีความสำคัญและมีผลต่อความถูกต้องแม่นยำในการสำรวจความคิดเห็นของประชาชน ซึ่งศูนย์วิจัยกรุงเทพโพลทำโพลในหัวข้อ เรื่องระดับประเทศ (เรื่อง บุคคลในสังคมไทยที่มีพฤติกรรมน่าชื่นชมยกย่องมากที่สุด) แต่ใช้เมืองหลวงเป็นตัวแทนเพื่อเก็บรวบรวมข้อมูล ทำให้กลุ่มตัวอย่างที่ได้ไม่เป็นตัวแทนประชากรของประเทศไทย ทำให้ผลโพลมีความคลาดเคลื่อน ซึ่งแตกต่างจากศูนย์วิจัยเอแบคโพลใช้วิธีการสุ่มแบบแบ่งกลุ่ม โดยแบ่งประเทศไทยออกเป็นจังหวัดและสุ่มจังหวัดจำนวน 32 จังหวัด ทำให้ผลโพลมีถูกต้องแม่นยำกว่า

3. ขนาดตัวอย่างมีความสำคัญน้อยกว่าความเป็นตัวแทนที่ดีของประชากร

สิ่งหนึ่งที่คนทั่วไปมักใช้ในการประเมินความแม่นยำของโพลคือขนาดตัวอย่าง โดยคิดว่าถ้ากลุ่มตัวอย่างมีขนาดใหญ่จะทำให้โพลมีความแม่นยำมาก แต่ถ้ากลุ่มตัวอย่างมีขนาดเล็กผลของโพลน่าจะมีความแม่นยำต่ำ แนวคิดแบบนี้มีส่วนถูกเพียงบางส่วนเท่านั้น ที่แท้จริงแล้วสิ่งที่สำคัญมากกว่านั้นคือความเป็นตัวแทนที่ดีของประชากร (วัฒนา สุนทรชัย, 2549, หน้า 102) แนวความคิดที่ว่าการสุ่มตัวอย่างจำนวนมาก ๆ เป็นแสนคนหรือล้านคน หรือเก็บข้อมูลจากทุกคนนั้นน่าจะทำให้โพลมีความแม่นยำสูง เนื่องจากการเก็บข้อมูลจากประชากรทั้งหมดจะไม่พบปัญหาความคลาดเคลื่อนจากการสุ่มตัวอย่าง (Sampling errors) แต่ปัญหาใหม่ที่เกิดขึ้นตามมา คือ ปัญหาความคลาดเคลื่อนที่ไม่เกี่ยวข้องกับการสุ่มตัวอย่าง (Non-sampling errors) เพราะการใช้กลุ่มขนาดใหญ่ ทำให้ปริมาณงานในการจัดการข้อมูลมาก จนทำให้พนักงานป้อนข้อมูลที่ผิดพลาดลงรหัสผิดพลาด และไม่ซื่อสัตย์ต่อหน้าที่ (การมั่วข้อมูล) (นพดล กรรณิกา, 2551)

การสุ่มตัวอย่างที่ล้นเกินไปหรือมากเกินไป (Oversampling) เป็นปัญหาสำคัญในอดีตของประเทศสหรัฐอเมริกา เมื่อนักการเมืองอเมริกาเรียกร้องให้สำนักโพลใช้กลุ่มตัวอย่างขนาดใหญ่เพราะผู้มีสิทธิ์ลงคะแนนเสียงเลือกตั้งมีจำนวนร้อยล้านคน แต่สำนักโพลทำการเก็บข้อมูล

กลุ่มตัวอย่างเป็นจำนวนหลักพันคน จนเป็นเหตุให้นักการเมืองเรียกร้องให้ใช้กลุ่มตัวอย่างขนาดใหญ่ จึงทำให้วารสารลิทเทอรารี่ ไดเจสท์ เก็บข้อมูลจากประชาชนเป็นจำนวนหลายล้านคน โดยมุ่งเน้นให้กระจายตัวครอบคลุมทุกสาขาอาชีพ ทุกเพศทุกวัย แต่ผิดหลักการของการสุ่มที่ว่า “การให้โอกาสในการถูกเลือกอย่างเท่าเทียมกัน” ทำให้ทำนายผลการเลือกตั้งประธานาธิบดีผิดพลาดและนำไปสู่การปิดกิจการของบริษัท (นพดล กรรณิกา, 2551) นอกจากนี้ การขาดจรรยาบรรณในการทำโพลส่งผลให้โพลผิดพลาดไปจากความเป็นจริง กล่าวคือ นักวิจัยบางคนอ้างทฤษฎีในการสุ่มตัวอย่าง เช่น การใช้ขนาดกลุ่มตัวอย่างตามตารางของยามาเน่ (Yamane) ซึ่งขนาดตัวอย่างตามตารางยามาเน่นั้นเหมาะกับการสุ่มตัวอย่างจากบัญชีรายชื่อโดยตรง แต่การเก็บข้อมูลในทางปฏิบัติกลับใช้วิธีเลือกตัวอย่างแบบเจาะจง แบบโควต้า หรือแบบสะดวก โดยไม่มีฐานข้อมูลและกรอบบัญชีรายชื่อประชากรเป้าหมาย

ตอนที่ 5 แนวคิดในการสร้างแบบจำลอง

1. แนวคิดในการสร้างแบบจำลองการสุ่ม

วัตถุประสงค์ของการสร้างแบบจำลองแผนที่แสดงค่าตัวแปรเชิงพื้นที่ เพื่อใช้ในการเปรียบเทียบประสิทธิภาพของการสุ่มตัวอย่างทั้ง 5 วิธี ได้แก่ การสุ่มอย่างง่าย, การสุ่มแบบมีระบบ, การสุ่มแบบชั้นภูมิ, การสุ่มแบบแบ่งกลุ่ม และการสุ่มแบบกริด เนื่องจากการสุ่มแบบกริดที่ใช้งานด้านวิทยาศาสตร์ถูกปรับปรุงและพัฒนาให้ใช้ในด้านสังคมและพฤติกรรมศาสตร์ ดังนั้น การทดสอบประสิทธิภาพของการสุ่มแบบกริดจึงจำเป็นต้องสร้างแบบจำลองสถานการณ์เพื่อใช้ตรวจสอบว่าการสุ่มแบบกริดมีความเหมาะสมในการนำมาประยุกต์ใช้หรือไม่ อย่างไร โดยเหตุผลและแนวคิดของการสร้างมีรายละเอียดดังนี้

1. กำหนดแผนที่รูปสี่เหลี่ยมขนาด m x n (ขนาด m แถว และ n หลัก) โดยกำหนดให้ประชากรที่อาศัยอยู่ในแต่ละกริดเซลล์มีจำนวนเท่ากัน จากภาพที่ 2-22 จำนวนประชากรในแต่ละพื้นที่มีขนาดเท่ากันหรือ $\mathbf{N_{11}=N_{12}=N_{13}=...=N_{33}}$

N_{11}	N_{12}	N_{13}
N_{21}	N_{22}	N_{23}
N_{31}	N_{32}	N_{33}

ภาพที่ 2-22 แบบจำลองแผนที่ขนาด 3 x 3 ช่อง

2. กำหนดค่าตัวแปรให้แผนที่ โดยตัวเลขที่อยู่ในแต่ละกริดเซลล์หมายถึงระดับของทัศนคติหรือเจตคติของประชากร และกำหนดให้ประชากรในแต่ละกริดเซลล์มีความคิดเห็นเหมือนกันทุกคน ยกตัวอย่างเช่น

5	5	5
4	4	4
3	3	3

ภาพที่ 2-23 แบบจำลองแผนที่แสดงค่าตัวแปร z

จากภาพที่ 2-23 ตัวเลขที่อยู่ในแผนที่หมายถึงค่าตัวแปรเชิงพื้นที่ (ค่าตัวแปร z) ซึ่งเป็นระดับความพึงพอใจในการปฏิบัติงานของเจ้าหน้าที่บรรณเทาสาธารณภัย เมือง A โดยมีระดับคะแนนความพึงพอใจตั้งแต่ 1 คะแนน (น้อยที่สุด) จนถึง 5 คะแนน (มากที่สุด) ดังนั้นค่าตัวแปรในแต่ละช่องหมายถึงค่าเฉลี่ยความพึงพอใจในการทำงานของเจ้าหน้าที่ ฯ ของพื้นที่แถบนั้น

3. หาค่าพารามิเตอร์ของประชากร เพื่อใช้เป็นค่าอ้างอิงในการเปรียบเทียบประสิทธิภาพของการสุ่มด้วยวิธีต่าง ๆ และเมื่อคำนวณค่าพารามิเตอร์จากแบบจำลองที่สร้างขึ้น พบว่า $\mu = 4$ และ $\sigma^2 = 0.67$

4. ทำการสุ่มกริดเซลล์ด้วยวิธีการสุ่มทั้ง 5 วิธี เพื่อศึกษาค่าสถิติที่ได้จากแต่ละวิธี และจากแบบจำลองที่สร้างขึ้นมีประชากร 9 ช่อง ถ้ากำหนดให้กลุ่มตัวอย่างเท่ากับ 4 ช่อง สามารถประยุกต์ใช้วิธีการสุ่มแบบต่าง ๆ กับแบบจำลองที่สร้างขึ้นดังนี้

4.1 การสุ่มอย่างง่าย-การสุ่มตัวอย่างกริดเซลล์ 4 ช่องใด ๆ จากแบบจำลองที่สร้างขึ้นอาจใช้วิธีจับฉลาก โดยใช้กระดาษสี่เหลี่ยมจัตุรัสเขียนตัวเลขตามแบบจำลองและตัดกระดาษเป็น 9 ชิ้น อย่างละเท่า ๆ กัน จากนั้นหลับตาหยิบกระดาษขึ้นมา 4 ใบ เพื่อคำนวณค่าสถิติจากกลุ่มตัวอย่างที่ได้และจดบันทึกค่าสถิติเก็บไว้ ยกตัวอย่างเช่น สุ่มตัวอย่างได้ระดับความพึงพอใจเท่ากับ 4, 5, 4 และ 4 คะแนน ดังนั้นค่าเฉลี่ย ($\bar{Z}$) เท่ากับ 4.25 และค่าความแปรปรวน (S^2) เท่ากับ 0.25

4.2. การสุ่มแบบมีระบบ–การสุ่มแบบมีระบบที่ใช้ในการวิจัยด้านสังคมและพฤติกรรมศาสตร์ ณ ปัจจุบันมีแนวคิดในการสุ่มเป็นแบบ 1 มิติ เท่านั้น ถึงแม้ว่าการปฏิบัติงานสนามบนพื้นผิวโลกเป็น 2 มิติ แต่การเตรียมกรอบตัวอย่างการสุ่มยังคงเป็นมิติเดียว สำหรับกรณีแบบจำลองนี้สามารถเขียนกรอบตัวอย่างการสุ่มได้ดังนี้

$X_1 = 5, X_2 = 5, X_3 = 5, X_4 = 4, X_5 = 4, X_6 = 4, X_7 = 3, X_8 = 3$ และ $X_9 = 3$

เมื่อประชากรมี 9 หน่วย และต้องการจำนวนตัวอย่าง 4 หน่วย ดังนั้นช่วงห่างของชุดตัวอย่างเท่ากับ 9/4 = 2.25 หน่วย หรือประมาณ 2 หน่วย สมมุติว่าตัวแรกสุ่มได้ $X_2 = 5$ อีก 3 จำนวนที่เหลือได้แก่ $X_4 = 4, X_6 = 4$ และ $X_8 = 3$ ดังนั้นค่าสถิติของวิธีการแบบมีระบบ คือ $\bar{Z}$ เท่ากับ 4 และ S^2 เท่ากับ 0.67

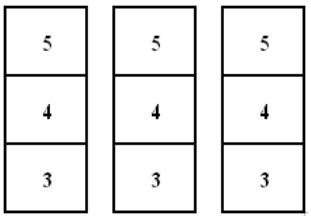
4.3 การสุ่มแบบชั้นภูมิ–หลักการของการสุ่มแบบชั้นภูมินั้นต้องแบ่งประชากรออกเป็นชั้น ๆ โดยแต่ละชั้นต้องพยายามให้มีความเป็นเอกพันธ์ให้ได้มากที่สุด จากแบบจำลองสามารถแบ่งออกเป็นชั้นได้ดังนี้

5	5	5
4	4	4
3	3	3

ภาพที่ 2-24 การแบ่งชั้นภูมิของแบบจำลอง

จากภาพที่ 2-24 ประชากรถูกแบ่งออกเป็น 3 ชั้น ในแต่ละชั้นมีความแปรปรวนเท่ากันคือ 0 ซึ่งแสดงให้เห็นว่าในแต่ละชั้นภูมิมีความเป็นเอกพันธ์อย่างสูง แต่โจทย์กำหนดให้สุ่มตัวอย่าง 4 หน่วย ซึ่งชั้นภูมิที่แบ่งได้มี 3 ชั้นภูมิ เพื่อไม่ให้เกิดความลำเอียงและอคติในการทดลองของตัวผู้วิจัย (ในการสร้างความได้เปรียบให้กับวิธีการสุ่มแบบกริดที่ตนเองสร้างขึ้น) จึงกำหนดให้สุ่มตัวอย่างในชั้นภูมิตรงกลาง 2 หน่วย ดังนั้นผลลัพธ์ที่ได้คือ 5, 4, 4 และ 3 ซึ่งมีค่าสถิติที่ได้จากแบบจำลอง $\bar{Z}$ เท่ากับ 4 และ S^2 เท่ากับ 0.67

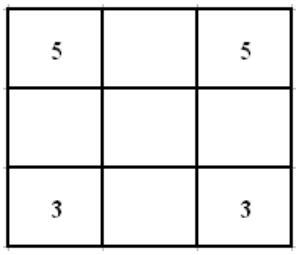
4.4. การสุ่มแบบแบ่งกลุ่ม–หลักการของวิธีการสุ่มแบบแบ่งกลุ่มนั้นต้องแบ่งประชากรออกเป็นกลุ่ม ๆ โดยความแปรปรวนภายในของแต่ละกลุ่มมีค่าสูง แต่ความแปรปรวนระหว่างกลุ่มมีค่าต่ำ ด้วยหลักการนี้เมื่อนำมาใช้กับแบบจำลองสามารถแบ่งกลุ่มตัวอย่างได้ดังนี้



ภาพที่ 2-25 การแบ่งกลุ่มของแบบจำลอง

จากภาพที่ 2-25 ประชากรถูกแบ่งออกเป็น 3 กลุ่ม แต่โจทย์กำหนดให้สุ่มตัวอย่าง 4 หน่วย ซึ่งปัญหานี้มีลักษณะคล้ายกับกรณีการสุ่มแบบแบ่งชั้นภูมิ เพื่อความเป็นมาตรฐานเดียวกัน กำหนดให้สุ่มตัวอย่างกลุ่มตรงกลาง 2 หน่วย และกลุ่มด้านซ้าย/ ขวาอย่างละ 1 หน่วย รวมเป็น 4 หน่วย เมื่อทำการสุ่มด้วยการจับฉลากได้ผลลัพธ์คือ 3, 5, 4 และ 3 ดังนั้น $\bar{Z}$ เท่ากับ 3.75 และ S^2 เท่ากับ 0.92

4.5 การสุ่มแบบกริด–แนวคิดของการสุ่มแบบกริดคือการสุ่มตัวอย่างที่ตำแหน่งต่าง ๆ บนแผนที่ โดยจุดกริดต้องกระจัดกระจายครอบคลุมพื้นที่ทั้งหมด ซึ่งแบบจำลองนี้เป็นแบบจำลองรูปร่างสี่เหลี่ยมจตุรัสขนาด 3 x 3 ช่อง ดังนั้นในแต่ละช่องต้องกระจายอยู่ตามทิศทั้งสี่ โดยมีค่า $\bar{Z}$ เท่ากับ 4 และ S^2 เท่ากับ 1.33 ดังภาพที่ 2-26 นี้



ภาพที่ 2-26 ตำแหน่งกริดเซลล์ที่ถูกสุ่มด้วยวิธีการสุ่มแบบกริด

5. ทำการเปรียบเทียบประสิทธิภาพการสุ่มตัวอย่างแต่ละวิธี โดยการนำค่าสถิติที่ได้ไปเทียบกับค่าพารามิเตอร์ของประชากร เช่น

$$e_{simple} = \left|\mu - \bar{Z}\right| = \left|4 - 4.25\right| = 0.25$$
$$\left|\Delta v_{simple}\right| = \left|\sigma^2 - S^2\right| = \left|0.67 - 0.25\right| = 0.42$$
$$e_{grid} = \left|\mu - \bar{Z}\right| = \left|4 - 4\right| = 0$$
$$\left|\Delta v_{grid}\right| = \left|\sigma^2 - S^2\right| = \left|0.67 - 1.33\right| = 0.66$$

จากผลลัพธ์ของการสุ่มอย่างง่ายและการสุ่มแบบกริด เมื่อเปรียบเทียบกันพบว่าวิธีการสุ่มตัวอย่างแบบกริดให้ประสิทธิภาพสูงกว่า เนื่องจากมีค่าเฉลี่ยของกลุ่มตัวอย่างเท่ากับค่าเฉลี่ยของประชากร อย่างไรก็ตามผลลัพธ์ที่ได้จากการสุ่มตัวอย่างอย่างง่าย มิใช่ว่าจะไม่ดีเสมอไป ซึ่งเป็นไปได้ว่าการสุ่มตัวอย่างอย่างง่ายในบางครั้ง อาจได้ผลลัพธ์ที่ดีกว่าการสุ่มตัวอย่างแบบกริด ยกตัวอย่าง เช่น ถ้ากลุ่มตัวอย่างที่ได้จากวิธีการสุ่มอย่างง่ายมีค่าเท่ากับ 5, 4, 3 และ 4 ดังนั้น $\bar{Z}$ เท่ากับ 4 และ S^2 เท่ากับ 0.67 และทำให้ประสิทธิภาพของการสุ่มอย่างง่ายเปลี่ยนแปลงไปดังนี้

$$e_{simple} = \left|\mu - \bar{Z}\right| = \left|4 - 4\right| = 0$$
$$\left|\Delta v_{simple}\right| = \left|\sigma^2 - S^2\right| = \left|0.67 - 0.67\right| = 0$$

เมื่อนำผลลัพธ์ที่เกิดขึ้นใหม่มาเปรียบเทียบประสิทธิภาพกับการสุ่มตัวอย่างแบบกริด พบว่า วิธีการสุ่มอย่างง่ายมีประสิทธิภาพในการสุ่มมากกว่าหรือกล่าวได้ว่ากลุ่มตัวอย่างที่ได้จากการสุ่มง่ายมีความเป็นตัวแทนของประชากรที่ดีกว่า เนื่องจากมีความแปรปรวนของกลุ่มตัวอย่างเท่ากับความแปรปรวนของประชากร

6. จากผลการเปรียบเทียบประสิทธิภาพในข้อที่ 5 ทำให้ทราบว่าจำนวนของการสุ่มตัวอย่างเพียงครั้งเดียวหรือน้อยครั้ง อาจทำให้เกิดข้อผิดพลาดในงานวิจัย ดังนั้นในการวิจัยเชิงทดลองครั้งนี้ได้ออกแบบให้มีจำนวนครั้งของการสุ่มตัวอย่างและจำนวนครั้งของการเปรียบเทียบเป็นจำนวนมาก เพื่อให้จำนวนครั้งของการสุ่มเข้าใกล้ค่าอนันต์ ($n \longrightarrow \infty$) โดยกำหนดให้จำนวนครั้งของการสุ่ม 10,000 ครั้ง ซึ่งแบ่งเป็นการทดลอง 10 ครั้ง และในแต่ละครั้งของการทดลองมีการสุ่มตัวอย่างจำนวน 1,000 ครั้ง

7. เนื่องด้วยเหตุผลตามข้อที่ 6 ทำให้ต้องมีการสุ่มทั้ง 5 วิธี (ได้แก่ การสุ่มอย่างง่าย, การสุ่มแบบมีระบบ, การสุ่มแบบชั้นภูมิ, การสุ่มแบบแบ่งกลุ่ม และการสุ่มแบบกริด) เป็นจำนวนมากถึง 10,000 ครั้ง ผลที่ตามมา คือ ความยุ่งยากในการวิจัย เนื่องจากต้องใช้แรงงานคนในการสุ่มและหาค่าสถิติเป็นจำนวนหลายหมื่นครั้ง ดังนั้นผู้วิจัยได้เขียนโปรแกรมการสุ่มตัวอย่างสำหรับแบบจำลอง เพื่อลดภาระในด้านแรงงาน นอกจากเหตุผลนี้ยังมีเหตุผลที่สำคัญอีกประการหนึ่ง คือ อาจมีการทักท้วงในเรื่องอคติหรือความลำเอียงของผู้วิจัย และทำให้ความน่าเชื่อถือของงานวิจัยลดลง ซึ่งการใช้งานโปรแกรมคอมพิวเตอร์แทนมนุษย์สามารถแก้ไขปัญหาเรื่องนี้ได้

8. เพื่อเพิ่มความน่าเชื่อถือของการทดลอง ผู้วิจัยได้สร้างแบบจำลองหลายรูปแบบหลายขนาด เพื่อศึกษาผลกระทบที่เกิดขึ้นจากแบบจำลอง เมื่อมีการเปลี่ยนแปลงทิศทางและขนาด ความสัมพันธ์ระหว่างค่าตัวแปรกับพิกัดในแผนที่ ในการทดลองครั้งนี้ได้สร้างแบบจำลอง จำนวน

6 แบบ ได้แก่ S0, S00, S1, S2, S3 และ S4 ซึ่งแบบจำลอง S0 และ S00 เป็นแบบจำลองกรณีพิเศษ เนื่องจากแบบจำลองทั้ง 2 นี้เป็นการจำลองสถานการณ์ในกรณีที่ธรรมชาติของกลุ่มประชากรไม่มีสหสัมพันธ์เชิงพื้นที่ ผลลัพธ์ที่เกิดขึ้นจากแบบจำลอง S0 และ S00 เป็นตัวบ่งชี้ประสิทธิภาพของการสุ่มแบบกริด สำหรับกรณีที่ธรรมชาติของกลุ่มประชากรไม่เป็นไปตามข้อตกลงเบื้องต้น

9. เมื่อทดลองสุ่มตัวอย่างครบทุกวิธีให้นำผลการทดลองพล็อตกราฟ เพื่อเปรียบเทียบประสิทธิภาพการสุ่มของทั้ง 5 วิธี ซึ่งการนำเสนอข้อมูลด้วยกราฟช่วยในการพิจารณาและตัดสินได้ง่าย เพราะมองเห็นภาพได้ชัดเจน

2. แนวคิดในการสร้างแบบจำลองวิธีการประมาณค่าเชิงพื้นที่

แบบจำลองการประมาณค่าเชิงพื้นที่มีลักษณะคล้ายกันกับแบบจำลองการสุ่มเกือบทุกประการ ทั้งในด้านหลักการ แนวคิด และวิธีการสร้าง แต่ความแตกต่างระหว่างแบบจำลองทั้ง 2 ประเภท คือ การนำแบบจำลองไปใช้งาน โดยแบบจำลองการสุ่มมีประโยชน์ในด้านการเปรียบเทียบประสิทธิภาพการสุ่มตัวอย่างทั้ง 5 วิธี แต่วัตถุประสงค์ในการสร้างและใช้งานแบบจำลองการประมาณค่าเชิงพื้นที่คือ การแสดงภาพ 3 มิติ ในการจำลองสถานการณ์การสุ่มแบบกริด เพื่อศึกษาค่าคลาดเคลื่อนร้อยละ (% Error) ที่เกิดจากการประมาณค่าเชิงพื้นที่ แนวคิดในการสร้างแบบจำลองมีรายละเอียด ดังนี้

1. กำหนดขนาดแผนที่รูปสี่เหลี่ยมขนาด m x n (ขนาด m แถว n หลัก) และกำหนดค่าตัวแปรในแต่ละช่อง โดยค่าตัวแปรในแต่ละช่องหมายถึง ระดับของทัศนคติหรือเจตคติของประชากรในท้องถิ่นหรือพื้นที่ ยกตัวอย่างแบบจำลองดังภาพที่ 2-27 ดังนี้

5	4	3
4	4	3
3	3	3

ภาพที่ 2-27 แบบจำลองแผนที่แสดงค่าตัวแปร z

2. สุ่มตัวอย่างกริดเซลล์จากแผนที่แสดงค่าตัวแปรเชิงพื้นที่เป็นจำนวน 4 ช่อง ด้วยวิธีการสุ่มแบบกริด ดังนั้นกริดเซลล์ที่ถูกสุ่มออกมาต้องกระจัดกระจายและมีระยะห่างเท่า ๆ กันทั้งในแนวตั้งและแนวนอน ดังภาพที่ 2-28

5		3
3		3

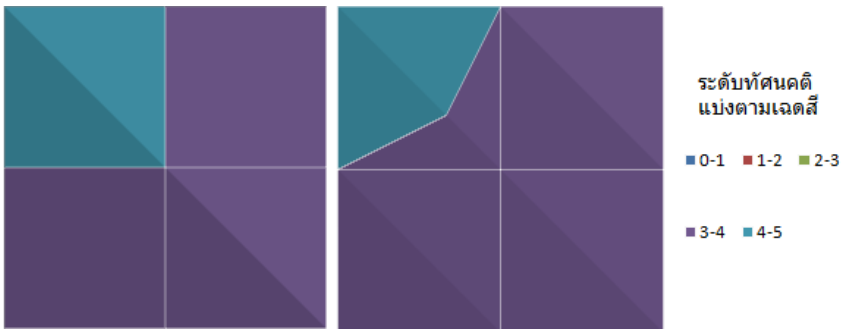
ภาพที่ 2-28 ตำแหน่งกริดเซลล์ที่ถูกสุ่มด้วยวิธีการสุ่มแบบกริด

3. ใช้วิธีการประมาณค่าเชิงพื้นที่กับแบบจำลองที่กำหนดขึ้นในข้อ 1 โดยใช้จุดกริดที่ทราบค่าตัวแปร (กริดเซลล์ที่ถูกสุ่มออกมา) ทำการประมาณค่าจุดกริดที่ไม่ทราบค่าตัวแปร (กริดเซลล์ที่ไม่ได้ถูกสุ่มออกมา) เมื่อกะประมาณค่าตัวแปร z จนครบทุกช่อง ผลลัพธ์ที่ได้ คือ แผนที่ที่เกิดจากการประมาณค่าเชิงพื้นที่ ดังภาพที่ 2-29

5	(5+3)/2	3
(5+3)/2	$\frac{5+3+3+3}{4}$	(3+3)/2
3	(3+3)/2	3

ภาพที่ 2-29 การประมาณค่าเชิงพื้นที่

4. นำแผนที่แสดงค่าตัวแปรเชิงพื้นที่ (หรือภาพที่ 2-30 ด้านซ้าย) กับแผนที่แสดงค่าประมาณเชิงพื้นที่ (หรือภาพที่ 2-30 ด้านขวา) มาเปรียบเทียบกัน เพื่อศึกษาประสิทธิภาพการสุ่มแบบกริด ในการเปรียบเทียบอาจใช้เฉดสีแทนค่าตัวแปร เพื่อช่วยสร้างความเข้าใจให้แก่บุคคลภายนอกอื่น ๆ นอกเหนือจากทีมนักวิจัย (เช่น ผู้รับผิดชอบโครงการหรือผู้มีส่วนได้ส่วนเสีย) สำหรับเครื่องมือวิจัยที่ใช้ในการแปลงค่าตัวแปรเป็นเฉดสี คือ โปรแกรม Microsoft excel ซึ่งเป็นโปรแกรมพื้นฐานและมีหลักการทำงานที่ง่ายต่อผู้ใช้งาน



ภาพที่ 2-30 แบบจำลองแสดงค่าตัวแปรเป็นเฉดสี

จากภาพที่ 2-30 แสดงให้เห็นว่าแผนที่ทั้ง 2 รูป เหมือนกันเกือบทั้งหมด ยกเว้นกริดเซลล์บริเวณตรงกลางเท่านั้น ถ้าแบบจำลองที่ยกตัวอย่างนี้เป็นแผนที่จริงและมีการเก็บข้อมูล ดังนั้นแผนที่แสดงค่าประมาณมีความเหมาะสมในการใช้งานและสามารถใช้งานได้จริง เพราะแผนที่ทั้ง 2 มีความแตกต่างกันน้อยมาก

5. การพิจารณาความถูกต้องและเหมาะสมของแผนที่ที่เกิดจากการประมาณค่าเชิงพื้นที่ สามารถคำนวณได้จากค่าคลาดเคลื่อนร้อยละ ซึ่งค่าคลาดเคลื่อนร้อยละนี้แสดงถึงความแตกต่างระหว่างค่าจริงกับค่าประมาณ โดยเทียบสัดส่วนเป็นร้อยละ

0%	0%	0%
0%	12.5%	0%
0%	0%	0%

ภาพที่ 2-31 การหาค่าคลาดเคลื่อนร้อยละ

ตามภาพที่ 2-31 เมื่อคำนวณค่าคลาดเคลื่อนร้อยละครบทุกช่องพบว่ามีกริดเซลล์ตรงกลางพื้นที่ (i = 2, j = 2) เท่านั้นที่แตกต่างจากกริดเซลล์อื่น ๆ ซึ่งค่าคลาดเคลื่อนร้อยละของจุดกริดนี้ไม่เท่ากับ 0 โดยมีรายละเอียดดังนี้

$$E = \frac{|e_{i'j'}|}{\mu_{i'j'}} \text{x} 100\% = \frac{|4-3.5|}{4} \text{x} 100\% = 12.5\%$$

นอกจากนี้ค่าคลาดเคลื่อนร้อยละสามารถแสดงผลในรูปแบบเฉดสี (เช่นเดียวกับแผนที่แสดงค่าตัวแปรและแผนที่แสดงค่าประมาณ) เพื่อก่อให้เกิดความเข้าใจและความชัดเจนในการทำรายงานและการนำเสนอข้อมูล

ตอนที่ 6 งานวิจัยในประเทศและต่างประเทศที่เกี่ยวข้อง

งานวิจัยต่างประเทศที่เกี่ยวข้อง

Dong-Sheng et al. (2011) ทำการศึกษาความหนาแน่นของการสุ่มที่ส่งผลต่อการตรวจหาสารคาร์บอนในดินที่อยู่ในรูปสารอินทรีย์ (Soil organic carbon : SOC) ตามมิติเชิงพื้นที่ ซึ่งการวิจัยที่เกี่ยวกับสารคาร์บอนในดินที่อยู่ในรูปสารอินทรีย์เป็นพื้นฐานของนโยบายของรัฐในด้านการบริหารจัดการเกษตรกรรมและและการป้องกันปัญหาสิ่งแวดล้อม ซึ่งประสิทธิในการตรวจหา SOC ขึ้นอยู่กับขนาดกลุ่มตัวอย่าง (จำนวนตำแหน่งของการสุ่มตัวอย่าง) โดยใช้ค่าสัมประสิทธิ์ของความแปรผัน (Coefficient of variation: CV) เป็นตัวบ่งชี้ที่สำคัญของ SOC

พื้นที่การสุ่มตัวอย่างในการวิจัยมีขนาด 927 ตารางกิโลเมตร ซึ่งตั้งอยู่ที่เมือง Yujiang ประเทศจีน โดยมีพื้นที่ปศุสัตว์ 13% พื้นที่ป่าไม้ 38% และพื้นที่เพาะปลูกประมาณ 39% พืชผลทางการเกษตรที่สำคัญ คือ ข้าว ถั่ว มั่นฝรั่ง งา และต้น Oilseed rape และลักษณะของเนื้อดินโดยส่วนใหญ่เกินกว่า 90% เป็นดินแดง ซึ่งมีส่วนประกอบจาก หินทราย ตมทะเลแดง หินดินดาน และตะกอนลำน้ำ

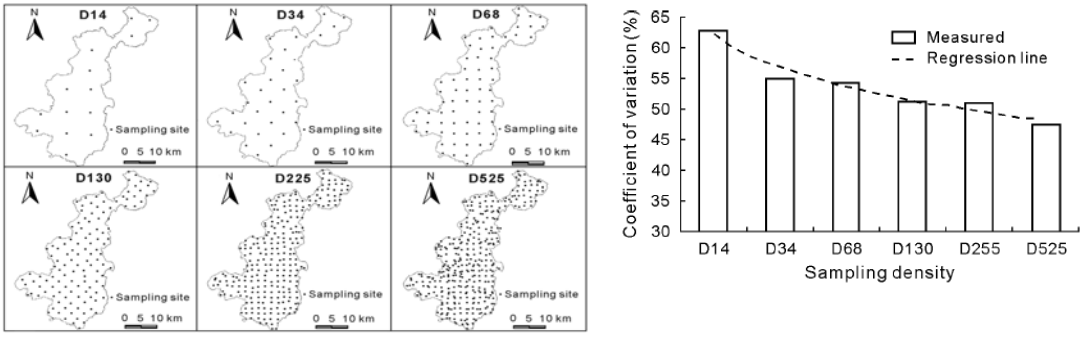
การสุ่มตัวอย่างใช้กริดขนาด 2 x 2 ตารางกิโลเมตร เพื่อศึกษาสภาพพื้นที่และประเภทของดิน โดยจำนวนจุดพิกัดในแต่ละกริดเซลล์นั้นจะมีจำนวนเท่าใดก็ขึ้นอยู่กับการใช้ประโยชน์ของพื้นที่ ในกรณีที่ 1 เมื่อภายในกริดเซลล์นั้นมีแบบแผนในการใช้ที่ดิน 1 แห่ง ซึ่งในที่แห่งนั้นมีการใช้ประโยชน์คิดเป็นเนื้อที่ 2/ 3 ของพื้นที่ทั้งหมด ให้สุ่มตัวอย่าง 1 จุดพิกัด กรณีที่ 2 เมื่อภายในกริดเซลล์นั้นมีแบบแผนในการใช้ที่ดิน 2 แห่ง ซึ่งในที่แต่ละแห่งมีการใช้ประโยชน์ที่ดินเกือบทั้งหมด ให้สุ่มตัวอย่าง 2 จุดพิกัด กรณีที่ 3 เมื่อภายในกริดเซลล์นั้นมีแบบแผนในการใช้ที่ดิน 3 แห่ง ซึ่งในที่แต่ละแห่งมีการใช้ประโยชน์ที่ดินทั้งหมดให้สุ่มตัวอย่าง 3 จุดพิกัด ซึ่งในแต่ละจุดพิกัดนั้น คือ ตำแหน่งของการเก็บตัวอย่างผิวหน้าดิน (ที่ระดับความลึก 0-20 เซ็นติเมตร) และทุกตำแหน่งของจุดพิกัดทำการอ้างตำแหน่งด้วยระบบ GPS (Global positioning system)

ในการทดลองของ Dong-Sheng et al. (2011) ได้กำหนดระดับความหนาแน่นของการสุ่ม 6 ระดับ ดังนี้

เมื่อขนาดของกริดเซลล์เท่ากับ 8 x 8 ตารางกิโลเมตร จะได้ตำแหน่งพิกัด 14 จุด และใช้สัญลักษณ์สำหรับความหนาแน่นของจุดพิกัดเมื่อแปรียบเทียบกับพื้นที่คือ D14 เมื่อขนาดของ

กริดเซลล์เท่ากับ 4 x 8 ตารางกิโลเมตร จะได้ตำแหน่งพิกัด 34 จุด และใช้สัญลักษณ์ คือ D34 เมื่อขนาดของกริดเซลล์เท่ากับ 4 x 4 ตารางกิโลเมตร จะได้ตำแหน่งพิกัด 68 จุด และใช้สัญลักษณ์ คือ D68 เมื่อขนาดของกริดเซลล์เท่ากับ 2 x 4 ตารางกิโลเมตร จะได้ตำแหน่งพิกัด 130 จุด และใช้สัญลักษณ์ คือ D130 เมื่อขนาดของกริดเซลล์เท่ากับ 2 x 2 ตารางกิโลเมตร จะได้ตำแหน่งพิกัด 255 จุด และใช้สัญลักษณ์ คือ D255 เมื่อขนาดของกริดเซลล์เท่ากับ 2 x 2 ตารางกิโลเมตร และมีการใช้ประโยชน์ของพื้นที่ในแบบกรณีต่าง ๆ (ซึ่งมีทั้งหมด 3 กรณี) จะได้ตำแหน่งพิกัด 525 จุด และใช้สัญลักษณ์ คือ D525

เมื่อพิจารณาจุดพิกัดตามระดับความหนาแน่นของการสุ่ม 6 ระดับ พบว่า แบบ D14 มีจุดพิกัดแบ่งเป็น พื้นที่เพาะปลูก 5 จุด พื้นที่ปศุสัตว์ 5 พื้นที่ป่าไม้ 4 แบบ D34 มีจุดพิกัดแบ่งเป็น พื้นที่เพาะปลูก 18 จุด พื้นที่ปศุสัตว์ 9 จุด พื้นที่ป่าไม้ 7 จุด, แบบ D68 มีจุดพิกัดแบ่งเป็น พื้นที่เพาะปลูก 36 จุด พื้นที่ปศุสัตว์ 19 จุด พื้นที่ป่าไม้ 13 จุด, แบบ D130 มีจุดพิกัดแบ่งเป็น พื้นที่เพาะปลูก 64 จุด พื้นที่ปศุสัตว์ 40 จุด พื้นที่ป่าไม้ 26 จุด, แบบ D130 มีจุดพิกัดแบ่งเป็น พื้นที่เพาะปลูก 115 จุด พื้นที่ปศุสัตว์ 87 จุด พื้นที่ป่าไม้ 53 จุด, และแบบ D130 มีจุดพิกัดแบ่งเป็นพื้นที่เพาะปลูก 294 จุด พื้นที่ปศุสัตว์ 152 จุด พื้นที่ป่าไม้ 79 จุด ซึ่งรายละเอียดของจำนวนและตำแหน่งจุดพิกัดเป็นไปตามภาพที่ 2-32 (ก)



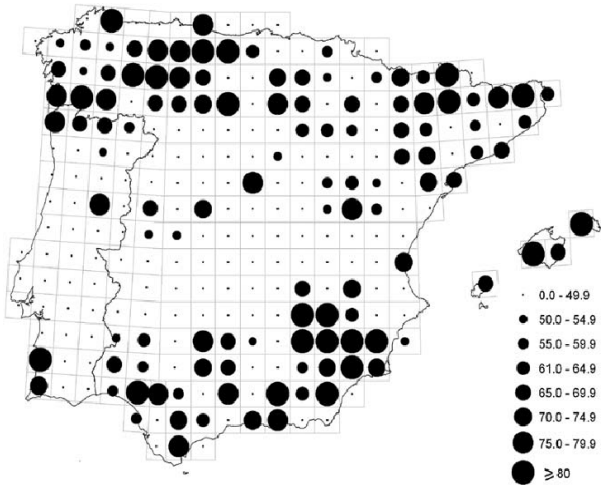
(ก) (ข)

ภาพที่ 2-32 ผลการทดลองที่ได้จากรูปแบบการสุ่ม 6 วิธี (Dong-Sheng et al., 2011)

จากภาพที่ 2-32 (ข) แสดงให้เห็นว่าค่า CV ของ SOC มีค่าอยู่ในช่วง 47.4%-62.8% เมื่อระดับความหนาแน่นเป็นแบบ D525 จะทำให้ค่า CV ของ SOC มีค่าน้อยที่สุด และเมื่อระดับความหนาแน่นเป็นแบบ D14 จะทำให้ค่า CV ของ SOC มีค่ามากที่สุด จากผลการทดลองวิเคราะห์ได้ว่าเมื่อเพิ่มระดับความหนาแน่นของการสุ่มจะทำให้ค่า CV ของ SOC มีค่าลดลง ซึ่งผลการทดลองที่ได้มีสอดคล้องกับงานวิจัยเรื่องพื้นที่ขนาดเล็กของ Onofiok (1993) เพราะการเพิ่มจำนวนของ

จุดพิกัดจะทำให้กริดเซลล์มีขนาดเล็กลง ซึ่งการเปลี่ยนระดับความหนาแน่นจากแบบ D14 มาเป็น D525 จะทำให้จำนวนจุดพิกัดต่อพื้นที่เปลี่ยนจาก 1.5 จุด/ 100 ตารางกิโลเมตร มาเป็น 57 จุด/ 100 ตารางกิโลเมตร โดยที่ค่า CVของ SOC ลดลง 15.4%.

Sánchez-Fernández, Lobo, Abellán, and Millán (2011) ทำการศึกษาโมเดลการทำนายความอุดมสมบูรณ์ของตัวด้วงน้ำ (Water beetles) ที่อาศัยอยู่ในเขตแหลมไอบีเรีย พื้นที่ที่ใช้ในการทดลอง คือ คาบสมุทรไอบีเรียและหมู่เกาะแบลีแอริก มีพื้นที่ครอบคลุม 585,644 ตารางเมตร ซึ่งเป็นพื้นที่ที่มีความหลากหลายทางชีวนิเวศ สภาพอากาศ และชนิดดิน โดยเฉพาะด้วงน้ำซึ่งเป็นสิ่งมีชีวิตเฉพาะถิ่นและมีจำนวนหลากหลายสายพันธ์ ข้อมูลของด้วงน้ำที่คาบไอบีเรียได้มาจากฐานข้อมูล ESACIB ฐานข้อมูลนี้เป็นฐานข้อมูลขนาดใหญ่ที่มีความสมบูรณ์ครบถ้วน ซึ่งมีการบันทึกข้อมูล มากกว่า 5000 ครั้ง และมีข้อมูลด้วงน้ำ 510 สปีชีส์ ฐานข้อมูล ESACIB มีขนาดกริดเซลล์เท่ากับ 10 x 10 ตารางกิโลเมตร แต่งานวิจัยของ Sánchez-Fernández et al. (2011) ได้ใช้ระบบพิกัดกริดแบบ UTM ขนาดกริดเซลล์เท่ากับ 50 x 50 ตารางกิโลเมตร จำนวนจุดกริดเท่ากับ 257 จุด ดังภาพที่ 2-33



ภาพที่ 2-33 แผนที่แสดงปริมาณตัวด้วงน้ำในเขตแหลมไอบีเรีย (Sánchez-Fernández et al., 2011)

ตัวแปรสิ่งแวดล้อมที่ใช้ทำนายความอุดมสมบูรณ์ของสปีชีส์มีอยู่ทั้งหมด 18 ตัวแปร โดยแบ่งเป็นตัวแปรสภาพอากาศสำหรับพื้นที่ขนาด 2,500 ตารางกิโลเมตร ซึ่งมีจำนวนเท่ากับ 9 ตัวแปร ได้แก่ อุณหภูมิต่ำสุดเฉลี่ยประจำเดือน, อุณหภูมิสูงสุดเฉลี่ยประจำเดือน, อุณหภูมิเฉลี่ยประจำปี, ปริมาณน้ำฝนเฉลี่ยประจำปี, หยาดน้ำฟ้าในฤดูร้อน, ร้อยละของเวลาสุริยะเฉลี่ยประจำปี, สภาพกรด, พิสัยของการเปลี่ยนแปลงอุณหภูมิประจำปี และหยาดน้ำฟ้าประจำปี ตัวแปรสภาพ

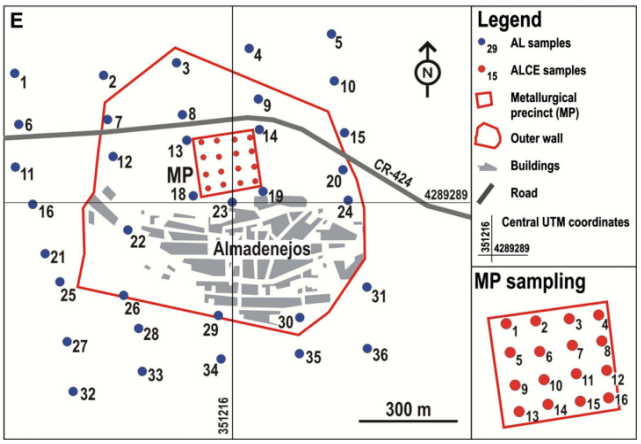
ภูมิประเทศมีจำนวนเท่ากับ 4 ตัวแปร ได้แก่ ค่าต่ำสุดของความสูงจากระดับน้ำทะเล, ค่าสูงสุดของความสูงจากระดับน้ำทะเล, ค่าเฉลี่ยของความสูงจากระดับน้ำทะเล และพิสัยของความสูงจากระดับน้ำทะเลปานกลาง ตัวแปรวิทยาหินมี 4 ตัวแปร ได้แก่ ร้อยละของพื้นที่ ๆ เป็นดินเหนียว, ร้อยละของพื้นที่ ๆ เป็นดินเนื้อปูน, ร้อยละของพื้นที่ ๆ เป็นทราย และร้อยละของพื้นที่ ๆ เป็นดินผสม และตัวแปรตัวสุดท้ายคือขนาดของผิวหน้าของพื้นน้ำ ซึ่งข้อมูลตัวแปรทั้งหมดได้จากฐานข้อมูล EDIT-geoplatform และเมื่อจำแนกกลุ่มตัวแปรโดยใช้เทคนิค PCA ด้วยการหมุนแกนแบบ Varimax พบว่า มีปัจจัย 5 ด้านที่มีค่า Eigenvalues มากกว่า 1 ได้แก่ ปัจจัยที่ 1 (Factor 1: F1) คือ ปัจจัยที่มีความสัมพันธ์ทางบวกต่อค่าเฉลี่ยระดับความสูง (น้ำหนักองค์ประกอบ = 0.91) และมีความสัมพันธ์ทางลบต่ออุณหภูมิต่ำสุดเฉลี่ยประจำเดือน (-0.90), ปัจจัยที่ 2 (F2) คือ ปัจจัยที่มีความสัมพันธ์ทางบวกต่อปริมาณน้ำฝนเฉลี่ยประจำปี (0.89), ปัจจัยที่ 3 (F3) คือ ปัจจัยที่มีความสัมพันธ์ทางลบต่อร้อยละของพื้นที่ ๆ เป็นดินเนื้อปูน (-0.76), ปัจจัยที่ 4 (F4) คือ ปัจจัยที่มีความสัมพันธ์ทางบวกต่ออุณหภูมิสูงสุดเฉลี่ยประจำเดือน (0.59) และปัจจัยที่ 5 (F5) คือ ปัจจัยมีความสัมพันธ์ทางบวกต่อร้อยละของพื้นที่ ๆ เป็นดินเหนียว (0.53)

ประสิทธิภาพของโมเดลการพยากรณ์พื้นที่ที่มีความอุดมสมบูรณ์ขึ้นอยู่กับตัวแปร ซึ่งค่าความสมบูรณ์ของฟังก์ชั่น Clench จะมีค่าระหว่าง 12.1% ถึง 57.5% ดังภาพที่ 2-23 เมื่อพิจารณาโมเดลที่ให้ค่า AIC น้อยที่สุดพบว่าให้ค่าความสมบูรณ์มากที่สุดเท่ากับ 80% ซึ่งโมเดลนั้นคือ

$$S = \mathrm{EXP}[0.11F1 - 0.01(F1)^2 - 0.03F2 + 0.04F3 + 0.05(F3)^2 - 0.07F4 + 0.05(F4)^2 - 0.08(F5)^2]$$

ซึ่งโมเดลนี้มีค่าเฉลี่ยร้อยละของความคลาดเคลื่อนเท่ากับ 29.9% (ช่วงความเชื่อมั่นที่ 95% มีค่าระหว่าง 23.13% ถึง 36.73%)

Martínez-Coronado, Oyarzun, Esbrí, Llanos, and Higueras (2012) ทำการศึกษาความเข้มข้นของปรอทบริเวณเหมืองแร่ The cerco de almadenejos (CDA) ในประเทศสเปน ซึ่งการศึกษาสภาพสิ่งแวดล้อมที่ถูกต้องนั้น ต้องให้ความสำคัญกับมลภาวะจากแหล่งอุตสาหกรรมเป็นหลัก ดังนั้น Martínez-Coronado et al. (2012) ได้ทำการเปรียบเทียบการปนเปื้อนระหว่างพื้นที่ถลุงแร่กับพื้นที่รอบนอก โดยการเก็บตัวอย่างจากดิน อากาศ และพืช เขตพื้นที่ถลุงแร่มีพื้นที่ขนาด 36,000 ตารางเมตร และสุ่มตัวอย่างออกมา 16 ตำแหน่ง ในส่วนเขตพื้นที่รอบนอกมีพื้นที่ขนาด 1,200,000 ตารางเมตร และสุ่มตัวอย่างออกมา 35 ตำแหน่ง (ภาพที่ 2-34) สำหรับพืชที่นำมาทดลองคือ ต้นหน่อไม้ฝรั่งจำนวน 13 ต้น ซึ่ง 10 ต้น นำมาจากเขตโรงงานและ 3 ต้นที่เหลือเอามาจากพื้นที่ภายนอก



ภาพที่ 2-34 ตำแหน่งที่ใช้ในการสุ่มตัวอย่างในเขต The cerco de almadenejos (Martínez-Coronado et al., 2012)

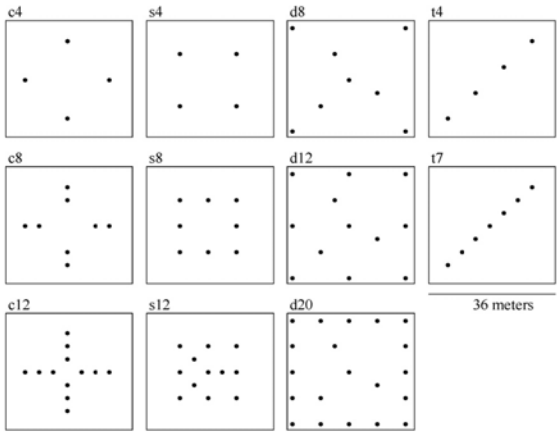
เมื่อเก็บตัวอย่างของดินและอากาศตามตำแหน่งต่าง ๆ ไปวิเคราะห์การปนเปื้อนของปรอท พบว่า พื้นที่ถลุงแร่มีค่าเฉลี่ยปริมาณปรอทในดิน 4,220 ug/g มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในดินเท่ากับ 4,926 ug/g มีค่าเฉลี่ยปริมาณปรอทในอากาศ 137 ng/m มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในอากาศ 230 ng/m ในส่วนพื้นที่รอบนอก มีค่าเฉลี่ยปริมาณปรอทในดิน 43 ug/g มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในดินเท่ากับ 40 ug/g มีค่าเฉลี่ยปริมาณปรอทในอากาศ 21 ng/m มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในอากาศ 21 ng/m

นอกจากนี้เมื่อวิเคราะห์สารปรอทจากต้นหน่อไม่ฝรั่งพบว่าต้นหน่อไม้ฝรั่งในพื้นที่ถลุงแร่มีค่าเฉลี่ยปริมาณปรอทในราก 96 ug/g และในใบ 30 ug/g มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในราก 151 ug/g และในใบ 49 ug/g ในส่วนพื้นที่รอบนอกมีค่าเฉลี่ยปริมาณปรอทในราก 0.7 ug/g และในใบ 0.7 ug/g มีส่วนเบี่ยงเบนเฉลี่ยปริมาณปรอทในราก 0.05 ug/g และในใบ 0.5 ug/g

Majasalmi et al. (2012) ทำการศึกษาแบบแผนการสุ่มที่เหมาะสมกับดัชนี LAI-2000 ของป่าเขตขั้วโลก ซึ่งดัชนีพื้นที่ใบไม้ (Leaf area index: LAI) คือพื้นที่ใบทั้งหมดของพืชหารด้วยพื้นที่ดินที่พืชปกคลุม ซึ่ง LAI มีความสัมพันธ์กับตัวแปรสภาพภูมิอากาศและ LAI เป็นตัวแปรที่แสดงความสัมพันธ์ระหว่างป่าไม้และสิ่งแวดล้อม เพราะใบไม้ทำให้เกิดปฏิกริยาการตรึงคาร์บอนไดออกไซด์และอัตราการผลิตปฐมภูมิสุทธิ (Net primary production: NNP) นอกจากนั้น LAI เป็นปัจจัยสำคัญที่ส่งผลต่อค่าสัดส่วนของการดูดซึมรังสีแสงอาทิตย์ (Absorbed photosynthetically active radiation: APAR) สำหรับอุปกรณ์ที่นำมาใช้ในการวิจัยคือ เครื่อง LAI-2000 PCA โดยการนำไปติดตั้งที่ยอดไม้ ซึ่งอุปกรณ์นี้เป็นออปติกเซนเซอร์ที่ใช้วัดการแผ่รังสีบนท้องฟ้า

พื้นที่การสุ่มตัวอย่างในการวิจัยมีขนาด 36 x 36 ตารางเมตร โดยแต่ละกริดเซลล์มีขนาด 4x4 ตารางเมตร ซึ่งวิธีการเก็บข้อมูลจะทำการวัดค่าบริเวณตรงกลางในแต่ละกริดเซลล์ตามพื้นราบ (Horizontal) และมีระยะตามแนวตั้ง (Vertical) สูงจากพื้นดิน 0.7 เมตร ดังนั้นอุปกรณ์ส่งสัญญาณที่ติดพื้นจะมีรัศมีการทำงาน 3.48 x (h-0.7) เมตร (ค่า 3.48 เป็นค่า Tangent ที่มากสุดของมุมของอุปกรณ์ส่งสัญญาณ)

ตัวแปรอิสระของงานวิจัยนี้ คือ ความหนาแน่นของการสุ่มตัวอย่างและตำแหน่งที่ใช้ในการเก็บข้อมูล เพราะความหนาแน่นของการสุ่มตัวอย่างมีความสัมพันธ์กับความถูกต้องของข้อมูล ได้แก่ จำนวน 4 จุด จำนวน 7 จุด จำนวน 8 จุด จำนวน 12 จุด และจำนวน 20 จุด ในส่วนตำแหน่งของการสุ่มนั้นต้องมีลักษณะรูปร่างที่สมมาตรและเป็นระบบ ได้แก่ แบบแนวขวาง แบบสี่เหลี่ยม แบบกระจาย และแบบตัดตามขวาง ทำให้เกิดกรณีของการสุ่มจำนวน 11 รูปแบบ ดังภาพที่ 2-35 แต่ในกรณีรูปแบบ d12 และ s12 ถูกออกแบบให้มีลักษณะรูปร่างโครงข่าย VALERI



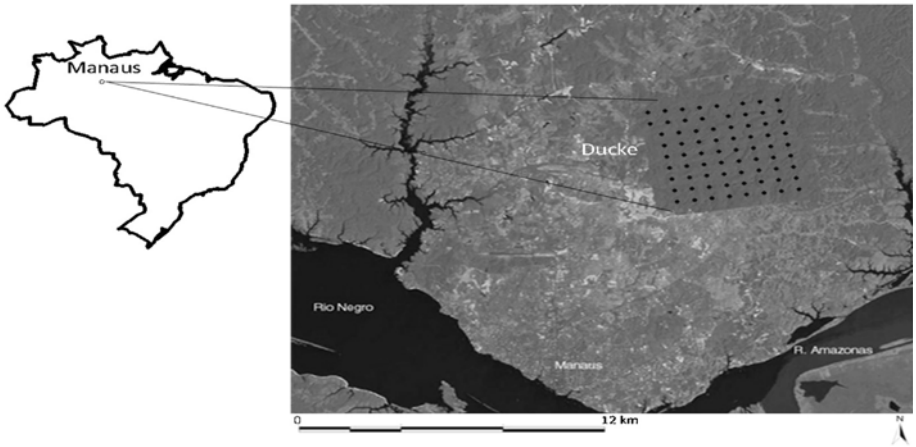
Scheme:	Averages:	
	s.e.	CV (%)
c4	0.10	4.31
c8	0.07	3.32
c12	0.06	2.83
s4	0.11	5.07
s8	0.08	3.63
s12	0.06	2.93
d8	0.09	4.08
d12	0.08	3.60
d20	0.06	2.56
t4	0.12	5.52
t7	0.10	4.63

หมายเหตุ ค่า s.e. หมายถึง ค่าความคลาดเคลื่อนมาตรฐานสมบูรณ์ และค่า CV(%) หมายถึง ค่าสัมประสิทธ์ความแปรผัน

ภาพที่ 2-35 ผลการทดลองที่ได้จากรูปแบบการสุ่ม 11 วิธี (Majasalmi, Rautiainen, Stenberg, & Rita, 2012)

จากผลการทดลองสรุปได้ว่าเมื่อความหนาแน่นของการสุ่มตัวอย่างอย่างมีค่ามากขึ้น จะทำให้ค่าความคลาดเคลื่อนมีค่าน้อยลง และตำแหน่งของการสุ่มมีผลต่อค่าความคลาดเคลื่อน กล่าวคือถ้าตำแหน่งที่สุ่มมีการกระจายตัวครอบคลุมพื้นที่มากขึ้น ค่าความคลาดเคลื่อนก็จะลดลง

Franklin et al. (2013) ทำการวิจัยสิ่งมีชีวิตขนาดเล็กจำพวกเห็บ/ ไร ที่ Reserve ducke ซึ่งเป็นพื้นที่ตรงกลางของป่าอะเมซอน ในประเทศบราซิล ซึ่ง Reserve ducke และใช้เทคนิคในการเก็บตัวอย่างด้วยวิธีการสุ่มแบบ RAPELD ซึ่งวิธีการสุ่มนี้มีหลักการว่าต้องทำการสร้างกริดขนาด 8 x 8 ตารางกิโลเมตร ภายในแต่ละกริดต้องสุ่มจุดกริดออกมา 72 จุดกริด โดยแต่ละจุดนั้นต้องห่างกันอย่างน้อย 1 กิโลเมตร (ภาพที่ 2-36) ซึ่ง Franklin et al. ได้เพิ่มรูปแบบการสุ่มอีกแบบหนึ่ง คือ การสร้างกริดขนาด 5 x 5 ตารางกิโลเมตร ภายในแต่ละกริดต้องสุ่มจุดกริดออกมา 30 จุดพิกัด และใน 30 จุดกริดนี้มี 16 จุดกริดที่ซ้อนทับกับกริดขนาดใหญ่ (72 จุดพิกัด) ซึ่งกริดขนาด 5 x 5 ตารางกิโลเมตร เป็นกริดขนาดมาตรฐานสำหรับพื้นที่อื่น ๆ ของป่าอะเมซอน



ภาพที่ 2-36 ตำแหน่งที่ใช้เก็บข้อมูลเห็บ/ ไรในพื้นที่ Reserve ducke ในป่าอะเมซอน (Franklin et al., 2013)

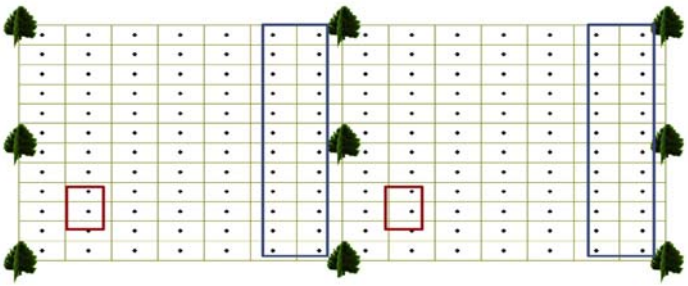
จากการเก็บข้อมูลพบว่ากลุ่มตัวอย่างตัวไรที่พบในกริดขนาดใหญ่มีไร 161 สายพันธุ์ และมี 79 สายพันธุ์ ที่ไม่มีความสัมพันธ์กันทางสปีชีส์ ในจำนวนตัวไรที่พบมีตัวไร 5 สปีชีส์ เท่านั้นที่มีจำนวนมากเกินกว่าร้อยละ 50 สำหรับกลุ่มตัวอย่างตัวไรที่พบในกริดขนาดเล็กทั้ง 16 จุดพิกัดมีค่าเฉลี่ย 112 สายพันธุ์ มีค่ามากที่สุด 121สายพันธุ์ และน้อยที่สุด 106 สายพันธุ์ จากผลการทดลองพบกริดขนาดใหญ่มีค่าเฉลี่ยของสายพันธุ์ที่ถูกค้นพบมากว่ากริดขนาดเล็กทำเท่ากับ 43 สปีชีส์ โดยที่จำนวนของสายพันธุ์ที่ถูกค้นพบระหว่างทั้งสองแบบมีความแตกต่างกัน 15 สปีชีส์

อย่างไรก็ตามเทคนิคการสุ่มแบบกริดขนาดเล็กมีข้อดีในการประหยัดค่าใช้จ่ายและลดระยะเวลาในการสำรวจ ซึ่งในปัจจุบันมีการเปลี่ยนแปลงทางสภาพภูมิอากาศและสิ่งแวดล้อมอย่างรวดเร็ว ดังนั้นการทำวิจัยจำเป็นต้องทำให้บ่อยครั้งมากขึ้น ดังนั้นวิธีการสุ่มนี้จึงมีความเหมาะสม

สำหรับการสำรวจข้อมูลมากกว่า เพราะเป้าหมายของการวิจัยคือต้องการศึกษาสภาพธรรมชาติและการอยู่อาศัยของตัวไร ไม่ใช่การค้นหาตัวไรสายพันธ์ที่หายาก

Lin, Li, Luo, Lin, and Li (2013) ได้ทำการศึกษาปริมาณน้ำยางกับแร่ธาตุในดิน โดยเลือกใช้พื้นที่สวนยางในจีน เนื่องจากประเทศจีนเป็นประเทศที่มีกำลังการผลิตยางธรรมชาติเป็นอันดับ 6 ของโลก ทำให้ต้นยางเป็นพืชเศรษฐกิจที่สำคัญของประเทศจีน ในแต่ละปีประเทศจีนต้องการยางธรรมชาติเป็นจำนวนมาก ซึ่งสวนยางขนาดใหญ่ตั้งอยู่ที่เกาะ Hainan ทางตอนใต้ของประเทศจีน การที่ต้นยางจะให้น้ำยางได้ดีขึ้นอยู่กับแร่ธาตุในดิน ในปัจจุบันมีการศึกษาโมเดลการให้ปุ๋ยแก่ต้นยางโดยการพิจารณาจากแร่ธาตุและคุณสมบัติของดิน

พื้นที่ใช้ในการทดลอง คือ สวนยางเมือง Yangjiang ตั้งอยู่ที่เกาะ Hainan ในประเทศจีน มีขนาดพื้นที่ 14 x 6 ตารางเมตร และความชันเฉลี่ย 4 องศา โดยที่ความสูงของพื้นที่ด้านตะวันออกสูงกว่าพื้นที่ด้านตะวันตก 0.5 เมตร ในพื้นที่ทดลองนี้มีต้นยางทั้งหมด 9 ต้น มีระยะระหว่างแถว 3 เมตร และระยะระหว่างคอลัมม์ 7 เมตร เมื่อกำหนดจุดกริดเท่ากับ 168 ตำแหน่ง จะได้กริดเซลล์รูปร่างสี่เหลี่ยมผืนผ้าขนาด 1 x 1.5 ตารางเมตร ตามภาพที่ 2-37 การเก็บตัวอย่างดินใช้เครื่องเจาะขนาดเส้นผ่าศูนย์กลางขนาด 8 ซ.ม. ที่ความลึก 0.2 เมตร จากนั้นนำตัวอย่างดินที่สุ่มได้มาทดสอบคุณสมบัติสารเคมีด้วยวิธีทดสอบแบบมาตรฐาน เพื่อศึกษาแร่ธาตุที่จำเป็นต่อต้นยางคือ Organic matter (OM), Total nitrogen (TN), Available phosphorus (AP), Available K (AK) และ pH ซึ่งระดับของTN, OM, AP และ AK มีความสัมพันธ์กับปริมาณน้ำยาง ดังนี้



ภาพที่ 2-37 ตำแหน่งที่ใช้เก็บตัวอย่างดินในสวนยางเมือง Yangjiang ประเทศจีน (Lin et al., 2013)

ถ้าดินชั้นบน (ดินที่ผิวหน้ามีความลึกไม่เกิน 20 ซ.ม.) มีค่า TN น้อยกว่า 0.8 g/ kg, ค่าOM น้อยกว่า 20 g/ kg, ค่าAP น้อยกว่า 5 mg/ kg และ ค่า AK น้อยกว่า 40 mg/ kg ต้นยางจะให้น้ำยางปริมาณน้อย

ถ้าดินชั้นบนมีค่า TN 0.8-1.4 g/ kg, ค่าOM 20-25 g /kg, ค่าAP 5-8 mg/ kg และ ค่า AK 40-60 mg/ kg ต้นยางจะให้น้ำยางปริมาณปานกลาง

ถ้าดินชั้นบนมีค่า TN มากกว่า 1.4 g/ kg ค่าOM มากกว่า 25 g/ kg, ค่าAP มากกว่า 8 mg/ kg และ ค่า AK มากกว่า60 mg/ kg ต้นยางจะให้น้ำยางมากและไม่จำเป็นต้องใส่ปุ๋ย

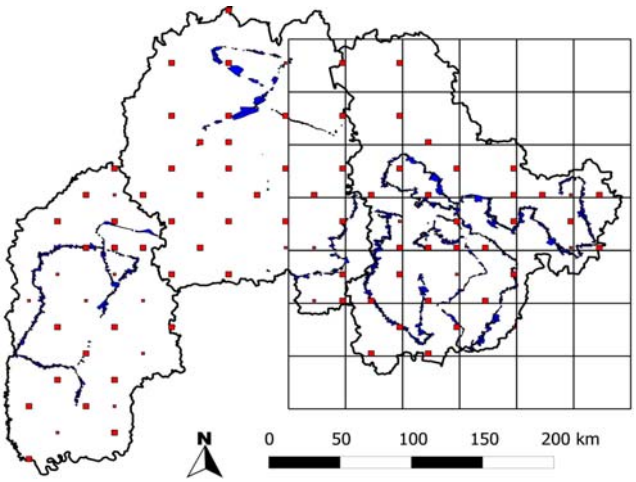
เมื่อตรวจสอบตัวอย่างดินที่สุ่มได้ พบว่า มีค่า pH อยู่ระหว่าง4.27-5.81 ค่าเฉลี่ย TN เท่ากับ 0.78 g/ kg และค่าเฉลี่ย OM เท่ากับ 14.18 g/ kg ซึ่งค่าทั้งสองนี้ต่ำกว่าขอบเขตล่างเล็กน้อย ค่าเฉลี่ยAK เท่ากับ 45.58 mg/ kg อยู่ในช่วงระดับปานกลาง ค่าเฉลี่ย AP เท่ากับ 8 mg/ kg ซึ่งมีค่าสูงกว่าขอบเขตบน และเมื่อพิจารณาค่าสัมประสิทธิ์ของความผันแปร(Coefficient of variation: CV) ของแร่ธาตุพบว่าค่า CV ของ pH มีค่าเพียงแค่ 5.82% แต่ค่า CV ของAP มีค่าสูงมาก (282.96%) ซึ่งมีค่า AP มีค่าระหว่าง 2 mg/ kg ถึง600 mg/ kg ในส่วนค่า CV ของ TN, AK and OM อยู่ในระดับปานกลาง (24.93-28.51%)

Gallego et al. (2014) ทำการประเมินประสิทธิภาพข้อมูลดาวเทียมในการประมาณค่าพื้นที่เพาะปลูกในประเทศยูเครน พื้นที่ที่ทำการศึกษาคือ เมืองKyivska, Khmelnytska และ Zhytomyrska ซึ่งมีพื้นที่รวมประมาณ 78,500 ตารางกิโลเมตร โดยแบ่งเป็น ข้าวสาลี 32% ข้าวบาเล่ย์ 21% ผัก 15% ข้าวโพด 12% ต้น Oilseed rape หัวบีท 6% ถั่วเหลือง 4 % ต้นทานตะวัน 2% โดยมีตำแหน่งที่ต้องสำรวจทั้งหมด 90 แห่ง แบ่งเป็น 34 แห่ง อยู่ในเขต Kyivska 26 แห่ง อยู่ในเขต Khmelnytska และ 30 แห่ง อยู่ในเขต Zhytomyrska โดยการสำรวจในแต่ละแห่งผู้สำรวจจะใช้ภาพถ่ายดาวเทียมเป็นเครื่องมือเพิ่มประสิทธิภาพ ซึ่งข้อมูลดาวเทียมที่นำมาใช้ได้แก่ MODIS, Landsat-5/ TM, AwiFS, LISS-III, และ RapidEye

การสุ่มตัวอย่างในงานวิจัยนี้ใช้วิธีการแบบแบ่งชั้นภูมิ โดยแบ่งออกเป็น 3 ชั้นภูมิ ได้แก่ ชั้นภูมิที่ 1 เป็นพื้นที่มีการทำเกษตรกรรมมากกว่า 50% ของพื้นที่ทั้งหมด ชั้นภูมิที่ 2 เป็นพื้นที่มีการทำเกษตรกรรมน้อยกว่า 50% ของพื้นที่ทั้งหมด และชั้นภูมิที่ 3 เป็นพื้นที่ที่ไม่มีการเพาะปลูก (ในพื้นที่นี้จะไม่มีการเก็บตัวอย่าง) และกำหนดกริดครอบคลุมพื้นที่ทั้งหมดไว้ โดยกริดเซลล์มีขนาด 40 x 40 ตารางกิโลเมตร ซึ่งภายในกริดเซลล์แต่ละอันจะมีแหล่งเพาะปลูกประมาณ 15-20 แห่ง และพนักงานสำรวจข้อมูลจะเก็บตัวอย่างในพื้นที่ขนาด 2x2 ตารางกิโลเมตร ตามภาพที่ 2-38

จากการเปรียบเทียบผลลัพธ์ที่ได้จากข้อมูลจากดาวเทียมและการสำรวจกลุ่มตัวอย่าง โดยใช้ค่าประสิทธิภาพสัมพัทธ์และค่าประสิทธิภาพราคา ซึ่งค่าประสิทธิภาพสัมพัทธ์เป็นตัวบ่งชี้ว่าภาพถ่ายดาวเทียมสามารถลดจำนวนครั้งของความคลาดเคลื่อนของการประมาณพื้นที่ และค่าประสิทธิภาพราคาเป็นตัวบ่งชี้ว่าภาพถ่ายดาวเทียมสามารถลดค่าใช้จ่ายในการสำรวจ จากการวิเคราะห์ พบว่า ค่าประสิทธิภาพสัมพัทธ์ของข้อมูลดาวเทียม เมื่อเรียงลำดับจากมากไปหาน้อยข้อมูลดาวเทียม MODIS เท่ากับ1.59 ข้อมูลดาวเทียม Landsat-5 เท่ากับ 1.54 และข้อมูลดาวเทียม AWIFS เท่ากับ 1.49 ในส่วนค่าประสิทธิภาพราคานั้น ข้อมูลดาวเทียม MODIS และข้อมูล

ดาวเทียม Landsat-5 มีค่าประสิทธิภาพราคาที่ยอมรับได้ แต่ข้อมูลดาวเทียม AWIFS, LISS-III, และ Rapid-Eye มีราคาที่แพงเกินไป



ภาพที่ 2-38 ตำแหน่งที่ใช้ในการเก็บข้อมูลในเมือง Kyivska, Khmelnytska และ Zhytomyrska (Gallego et al., 2014)

งานวิจัยในประเทศที่เกี่ยวข้อง

ประชา สุวัฒนพันธุ์กุล (2534) ทำการเปรียบเทียบประสิทธิภาพของการเลือกตัวอย่างแบบโควต้ากับการสุ่มตัวอย่างแบบชั้นภูมิ โดยใช้ อัตราส่วนของผลต่างของค่าเฉลี่ยตัวอย่าง (RDAM) เป็นเกณฑ์ในการเปรียบเทียบ โดยปัจจัยที่มีผลกระทบต่อประสิทธิภาพของการเลือกตัวอย่างทั้ง 2 วิธี ได้แก่

ขนาดตัวอย่าง จำนวนกลุ่ม ค่าสัมประสิทธิ์ความแปรผัน ระดับความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษากับปัจจัยที่ใช้ในการกำหนด ซึ่งผลของการศึกษาสามารถสรุปได้ว่า วิธีการสุ่มตัวอย่างแบบชั้นภูมิมีประสิทธิภาพดีกว่าวิธีการเลือกตัวอย่างแบบโควต้าในทุกกรณีที่ศึกษา แต่ในกรณีที่ค่าสัมประสิทธิ์ความแปรผันต่ำ และระดับความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษากับปัจจัยที่ใช้ในการกำหนดกลุ่มสูง อาจใช้วิธีเลือกตัวอย่างแบบโควต้าได้ ทั้งนี้เพราะสะดวกในการเก็บข้อมูลมากกว่าการสุ่มตัวอย่างแบบชั้นภูมิ

ดวงใจ ปวีณอภิชาต (2535) ทำการศึกษาเปรียบเทียบค่าประมาณมัชฌิมเลขคณิต และค่าประมาณความแปรปรวน ที่ได้จากการสุ่มแบบชั้นภูมิ โดยใช้แผนการสุ่มทั้งหมด 6 วิธี ซึ่งได้จากตัวแปรจำแนกชั้นภูมิที่แตกต่างกัน 3 ลักษณะ (ประเภทโรงเรียน ขนาดโรงเรียน และเ

ขตอำเภอ) กับวิธีกำหนดขนาดกลุ่มตัวอย่างย่อยที่แตกต่างกัน 2 ลักษณะ (วิธีกำหนดแบบนีย์แมน และวิธีกำหนดแบบสัดส่วน) แต่ละแผนการสุ่มใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่นต่างกัน 3 ระดับ คือ 90%, 95% และ 99% ประชากรที่ใช้ในการวิจัย คือ นักเรียนชั้นมัธยมศึกษาปีที่ 2 ในเขตท้องที่การศึกษาที่ 5 กรุงเทพมหานคร สังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ปีการศึกษา 2534 จำนวน 9,227 คน ดำเนินการสุ่มตัวอย่างในแต่ละวิธี ขนาด และระดับความเชื่อมั่น โดยใช้คอมพิวเตอร์กระทำการสุ่มซ้ำ1,000 ครั้ง เปรียบเทียบค่าประมาณของคะแนนผลสัมฤทธิ์วิชาคณิตศาสตร์ของประชากร โดยใช้เกณฑ์ในการเปรียบเทียบความมีประสิทธิภาพ 4 เกณฑ์ คือ ความใกล้เคียง ความแปรปรวน ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองและค่าประสิทธิภาพสัมพันธ์ของค่าประมาณมัชฌิมเลขคณิต และค่าประมาณความแปรปรวนของประชากร ผลการวิจัยที่สำคัญมีดังนี้

1. ขนาดโรงเรียน เป็นตัวแปรจำแนกชั้นภูมิที่มีประสิทธิภาพมากที่สุดในการประมาณค่ามัชฌิมเลขคณิตทุกขนาดกลุ่มตัวอย่าง และประมาณค่าความแปรปรวนเมื่อใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่น 95% โดยที่เขตอำเภอ เป็นตัวแปรจำแนกชั้นภูมิที่มีประสิทธิภาพมากที่สุดในการประมาณค่าความแปรปรวนเมื่อใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่น 90% และ 99%

2. วิธีการสุ่มที่ใช้ขนาดโรงเรียนเป็นตัวแปรจำแนกชั้นภูมิและกำหนดขนาดกลุ่มตัวอย่างย่อยแบบนีย์แมนมีประสิทธิภาพมากที่สุดในการประมาณค่ามัชฌิมเลขคณิต และประมาณค่าความแปรปรวน เมื่อใช้ขนาดกลุ่มตัวอย่างที่ระดับความเชื่อมั่น 99% และ 95% ตามลำดับ โดยส่วนใหญ่ของวิธีการสุ่มทั้งหมด วิธีกำหนดแบบนีย์แมนมีประประสิทธิภาพ ในการประมาณค่ามากกว่าวิธีกำหนดแบบสัดส่วน ยกเว้นกรณีการประมาณค่าความแปรปรวน โดยใช้ขนาดโรงเรียนเป็นตัวแปรจำแนกชั้นภูมิ ที่ระดับความเชื่อมั่น 90% หรือ 95%

ธวัช สุขแซว (2539) ศึกษาการกำหนดขอบเขตของชั้นภูมิชนิดสองทางที่เหมาะสมการประมาณและเปรียบ เทียบประสิทธิภาพการประมาณจำนวนวันลาของข้าราชการ ครูในโรงเรียน โดยใช้แผนแบบการสุ่มตัวอย่าง 2 แผนแบบ คือ 1) แผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสองทาง โดยใช้ตัวประมาณที่ไม่เอนเอียงและเอนเอียง 2) แผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดทางเดียว ผลการศึกษาพบว่า ขอบเขตของชั้นภูมิชนิดสองทางที่เหมาะสม คือ การกำหนดขอบเขตของชั้นภูมิชนิดแรกเป็น ขนาดของโรงเรียน ประกอบด้วย 3 ขนาด คือ ขนาด เล็กและกลาง, ขนาดใหญ่ และขนาดใหญ่พิเศษ ขอบเขตของชั้นภูมิชนิดที่สองเป็นลักษณะพื้นที่ที่โรงเรียนตั้งอยู่ ประกอบด้วย 5 ลักษณะ คือ พื้นที่อิ่มตัว,พื้นที่ขยายตัวช้า, พื้นที่ขยายตัวปานกลาง, พื้นที่ขยายตัวเร็ว (วงแหวนชั้นกลาง) และพื้นที่ขยายตัวเร็ว (วงแหวนชั้นนอก) ดังนั้นประชากรประกอบด้วยชั้นภูมิทั้งหมด 15 ชั้นภูมิ ในการเปรียบเทียบประสิทธิภาพการประมาณของ

แผนแบบการสุ่มตัวอย่างทั้งสองแบบ โดยพิจารณาจากค่าความแตกต่างระหว่างค่าจริงและประมาณ และค่าสัมประสิทธ์ความผันแปร พบว่า แผนแบบการสุ่มตัวอย่างแบบ แบ่งชั้นภูมิชนิดสองทาง โดยใช้ตัวประมาณที่เอนเอียง มีประสิทธิภาพในการประมาณสูงสุด รองลงมา คือ แผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสองทาง โดยใช้ตัวประมาณที่ไม่เอนเอียง และแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดทางเดียวตามลำดับ

ศิรประภา มโนมัธย์ (2539) ทำการเปรียบเทียบประสิทธิภาพความแม่นยำของแผนการสุ่มแบบ Adaptive cluster sampling เมื่อตัวอย่างขั้นต้นสุ่มแบบง่าย แบบชั้นภูมิและแบบมีระบบ โดยใช้ร้อยละของอัตราส่วนความแปรปรวนของค่าเฉลี่ย (%eff) ระหว่างการสุ่มตัวอย่างขั้นต้นแบบมีชั้นภูมิและแบบมีระบบเทียบกับการสุ่มตัวอย่างขั้นต้นแบบง่ายของ Thompson (1990; 1991 a; 1991 b อ้างถึงใน ศิรประภา มโนมัธย์, 2539) สำหรับตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Hansen-hurwitz การดำเนินงานวิจัยประกอบด้วยการจำลองแบบประชากรที่มีลักษณะหายากสร้างขึ้นทั้งหมด 5 ประชากร แต่ละกรณีเป็นประชากรที่มีลักษณะต่างกัน โดยใช้กระบวนการพัวซอง คลัสเตอร์ (Poisson cluster process) ซึ่งตำแหน่งและจำนวนหลักสร้างด้วยกระบวนการพัวซองที่มีพารามิเตอร์เป็น 30 ส่วนตำแหน่งของบริวารสร้างจากตัวแปรสุ่มที่มีการแจกแจงแบบปกติสองตัวแปร (Bivariate normal) ขนาดตัวอย่างที่ใช้ในแต่ละวิธี คือ 4 8 16 32 และ 64 หน่วย กระบวนการสุ่มขั้นต้นใช้แบบง่าย แบบมีชั้นภูมิ และแบบมีระบบ การแบ่งชั้นภูมิแบ่ง เป็น 4 ชั้นภูมิที่มีขนาดเท่า ๆ กัน ส่วนการสุ่มขั้นต้นแบบมีระบบ ใช้แผนการเลือกตัวอย่าง ที่มีหน่วยปฐมภูมิเป็นพื้นที่ 5 x 5 ตารางหน่วย และมีหน่วยทุติยภูมิเป็น 4 หน่วย ผลการวิจัย พบว่า ในการเปรียบเทียบ %eff สำหรับตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจาก ตัวประมาณ Hansen-hurwitz วิธีการสุ่มขั้นต้นแบบมีระบบมีความแม่นยำดีที่สุดในทุกขนาดตัวอย่าง โดยเมื่อตัวอย่างมีขนาด 4 8 16 32 และ 64 ให้ค่า %eff(,SYS) เฉลี่ยหรือ %eff(,SYS) จากประชากรทั้ง 5 กรณี เป็น 50.154 50.156 50.158 50.166 50.070 ตามลำดับ รองลงมา คือ การสุ่มขั้นต้นแบบมีชั้นภูมิ โดยเมื่อตัวอย่างมีขนาด 4 8 16 32 และ 64 ให้ค่า %eff(,STR) เป็น 77.936 77.934 77.142 79.082 ตามลำดับ ยกเว้นกรณีที่ประชากรเป็นกลุ่มเล็กมีกลุ่มเดียวนั้น เมื่อแบ่งเป็นชั้นภูมิแล้วทำให้ประสิทธิภาพของการสุ่มขั้นต้นแบบมีระบบเท่ากับการสุ่มขั้นต้นแบบมีชั้นภูมิ สำหรับตัวประมาณค่าเฉลี่ยที่ดัดแปลงมาจากตัวประมาณ Horvitz-thompson เมื่อตัวอย่างมีขนาดเล็กและเพิ่มขึ้นจาก 4 เป็น 8 หน่วย และจาก 8 เป็น 16 หน่วย พบว่า ประสิทธิภาพในแง่ของความแม่นยำของการสุ่มขั้นต้นแบบมีระบบจะดีที่สุด โดยเมื่อตัวอย่างมีขนาด 4 8 16 ให้ค่า %eff(,SYS) เป็น 63.364 56.024 45.062 ตามลำดับ รองลงมา คือ การสุ่มขั้นต้นแบบมีชั้นภูมิโดยให้ค่า %eff(,STR) เป็น 84.334 77.930 68.386 ตามลำดับ ชี้ให้เห็นว่าประสิทธิภาพเพิ่มขึ้นตามขนาดตัวอย่างที่เพิ่มขึ้นด้วย ยกเว้นกรณีที่ประชากรเป็นกลุ่มเล็ก

มีกลุ่มเดียวนั้น เมื่อแบ่งเป็นชั้นภูมิแล้วทำให้ประสิทธิภาพของการสุ่มขั้นต้นแบบมีระบบเท่ากับการสุ่มขั้นต้นแบบมีชั้นภูมิ ส่วนตัวอย่างที่มีขนาดใหญ่และเพิ่มขึ้นจาก 32 เป็น 64 หน่วย ประสิทธิภาพของการสุ่มขั้นต้น แบบมีระบบและแบบมีชั้นภูมิกลับลดลงจนเท่ากับการสุ่มตัวอย่างขั้นต้นแบบง่าย โดยขนาดตัวอย่าง เป็น 32 ให้ค่า %eff(,SYS) = 89.074 และ %eff(,STR) = 95.102 ตามลำดับ และขนาดตัวอย่าง เป็น 64 ให้ค่า %eff จากการสุ่มตัวอย่างขั้นต้นทั้ง 3 วิธี เท่ากัน ในทุกกรณีพบว่า แผนแบบ Adaptive cluster sampling สามารถใช้ได้กับประชากรที่มี ลักษณะหายากและอยู่รวมกันเป็นกลุ่ม ประสิทธิภาพของแผนแบบแตกต่างกันขึ้นอยู่กับลักษณะ การกระจายของประชากรที่ทำการศึกษา วิธีการสุ่มตัวอย่างขั้นต้น ตัวประมาณที่ใช้ และขนาด ตัวอย่างที่ทำการสุ่มในขั้นต้น และยังพบว่า ประสิทธิภาพของแผนการสุ่มแบบ Adaptive cluster sampling ที่มีการสุ่มตัวอย่างขั้นต้นแบบมีระบบใช้ได้ดีกว่าแบบอื่น ๆ รองลงมา คือ แบบมีชั้นภูมิและแบบง่ายตามลำดับ

แก้วใจ คำสุข (2543) ศึกษาความเที่ยงสัมพัทธ์ของแผนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิกับแผนการสุ่มตัวอย่างแบบธรรมดา ในกรณีความชุกของโรคและอาการต่าง ๆ โดยสุ่มตัวอย่างจากข้อมูลการสำรวจสถานะสุขภาพอนามัยของประชาชนไทยที่มีอายุตั้งแต่ 15 ปีขึ้นไป จำนวน 15,042 คน ด้วยการสอบถาม และตรวจร่างกายทั่วประเทศ ปี พ.ศ. 2534-2535 ของสถาบันวิจัยสาธารณสุขไทย เพื่อศึกษา อาการปวดหลัง การสูบบุหรี่ ภาวะโลหิตจาง ภาวะไขมันในเลือดสูง และภาวะความดันโลหิตสูง ซึ่งมีสัดส่วนการเกิดโรค/ อาการเท่ากับ 0.399 0.297 0.219 0.101 และ 0.054 ตามลำดับ ด้วยแผนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิด้วยการจัดสรรแบบสัดส่วน และจัดสรรแบบเนย์แมน ตัวแปรที่ใช้กำหนดชั้นภูมิคือ เพศ เขตที่อยู่อาศัย และอายุ (อายุใช้ 25 30 35 40 45 50 55 และ 60 ปี เป็นจุดตัดในการ กำหนดชั้นภูมิ) เกณฑ์กำหนดชั้นภูมิแบ่งเป็น 5 ช่วง คือ 0-10 11-30 31-60 61-100 และมากกว่า 100 ดังนั้น แผนการสุ่มเท่ากับ 30 แผน ในแต่ละแผนสุ่มซ้ำ 1,000 ครั้ง ผลการศึกษา พบว่า แผนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิมีความเที่ยงสูงกว่าแผนการสุ่มตัวอย่างแบบธรรมดา โดยความเที่ยงจะสูงขึ้นตามความแตกต่างของสัดส่วนระหว่าง ชั้นภูมิที่เพิ่มขึ้น และแผนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิด้วยการจัดสรรแบบเนย์แมนมีความเที่ยงสูงกว่าการจัดสรรแบบสัดส่วน ความเที่ยงที่ได้รับ (Percent gain) สูงที่สุดอยู่ในช่วงความแตกต่างของสัดส่วนมากกว่าร้อยละ 100 ของการสูบบุหรี่ด้วยวิธีกำหนดขนาดตัวอย่างกลุ่มย่อยแบบเนย์แมน มีค่า เท่ากับร้อยละ 58.61 ซึ่งคิดเป็นขนาดความแตกต่างของความแปรปรวนเพียงร้อยละ 0.09 ของค่าสัดส่วน แต่ค่าเฉลี่ยของค่าประมาณสัดส่วนมีค่าสูงกว่าค่าพารามิเตอร์ร้อยละ 29.17

นิภาพร โพธิ์ชัย (2544) ทำการศึกษาเปรียบเทียบแผนแบบการสุ่มตัวอย่างแบบสองชั้นและแผนแบบการสุ่มตัวอย่างแบบง่ายสำหรับข้อมูลทวินาม โดยการวิเคราะห์ข้อมูลปริมาตร

ไม้ยูคาลิปตัส 2 แบบ คือ การวัดค่าจริงและการวัดค่าโดยประมาณ ซึ่งการสุ่มตัวอย่างแบบสองชั้น ได้วัดค่าจริงและค่าโดยประมาณ ในส่วนแผนแบบการสุ่มตัวอย่างแบบง่ายนั้นวัดค่าจริง โดยมีเกณฑ์ที่ใช้ในการแบ่งกลุ่ม คือ 0.020 0.030 และ 0.040 ลูกบาศก์เมตร ซึ่งในแต่ละกรณีจะทำการสุ่มซ้ำ 30 รอบ ในการศึกษา เปรียบเทียบประสิทธิภาพของแผนแบบการสุ่มตัวอย่างแบบสองชั้น และการแผนแบบการสุ่ม ตัวอย่างแบบง่ายสำหรับข้อมูลทวินามนั้น จะพิจารณาค่า Relative efficiency (R.E.) ผลการศึกษาทั้ง 3 กรณีสรุปได้ว่า แผนแบบการสุ่มตัวอย่างแบบสองชั้นมีประสิทธิภาพสูงกว่าแผนแบบการสุ่มตัวอย่างแบบง่าย

สุทิน ชนะบุญ (2544) ศึกษาแบบจำลองเหตุการณ์ เพื่อเปรียบเทียบผลการสุ่มตัวอย่างแบบง่ายกับการสุ่มตัวอย่างแบบแบ่งชั้นภูมิ โดยใช้ข้อมูลความจำเป็นพื้นฐานเป็นตัวแปรแบ่งชั้นภูมิ และเปรียบเทียบผลการใช้แผนการสุ่มตัวอย่างเดียวกับ การใช้แผนการสุ่มตัวอย่างคู่ ในกรณีการชั่งน้ำหนักในเด็กอายุต่ำกว่า 5 ปี จำนวน 124 คน ในเขตรับผิดชอบของ สถานีอนามัยดงบัง จังหวัดขอนแก่น ในหมู่บ้านขนาดใหญ่ กลาง และเล็ก จำนวน 57 คน 40 คน และ 27 คน ตามลำดับ โดยปัจจัยความจำเป็นพื้นฐานที่มีความสัมพันธ์กับการได้รับการชั่งน้ำหนักในเด็กอายุต่ำกว่า 5 ปี คือ ครัวเรือนได้รับข่าวสารที่เป็นประโยชน์อย่างน้อยสัปดาห์ละ 3 ครั้ง และครอบครัวมีความอบอุ่น เมื่อนำปัจจัยมาเป็นตัวแปรแบ่งชั้นภูมิ และจำลองสถานการณ์สุ่มตัวอย่างในแต่ละหมู่บ้านเท่ากับ 50,000 ครั้ง โดยใช้ความครอบคลุมร้อยละ 80 เป็นเกณฑ์ยอมรับความสำเร็จของกิจกรรมชั่งน้ำหนัก พบว่า อัตราการยอมรับความสำเร็จของการสุ่มตัวอย่างแบบง่าย ในหมู่บ้านขนาดใหญ่ กลาง และเล็ก เป็นร้อยละ 96 ร้อยละ 93 และร้อยละ 100 ของการจำลองเหตุการณ์ ตามลำดับ สำหรับการสุ่มตัวอย่างแบบแบ่งชั้นภูมิผ่านเกณฑ์ยอมรับความสำเร็จลดลงเล็กน้อยทั้งในหมู่บ้านขนาดใหญ่และขนาดกลาง คือ ร้อยละ 2 และร้อยละ 4 ตามลำดับ ส่วนหมู่บ้านขนาดเล็กยังคงเท่าเดิม จึงสรุปได้ว่า แผนการสุ่มตัวอย่าง ทั้งสองแบบให้ผลไม่แตกต่างกัน และในการเปรียบเทียบผลการสุ่มตัวอย่างแบบง่ายโดยใช้แผนการสุ่มตัวอย่างเดียวกับการใช้แผนการสุ่มตัวอย่างคู่ พบว่า ในการสุ่มครั้งที่ 1 การใช้ แผนการสุ่มตัวอย่างคู่จะยอมรับหมู่บ้านน้อยกว่าการใช้แผนการสุ่มตัวอย่างเดียวเพียงร้อยละ 7 ถึงร้อยละ 12 และเมื่อสุ่มครบสองครั้ง จะยอมรับหมู่บ้านไม่แตกต่างจากการใช้แผนการสุ่มตัวอย่างเดียว

กรรณิการ์ มิ่งสอน (2547) ศึกษาแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มสองขั้นและแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มแบบง่ายทั้งวิธีการประมาณโดยใช้อัตราส่วนรวมและวิธีการประมาณโดยใช้อัตราส่วนแยก และเปรียบเทียบประสิทธิภาพระหว่างวิธีการประมาณ 2 วิธี และระหว่างแผนแบบการสุ่มตัวอย่าง 2 แผนแบบ โดยใช้ข้อมูลจำนวนนักเรียนและจำนวนครู ในโรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษา ประจำปีการศึกษา

2545 เพื่อประมาณอัตราส่วนของนักเรียนต่อครู ซึ่งการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มสองขั้น กำหนดให้ชั้นภูมิ คือ ภาค หน่วยตัวอย่างขั้นแรก คือ จังหวัด และหน่วยตัวอย่างขั้นที่สอง คือ โรงเรียน ในส่วนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มแบบง่ายกำหนดให้ชั้นภูมิ คือภาค และหน่วยตัวอย่าง คือ โรงเรียน ผลการศึกษาแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิทั้งชนิดสุ่มสองขั้นและสุ่มแบบง่ายในทุกวิธีการประมาณ พบว่า เมื่อจำนวนจังหวัดตัวอย่างเพิ่มขึ้น การประมาณมีประสิทธิภาพสูงขึ้น และเมื่อจำนวนโรงเรียนตัวอย่างเพิ่มขึ้น การประมาณมีประสิทธิภาพสูงขึ้น การเปรียบเทียบประสิทธิภาพในการประมาณระหว่างวิธีการประมาณสองวิธี พบว่า วิธีการประมาณโดยใช้อัตราส่วนแยกส่วนใหญ่มีประสิทธิภาพสูงกว่าวิธีการประมาณโดยใช้อัตราส่วนรวมทั้งสองแผนแบบ ส่วนการเปรียบเทียบประสิทธิภาพในการประมาณระหว่างแผนแบบการสุ่มตัวอย่างสองแผนแบบพบว่า แผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มแบบง่ายมีประสิทธิภาพสูงกว่าแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มสองขั้นในทุกวิธีการประมาณ

จากการศึกษางานวิจัยในประเทศและต่างประเทศที่เกี่ยวข้องสามารถสรุปเป็นประเด็นที่สำคัญ ดังนี้

1. วิธีการสุ่มแบบกริดเป็นวิธีที่ได้รับการยอมรับและนำไปใช้ในงานวิจัยวิทยาศาสตร์กายภาพเป็นจำนวนมากและหลากหลายสาขาวิชา ยกตัวอย่างเช่น Sánchez-Fernández et al. (2011) ใช้วิธีสุ่มแบบกริด เพื่อสร้างโมเดลการทำนายปริมาณตัวด้วงน้ำ (Water beetles) ที่อาศัยอยู่ในเขตแหลมไอบีเรีย Martínez-Coronado et al. (2011) ใช้วิธีการสุ่มแบบกริด เพื่อศึกษาความเข้มข้นของปรอทบริเวณเหมืองแร่ The Cerco de Almadenejos ในประเทศสเปน โดยใช้จุดกริด 16 ตำแหน่ง และ 35 ตำแหน่ง Lin (2013) ใช้วิธีการสุ่มแบบกริด เพื่อศึกษาปริมาณน้ำยางพารากับแร่ธาตุในดิน ที่สวนยางเมือง Yangjiang โดยใช้จุดกริดเท่ากับ 168 จุดกริด และขนาดกริดเซลล์เท่ากับ 1 x 1.5 ตารางเมตร เป็นต้น

2. งานวิจัยของ Dong-Sheng et al. (2011) และงานวิจัยของ Majasalmi et al. (2012) เป็นการวิจัยความสัมพันธ์ระหว่างความหนาแน่นของการสุ่ม (Sampling density) และค่าสัมประสิทธิ์ของความแปรผัน (CV) ซึ่งให้ผลการทดลองตรงกันคือ เมื่อความหนาแน่นของการสุ่มมากขึ้น จะทำให้ค่าสัมประสิทธิ์ของความแปรผันน้อยลง และถ้าตำแหน่งที่สุ่มมีการกระจายตัวครอบคลุมพื้นที่มากขึ้น ค่าความคลาดเคลื่อน s.e. จะลดลง

3. มีงานวิจัยหลายเรื่องที่ยืนยันว่าวิธีการสุ่มตัวอย่างแบบชั้นภูมิเป็นวิธีการที่มีประสิทธิภาพสูง ซึ่งเป็นไปตามคุณลักษณะและจุดเด่นของวิธีการสุ่มนี้ เช่น งานวิจัยของประชา สุวัฒนพันธุ์กุล (2534) และงานวิจัยของแก้วใจ คำสุข (2543) แต่ประชา สุวัฒนพันธุ์กุลได้เสนอแนะว่าในกรณีที่ค่าสัมประสิทธิ์ความแปรผันต่ำ และระดับความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษา

กับปัจจัยที่ใช้ในการกำหนดกลุ่มสูง อาจใช้วิธีเลือกตัวอย่างแบบโควต้าได้ ทั้งนี้เพราะสะดวกในการเก็บข้อมูลมากกว่าการสุ่มตัวอย่างแบบชั้นภูมิ ซึ่งปัญหาค่าใช้จ่ายในการสำรวจและการควบคุมงานสนามเป็นจุดด้อยของวิธีการนี้ นอกจากนั้น ดวงใจ ปวีณอภิชาต (2535) ทำการศึกษาตัวแปรจำแนกชั้นภูมิที่แตกต่างกัน 3 ลักษณะได้แก่ ประเภทโรงเรียน ขนาดโรงเรียน และเขตอำเภอ พบว่าขนาดโรงเรียน เป็นตัวแปรจำแนกชั้นภูมิที่มีประสิทธิภาพมากที่สุดในการประมาณค่ามัชฌิมเลขคณิตทุกขนาดกลุ่มตัวอย่าง

4. แผนการสุ่มแบบ Adaptive cluster sampling สามารถใช้ได้กับประชากรที่มีลักษณะหายากและอยู่รวมกันเป็นกลุ่ม ซึ่งประสิทธิภาพที่แตกต่างกันขึ้นอยู่กับลักษณะการกระจายของประชากรที่ทำการศึกษา

5. ในการศึกษาเปรียบเทียบประสิทธิภาพของแผนแบบการสุ่มตัวอย่างแบบสองชั้นและแผนแบบการสุ่มตัวอย่างแบบง่ายสำหรับข้อมูลทวินามนั้น โดยใช้ค่า Relative efficiency (R.E.) สรุปได้ว่า แผนแบบการสุ่มตัวอย่างแบบสองชั้นมีประสิทธิภาพสูงกว่าแผนแบบการสุ่มตัวอย่างแบบง่าย

บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อตรวจสอบความแม่นยำในการประมาณค่าและเปรียบเทียบประสิทธิภาพของการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ ซึ่งวิธีการสุ่มมีทั้งหมด 5 วิธี ได้แก่ 1) วิธีการสุ่มอย่างง่าย 2) วิธีการสุ่มแบบมีระบบ 3) วิธีการสุ่มแบบแบ่งชั้นภูมิ 4) วิธีการสุ่มแบบกลุ่ม และ 5) วิธีการสุ่มแบบกริด โดยทำการทดลองกับแบบจำลองที่สร้างขึ้น ซึ่งการตรวจสอบความแม่นยำเป็นการทดสอบเบื้องต้น เพื่อศึกษาประสิทธิผลของการสุ่มแบบกริดเมื่อเทียบกับวิธีการสุ่มอื่น ๆ ในส่วนการทดสอบวิธีการสุ่มแบบกริดกับฐานข้อมูลคะแนน O-NET เป็นการทดสอบ เพื่อพิสูจน์ยืนยันว่าวิธีการสุ่มแบบกริด (ซึ่งเป็นวิธีการเก็บตัวอย่างของวิทยาศาสตร์กายภาพ) มีความเหมาะสมในการนำมาประยุกต์ใช้กับงานวิจัยด้านพฤติกรรมศาสตร์และสังคมศาสตร์หรือไม่ อย่างไร โดยตัวบ่งชี้ คือ สารสนเทศที่เกิดจากวิธีการสุ่มแบบกริด วิธีการสุ่มแบบกริด-ชั้นภูมิ และวิธีการสุ่มแบบหลายขั้นตอน รายละเอียดการดำเนินการวิจัยมีดังต่อไปนี้

ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรของการวิจัยมี 2 รูปแบบ คือ ประชากรของแบบจำลองที่สร้างขึ้น และประชากรจากข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ประชากรที่ใช้ในการศึกษามีดังนี้

1. ประชากรสำหรับกรณีแบบจำลองการสุ่ม คือ ค่าตัวแปรทุก ๆ ค่าที่อยู่ในตารางแบบจำลองนั้น ๆ เช่น แบบจำลอง S0 เป็นตารางแบบ 9 x 9 ตารางหน่วย ดังนั้นมีประชากรทั้งหมด 81 หน่วย เป็นต้น

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-1 แบบจำลอง Sampling 0 (S0)

ค่าตัวแปรตามภาพที่ 3-1 คือ ระดับทัศนคติมีทั้งหมด 9 ระดับ โดยเรียงจากค่าน้อยไปหาค่ามาก (จาก 1 จนถึง 9) เช่น หมายเลข 1 หมายความว่าน้อยที่สุด หรือหมายเลข 9 หมายความว่ามากที่สุด ดังนั้นในแต่ละแบบจำลองจะมีขนาดประชากรและขนาดกลุ่มตัวอย่างไม่เท่ากัน ซึ่งขนาดประชากรจะมีจำนวนเท่าใดนั้นขึ้นอยู่กับความหมายและเหตุผลในการสร้างแบบจำลอง

2. ประชากรสำหรับกรณีศึกษา คือโรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 ทั้งหมด 2,584 โรง

กลุ่มตัวอย่าง

เนื่องจากประชากรของการวิจัยมี 2 รูปแบบ ดังนั้นกลุ่มตัวอย่างที่ถูกสุ่มออกมาเพื่อศึกษาหรือทดลองต้องมี 2 รูปแบบด้วย คือ กลุ่มตัวอย่างของแบบจำลองที่สร้างขึ้น และกลุ่มตัวอย่างจากฐานข้อมูลคะแนน O-NET ดังนี้

1. กลุ่มตัวอย่างสำหรับกรณีแบบจำลองการสุ่ม คือ กลุ่มตัวอย่างที่ถูกสุ่มมาจากแบบจำลองแผนที่ขนาดต่าง ๆ เช่น แบบจำลอง S0 เป็นตารางแบบ 9 x 9 ตารางหน่วย กลุ่มตัวอย่างเท่ากับ 9 หน่วย (ตามภาพที่ 3-2) เป็นต้น ซึ่งหลักการและเหตุผลในการกำหนดขนาดตัวอย่างได้อธิบายในหัวข้อแบบจำลองการสุ่มและแบบจำลองการประมาณค่าเชิงพื้นที่

1	7	4	2	9	5	6	**4**	5
8	5	8	3	1	5	6	2	6
2	**4**	**5**	8	5	6	3	7	3
5	8	9	4	7	**7**	2	9	**4**
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	**7**	6	4
5	6	2	6	4	9	3	8	**2**
4	1	5	9	**9**	2	4	6	5
3	6	8	5	5	4	**9**	7	8

ภาพที่ 3-2 การสุ่มอย่างง่ายของแบบจำลอง S0

จากภาพที่ 3-2 กลุ่มตัวอย่างที่ถูกสุ่มด้วยวิธีการสุ่มอย่างง่ายได้แก่ 4, 4, 5, 7, 4, 7, 2, 9, และ 9 ซึ่งค่าตัวแปรทั้ง 9 ตัว จะถูกนำไปวิเคราะห์หาค่าเฉลี่ยและค่าความแปรปรวน

2. กลุ่มตัวอย่างสำหรับกรณีศึกษา คือ โรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 จำนวน 335 โรงเรียน ซึ่งขนาดกลุ่มตัวอย่างดังกล่าวได้จากการวิเคราะห์และเปรียบเทียบวิธีการหาขนาดกลุ่มตัวอย่าง

จำนวน 5 วิธี ได้แก่ การกำหนดเกณฑ์ ตารางสำเร็จรูปของ Taro Yamane ตารางสำเร็จรูปของ Krejcie และ Morgan การประมาณค่าประชากรด้วยสูตร และการวิเคราะห์อำนาจการทดสอบ ดังนี้

2.1 การกำหนดเกณฑ์

ในกรณีที่ทราบจำนวนประชากรที่แน่นอนสามารถใช้เกณฑ์ในการกำหนดร้อยละของประชากร ถ้าขนาดประชากรเป็นหลักร้อย ต้องใช้กลุ่มตัวอย่างอย่างน้อย 25% ถ้าขนาดประชากรเป็นหลักพัน ต้องใช้กลุ่มตัวอย่างอย่างน้อย 10% ถ้าขนาดประชากรเป็นหลักหมื่น ต้องใช้กลุ่มตัวอย่างอย่างน้อย 5% ถ้าขนาดประชากรเป็นหลักแสน ต้องใช้กลุ่มตัวอย่างอย่างน้อย 1% (ธีรวุฒิ เอกะกุล, 2543) เมื่อประชากรโรงเรียนเท่ากับ 2584 โรงเรียน ดังนั้นต้องใช้กลุ่มตัวอย่างน้อยที่สุดเท่ากับ 10% หรือเท่ากับ 259 โรงเรียน

2.2 การใช้ตารางสำเร็จรูปของ Taro Yamane

ตารางสำเร็จรูปของ Taro Yamane เป็นตารางที่ใช้หาขนาดของกลุ่มตัวอย่างเพื่อประมาณค่าสัดส่วนของประชากร ที่ระดับความเชื่อมั่น 95% และความคลาดเคลื่อน 5% ต้องใช้ขนาดกลุ่มตัวอย่างน้อยที่สุดเท่ากับ 353 โรงเรียน (ขนาดประชากร 3,000 หน่วย ต้องใช้กลุ่มตัวอย่างน้อยที่สุด 353 หน่วย)

2.3 การใช้ตารางสำเร็จรูปของ Krejcie และ Morgan

ตารางสำเร็จรูปของ Krejcie และ Morgan เป็นตารางที่ใช้หาขนาดของกลุ่มตัวอย่างเพื่อประมาณค่าสัดส่วนของประชากร ที่ระดับความเชื่อมั่น 95% และความคลาดเคลื่อน 5% ต้องใช้ขนาดกลุ่มตัวอย่างน้อยที่สุดเท่ากับ 335 โรงเรียน (ขนาดประชากร 2,600 หน่วย ต้องใช้กลุ่มตัวอย่างน้อยที่สุด 335 หน่วย)

2.4 การประมาณค่าประชากรด้วยสูตร

สูตรการหาขนาดกลุ่มตัวอย่าง เพื่อประมาณค่าประชากร ในกรณีที่ทราบจำนวนประชากร (มารยาท โยทองยศ และ ปราณี สวัสดิสรรพ์, 2551)

$$n = \frac{N\sigma^2}{\frac{Ne^2}{Z^2} + \sigma^2}$$

n คือ จำนวนตัวอย่าง

N คือ จำนวนประชากรเท่ากับ 2,584 โรงเรียน

σ^2 คือ ความแปรปรวนของประชากรเท่ากับ 472.43

e คือ ความคลาดเคลื่อนเท่ากับ 5 คะแนน

Z คือ คะแนนมาตรฐานที่เกี่ยวข้องกับระดับความเชื่อมั่น (จากตารางโค้งปกติ)

เมื่อประชากรเท่ากับ 2,584 โรงเรียน ที่ระดับความเชื่อมั่น 95% ค่า Z จะเท่ากับ 1.96 ต้องใช้ขนาดกลุ่มตัวอย่างน้อยที่สุดเท่ากับ 71 โรงเรียน และที่ระดับความเชื่อมั่น 99% ค่า Z จะเท่ากับ 2.58 ต้องใช้ขนาดกลุ่มตัวอย่างน้อยที่สุดเท่ากับ 120 โรงเรียน

2.5 การวิเคราะห์อำนาจการทดสอบ

การวิจัยนี้ได้กำหนดอำนาจในการทดสอบที่ระดับ 0.95 กำหนดระดับนัยสำคัญทางสถิติที่ระดับ .05 และกำหนดขนาดอิทธิพลเท่ากับ 0.2299908 (ซึ่งคำนวณจากสมมติฐานฐานหลัก (Ho) = 161.67 คะแนน, สมมติฐานรอง (H1) = 156.67 คะแนน และ ค่าเบี่ยงเบนมาตรฐาน (**SD**) = 21.74 คะแนน) ได้ขนาดกลุ่มตัวอย่างอย่างน้อย 248 โรงเรียน

เมื่อวิเคราะห์วิธีการหาขนาดกลุ่มตัวอย่างทั้งหมดพบว่าวิธีการใช้ตารางสำเร็จรูปของ Taro Yamane ให้ขนาดกลุ่มตัวอย่างมากที่สุด คือ 347 โรงเรียน ดังนั้นถ้าเลือกใช้วิธีการนี้จะผ่านเกณฑ์การหาขนาดกลุ่มตัวอย่างในทุกเงื่อนไข (> 259, > 335, > 71, > 120 และ > 248) แต่แนวคิดของ Taro Yamane เป็นการกำหนดขนาดกลุ่มตัวอย่างที่มีขนาดใหญ่ที่สุดเท่าที่จะเป็นไปได้ ซึ่งบางงานวิจัยจึงได้ขนาดใหญ่เกินความจำเป็น (องอาจ นัยพัฒน์, 2551, หน้า 124, 126) ทำให้สิ้นเปลืองค่าใช้จ่ายในการเก็บรวบรวมข้อมูลได้ ดังนั้นผู้วิจัยจึงเลือกใช้วิธีตารางสำเร็จรูปของ Krejcie และ Morgan ซึ่งใช้ขนาดกลุ่มตัวอย่างเป็นอันดับที่ 2 ผ่านเกณฑ์ในทุกเงื่อนไข

ตัวแปรที่ใช้ในการวิจัย

การวิจัย เรื่อง “แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด” เป็นการวิจัยที่มุ่งทดสอบประสิทธิภาพของการสุ่มแบบกริด เมื่อเทียบกับวิธีการสุ่มรูปแบบอื่น ๆ โดยพิจารณาจากค่าสถิติที่ได้กับค่าพารามิเตอร์ ถ้าค่าสถิติที่ได้มีความใกล้เคียงกับค่าพารามิเตอร์แสดงว่ากลุ่มตัวอย่างเป็นตัวแทนที่ดีของประชากร ดังนั้นตัวแปรอิสระของการวิจัยคือวิธีการสุ่มตัวอย่าง และตัวแปรตามของการวิจัย คือ ค่าสถิติ ซึ่งการวิจัยครั้งนี้แบ่งการทดลองได้เป็น 2 ขั้นตอน ดังนั้นตัวแปรการวิจัยจึงมี 2 รูปแบบเช่นกัน ดังนี้

ขั้นตอนที่ 1: การทดลองการสุ่มแบบกริดโดยอาศัยแบบจำลองที่สร้างขึ้น ตัวแปรอิสระ คือ วิธีการสุ่มตัวอย่าง ซึ่งมี 5 วิธี ได้แก่ การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีระบบ การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ การสุ่มตัวอย่างแบบกลุ่ม และการสุ่มตัวอย่างแบบกริด ตัวแปรตาม คือ ค่าสถิติ ประกอบด้วยค่าเฉลี่ยของความคลาดเคลื่อนและผลรวมของค่าความแตกต่างความแปรปรวน

ขั้นตอนที่ 2: การทดลองการสุ่มแบบกริดโดยอาศัยข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ตัวแปรอิสระคือวิธีการสุ่มตัวอย่าง ซึ่งมี 3 วิธี

ได้แก่ การสุ่มแบบหลายขั้นตอน การสุ่มตัวอย่างแบบกริด และการสุ่มแบบกริด-ชั้นภูมิ ตัวแปรตาม คือ ค่าสถิติ ประกอบด้วยค่าเฉลี่ยของความคลาดเคลื่อนและค่าความแปรปรวนของคะแนน O-NET

ขั้นตอนการวิจัย

ขั้นตอนการวิจัยสามารถแบ่งออกได้เป็น 2 ขั้นตอนใหญ่ ๆ คือ ขั้นตอนที่ 1 เป็นการทดลองการสุ่มแบบกริดโดยอาศัยแบบจำลองที่สร้างขึ้น และขั้นตอนที่ 2 เป็นการทดลองการสุ่มแบบกริดโดยอาศัยข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ซึ่งประชากรที่ใช้ในการศึกษามีดังนี้

ขั้นตอนที่ 1

1. สร้างแบบจำลองการสุ่มเป็นแผนที่รูปสี่เหลี่ยมขนาด m x n (ขนาด m แถว และ n หลัก) ซึ่งการกำหนดให้ขนาดแผนที่มีความแตกต่างกัน เพื่อทดสอบความสัมพันธ์ระหว่างขนาดแผนที่กับระดับความแม่นยำในการประมาณค่า และกำหนดค่าตัวแปรเชิงพื้นที่ให้มีสหสัมพันธ์เชิงพื้นที่เป็นศูนย์ (สำหรับแบบจำลอง S0 และ S00) หรือมีสหสัมพันธ์เชิงพื้นที่เป็นบวก (สำหรับแบบจำลอง S1 S2 S3 และ S4)

2. ใช้โปรแกรม C เขียนคำสั่งควบคุมโปรแกรม เพื่อหาค่าพารามิเตอร์และค่าสถิติ ซึ่งการหาค่าพารามิเตอร์ของแต่ละแบบจำลองนั้นมีเพียงค่าเดียว คือ ผลรวมของค่าตัวแปรทุกตัวหารด้วยจำนวนทั้งหมด แต่ในส่วนของค่าสถิติขึ้นอยู่กับวิธีการสุ่ม และจำนวนครั้งของการสุ่ม ดังนั้น เพื่อให้การทดลองมีความคลาดเคลื่อนต่ำและมีความน่าเชื่อถือสูงจึงกำหนดให้จำนวนครั้งของการสุ่ม เท่ากับ 10,000 ครั้ง เพื่อให้จำนวนครั้งของการสุ่มเข้าใกล้ค่าอนันต์ ($n \longrightarrow \infty$) โดยแบ่งเป็นการทดลอง 10 ครั้ง และในแต่ละครั้งของการทดลองมีการสุ่มตัวอย่างจำนวน 1,000 ครั้ง โปรแกรมภาษา C จะคำนวณค่าสถิติและเปรียบเทียบกับค่าพารามิเตอร์ เพื่อหาค่าเฉลี่ยของความคลาดเคลื่อนและผลรวมของค่าความแตกต่างความแปรปรวน

3. เริ่มต้นทดลองจากแบบจำลอง S0 โดยการเรียกโปรแกรมทำงาน (Run program) และจดบันทึกข้อมูลค่าสถิติ (ค่าเฉลี่ยของความคลาดเคลื่อน และผลรวมของค่าความแตกต่างความแปรปรวน) จากจอภาพ (Monitor) ในการทดลองแต่ละครั้งจนครบ 10 ครั้ง โปรแกรมที่ใช้นำเสนอข้อมูล คือ Microsoft excel โดยกรอกข้อมูลผลการทดลอง 10 ครั้ง และเลือกใช้ฟังก์ชั่นในการพล็อตกราฟ จากนั้นทดลองแบบจำลอง S00, S1, S2, S3 และ S4 จนครบทุกแบบ

4. เปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อน และผลรวมของค่าความแตกต่างความแปรปรวนประสิทธิภาพการสุ่มในแต่รูปภาพ/กราฟ เพื่อหาวิธีการสุ่มตัวอย่างที่ดีที่สุดในแต่ละแบบจำลอง

5. สร้างแบบจำลองการประมาณค่าเชิงพื้นที่ (Spatial interpolation/ SI) เป็นแผนที่รูปสี่เหลี่ยมขนาด m x n (ขนาด m แถว และ n หลัก) เพื่อวิเคราะห์ความคลาดเคลื่อนที่เกิดขึ้น โดยใช้ค่าคลาดเคลื่อนร้อยละ (Percent error) เป็นตัวบ่งชี้ประสิทธิภาพของการประมาณค่าเชิงพื้นที่ ซึ่งแบบจำลองมีทั้งหมด 3 แบบ คือ SI1, SI2 และ SI3 โดยแบบจำลอง SI1 เป็นแผนที่ขนาด 12 x 12 ช่อง และมีค่าตัวแปร 7 ระดับ แบบจำลอง SI1 เป็นแผนที่ขนาด 12 x 12 ช่อง และมีค่าตัวแปรตั้งแต่ 7 ถึง 100 คะแนน แบบจำลอง SI1 เป็นแผนที่ขนาด 13 x 15 ช่อง และมีค่าตัวแปร 5 ระดับ

6. ใช้งาน โปรแกรม Microsoft excel ในการประมาณค่าเชิงพื้นที่ โดยการสร้างตารางคำนวนจากสูตรคำนวน ยกตัวอย่างเช่น การประมาณค่าเชิงพื้นที่สามารถสร้างตารางคำนวนได้ดังภาพที่ 3-3 และ 3-4 เป็นต้น

5	(5+3)/2	3
(5+3)/2	$\frac{5+3+3+3}{4}$	(3+3)/2
3	(3+3)/2	3

ภาพที่ 3-3 การประมาณค่าเชิงพื้นที่

A1	=(A1+C1)/2	C1
=(A1+A3)/2	=(A1+C1+A3+C3)/4	=(C1+C3)/2
A3	=(A3+C3)/2	C3

ภาพที่ 3-4 การสร้างตารางคำนวณ

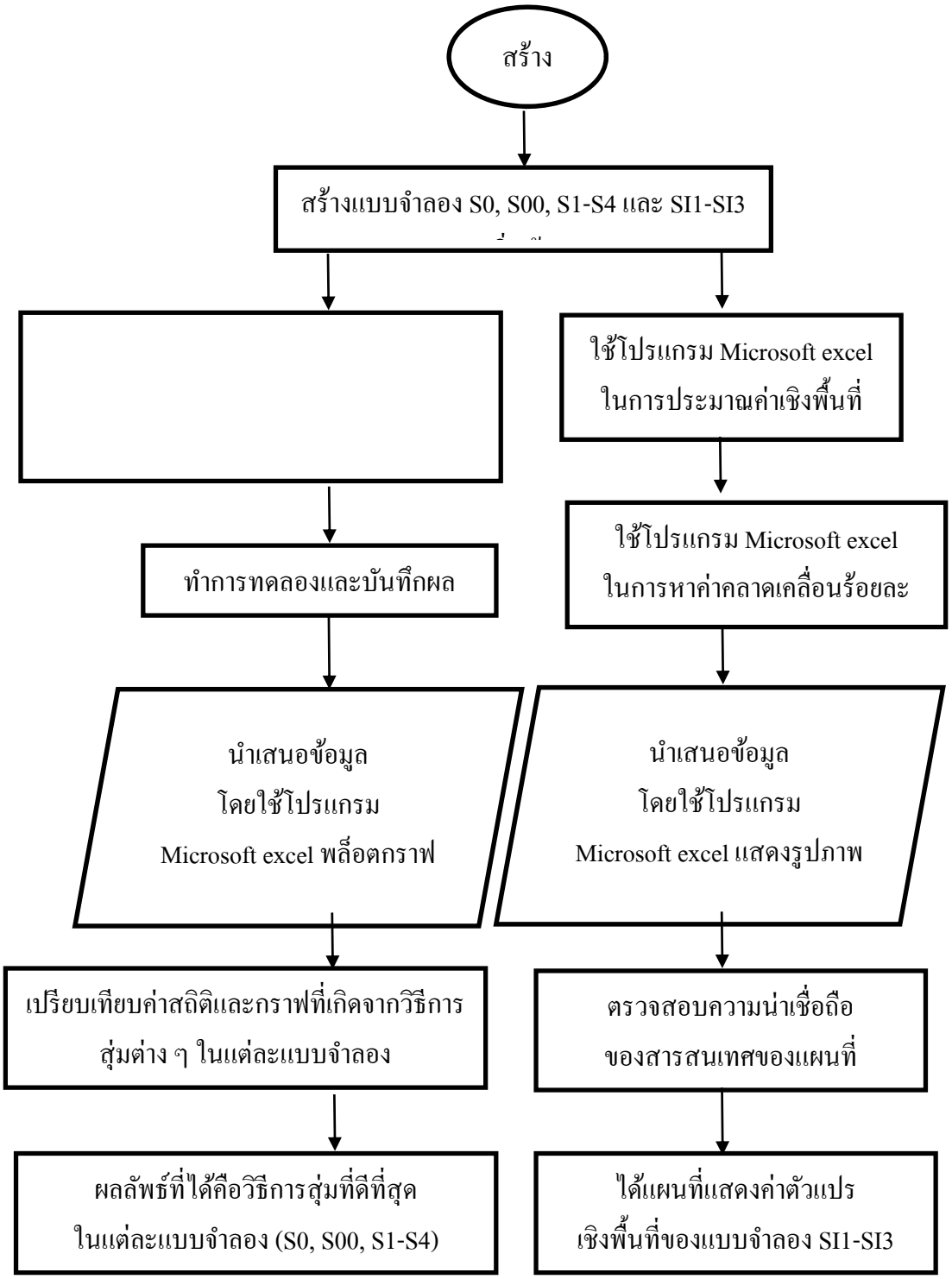
7. ใช้งานโปรแกรม Microsoft excel เพื่อเปรียบเทียบค่าจริง (ค่าพารามิเตอร์) กับค่าประมาณในลักษณะตำแหน่งต่อตำแหน่ง โดยใช้สูตรค่าคลาดเคลื่อนร้อยละที่ผู้วิจัยสร้างขึ้น (ในบทที่ 2) ดังภาพที่ 3-5

0%	0%	0%
0%	12.5%	0%
0%	0%	0%

ภาพที่ 3-5 การหาค่าคลาดเคลื่อนร้อยละ

8. เริ่มต้นทดลองจากแบบจำลอง SI1 โดยกรอกข้อมูลค่าตัวแปรลงในตารางจนครบตามแบบจำลอง จากนั้นแปลงข้อมูลตัวเลขให้เป็นรูปภาพด้วยการใช้งานฟังก์ชั่นของโปรแกรม Microsoft excel ดังนี้ 8.1) เลือกคำสั่งแทรกจากแถบคำสั่ง (Menu bar) 8.2) คำสั่งแผนภูมิ 8.3) เลือกชนิดของแผนภูมิเป็นแบบพื้นผิว 8.4) เลือกฟังก์ชั่นการหมุน 3 มิติ 8.5) หมุนแกน x ไปที่ 0 องศา หมุนแกน Y ไปที่ -90 องศา หมุนเปอร์สเปคทีฟไปที่ 0 องศา เมื่อทำตามขั้นตอนจนครบจะปรากฏแผนที่แสดงค่าตัวแปรเชิงพื้นที่เกิดขึ้น ให้ทดลองซ้ำกับแบบจำลอง SI2 และ SI3 จนครบทุกแบบ

9. ตรวจสอบความน่าเชื่อถือของสารสนเทศของแผนที่ที่ได้จากวิธีการสุ่มแบบกริด และจากรายละเอียดของขั้นตอนที่ 1 (ตั้งแต่ข้อ 1 ถึง 9) สามารถแสดงเป็นผังงานได้ตามภาพที่ 3-6



ภาพที่ 3-6 ผังงานการวิจัยขั้นตอนที่ 1

ขั้นตอนที่ 2

10. จัดหาข้อมูลและตำแหน่งพิกัดภูมิศาสตร์ โดยใช้ฐานข้อมูลของคะแนน O-NET ปีการศึกษา 2548 ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ซึ่งมีรายละเอียดที่ครบถ้วนสมบูรณ์ เช่น รายชื่อโรงเรียน รหัสโรงเรียน จังหวัด คะแนนเฉลี่ย คะแนนสูงสุด คะแนนต่ำสุด ฯลฯ สิ่งที่ต้องจัดหาเพิ่มเติม คือ พิกัดภูมิศาสตร์ของโรงเรียน ในการวิจัยครั้งนี้ได้ใช้ โปรแกรม Google map ในการหาค่าละติจูดและลองจิจูดของโรงเรียน

11. จัดเตรียมข้อมูล เพื่อเตรียมการนำข้อมูลเข้าระบบ ซึ่งข้อมูลในระบบภูมิสารสนเทศแบ่งออกเป็น 2 ประเภท คือ ข้อมูลเชิงพื้นที่ (Spatial data) และข้อมูลเชิงคุณลักษณะ (Attribute data) โดยข้อมูลเชิงพื้นที่ประกอบด้วย ข้อมูลรูปแบบปิด (Polygon) คือ แผนที่ประเทศไทย ชั้นข้อมูลเชิงตำแหน่ง (Point) คือ พิกัดภูมิศาสตร์ของโรงเรียน ในส่วนข้อมูลเชิงคุณลักษณะเป็นข้อมูลที่ใช้อธิบายประกอบข้อมูลเชิงพื้นที่นั้น เช่น ค่าละติจูด ค่าลองจิจูด คะแนนเฉลี่ย ฯลฯ ดังภาพที่ 3-7

Id	Name	Lat	Long	Mean_Onet
1	00100701:เตรียมอุดมศึกษา	13.7404559	100.5309159	355.77
2	00100102:สตรีวิทยา	13.7575675	100.5013876	279.63
3	00100101:สวนกุหลาบวิทยาลัย	13.7427956	100.4981714	270.92

ภาพที่ 3-7 ข้อมูลเชิงคุณลักษณะ

12. ใช้โปรแกรม QGIS นำเสนอข้อมูลในรูปแบบแผนที่พร้อมทั้งรายละเอียด เพื่อศึกษาลักษณะและรายละเอียดของประชากร เช่น ตำแหน่งที่ตั้งของโรงเรียน ความหนาแน่นของโรงเรียนในพื้นที่ต่าง ๆ การกระจายของคะแนน O-NET เป็นต้น

13. หาค่าพารามิเตอร์คะแนน O-NET ของโรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้าสอบ O-NET เป็นจำนวน 2,584 โรงเรียน โดยใช้โปรแกรม Microsoft excel ซึ่งค่าพารามิเตอร์เป็นตัวชี้วัดความแม่นยำระหว่างวิธีการสุ่มหลายขั้นตอนกับวิธีการสุ่มแบบกริด

14. ออกแบบการสุ่มตัวอย่างด้วยวิธีการสุ่มหลายขั้นตอน วิธีการสุ่มแบบกริด และวิธีการสุ่มแบบกริด-ชั้นภูมิ มีรายละเอียด ดังนี้

14.1 ออกแบบการสุ่มตัวอย่างแบบหลายขั้นตอน เริ่มต้นจากหาจำนวนโรงเรียนและจังหวัดในแต่ละภูมิภาค ซึ่งการคำนวณหาจำนวนโรงเรียนในแต่ละภูมิภาคต้องพิจารณาจากสัดส่วนระหว่างจำนวนโรงเรียนในแต่ละภูมิภาคต่อจำนวนโรงเรียนทั้งหมด โดยใช้ตาราง Krejcie และ Morgan สำหรับกลุ่มตัวอย่าง จำนวน 335 โรงเรียน และใช้กฎแห่งความชัดเจน (Rule of tumb)

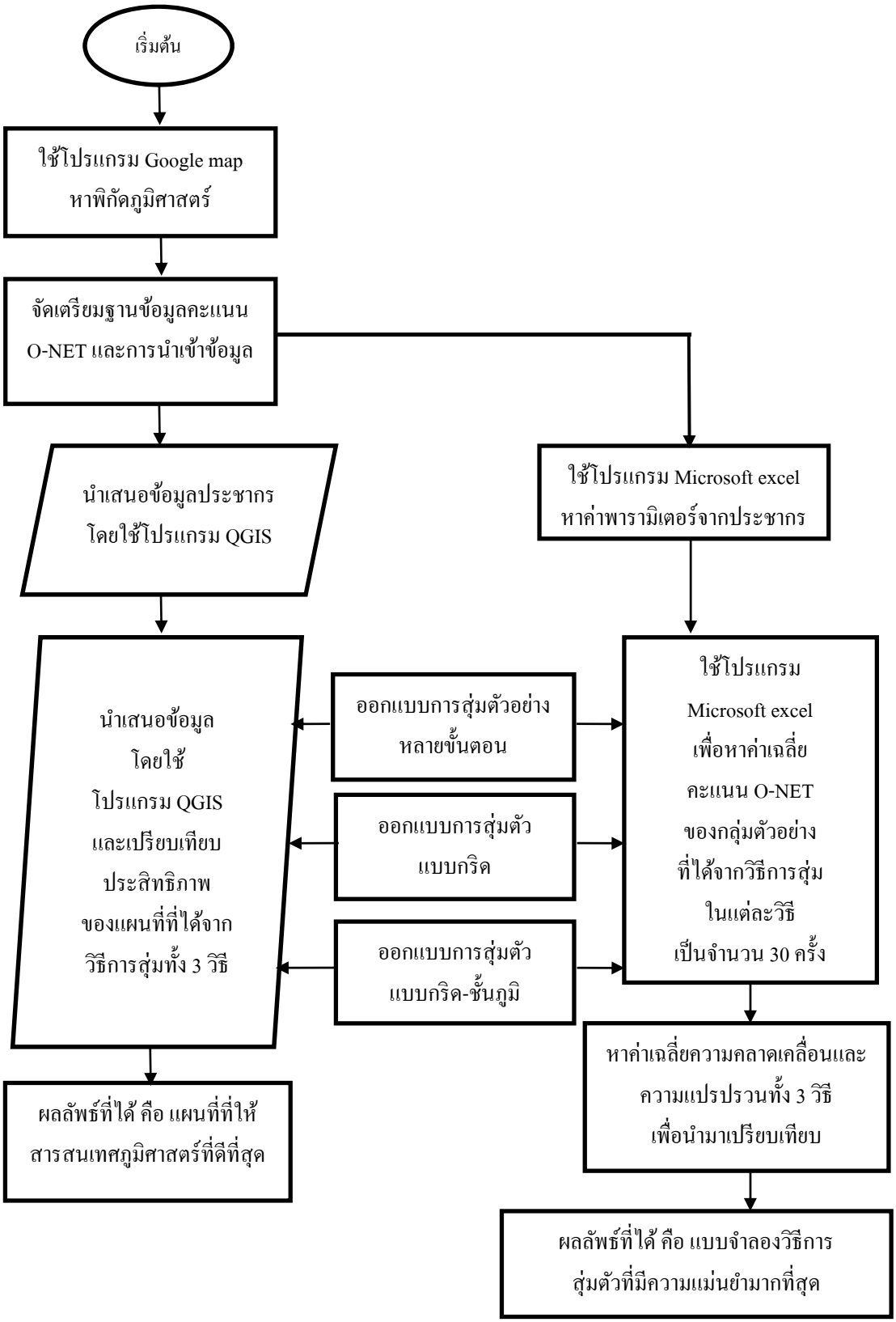
ของ Neuman ในการคำนวณหาจำนวนจังหวัดในแต่ละภูมิภาค จากนั้นเขียนโปรแกรมในการสุ่มจังหวัดและโรงเรียน

14.2 ออกแบบการสุ่มตัวอย่างแบบกริดมีขั้นตอนเริ่มต้นจากการหาตำแหน่งจุดกริดบนแผนที่ โดยใช้โปรแกรม QGIS และทำการกำหนดขนาดตัวอย่างในแต่ละจุดกริดตามสัดส่วน จากนั้นใช้วงกลมการสุ่มที่สร้างขึ้นจากโปรแกรม QGIS สุ่มตัวอย่างโรงเรียนในแต่ละจุดกริด

14.3 การออกแบบการสุ่มตัวอย่างแบบกริด-ชั้นภูมิ มีวิธีการทำเหมือนกับการออกแบบการสุ่มตัวอย่างแบบกริด โดยเริ่มจากการหาตำแหน่งจุดกริดบนแผนที่ด้วยโปรแกรม QGIS และทำการกำหนดขนาดตัวอย่างในแต่ละจุดกริดตามสัดส่วน จากนั้นใช้วงกลมการสุ่มที่สร้างขึ้นจากโปรแกรม QGIS สุ่มตัวอย่างโรงเรียนในแต่ละจุดกริด แต่สิ่งที่แตกต่างกัน คือ วิธีการนี้ต้องขยายรัศมีวงกลมให้ไกลที่สุด (รัศมีวงกลมเท่ากับครึ่งหนึ่งของระยะห่างระหว่างจุดกริด) ขั้นตอนต่อไป คือ นำกลุ่มตัวอย่างทั้งหมดที่อยู่ภายในรัศมีของวงกลมมาแบ่งเป็นชั้นภูมิและทำการสุ่มตัวอย่างในแต่ละชั้นภูมิตามสัดส่วน

15. หาค่าสถิติที่ได้จากการสุ่มทั้ง 3 วิธี ได้แก่ ค่าเฉลี่ยและความแปรปรวนของคะแนน O-NET จากนั้นนำค่าเฉลี่ยของวิธีการสุ่มทั้ง 3 วิธี มาวิเคราะห์ความแม่นยำและประสิทธิภาพโดยใช้ค่าเฉลี่ยความคลาดเคลื่อนและความแปรปรวนเป็นเกณฑ์

16. ใช้โปรแกรม QGIS นำเสนอข้อมูลในรูปแบบแผนที่ที่ได้จากแบบจำลองทั้ง 3 วิธี และนำแผนที่แสดงค่าตัวแปรระดับคะแนน O-NET ที่ได้จากวิธีการสุ่มทั้ง 3 วิธีมาเปรียบเทียบกัน โดยใช้ภาพแผนที่แสดงลักษณะและรายละเอียดของประชากรเป็นตัวตัดสิน และจากรายละเอียดของขั้นตอนที่ 2 (ตั้งแต่ข้อ 10 ถึง 16) สามารถแสดงเป็นผังงานได้ตามภาพที่ 3-8



ภาพที่ 3-8 ผังงานการวิจัยขั้นตอนที่ 2

เครื่องมือที่ใช้ในการวิจัย

1. เครื่องคอมพิวเตอร์ CPU: Intel Core i5-6200U RAM: 4 GB DDR3L HDD: 1 TB 5400 RPM

2. ข้อมูลที่ใช้ในการวิจัย ประกอบด้วยฐานข้อมูลคะแนน O-NET จากศูนย์ปฏิบัติการ GPA สพฐ. ประจำปี พ.ศ. 2548 และข้อมูลสารสนเทศภูมิศาสตร์ของประเทศไทย

3. โปรแกรมคอมพิวเตอร์ที่ใช้ในการจัดกระทำข้อมูลและวิเคราะห์ข้อมูล ได้แก่ โปรแกรมภาษา C โปรแกรม Microsoft excel โปรแกรม Google map โปรแกรม G*Power และโปรแกรม Quantum GIS (QGIS) การที่ผู้วิจัยเลือกใช้โปรแกรม QGIS เนื่องจากโปรแกรมนี้เป็นโปรแกรมรหัสเปิด (Free open source program)

ผู้วิจัยได้ใช้โปรแกรมภาษา C เพื่อจำลองสถานการณ์ในการสุ่มตัวอย่างจากแบบจำลอง โดยเริ่มต้นจากหาค่าเฉลี่ยและความแปรปรวนของประชากร จากนั้นหาค่าเฉลี่ยและความแปรปรวนของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบต่าง ๆ และในขั้นตอนสุดท้าย คือ การเปรียบเทียบค่าสถิติกับค่าพารามิเตอร์ เพื่อเปรียบเทียบประสิทธิภาพของการสุ่มแบบต่าง ๆ โดยมีรายละเอียดและขั้นตอนดังนี้

1. กำหนดขนาดแผนที่และค่าตัวแปรที่วัดจากประชากร ยกตัวอย่างเช่นกำหนดให้ประชากรมีจำนวน 144 ตำแหน่ง และตัวเลขในแต่ละตำแหน่งคือระดับสเกลทัศนคติ (Likert scale) ซึ่งมีทั้งหมด 7 ระดับ ดังนี้

```
float ij[12][12]= {6,6,6,5,5,5,5,5,5,6,6,6,
                   6,6,6,5,5,5,5,5,5,6,6,6,
                   4,4,4,3,3,3,3,3,3,4,4,4,
                   4,4,4,3,3,3,3,3,3,4,4,4,
                   2,2,2,1,1,1,1,1,1,2,2,2,
                   2,2,2,1,1,1,1,1,1,2,2,2,
                   3,3,3,2,2,2,2,2,2,3,3,3,
                   3,3,3,2,2,2,2,2,2,3,3,3,
                   5,5,5,4,4,4,4,4,4,5,5,5,
                   5,5,5,4,4,4,4,4,4,5,5,5,
                   7,7,7,6,6,6,6,6,6,7,7,7,
                   7,7,7,6,6,6,6,6,6,7,7,7};
```

2. หาค่าเฉลี่ยของประชากร (average) และความแปรปรวนของประชากร (var)

```
for (i=0;i<row;i++)
{
for (j=0;j<column;j++)
{
total = total+ij[i][j];
}
}
average = total/P_size;
printf("average= %f\n",average);
for (i=0;i<row;i++)
{
for (j=0;j<column;j++)
{
sum=sum+pow((average-ij[i][j]),2);
}
}
var = sum/P_size;
printf("var= %f\n",var);
```

3. ใช้วิธีการสุ่มแบบกริดหาค่าเฉลี่ยของกลุ่มตัวอย่าง (average5) และค่าเฉลี่ยของความคลาดเคลื่อน (error5)

```
for (k=0;k<N_point;k++)
{
x=position_i[k];
for (i=0;i<grid_row;i++)
{
y=position_j[k];
for (j=0;j<grid_column;j++)
{
```

```
total = total + ij[x][y];
y=y+1;
}
x=x+1;
}
}
average5=total/S_size;
printf("average5= %f\n",average5);
error5 = fabs(average-average5);
printf("error5= %f\n",error5);
```

4. ใช้วิธีการสุ่มแบบกริดหาค่าผลรวมความแตกต่างของความแปรปรวน (dv5)

```
for (k=0;k<N_point;k++)
{
x=position_i[k];
for (i=0;i<grid_row;i++)
{
y=position_j[k];
for (j=0;j<grid_column;j++)
{
sum=sum+pow((average5-ij[x][y]),2);
y=y+1;
}
x=x+1;
}
}
var5 = sum/(S_size-1);
dv5 = 1000*fabs(var-var5);
printf("dv5= %f\n",dv5);
```

5. ใช้วิธีการสุ่มตัวอย่างอย่าง่ายหาค่าเฉลี่ยของกลุ่มตัวอย่าง (average1)

```
for (i=0;i<1000;i++)
{
for (j=0;j<S_size;j++)
{
x = rand()%row;
y = rand()%column;
total = total + ij[x][y];
sampling[j]=ij[x][y];
}
average1 = total/S_size;
```

6. ใช้วิธีการสุ่มตัวอย่างอย่าง่ายหาค่าเฉลี่ยของความคลาดเคลื่อน (error1) และค่าผลรวมของความแตกต่างของความแปรปรวน (dv1)

```
for (k=0;k<S_size;k++)
{
sum=sum+pow((average1-sampling[k]),2);
}
var1 = sum/(S_size-1);
dv1 = dv1+fabs(var-var1);
error1 = error1+fabs(average-average1);
}
error1=error1/1000;
printf("error1= %f\n",error1);
printf("dv1= %f\n",dv1);
```

7. เนื่องจากการสุ่มตัวอย่างแบบมีระบบมีลักษณะการสุ่มเป็นแบบ 1 มิติ ดังนั้นต้องลดขนาดมิติจากเดิม 2 มิติ ให้เหลือ 1 มิติ (one_D[k]) พร้อมทั้งหาช่วงห่างของชุดตัวอย่าง (a)

```
a=P_size/S_size;
printf("gap = %d\n",a);
```

```
k=-1;
for (i=0;i<row;i++)
{
for (j=0;j<column;j++)
{
k=k+1;
one_D[k] = ij[i][j];
}
}
```

8. ใช้วิธีการสุ่มตัวอย่างแบบมีระบบหาค่าเฉลี่ยของกลุ่มตัวอย่าง (average2)

```
for (i=0;i<1000;i++)
{
b = rand()%a;
for (j=0;j<S_size;j++)
{
total = total + one_D[b];
sampling[j]=one_D[b];
b=b+a;
}
average2 = total/S_size;
printf("average2= %f\n",average2);
```

9. ใช้วิธีการสุ่มตัวอย่างแบบมีระบบหาค่าเฉลี่ยของความคลาดเคลื่อน (error2) และค่าผลรวมของความแตกต่างของความแปรปรวน (dv2)

```
for (k=0;k<S_size;k++)
{
sum=sum+pow((average2-sampling[k]),2);
}
var2 = sum/(S_size-1);
```

```
dv2 = dv2+fabs(var-var2);
total=0;
error2 = error2+fabs(average-average2);
}
error2=error2/1000;
printf("error2= %f\n",error2);
printf("dv2= %f\n",dv2);
```

10. ในกรณีการสุ่มแบบแบ่งกลุ่มต้องแบ่งแผนที่ออกเป็นส่วน ๆ และสำหรับกรณีนี้แบ่งกลุ่มได้ 3 กลุ่ม ดังนี้

```
float cluster1[12][4]={6,6,6,5,
        6,6,6,5,
                        4,4,4,3,
                        4,4,4,3,
        2,2,2,1,
                        2,2,2,1,
                        3,3,3,2,
        3,3,3,2,
        5,5,5,4,
        5,5,5,4,
        7,7,7,6,
        7,7,7,6};
float cluster2[12][4]={5,5,5,5,
        5,5,5,5,
        3,3,3,3,
        3,3,3,3,
        1,1,1,1,
        1,1,1,1,
        2,2,2,2,
        2,2,2,2,
```

```
        4,4,4,4,
        4,4,4,4,
        6,6,6,6,
        6,6,6,6};
float cluster3[12][4]={5,6,6,6,
        5,6,6,6,
        3,4,4,4,
                        3,4,4,4,
                        1,2,2,2,
                        1,2,2,2,
                        2,3,3,3,
        2,3,3,3,
        4,5,5,5,
        4,5,5,5,
        6,7,7,7,
        6,7,7,7};
```

11. การเขียนคำสั่งในโปรแกรม (Source code) ของการหาค่าเฉลี่ยด้วยวิธีการแบ่งกลุ่ม มีหลักการเขียน Source code คล้ายกับวิธีการสุ่มแบบอื่น ๆ แต่เนื่องจากข้อมูลที่นำเข้ามามีจำนวนเท่ากับ 3 ส่วน ดังนั้นต้องเพิ่มการควบคุมทิศทางแบบวนรอบ (วน loop) ให้เท่ากับ 3 ส่วน

```
for (k=0;k<1000;k++)
{
for (j=0;j<12;j++)
{
a = rand()%12;
b = rand()%4;
total = total+cluster1[a][b];
sampling1[j]=cluster1[a][b];
}
for (j=0;j<12;j++)
```

```
{
a = rand()%12;
b = rand()%4;
total = total+cluster2[a][b];
sampling2[j]=cluster2[a][b];
}
for (j=0;j<12;j++)
{
a = rand()%12;
b = rand()%4;
total = total+cluster3[a][b];
sampling3[j]=cluster3[a][b];
}
average4 = total/36;
printf("average4= %f\n",average4);
```

12. หาค่าเฉลี่ยความคลาดเคลื่อน (error4) และค่าผลรวมความแตกต่างของความแปรปรวน (dv4) ด้วยวิธีการสุ่มตัวอย่างแบบแบ่งกลุ่ม

```
for (i=0;i<12;i++)
{
sum=sum+pow((average4-sampling1[i]),2);
}
for (i=0;i<12;i++)
{
sum=sum+pow((average4-sampling2[i]),2);
}
for (i=0;i<12;i++)
{
sum=sum+pow((average4-sampling3[i]),2);
}
```

```
var4 = sum/35;
dv4 = dv4+fabs(var-var4);
error4 = error4+fabs(average-average4);
}
error4=error4/1000;
printf("error4= %f\n",error4);
printf("dv4= %f\n",dv4);
```

13. ในกรณีการสุ่มแบบชั้นภูมิต้องแบ่งชุดข้อมูลออกเป็นชั้น โดยให้ภายในกลุ่มมีความแปรปรวนต่ำ แต่ความแปรปรวนระหว่างกลุ่มต้องสูง ซึ่งกรณีนี้สามารถแบ่งได้ 6 ชั้นภูมิ ดังนี้

```
float stratum1[2][12]= {6,6,6,5,5,5,5,5,5,6,6,6,
                        6,6,6,5,5,5,5,5,5,6,6,6};
float stratum2[2][12]= {4,4,4,3,3,3,3,3,3,4,4,4,
                        4,4,4,3,3,3,3,3,3,4,4,4};
float stratum3[2][12]= {2,2,2,1,1,1,1,1,1,2,2,2,
                        2,2,2,1,1,1,1,1,1,2,2,2};
float stratum4[2][12]= {3,3,3,2,2,2,2,2,2,3,3,3,
                        3,3,3,2,2,2,2,2,2,3,3,3};
float stratum5[2][12]= {5,5,5,4,4,4,4,4,4,5,5,5,
                        5,5,5,4,4,4,4,4,4,5,5,5};
float stratum6[2][12]= {7,7,7,6,6,6,6,6,6,7,7,7,
                        7,7,7,6,6,6,6,6,6,7,7,7};
```

14. การหาค่าเฉลี่ย (average3) ด้วยวิธีการแบ่งชั้นภูมิเป็นจำนวน 6 ชั้นภูมิ

```
for (k=0;k<1000;k++)
{
for (j=0;j<6;j++)
{
a = rand()%2;
```

```
b = rand()%12;
total = total+stratum1[a][b];
sampling1[j]=stratum1[a][b];
}
for (j=0;j<6;j++)
{
a = rand()%2;
b = rand()%12;
total = total+stratum2[a][b];
sampling2[j]=stratum2[a][b];
}
for (j=0;j<6;j++)
{
a = rand()%2;
b = rand()%12;
total = total+stratum3[a][b];
sampling3[j]=stratum3[a][b];
}
for (j=0;j<6;j++)
{
a = rand()%2;
b = rand()%12;
total = total+stratum6[a][b];
sampling6[j]=stratum6[a][b];
}
average3 = total/36;
printf("average3= %f\n",average3);
```

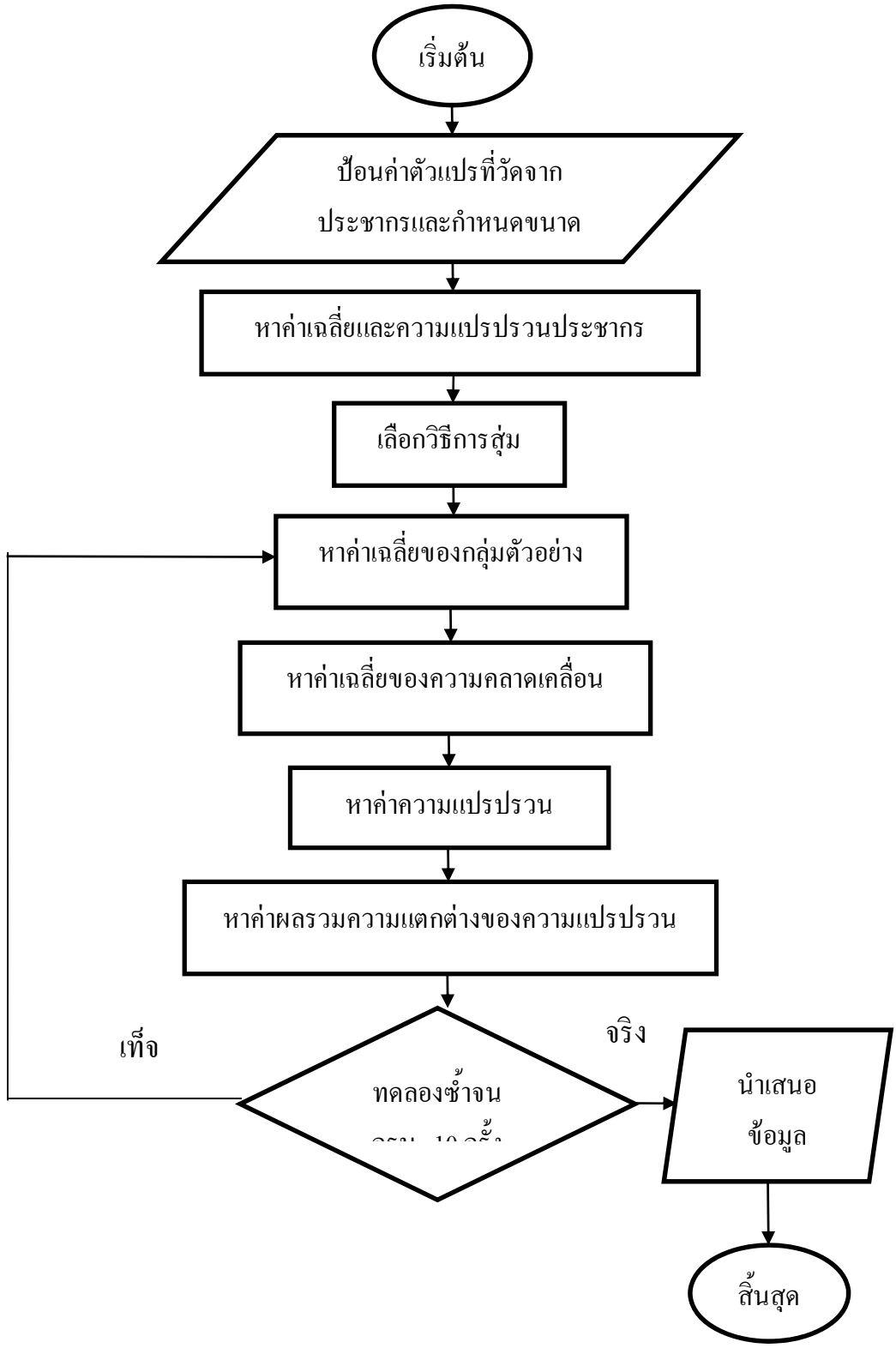
15. ใช้วิธีการสุ่มตัวอย่างแบบมีระบบหาค่าเฉลี่ยของความคลาดเคลื่อน (error3) และค่าผลรวมของความแตกต่างของความแปรปรวน (dv3)

```
for (i=0;i<6;i++)
{
sum=sum+pow((average3-sampling1[i]),2);
}
for (i=0;i<6;i++)
{
sum=sum+pow((average3-sampling2[i]),2);
}
for (i=0;i<6;i++)
{
sum=sum+pow((average3-sampling3[i]),2);
}
for (i=0;i<6;i++)
{
sum=sum+pow((average3-sampling6[i]),2);
}
var3 = sum/35;
dv3 = dv3+fabs(var-var3);
error3 = error3+fabs(average-average3);
}
error3=error3/1000;
printf("error3= %f\n",error3);
printf("dv3= %f\n",dv3);
```

หมายเหตุ การเขียนคำสั่งในโปรแกรม (Source code) สำหรับกรณีสุ่มแบบแบ่งกลุ่มและการสุ่มแบบชั้นภูมิไม่สามารถเขียนให้เป็นแบบทั่วไปหรือแบบสากลได้ เพราะการแบ่งประชากรออกเป็นกลุ่มก้อนนั้นไม่มีความแน่นอน เนื่องจากการแบ่งประชากรในแต่ละกรณีต้องยึดถือทิศทางและระดับความเข้มของค่าตัวแปรเป็นหลัก ดังนั้นการเขียนคำสั่งในโปรแกรม (Source code) จึงต้องเขียนในเชิงลักษณะเฉพาะตามแต่ละกรณี

จากหลักการในการเขียนคำสั่งในโปรแกรมทั้ง 15 ข้อ สามารถสรุปเป็นแผนผังงาน ดังนี้

ผังงานการเปรียบเทียบประสิทธิภาพวิธีการสุ่ม



ภาพที่ 3-9 ผังงานการเปรียบเทียบประสิทธิภาพวิธีการสุ่ม

แบบจำลองการสุ่ม

แบบจำลองการสุ่มเป็นเครื่องมือสำคัญในการเปรียบเทียบประสิทธิภาพวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ ซึ่งแบบจำลองการสุ่มคือแผนที่แสดงค่าตัวแปรเชิงพื้นที่ โดยมีเส้นแนวนอนเป็นเส้นละติจูดและมีเส้นแนวตั้งเป็นเส้นลองจิจูด ดังนั้นพื้นที่ทั้งหมดจึงถูกแบ่งด้วยเส้นแนวตั้งและแนวนอนออกเป็นพื้นที่ขนาดเล็ก (กริดเซลล์) ซึ่งแต่ละกริดเซลล์มีประชาชนอาศัยอยู่รวมกันเป็นสังคม ประชากรภายในสังคมนั้นมีความเชื่อ ทัศนคติและความคิดเห็นสอดคล้องกับขนบธรรมเนียมประเพณีและวิถีชีวิต เมื่อวัดค่าตัวแปรในแต่ละกริดเซลล์ พบว่าประชาชนมีความรู้สึกนึกคิดต่อเรื่องใดเรื่องหนึ่งคล้ายกัน (สมมุติว่าค่าตัวแปรที่วัดได้เท่ากัน) แบบจำลองการสุ่มมีทั้งหมด 6 แบบ ได้แก่ S0, S00, S1, S2, S3 และ S4 ซึ่งแต่ละแบบมีเนื้อหาและลักษณะที่ต่างกัน เนื่องจากถูกออกแบบและพัฒนาด้วยแนวคิดที่ต่างกัน โดยแต่ละแบบมีรายละเอียด ดังนี้

แบบจำลอง Sampling 0 (S0)

<u>การประมาณค่าเชิงพื้นที่</u>

แบบจำลอง S0 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อศึกษาค่าสถิติที่ได้จากการทดลองกรณีที่ธรรมชาติของกลุ่มประชากรไม่มีความสัมพันธ์ทางพื้นที่ (ไม่เป็นไปตามข้อตกลงเบื้องต้น) โดยกำหนดให้ธรรมชาติของกลุ่มประชากรมีสหสัมพันธ์เป็นศูนย์ โดยระดับทัศนคติของแบบจำลอง S0 แบ่งได้ทั้งหมด 9 ระดับ โดยเรียงจากค่าน้อยไปหาค่ามาก (จาก 1 จนถึง 9) เช่น หมายเลข 1 หมายความว่าน้อยที่สุด หรือหมายเลข 9 หมายความว่ามากที่สุด ดังภาพที่ 3-10

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-10 แบบจำลอง Sampling 0 (S0)

<u>ประชากร</u>

ประชากรของแบบจำลอง S0 เท่ากับ 81 ช่อง

<u>กลุ่มตัวอย่าง</u>

กลุ่มตัวอย่างของแบบจำลอง S0 เท่ากับ 9 ช่อง

วิธีการสุ่ม 5 วิธีที่นำมาใช้กับแบบจำลอง S0 คือ การสุ่มอย่างง่าย การสุ่มแบบมีระบบ การสุ่มแบบชั้นภูมิ การสุ่มแบบแบ่งกลุ่ม และการสุ่มแบบจุดกริด ซึ่งวิธีการสุ่มเกือบทั้งหมดสามารถเขียนโปรแกรมสร้างแบบจำลอง S0 ยกเว้นการสุ่มแบบชั้นภูมิเท่านั้น เพราะการสุ่มแบบชั้นภูมิมีหลักการว่า ภายในแต่ละชั้นภูมิต้องมีลักษณะเป็นเอกพันธ์ แต่ประชากรของแบบจำลอง S0 มีค่าสหสัมพันธ์เป็น 0 จึงไม่สามารถทำการสุ่มแบบชั้นภูมิได้ ทำให้เหลือการวิธีสุ่มเพียง 4 วิธี ซึ่งมีรายละเอียด ดังนี้

<u>การสุ่มอย่างง่ายของแบบจำลอง S0</u>

ทำการสุ่มกลุ่มตัวอย่าง 9 ช่อง จากประชากรทั้งหมด 81 ช่อง ซึ่งวิธีการสุ่มนี้ไม่มีกฎหรือกติกาใด ๆ ดังนั้นทุก ๆ ช่อง มีโอกาสที่จะถูกเลือกอย่างละเท่า ๆ กัน ที่ความน่าจะเป็นเท่ากับ 1/9

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-11 การสุ่มอย่างง่ายของแบบจำลอง S0

จากภาพที่ 3-11 ช่องสีเหลือง คือ สมาชิกของกลุ่มตัวอย่างทั้ง 9 ตัว ที่ถูกสุ่มมาเป็นตัวแทนของประชากร เพื่อศึกษาค่าเฉลี่ยกับความแปรปรวนและใช้เปรียบเทียบค่าพารามิเตอร์ โดยต้องทำการทดลองทั้งหมด 10,000 ครั้ง (ซึ่งตามรูปภาพคือผลของการสุ่มกลุ่มตัวอย่างเพียงครั้งเดียว)

<u>การสุ่มแบบมีระบบของแบบจำลอง S0</u>

เนื่องจากประชากรมีทั้งหมด 81 ช่อง และกลุ่มตัวอย่างมี 9 ช่อง ดังนั้นระยะห่างระหว่างสมาชิก คือ 9 ช่อง (= 81/9) โดยการเลือกสมาชิกทั้ง 9 ช่อง จะเลือกจากตำแหน่งด้านซ้ายไปทางด้านขวาและจากตำแหน่งด้านบนลงสู่ด้านล่าง (คล้ายการอ่านหนังสือ)

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-12 การสุ่มแบบมีระบบของแบบจำลอง S0

จากภาพที่ 3-12 ช่องสีเหลือง คือ สมาชิกของกลุ่มตัวอย่างทั้ง 9 ช่อง ที่ถูกเลือกให้เป็นตัวแทนประชากร โดยเริ่มจากแถวที่ 1 คอลัมน์ที่ 2 เป็นจุดเริ่มต้น จากนั้นนับไปทางขวาจากบนลงล่างอีก 9 ช่อง จะได้สมาชิกตัวที่ 2 คือ แถวที่ 2 คอลัมน์ที่ 2 และกระทำซ้ำ ๆ จนกว่าจะครบ 9 ช่อง สำหรับการกำหนดตำแหน่งจุดเริ่มต้นหรือกริดเซลล์แรกใช้วิธีการสุ่มอย่างง่าย

<u>การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S0</u>

แบบจำลอง S0 สามารถแบ่งประชากรออกเป็น 3 กลุ่ม ตามแนวคอลัมน์ ได้แก่ กลุ่มที่ 1 คือ คอลัมน์ที่ 1-3 (มีความแปรปรวน 5.41), กลุ่มที่ 2 คือ คอลัมน์ที่ 4-6 (มีความแปรปรวน 5.26) และกลุ่มที่ 3 คือ คอลัมน์ที่ 7-9 (มีความแปรปรวน 4.66) จากการแบ่งกลุ่มพบว่าความแปรปรวนภายในสูงแต่ความแปรปรวนระหว่างกลุ่มใกล้เคียงกัน ตามภาพที่ 3-13 ดังนั้นการแบ่งกลุ่มเป็นไปตามทฤษฎีการสุ่มแบบแบ่งกลุ่ม โดยประชากรกลุ่มย่อย (Sub-population) ต้องถูกสุ่มออกมา 3 ช่อง รวมทั้งสิ้น 9 ช่อง

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-13 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S0

<u>การสุ่มแบบจุดกริดของแบบจำลอง S0</u>

ตามทฤษฎีการสุ่มแบบกริดกำหนดตำแหน่งจุดกริดต้องกระจัดกระจายและครอบคลุมพื้นที่ทั้งหมด จากภาพที่ 3-14 สมาชิกของกลุ่มตัวอย่างทั้ง 9 ช่อง มีตำแหน่งกระจายอยู่ทั่วทุกสารทิศ มีระยะห่างตามแนวนอนเท่ากัน คือ 2 ช่อง และมีระยะห่างตามแนวตั้งเท่ากัน คือ 2 ช่อง

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	2	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-14 การสุ่มแบบจุดกริดของแบบจำลอง S0

แบบจำลอง Sampling 00 (S00)

<u>การประมาณค่าเชิงพื้นที่</u>

วัตถุประสงค์การสร้างแบบจำลอง S00 เพื่อศึกษาแบบจำลองเหตุการณ์ในกรณีที่มีโชคดีในการสุ่ม ทำให้ค่าเฉลี่ยกลุ่มตัวอย่างที่สุ่มด้วยวิธีการสุ่มแบบกริดมีค่าใกล้เคียงค่าเฉลี่ยของประชากร ดังนั้นแบบจำลอง S00 จึงเหมือนกับแบบจำลอง S0 เกือบทั้งหมด ยกเว้นแถวที่ 2 คอลัมน์ที่ 8 ให้เปลี่ยนจาก 2 เป็น 8 ดังภาพที่ 3-15 เนื่องจากการเป็นตัวแทนที่ดีของกลุ่มประชากรนั้นวัดได้จากค่าเฉลี่ยและความแปรปรวน เมื่อค่าเฉลี่ยของกลุ่มตัวอย่างมีค่าใกล้เคียงกับค่าพารามิเตอร์ ดังนั้นค่าความแปรปรวนของกลุ่มตัวอย่างที่ได้จากวิธีการสุ่มแบบกริดจะเป็นตัวบ่งชี้ว่ากลุ่มตัวอย่างที่สุ่มได้นั้นเป็นตัวแทนที่ดีจริงหรือไม่ ในส่วนการสุ่มตัวอย่างทั้ง 5 วิธีของแบบจำลอง S00 นั้น มีวิธีการทำเหมือนกันกับแบบจำลอง S0 ทุกประการ

1	7	4	2	9	5	6	4	5
8	5	8	3	1	5	6	8	6
2	4	5	8	5	6	3	7	3
5	8	9	4	7	7	2	9	4
1	5	6	7	4	3	5	8	3
3	7	4	5	8	6	7	6	4
5	6	2	6	4	9	3	8	2
4	1	5	9	9	2	4	6	5
3	6	8	5	5	4	9	7	8

ภาพที่ 3-15 แบบจำลอง Sampling 00 (S00)

<u>ประชากร</u>

ประชากรของแบบจำลอง S00 เท่ากับ 81 ช่อง

<u>กลุ่มตัวอย่าง</u>

กลุ่มตัวอย่างของแบบจำลอง S00 เท่ากับ 9 ช่อง

แบบจำลอง Sampling 1 (S1)

การประมาณค่าเชิงพื้นที่

แบบจำลอง S1 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อศึกษาค่าสถิติที่ได้จากการทดลอง กรณีที่ธรรมชาติของกลุ่มประชากรเป็นมีการประมาณค่าเชิงพื้นที่ ซึ่งระดับทัศนคติของแบบจำลอง S1 แบ่งได้ทั้งหมด 5ระดับ ตั้งแต่ 1 จนถึง 5 โดยค่าตัวแปรเท่ากับ 1 หมายถึงน้อยที่สุด และค่าตัวแปรเท่ากับ 5 หมายถึง มากที่สุด จากภาพที่ 3-16 แสดงให้เห็นว่าแบบจำลอง S1 มีค่าลดลงในแนวแกนตั้ง (y) และมีค่ามากขึ้นจากแนวกึ่งกลางแผนที่ไปยังด้านซ้ายและด้านขวา

3	2	2	2	2	2	2	2	3
3	3	3	2	2	2	3	3	3
3	3	3	3	2	3	3	3	3
4	3	3	3	3	3	3	3	4
4	4	4	3	3	3	4	4	4
4	4	4	4	3	4	4	4	4
5	4	4	4	4	4	4	4	5
5	5	5	4	4	4	5	5	5
5	5	5	5	4	5	5	5	5

ภาพที่ 3-16 แบบจำลอง Sampling 1 (S1)

ประชากร

ประชากรของแบบจำลอง S1 เท่ากับ 81 ช่อง

กลุ่มตัวอย่าง

กลุ่มตัวอย่างของแบบจำลอง S1 เท่ากับ 9 ช่อง

ในส่วนการสุ่มอย่างง่าย การสุ่มแบบมีระบบ และการสุ่มแบบจุดกริดของแบบจำลอง S1 นั้น มีวิธีการทำเหมือนกันกับแบบจำลอง S0 และ S00 แต่ในส่วนที่แตกต่างกันคือ การสุ่มแบบชั้นภูมิ และการสุ่มแบบแบ่งกลุ่ม เมื่อนำการสุ่มทั้ง 2 วิธี ไปประยุกต์ใช้กับแบบจำลอง S1 ได้ผลลัพธ์ ดังนี้

การสุ่มแบบชั้นภูมิของแบบจำลอง S1

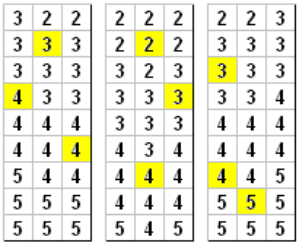
แบบจำลอง S1สามารถแบ่งประชากรออกเป็น 4 ชั้นภูมิตามแถว ได้แก่ ชั้นภูมิที่ 1 คือ แถวที่ 1 (มีความแปรปรวน 0.194), ชั้นภูมิที่ 2 คือ แถวที่ 2-4 (มีความแปรปรวน 0.225), ชั้นภูมิที่ 3 คือ แถวที่ 5-7 (มีความแปรปรวน 0.225) และชั้นภูมิที่ 4 คือ แถวที่ 8-9 (มีความแปรปรวน 0.183) จากการแบ่งชั้นภูมิ พบว่า ความแปรปรวนภายในต่ำ แต่ความแปรปรวนระหว่างกลุ่มแตกต่างกันสูง ดังภาพที่ 3-17 ซึ่งการสุ่มตัวอย่างจากประชากรกลุ่มย่อยให้จัดสรรตัวอย่างตามสัดส่วน โดยชั้นภูมิที่ 1 ต้องสุ่มตัวอย่างสมาชิกออกมา 1 ช่อง, ชั้นภูมิที่ 2 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่อง, ชั้นภูมิที่ 3 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่อง และชั้นภูมิที่ 4 ต้องสุ่มตัวอย่างสมาชิกออกมา 2 ช่อง

3	2	2	2	**2**	2	2	2	3
3	**3**	3	2	2	2	**3**	3	3
3	3	3	3	2	3	3	3	3
4	3	3	3	3	**3**	3	3	4
4	4	4	3	3	3	4	4	4
4	**4**	4	4	**3**	**4**	4	4	4
5	4	4	4	4	4	4	4	5
5	5	5	4	4	4	**5**	5	5
5	5	5	**5**	4	5	5	5	5

ภาพที่ 3-17 การสุ่มแบบชั้นภูมิของแบบจำลอง S1

การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S1

แบบจำลอง S1 สามารถแบ่งประชากรออกเป็น 3 กลุ่ม ตามแนวคอลัมน์ ได้แก่ กลุ่มที่ 1 คือ คอลัมน์ที่ 1-3 (มีความแปรปรวน 0.872), กลุ่มที่ 2 คือ คอลัมน์ที่ 4-6 (มีความแปรปรวน 0.872) และกลุ่มที่ 3 คือ คอลัมน์ที่ 7-9 (มีความแปรปรวน 0.872) จากการแบ่งกลุ่มพบว่าความแปรปรวนภายในสูงแต่ความแปรปรวนระหว่างกลุ่มมีค่าเท่ากัน ดังภาพที่ 3-18 ซึ่งการสุ่มตัวอย่างจากประชากรกลุ่มย่อยให้จัดสรรตัวอย่างตามสัดส่วน โดยกลุ่มที่ 1 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่อง กลุ่มที่ 2 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่องและกลุ่มที่ 3 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่อง

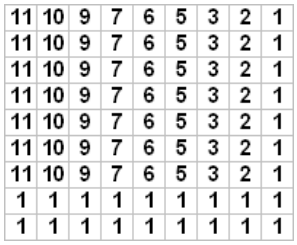


ภาพที่ 3-18 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S1

แบบจำลอง Sampling 2 (S2)

การประมาณค่าเชิงพื้นที่

แบบจำลอง S2 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อเปรียบเทียบประสิทธิภาพของการสุ่มทั้ง 5 แบบ ซึ่งระดับทัศนคติของแบบจำลอง S2 แบ่งได้ทั้งหมด 11 ระดับ โดยเริ่มตั้งแต่ 1 จนถึง 11 ซึ่งหมายเลข 1 หมายความว่าน้อยที่สุด และหมายเลข 11 หมายความว่ามากที่สุด จากภาพที่ 3-19 แสดงให้เห็นว่าแบบจำลอง S2 มี 2 แถวล่างสุดเท่านั้นที่คงที่ ดังนั้นพื้นที่ที่มีการเปลี่ยนแปลงจะมีขนาด 7 x 9 ช่อง ซึ่งแนวแกนนอน (x) มีค่าลดลงจากซ้ายไปขวา แต่แนวแกนตั้ง (y) มีค่าคงที่



ภาพที่ 3-19 แบบจำลอง Sampling 2 (S2)

ประชากร

ประชากรของแบบจำลอง S2 เท่ากับ 81 ช่อง

กลุ่มตัวอย่าง

กลุ่มตัวอย่างของแบบจำลอง S2 เท่ากับ 9 ช่อง

ในส่วนการสุ่มอย่างง่าย การสุ่มแบบมีระบบ และการสุ่มแบบจุดกริดของแบบจำลอง S2 นั้น มีวิธีการทำเหมือนกันกับแบบจำลอง S0, S00 และ S1 แต่ในส่วนที่แตกต่างกันคือ การสุ่มแบบชั้นภูมิและการสุ่มแบบแบ่งกลุ่ม เมื่อนำการสุ่มทั้ง 2 วิธีไปประยุกต์ใช้กับแบบจำลอง S2 ได้ผลลัพธ์ ดังนี้

การสุ่มแบบชั้นภูมิของแบบจำลอง S2

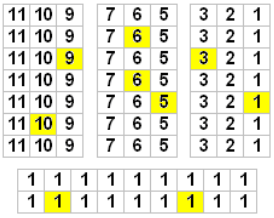
จากภาพที่ 3-20 แสดงให้เห็นการแบ่งประชากรออกเป็น 4 ชั้นภูมิตามแถว ได้แก่ ชั้นภูมิที่ 1 คือ แถวล่างที่สุดสองแถว, ชั้นภูมิที่ 2 คือ คอลัมน์ที่ 1-3, ชั้นภูมิที่ 3 คือ คอลัมน์ที่ 4-6 และชั้นภูมิที่ 4 คือ คอลัมน์ที่ 7-9 จากการแบ่งชั้นภูมิ พบว่า มีลักษณะภายในคล้ายกันหรือเป็นเอกพันธ์ (Homogeneous) แต่มีความแตกต่างกันระหว่างชั้นภูมิ ตามแผนการสุ่มของแบบจำลอง S2 มีขนาดกลุ่มตัวอย่างเท่ากับ 9 ช่อง ซึ่งมีรายละเอียดการสุ่ม ดังนี้

ชั้นภูมิที่ 1 ต้องสุ่มตัวอย่างสมาชิกออกมา 2 ช่อง

ชั้นภูมิที่ 2 ต้องสุ่มตัวอย่างสมาชิกออกมา 2 ช่อง

ชั้นภูมิที่ 3 ต้องสุ่มตัวอย่างสมาชิกออกมา 3 ช่อง

ชั้นภูมิที่ 4 ต้องสุ่มตัวอย่างสมาชิกออกมา 2 ช่อง

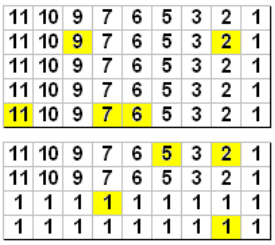


ภาพที่ 3-20 การสุ่มแบบชั้นภูมิของแบบจำลอง S2

การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S2

การแบ่งกลุ่มของแบบจำลอง S2 สามารถแบ่งประชากรออกเป็น 2 กลุ่ม ตามแถว ได้แก่ กลุ่มที่ 1 คือ แถวที่ 1-5 (มีความแปรปรวน 11.591), กลุ่มที่ 2 คือ แถวที่ 6-9 (มีความแปรปรวน 12.257) ดังภาพที่ 3-21 จากการแบ่งกลุ่ม พบว่า ความแปรปรวนภายในสูงแต่ความแปรปรวน

ระหว่างใกล้เคียงกัน เพื่อให้กลุ่มตัวอย่างครบตามแผนการสุ่ม ดังนั้นกลุ่มที่ 1 ต้องสุ่มตัวอย่างสมาชิกออกมา 5 ช่อง และกลุ่มที่ 2 ต้องสุ่มตัวอย่างสมาชิกออกมา 4 ช่อง

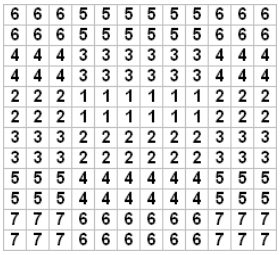


ภาพที่ 3-21 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S2

แบบจำลอง Sampling 3 (S3)

<u>การประมาณค่าเชิงพื้นที่</u>

แบบจำลอง S3 เป็นแบบจำลองที่ถูกออกแบบให้แผนที่มีขนาดใหญ่ขึ้น เพื่อศึกษาผลกระทบที่เกิดขึ้นในกรณีที่มีการเปลี่ยนแปลงขนาดของแผนที่ ซึ่งระดับทัศนคติของแบบจำลอง S3 แบ่งได้ทั้งหมด 7 ระดับ โดยเรียงจากค่าน้อยที่สุด (หมายเลข 1) ไปหาค่าที่มากที่สุด (หมายเลข 7) จากภาพที่ 3-22 แสดงให้เห็นว่าแบบจำลอง S3 พื้นที่ภายในมีค่าตัวแปรน้อยและน้อยที่สุดที่บริเวณใกล้ ๆ กับจุดกึ่งกลางแผนที่ (ซึ่งอยู่ทางทิศเหนือ) แต่พื้นที่ภายนอกมีค่าตัวแปรสูงกว่าและมีค่าสูงที่สุดที่บริเวณห่างไกลจากจุดกึ่งกลางแผนที่ ซึ่งอยู่ทางทิศตะวันออกเฉียงใต้ (มุมขวาด้านล่างของแผนที่) และทางทิศตะวันตกเฉียงใต้ (มุมซ้ายด้านล่างของแผนที่)



ภาพที่ 3-22 แบบจำลอง Sampling 3 (S3)

<u>ประชากร</u>

ประชากรของแบบจำลอง S3 เท่ากับ 144 ช่อง

กลุ่มตัวอย่าง

กลุ่มตัวอย่างของแบบจำลอง S3 เท่ากับ 36 ช่อง

สำหรับการสุ่มอย่างง่ายของแบบจำลอง S3 นั้น มีวิธีการทำเหมือนกันกับแบบจำลอง S2 แต่อีก 4 วิธีนั้นมีรายละเอียดที่แตกต่างจากแบบจำลองอื่น ๆ เล็กน้อย ดังนี้

การสุ่มแบบมีระบบของแบบจำลอง S3

เนื่องจากประชากรมีทั้งหมด 144 ช่อง และกลุ่มตัวอย่างมี 36 ช่อง ดังนั้นระยะห่างระหว่างสมาชิก คือ 4 ช่อง (= 144/36) โดยการเลือกสมาชิกทั้ง 36 ช่อง จะเลือกจากตำแหน่งด้านซ้ายไปทางด้านขวาและจากตำแหน่งด้านบนลงสู่ด้านล่าง (คล้ายการอ่านหนังสือ)

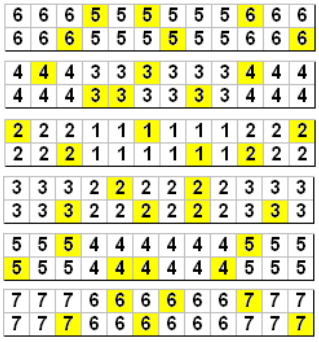
6	6	6	5	5	5	5	5	6	6	6
6	6	6	5	5	5	5	5	6	6	6
4	4	4	3	3	3	3	3	4	4	4
4	4	4	3	3	3	3	3	4	4	4
2	2	2	1	1	1	1	1	2	2	2
2	2	2	1	1	1	1	1	2	2	2
3	3	3	2	2	2	2	2	3	3	3
3	3	3	2	2	2	2	2	3	3	3
5	5	5	4	4	4	4	4	5	5	5
5	5	5	4	4	4	4	4	5	5	5
7	7	7	6	6	6	6	6	7	7	7
7	7	7	6	6	6	6	6	7	7	7

ภาพที่ 3-23 การสุ่มแบบมีระบบของแบบจำลอง S3

จากภาพที่ 3-23 ช่องที่มีสีทั้งหมด คือ สมาชิกของกลุ่มตัวอย่างทั้ง 36 ช่อง ที่ถูกเลือกให้เป็นตัวแทนประชากร โดยช่องแรกที่ถูกสุ่ม คือ แถวที่ 1 คอลัมน์ที่ 6 (ช่องสีฟ้า) เป็นจุดเริ่มต้น จากนั้นนับไปทางขวาจากบนลงล่างอีก 4 ช่อง (วิธีการนับคล้ายกับการอ่านหนังสือ) จะได้สมาชิกตัวที่ 2 คือ แถวที่ 1 คอลัมน์ที่ 10 และกระทำซ้ำ ๆ จนกว่าจะครบ 36 ช่อง แต่ในช่องที่ 35 เป็นแถวสุดท้ายแล้ว (แถวที่ 12 คอลัมน์ที่ 10) ดังนั้นต้องนับย้อนกลับขึ้นมาใหม่ ดังนั้นช่องที่ 36 คือ แถวที่ 1 คอลัมน์ที่ 2 หรือช่องสีม่วง สำหรับการกำหนดตำแหน่งจุดเริ่มต้นหรือกริดเซลล์แรกใช้วิธีการสุ่มอย่างง่าย

การสุ่มแบบชั้นภูมิของแบบจำลอง S3

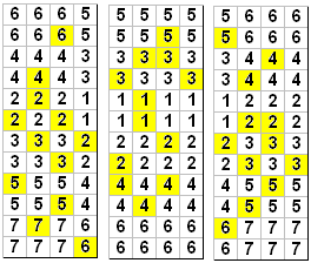
ตามทฤษฎีการสุ่มแบบแบ่งชั้นภูมิสามารถแบ่งประชากรของแบบจำลอง S3 ออกเป็น 6 ชั้นภูมิ ได้แก่ ชั้นภูมิที่ 1 คือ แถวที่ 1-2, ชั้นภูมิที่ 2 คือ แถวที่ 3-4, ชั้นภูมิที่ 3 คือ แถวที่ 5-6, ชั้นภูมิที่ 4 คือ แถวที่ 7-8, ชั้นภูมิที่ 5 คือ แถวที่ 9-10 และชั้นภูมิที่ 6 คือ แถวที่ 11-12 จากการแบ่งชั้นภูมิพบว่า มีลักษณะภายในคล้ายกันหรือเป็นเอกพันธ์ (Homogeneous) แต่มีความแตกต่างกันระหว่างชั้นภูมิ ดังภาพที่ 3-24 เพื่อให้กลุ่มตัวอย่างมีจำนวนครบ 36 หน่วย ตามแผนการสุ่ม ดังนั้นในแต่ละชั้นภูมิต้องสุ่มตัวอย่างออกมา 6 ช่อง จึงครบ 36 ช่อง



ภาพที่ 3-24 การสุ่มแบบชั้นภูมิของแบบจำลอง S3

การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S3

ตามทฤษฎีการสุ่มแบบแบ่งกลุ่มสามารถแบ่งประชากรของแบบจำลอง S3 ออกเป็น 3 กลุ่ม ได้แก่ กลุ่มที่ 1 คือ คอลัมน์ที่ 1-3, กลุ่มที่ 2 คือ คอลัมน์ที่ 4-6 และกลุ่มที่ 3 คือ คอลัมน์ที่ 7-9 จากการแบ่งกลุ่ม พบว่า ความแปรปรวนภายในสูงแต่ความแปรปรวนระหว่างกลุ่มใกล้เคียงกัน ดังภาพที่ 3-25 เพื่อให้กลุ่มตัวอย่างมีจำนวนครบตามแผนการสุ่ม ดังนั้นในแต่ละกลุ่มประชากรย่อยต้องสุ่มตัวอย่างออกมา 12 ช่อง จึงจะครบ 36 ช่อง



ภาพที่ 3-25 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S3

การสุ่มแบบจุดกริดของแบบจำลอง S3

เมื่อประยุกต์ใช้แนวคิดของการสุ่มแบบจุดกริดกับแบบจำลอง S3 ผลลัพธ์ที่เกิดขึ้นเป็นไปตามภาพที่ 3-26 ซึ่งสมาชิกของกลุ่มตัวอย่างทั้ง 36 ช่อง อยู่อย่างกระจัดกระจายและครอบคลุมพื้นที่ทั้งหมดของแผนที่ โดยมีระยะห่างตามแนวนอนเท่ากัน คือ 2 ช่อง และมีระยะห่างตามแนวตั้งเท่ากัน คือ 2 ช่อง

6	6	6	5	5	5	5	5	5	6	6	6
6	6	6	5	5	5	5	5	5	6	6	6
4	4	4	3	3	3	3	3	3	4	4	4
4	4	4	3	3	3	3	3	3	4	4	4
2	2	2	1	1	1	1	1	1	2	2	2
2	2	2	1	1	1	1	1	1	2	2	2
3	3	3	2	2	2	2	2	2	3	3	3
3	3	3	2	2	2	2	2	2	3	3	3
5	5	5	4	4	4	4	4	4	5	5	5
5	5	5	4	4	4	4	4	4	5	5	5
7	7	7	6	6	6	6	6	6	7	7	7
7	7	7	6	6	6	6	6	6	7	7	7

ภาพที่ 3-26 การสุ่มแบบจุดกริดของแบบจำลอง S3

แบบจำลอง Sampling 4 (S4)

<u>การประมาณค่าเชิงพื้นที่</u>

แบบจำลอง S4 เป็นแบบจำลองที่ถูกพัฒนาจากแบบจำลอง S3 โดยออกแบบให้แผนที่ค่าตัวแปรที่หลากหลาย เพื่อศึกษาประสิทธิภาพของการสุ่มแต่ละแบบ โดยตัวเลขในแต่ละช่องแทนคะแนนที่วัดได้จากแบบทดสอบ ซึ่งมีคะแนนเต็ม 100 คะแนน จากภาพที่ 3-27 แสดงให้เห็นว่าพื้นที่ภายในของแบบจำลอง S4 มีค่าตัวแปรน้อยที่บริเวณใกล้ ๆ กับจุดกึ่งกลางแผนที่ แต่พื้นที่ภายนอกมีค่าตัวแปรสูงที่บริเวณห่างไกลจากจุดกึ่งกลางแผนที่ ดังนั้นการเพิ่มของค่าตัวแปรมีทิศทางจากด้านในไปยังด้านนอก

80	85	76	69	67	74	70	76	79	88	92	91
78	86	75	64	68	70	69	74	66	85	88	90
57	59	50	41	45	44	38	40	45	58	49	52
53	58	60	40	44	39	46	38	42	58	54	60
27	25	32	7	16	14	18	9	8	30	32	24
26	24	23	10	15	11	9	16	14	26	29	23
41	45	38	30	26	25	32	30	33	41	45	39
40	42	38	35	34	29	26	31	28	46	40	41
72	69	75	60	58	53	56	54	50	74	75	68
72	77	68	55	58	60	59	55	50	70	70	70
100	98	90	88	86	80	78	76	82	91	94	96
94	90	99	82	84	84	88	80	78	94	100	92

ภาพที่ 3-27 แบบจำลอง Sampling 4 (S4)

<u>ประชากร</u>

ประชากรของแบบจำลอง S4 เท่ากับ 144 ช่อง

<u>กลุ่มตัวอย่าง</u>

กลุ่มตัวอย่างของแบบจำลอง S4 เท่ากับ 36 ช่อง

ในส่วนการสุ่มอย่างง่าย การสุ่มแบบมีระบบ และการสุ่มแบบจุดกริดของแบบจำลอง S4 นั้น มีวิธีการทำเหมือนกันกับแบบจำลอง S3 แต่ในส่วนที่แตกต่างกันคือ การสุ่มแบบชั้นภูมิและการสุ่มแบบแบ่งกลุ่ม เมื่อนำการสุ่มทั้ง 2 วิธีไปประยุกต์ใช้กับแบบจำลอง S4 ได้ผลลัพธ์ ดังนี้

การสุ่มแบบชั้นภูมิของแบบจำลอง S4

ประชากรของแบบจำลอง S4 ถูกแบ่งออกเป็น 6 ชั้นภูมิตามแถว ได้แก่ ชั้นภูมิที่ 1 คือ แถวที่ 1-2, ชั้นภูมิที่ 2 คือ แถวที่ 3-4, ชั้นภูมิที่ 3 คือ แถวที่ 5-6, ชั้นภูมิที่ 4 คือ แถวที่ 7-8, ชั้นภูมิที่ 5 คือแถวที่ 9-10 และชั้นภูมิที่ 6 คือ แถวที่ 11-12 (ดังรูปภาพที่ 3-28) จากการแบ่งชั้นภูมิ พบว่า มีลักษณะภายในคล้ายกันหรือเป็นเอกพันธ์ (Homogeneous) แต่มีความแตกต่างกันระหว่างชั้นภูมิ สำหรับการจัดสรรตัวอย่างในแต่ละชั้นภูมิเป็นแบบสัดส่วน โดยในแต่ละชั้นภูมิต้องสุ่มตัวอย่างออกมา 6 ช่อง จึงจะครบ 36 ตามที่วางแผนไว้

80	85	76	69	67	74	70	76	79	88	92	91
78	86	75	64	68	70	69	74	66	85	88	90
57	59	50	41	45	44	38	40	45	58	49	52
53	58	60	40	44	39	46	38	42	58	54	60
27	25	32	7	16	14	18	9	8	30	32	24
26	24	23	10	15	11	9	16	14	26	29	23
41	45	38	30	26	25	32	30	33	41	45	39
40	42	38	35	34	29	26	31	28	46	40	41
72	69	75	60	58	53	56	54	50	74	75	68
72	77	68	55	58	60	59	55	50	70	70	70
100	98	90	88	86	80	78	76	82	91	94	96
94	90	99	82	84	84	88	80	78	94	100	92

ภาพที่ 3-28 การสุ่มแบบชั้นภูมิของแบบจำลอง S4

การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S4

ประชากรของแบบจำลอง S4 ถูกแบ่งออกเป็น 3 กลุ่ม ตามแนวคอลัมน์ได้แก่ กลุ่มที่ 1 คือ คอลัมน์ที่ 1-3, กลุ่มที่ 2 คือ คอลัมน์ที่ 4-6 และกลุ่มที่ 3 คือคอลัมน์ที่ 7-9 (ดังภาพที่ 3-29) จากการแบ่งกลุ่ม พบว่า ความแปรปรวนภายในสูง แต่ความแปรปรวนระหว่างกลุ่มใกล้เคียงกัน สำหรับการจัดสรรตัวอย่างในแต่ละกลุ่มเป็นแบบสัดส่วน โดยในแต่ละชั้นภูมิต้องสุ่มตัวอย่างออกมา 12 ช่อง จึงจะครบ 36 ตามที่วางแผนไว้

80	85	76	69	67	74	70	76	79	88	92	91
78	86	75	64	68	70	69	74	66	85	88	90
57	59	50	41	45	44	38	40	45	58	49	52
53	58	60	40	44	39	46	38	42	58	54	60
27	25	32	7	16	14	18	9	8	30	32	24
26	24	23	10	15	11	9	16	14	26	29	23
41	45	38	30	26	25	32	30	33	41	45	39
40	42	38	35	34	29	26	31	28	46	40	41
72	69	75	60	58	53	56	54	50	74	75	68
72	77	68	55	58	60	59	55	50	70	70	70
100	98	90	88	86	80	78	76	82	91	94	96
94	90	99	82	84	84	88	80	78	94	100	92

ภาพที่ 3-29 การสุ่มแบบแบ่งกลุ่มของแบบจำลอง S4

แบบจำลองการประมาณค่าเชิงพื้นที่ (Spatial interpolation/ SI)

แบบจำลองการประมาณค่าเชิงพื้นที่มีลักษณะคล้ายกันกับแบบจำลองการสุ่มเกือบทุกประการ ทั้งในด้านหลักการ แนวคิด และวิธีการสร้าง แต่แตกต่างกันตรงการนำไปใช้เท่านั้น ในการทดลองครั้งนี้ได้สร้างแบบจำลองหลายรูปแบบหลายขนาด เพื่อศึกษาผลกระทบที่เกิดขึ้นจากแบบจำลอง เมื่อมีการเปลี่ยนแปลงทิศทางและขนาดความสัมพันธ์ระหว่างค่าตัวแปรกับพิกัดในแผนที่ โดยแบบจำลองการประมาณเชิงพื้นที่มีทั้งหมดจำนวน 3 แบบ ได้แก่ SI1, SI2 และ SI3 ซึ่งแบบจำลอง SI1 เป็นการจำลองสถานการณ์พื้นที่ ซึ่งเก็บรวบรวมข้อมูลด้วยแบบสอบถามความคิดเห็นด้วยมาตรวัดแบบสเกลทัศนคติ (Likert scale) จำนวน 7 ระดับ, แบบจำลอง SI2 เป็นการจำลองสถานการณ์พื้นที่ ซึ่งเก็บรวบรวมข้อมูลด้วยแบบทดสอบที่มีคะแนนเต็ม 100 คะแนน และแบบจำลอง SI3 เป็นการจำลองสถานการณ์พื้นที่ ซึ่งเก็บรวบรวมข้อมูลด้วยแบบสอบถามความคิดเห็นด้วยมาตรวัดแบบสเกลทัศนคติ (Likert scale) จำนวน 5 ระดับ ซึ่งรายละเอียดของแบบจำลองมีดังนี้

แบบจำลอง Spatial interpolation 1 (SI1)

แบบจำลอง SI1 เป็นแผนที่ขนาด 12 x 12 ช่อง โดยมีระดับของทัศนคติหรือเจตคติของประชากร ณ ท้องถิ่นนั้น ๆ แบ่งได้ทั้งหมด 7 ระดับ โดยเรียงจากค่าน้อยไปหาค่ามาก (จาก 1 จนถึง 7) เช่น หมายเลข 1 หมายความว่าน้อยที่สุดหรือหมายเลข 7 หมายความว่ามากที่สุด จากภาพที่ 3-30 แสดงถึงทิศทางและระดับความเข้มของค่าตัวแปรแบบจำลอง SI1 ซึ่งพื้นที่ภายในมีค่าตัวแปรน้อยและน้อยที่สุดที่บริเวณใกล้ ๆ กับจุดกึ่งกลางแผนที่ (ซึ่งอยู่ทางทิศเหนือ) แต่พื้นที่ภายนอกมีค่าตัวแปรสูงกว่าและมีค่าสูงที่สุดที่บริเวณห่างไกลจากจุดกึ่งกลางแผนที่ (ซึ่งอยู่ทางทิศตะวันออกเฉียงใต้และทางทิศตะวันตกเฉียงใต้)

6	6	6	5	5	5	5	5	6	6	6
6	6	6	5	5	5	5	5	6	6	6
4	4	4	3	3	3	3	3	4	4	4
4	4	4	3	3	3	3	3	4	4	4
2	2	2	1	1	1	1	1	2	2	2
2	2	2	1	1	1	1	1	2	2	2
3	3	3	2	2	2	2	2	3	3	3
3	3	3	2	2	2	2	2	3	3	3
5	5	5	4	4	4	4	4	5	5	5
5	5	5	4	4	4	4	4	5	5	5
7	7	7	6	6	6	6	6	7	7	7
7	7	7	6	6	6	6	6	7	7	7

ภาพที่ 3-30 แบบจำลอง Spatial Interpolation 1 (SI1)

เมื่อนำแนวคิดของการสุ่มแบบจุดกริดมาประยุกต์ใช้กับแบบจำลอง SI1 ซึ่งประชากรของแบบจำลองเท่ากับ 144 ช่อง และถ้ากำหนดให้สุ่มตัวอย่าง 36 ช่อง ผลลัพธ์ที่ได้เป็นไปตาม

ภาพที่ 3-31 โดยตำแหน่งจุดกริด 9 จุด อยู่อย่างกระจัดกระจายและครอบคลุมพื้นที่ทั้งหมด โดยมีระยะห่างตามแนวนอนเท่ากัน คือ 2 ช่อง และมีระยะห่างตามแนวตั้งเท่ากันคือ 2 ช่อง ซึ่งค่าสถิติและค่าพารามิเตอร์ของแบบจำลอง SI1 คือ $\mu = 4$, $\sigma^2 = 3.167$, $\bar{z} = 4.167$ และ $s^2 = 3.229$

6	6	6	5	5	5	5	5	5	6	6	6
6	6	6	5	5	5	5	5	5	6	6	6
4	4	4	3	3	3	3	3	3	4	4	4
4	4	4	3	3	3	3	3	3	4	4	4
2	2	2	1	1	1	1	1	1	2	2	2
2	2	2	1	1	1	1	1	1	2	2	2
3	3	3	2	2	2	2	2	2	3	3	3
3	3	3	2	2	2	2	2	2	3	3	3
5	5	5	4	4	4	4	4	4	5	5	5
5	5	5	4	4	4	4	4	4	5	5	5
7	7	7	6	6	6	6	6	6	7	7	7
7	7	7	6	6	6	6	6	6	7	7	7

ภาพที่ 3-31 ตำแหน่งจุดกริดของแบบจำลอง SI1

แบบจำลอง SI1 เป็นแบบจำลองที่ถูกพัฒนาจากแบบจำลอง S3 เพื่อศึกษาการประมาณค่าเชิงพื้นที่จากวิธีการสุ่มแบบกริด โดยการเปรียบเทียบค่าประมาณกับค่าตัวแปรที่ได้กำหนดไว้ขั้นต้นเพื่อวิเคราะห์ค่าความคลาดเคลื่อนร้อยละ เมื่อนำวิธีการประมาณค่าเชิงพื้นที่มาประยุกต์ใช้กับแบบจำลอง SI1 โดยการหาค่าเฉลี่ยที่จุดกริดต่าง ๆ ซึ่งการประมาณค่าจุดกริดแบ่งออกเป็น 2 รูปแบบ คือ การประมาณค่าจุดกริดสีเหลืองและจุดกริดสีขาว ซึ่งจุดกริดสีเหลืองเป็นจุดกริดที่ทำการสำรวจข้อมูล และสามารถหาค่าเฉลี่ย ($\bar{z}i, j$) ได้จากสมาชิกทั้ง 4 ช่อง ยกตัวอย่างเช่น จุดกริดตะวันตกเฉียงใต้ $(i = 3, j = 1)$ ซึ่งมี $a_{3,1}^{(1)} = 5$, $a_{3,1}^{(2)} = 5$, $a_{3,1}^{(3)} = 7$ และ $a_{3,1}^{(4)} = 7$ ทำให้ $\bar{z}_{3,1} = 6$ แต่ในส่วนจุดกริดสีขาวต้องนำวิธีการประมาณค่าเชิงพื้นที่มาใช้คำนวณจุดกริดเหล่านั้น เมื่อทำการประมาณค่าจนครบทุกจุด ผลลัพธ์ที่เกิดขึ้นเป็นไปตามภาพที่ 3-32

5	4	4	4	5
3	2	2	2	3
3	2	2	2	3
4	3	3	3	4
6	5	5	5	6

ภาพที่ 3-32 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI1

แบบจำลอง Spatial interpolation 2 (SI2)

แบบจำลอง SI2 เป็นแผนที่นี้มีขนาด 12 x 12 ช่อง โดยตัวเลขในแต่ละช่องแทนคะแนนที่วัดได้จากแบบทดสอบ คะแนนเต็ม 100 คะแนน ซึ่งตัวเลขนี้ คือ ค่าเฉลี่ยของประชากร ณ ท้องถิ่นนั้น ๆ จากภาพที่ 3-33 แสดงถึงทิศทางและระดับความเข้มของค่าตัวแปรแบบจำลอง SI2 ซึ่งพื้นที่ภายในมีค่าตัวแปรแปรน้อยและน้อยที่สุดที่บริเวณใกล้ ๆ กับจุดกึ่งกลางแผนที่ (ซึ่งอยู่ทางทิศเหนือ) แต่พื้นที่ภายนอกมีค่าตัวแปรสูงกว่าและมีค่าสูงที่สุดที่บริเวณห่างไกลจากจุดกึ่งกลางแผนที่ ซึ่งอยู่ทางทิศตะวันออกเฉียงใต้และทางทิศตะวันตกเฉียงใต้

80	85	76	69	67	74	70	76	79	88	92	91
78	86	75	64	68	70	69	74	66	85	88	90
57	59	50	41	45	44	38	40	45	58	49	52
53	58	60	40	44	39	46	38	42	58	54	60
27	25	32	7	16	14	18	9	8	30	32	24
26	24	23	10	15	11	9	16	14	26	29	23
41	45	38	30	26	25	32	30	33	41	45	39
40	42	38	35	34	29	26	31	28	46	40	41
72	69	75	60	58	53	56	54	50	74	75	68
72	77	68	55	58	60	59	55	50	70	70	70
100	98	90	88	86	80	78	76	82	91	94	96
94	90	99	82	84	84	88	80	78	94	100	92

ภาพที่ 3-33 แบบจำลอง Spatial Interpolation 2 (SI2)

เมื่อนำแนวคิดของการสุ่มแบบจุดกริดมาประยุกต์ใช้กับแบบจำลอง SI2 ซึ่งประชากรของแบบจำลองเท่ากับ 144 ช่อง และถ้ากำหนดให้สุ่มตัวอย่าง 36 ช่อง ผลลัพธ์ที่ได้เป็นไปตามภาพที่ 3-34 โดยตำแหน่งจุดกริด 9 จุด อยู่อย่างกระจัดกระจายและครอบคลุมพื้นที่ทั้งหมด โดยมีระยะห่างตามแนวนอนเท่ากัน คือ 2 ช่อง และมีระยะห่างตามแนวตั้งเท่ากัน คือ 2 ช่อง ซึ่งค่าสถิติและค่าพารามิเตอร์ของแบบจำลอง SI2 คือ $\mu = 55.590$, $\sigma^2 = 623.409$, $\bar{z} = 57.056$ และ $s^2 = 627.768$

80	85	76	69	67	74	70	76	79	88	92	91
78	86	75	64	68	70	69	74	66	85	88	90
57	59	50	41	45	44	38	40	45	58	49	52
53	58	60	40	44	39	46	38	42	58	54	60
27	25	32	7	16	14	18	9	8	30	32	24
26	24	23	10	15	11	9	16	14	26	29	23
41	45	38	30	26	25	32	30	33	41	45	39
40	42	38	35	34	29	26	31	28	46	40	41
72	69	75	60	58	53	56	54	50	74	75	68
72	77	68	55	58	60	59	55	50	70	70	70
100	98	90	88	86	80	78	76	82	91	94	96
94	90	99	82	84	84	88	80	78	94	100	92

ภาพที่ 3-34 ตำแหน่งจุดกริดของแบบจำลอง SI2

แบบจำลอง SI2 เป็นแบบจำลองที่ถูกพัฒนาจากแบบจำลอง SI1 โดยออกแบบตัวแปรให้มีค่ามากและหลากหลาย เพื่อเปรียบเทียบผลการทดลองระหว่างแบบจำลอง SI1 ที่มีค่าตัวแปรน้อย ๆ กับแบบจำลอง SI2 ที่มีค่าตัวแปรมาก ๆ เพื่อวิเคราะห์ค่าความคลาดเคลื่อนร้อยละรวมทั้งประสิทธิภาพของการสุ่มแบบกริด เมื่อนำวิธีการประมาณค่าเชิงพื้นที่มาประยุกต์ใช้กับแบบจำลอง SI1 โดยการหาค่าเฉลี่ยทุกตำแหน่งของแผนที่จนครบทุกจุดกริด ผลลัพธ์ที่เกิดขึ้นเป็นไปตามภาพที่ 3-35

67.5	54.5	55.25	56.25	70
43.75	26.75	29.25	24.25	43.5
32.5	20.25	19.25	23.25	35.25
56	46.75	41	40.75	58.75
83.25	71.75	69.25	65.75	81.25

ภาพที่ 3-35 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI2

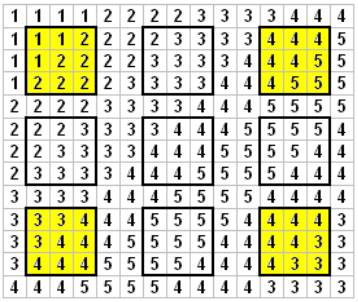
แบบจำลอง Spatial interpolation 3 (SI3)

แบบจำลอง SI3 เป็นแผนที่ขนาด 13 x 15 ช่อง โดยมีค่าระดับของทัศนคติหรือเจตคติของประชากร ณ ท้องถิ่นนั้น ๆ ซึ่งระดับทัศนคติของแบบจำลอง SI3 แบ่งได้ทั้งหมด 5ระดับ โดยเรียงจากค่าน้อยไปหาค่ามาก (จาก 1 จนถึง 5) เช่น หมายเลข 1 หมายความว่าน้อยที่สุดหรือหมายเลข 5 หมายความว่ามากที่สุด จากภาพที่ 3-36 แสดงให้เห็นว่าแบบจำลอง SI3 มีค่าเพิ่มมากขึ้นจากมุมบนด้านซ้ายของแผนที่ไปตามแนวทแยงมุมจนกระทั่งมีค่าตัวแปรมากที่สุดบริเวณตรงกลาง (แนวทแยงมุมลากจากขวาบนไปซ้ายล่าง) จากนั้นค่าตัวแปรเริ่มลดต่ำลงจนถึงมุมล่างด้านขวา

1	1	1	1	2	2	2	2	3	3	3	4	4	4
1	1	1	2	2	2	2	3	3	3	4	4	4	5
1	1	2	2	2	2	3	3	3	3	4	4	5	5
1	2	2	2	2	3	3	3	3	4	4	4	5	5	5
2	2	2	2	3	3	3	3	4	4	4	5	5	5	5
2	2	2	3	3	3	3	4	4	4	5	5	5	5	4
2	2	3	3	3	3	4	4	4	5	5	5	5	4	4
2	3	3	3	3	4	4	4	5	5	5	5	4	4	4
3	3	3	3	4	4	4	5	5	5	5	4	4	4	4
3	3	3	4	4	4	5	5	5	5	4	4	4	4	3
3	3	4	4	4	5	5	5	5	4	4	4	4	3	3
3	4	4	4	5	5	5	5	4	4	4	4	3	3	3
4	4	4	5	5	5	5	4	4	4	4	3	3	3	3

ภาพที่ 3-36 แบบจำลอง Spatial interpolation 3 (SI3)

เมื่อนำแนวคิดของการสุ่มแบบจุดกริดมาประยุกต์ใช้กับแบบจำลอง SI3 ซึ่งประชากรของแบบจำลองเท่ากับ 195 ช่อง และถ้ากำหนดให้สุ่มตัวอย่าง 36 ช่อง ผลลัพธ์ที่ได้เป็นไปตามภาพที่ 3-37 โดยตำแหน่งจุดกริด 4 จุด อยู่อย่างกระจัดกระจายและครอบคลุมพื้นที่ทั้งหมด โดยมีระยะห่างตามแนวนอนเท่ากันคือ 7 ช่อง และมีระยะห่างตามแนวตั้งเท่ากันคือ 5 ช่อง ซึ่งค่าสถิติและค่าพารามิเตอร์ของแบบจำลอง SI3 คือ $\mu = 3.528$, $\sigma^2 = 1.254$, $\bar{z} = 3.333$ และ $s^2 = 1.257$



ภาพที่ 3-37 ตำแหน่งจุดกริดของแบบจำลอง SI3

แบบจำลอง SI3 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อศึกษาผลกระทบในกรณีที่ค่าตัวแปรมีการเรียงตัวไม่เป็นไปตามแนวแกน x และแนวแกน y และเพื่อวิเคราะห์ค่าความคลาดเคลื่อนร้อยละรวมทั้งประสิทธิภาพของการสุ่มแบบกริด ในการประมาณค่าแบบจำลอง SI3 ต้องหาค่าเฉลี่ยตัวแปรในแต่ละพื้นที่ ในกรณีพื้นที่สีเหลือง (จุดกริด)ให้นำค่าตัวแปรของทุกกริดเซลล์ (9 ช่อง) มาคำนวณหาค่าเฉลี่ย ยกตัวอย่างเช่น จุดกริดตะวันตกเฉียงใต้ $(i = 3, j = 1)$ ซึ่งมี $a_{3,1}^{(1)} = 3$, $a_{3,1}^{(2)} = 3$, $a_{3,1}^{(3)} = 4$, $a_{3,1}^{(4)} = 3$ $a_{3,1}^{(5)} = 4$, $a_{3,1}^{(6)} = 4$, $a_{3,1}^{(7)} = 4$, $a_{3,1}^{(8)} = 4$ และ $a_{3,1}^{(9)} = 4$ ทำให้ $\bar{z}_{3,1} = 3.667$ แต่ในส่วนจุดกริดสีขาวต้องนำวิธีการประมาณค่าเชิงพื้นที่มาใช้คำนวณจุดกริดเหล่านั้น เมื่อทำการประมาณค่าจนครบทุกจุด ผลลัพธ์ที่เกิดขึ้นเป็นไปตามภาพที่ 3-38

1.667	2.889	4.333
2.667	4	4.667
3.667	4.889	3.667

ภาพที่ 3-38 ค่าเฉลี่ยของจุดกริดต่าง ๆ ในแบบจำลอง SI3

การทดสอบทางการศึกษาระดับชาติขั้นพื้นฐาน (Ordinary national educational test/ O–NET)

การทดสอบทางการศึกษาระดับชาติขั้นสามัญ (Ordinary national educational test: O-NET) เป็นการวัดผลสัมฤทธิ์รวบยอดของผู้เรียนสำหรับในช่วงชั้นที่ 4 โดยสอบชั้นประถมศึกษาปีที่ 6, มัธยมศึกษาปีที่ 3 และ 6 ตามมาตรฐานการเรียนรู้ของหลักสูตรการศึกษาขั้นพื้นฐานในกลุ่มสาระต่าง ๆ รวม 8 กลุ่มสาระการเรียนรู้ ได้แก่ 1) ภาษาไทย 2) สังคมศึกษา ศาสนาและวัฒนธรรม 3) ภาษาอังกฤษ 4) คณิตศาสตร์ 5) วิทยาศาสตร์ 6) สุขศึกษาและพละศึกษา 7) ศิลปะ และ 8) การงานอาชีพและเทคโนโลยี ซึ่งการสอบ O-NET เป็นการสอบประจำปีเพียงครั้งเดียวและมีการจัดสอบในช่วงเดือนกุมภาพันธ์ของทุกปี โดยมีวัตถุประสงค์ เพื่อทดสอบความรู้และความคิดของนักเรียน ป.6, ม.3 และ ม.6 และนำผลการสอบไปใช้ในการปรับปรุงระบบการศึกษาในด้านต่าง ๆ เช่น ด้านคุณภาพการเรียนการสอนของโรงเรียน ด้านการประกันคุณภาพการศึกษา ด้านการประเมินผลการเรียนรู้ของนักเรียนระดับชาติ นอกจากนี้ผู้เข้ารับการทดสอบสามารถนำผลการสอบไปใช้ในการคัดเลือกเข้าศึกษาต่อในระดับมัธยมศึกษาตอนต้น มัธยมศึกษาตอนปลาย และสถาบันอุดมศึกษา

การใช้งานโปรแกรม QGIS กับฐานข้อมูลของคะแนน O-NET ปีการศึกษา 2548 จะเริ่มต้นจากการนำเสนอข้อมูลในรูปแบบแผนที่ เพื่อให้เขาใจถึงลักษณะและรายละเอียดของประชากร เช่น ผลการเรียนเฉลี่ยสะสม (เกรดเฉลี่ย Gpax) คะแนน O-NET ค่าสัมประสิทธิ์สหสัมพันธ์ ขั้นต่อมา คือ นำเสนอข้อมูลในรูปแบบแผนที่ที่ได้จากแบบจำลองการสุ่มแบบหลายขั้นตอน เช่น ตำแหน่งที่ตั้งของกลุ่มตัวอย่าง การกระจายตัวของตัวอย่างตามแผนที่จริง เป็นต้น ในขั้นสุดท้าย คือ การนำโปรแกรม QGIS มาประยุกต์กับแบบจำลองการสุ่มแบบจุดกริด เช่น การหาตำแหน่งจุดกริดของการสุ่ม การกำหนดรัศมีวงกลมการสุ่ม การนำเสนอข้อมูลในรูปแบบแผนที่ให้เป็นภาพเฉดสี ฯลฯ

แผนที่ประเทศไทยแสดงผลการเรียนเฉลี่ยสะสม (เกรดเฉลี่ย Gpax)

การประเมินผลการเรียนรู้ของรายวิชา ได้กำหนดไว้ในหลักสูตรการศึกษาขั้นพื้นฐาน ในการตัดสินระดับผลการเรียนรายวิชา ซึ่งคิดจากคะแนนเต็ม 100 คะแนน โดยใช้ตัวเลขแสดงระดับผลการเรียนเป็น 8 ระดับ ดังนี้

“4”	หมายถึง	ผลการเรียนดีเยี่ยม
“3.5”	หมายถึง	ผลการเรียนดีมาก
“3”	หมายถึง	ผลการเรียนดี
“2.5”	หมายถึง	ผลการเรียนค่อนข้างดี
“2”	หมายถึง	ผลการเรียนน่าพอใจ

“1.5” หมายถึง ผลการเรียนพอใช้

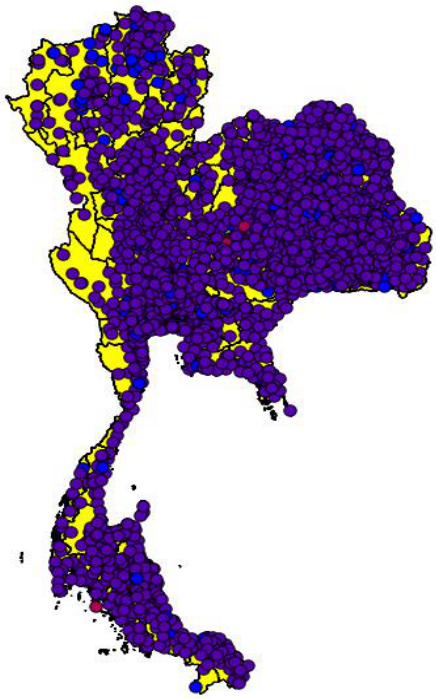
“1” หมายถึง ผลการเรียนผ่านเกณฑ์ขั้นต่ำ

“0” หมายถึง ผลการเรียนต่ำกว่าเกณฑ์ขั้นต่ำ

ซึ่งการประเมินผลทั้ง 8 ระดับนั้น คิดเป็นรายบุคคล แต่เนื่องจากฐานข้อมูลที่ในการวิเคราะห์นั้น เป็นค่าเฉลี่ยที่ได้จากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ของแต่ละโรงเรียน ดังนั้นค่าที่แสดงในภาพจึงเป็นค่าเฉลี่ยในภาพรวมของแต่ละโรงเรียน โดยแบ่งเป็นช่วงเกรดเฉลี่ยดังนี้

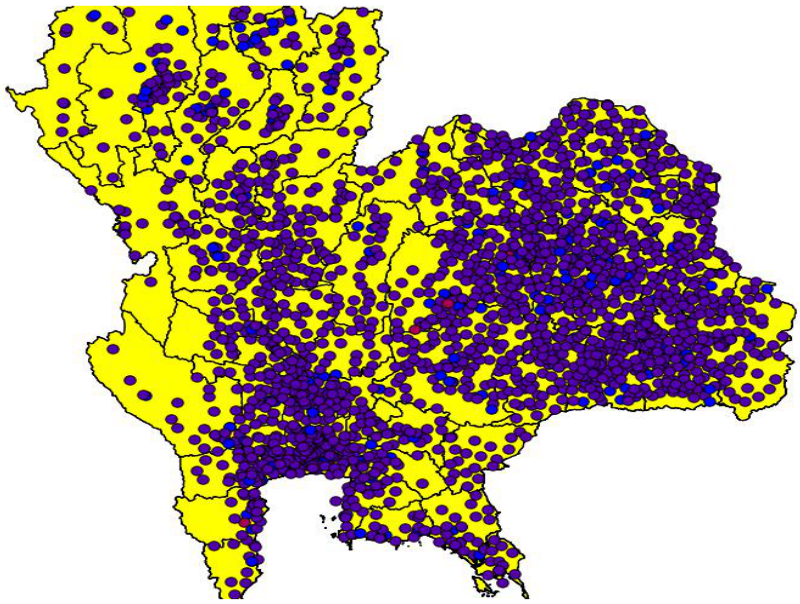
ตารางที่ 3-1 ความหมายของสีตามช่วงเกรดเฉลี่ยสะสม

ช่วงเกรดเฉลี่ย Gpax	สี	ความหมาย
0.00-0.99	สีแดง	ผลการเรียนต่ำกว่าเกณฑ์ขั้นต่ำ
1.00-1.99	สีแดงปนม่วง	ผลการเรียนพอใช้
2.00-2.99	สีน้ำเงินปนม่วง	ผลการเรียนค่อนข้างดี
3.00-4.00	สีน้ำเงิน	ผลการเรียนดีมาก

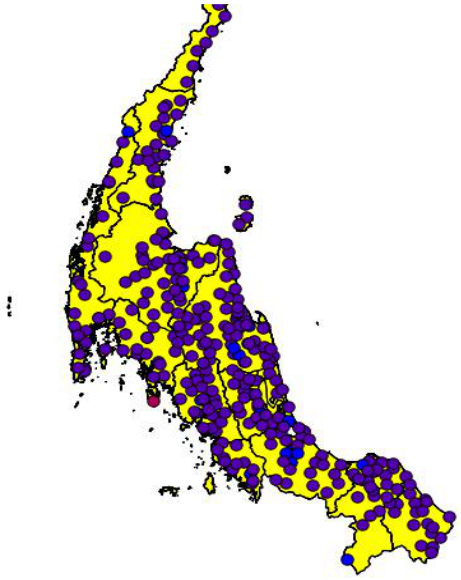


ภาพที่ 3-39 ผลการเรียนเฉลี่ยสะสมของโรงเรียนทั้งหมด

จากภาพที่ 3-39 ให้เห็นว่านักเรียนโดยส่วนใหญ่ มีผลการเรียนค่อนข้างดี หรือกล่าวได้อีกอย่างหนึ่งว่า โดยปกติแล้วโรงเรียนและครูอาจารย์จะให้เกรดนักเรียนตั้งแต่ 2 ขึ้นไป



ภาพที่ 3-40 ผลการเรียนเฉลี่ยสะสมของโรงเรียนในเขตตอนบนของประเทศ



ภาพที่ 3-41 ผลการเรียนเฉลี่ยสะสมของโรงเรียนในเขตตอนล่างของประเทศ

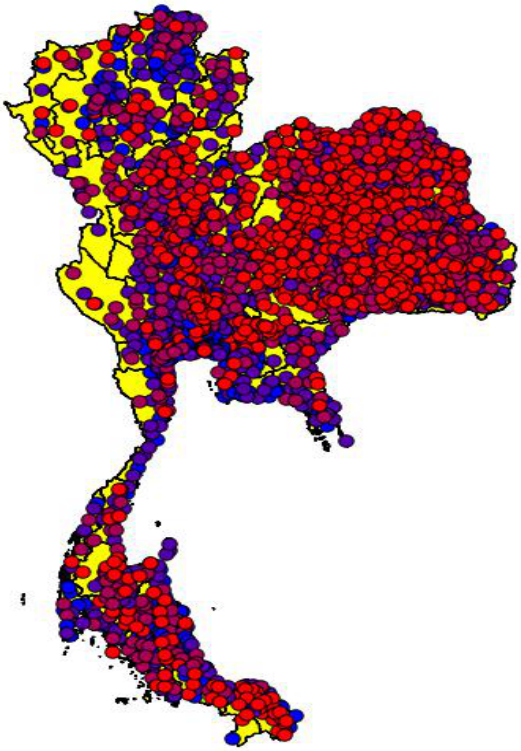
เมื่อทำการขยายภาพ เพื่อวิเคราะห์ข้อมูล พบว่า ในภาพที่ 3-40 และ 3-41 นักเรียนที่มีผลการเรียนดีมาก หรือโรงเรียนและครูอาจารย์จะให้เกรดนักเรียนตั้งแต่ 3 ขึ้นไป อยู่อย่างกระจัดกระจายครอบพื้นที่ทั่วประเทศ จากข้อมูลของภาพทั้ง 3 รูป สรุปได้ว่านักเรียนระดับชั้นมัธยมศึกษาปีที่ 6 ของประเทศไทย ในปีการศึกษา 2548 มีผลการเรียนผลการเรียนเฉลี่ยสะสมค่อนข้างดี

แผนที่ประเทศไทยแสดงผลคะแนน O-NET

ในการสอบ O-NET ปีการศึกษา 2548 นั้นมีวิชาที่ใช้ในการทดสอบ 5 วิชาหลัก คือ ภาษาไทย คณิตศาสตร์ วิทยาศาสตร์ สังคมศึกษา และภาษาอังกฤษ ถ้านำคะแนนเฉลี่ย ONET ของแต่ละโรงเรียนมาจัดอันดับคะแนน โดยเรียงต่อกันจากน้อยไปหามาก จากนั้นทำการแบ่งข้อมูลออกเป็น 4 ส่วน เท่า ๆ กัน จะได้ผลลัพธดังนี้

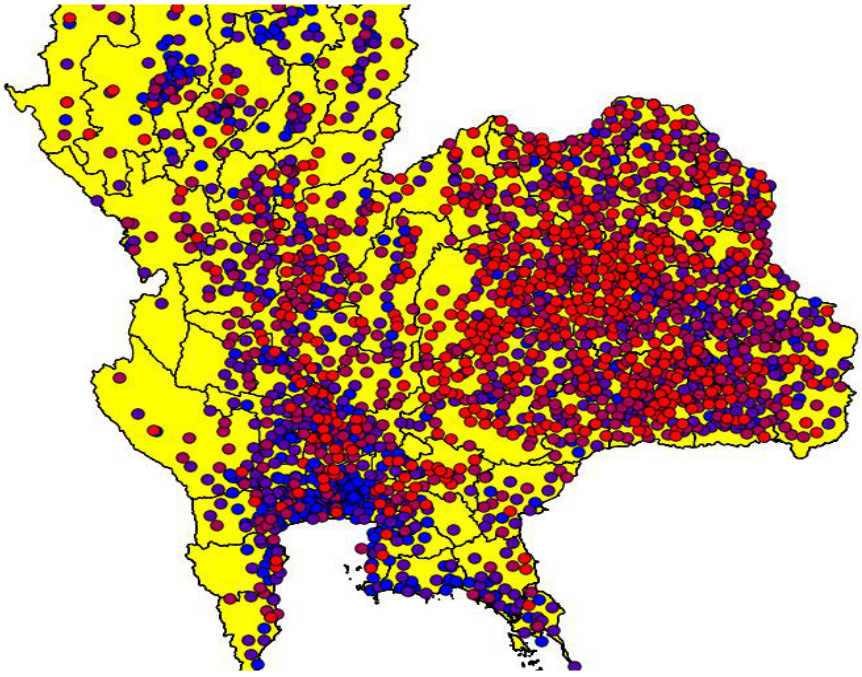
ตารางที่ 3-2 ความหมายของสีตามช่วงคะแนน O-NET

ช่วงคะแนน O-NET	สี	ความหมาย
121.25-148.64	สีแดง	ช่วงคะแนนสอบของควอร์ไทล์ที่ 1
148.64-155.92	สีแดงปนม่วง	ช่วงคะแนนสอบของควอร์ไทล์ที่ 2
155.92-166.35	สีน้ำเงินปนม่วง	ช่วงคะแนนสอบของควอร์ไทล์ที่ 3
166.35-355.77	สีน้ำเงิน	ช่วงคะแนนสอบของควอร์ไทล์ที่ 4

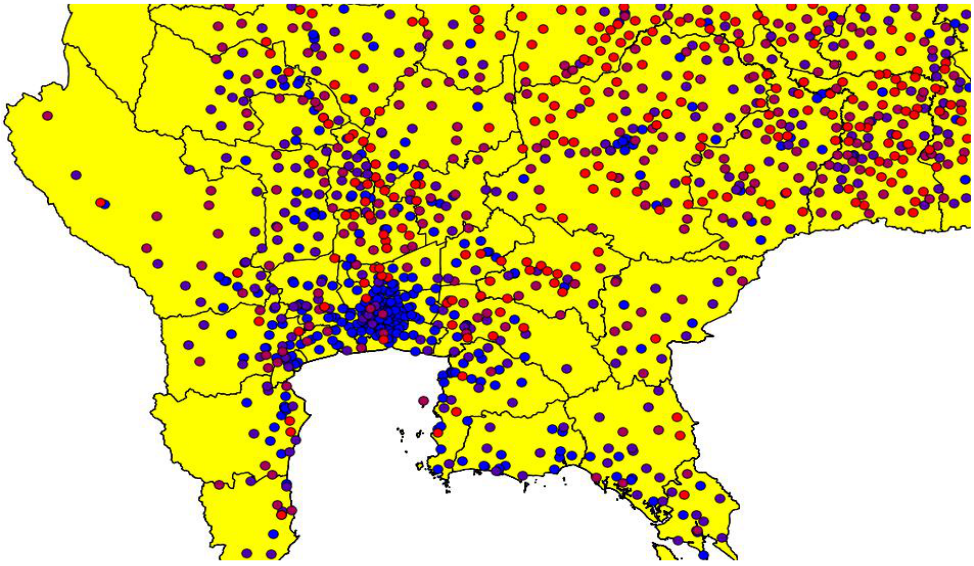


ภาพที่ 3-42 คะแนน O-NET ของโรงเรียนทั้งหมด

จากภาพที่ 3-43 แสดงให้เห็นว่านักเรียนโดยส่วนใหญ่สอบได้คะแนน O-NET น้อย ซึ่งนักเรียนส่วนใหญ่ทำคะแนนได้น้อยกว่า 200 คะแนน จากทั้ง 500 คะแนน สรุปได้ว่า นักเรียนส่วนใหญ่ของประเทศสอบไม่ผ่านเกณฑ์ร้อยละ 50



ภาพที่ 3-43 คะแนน O-NETของโรงเรียนในเขตตอนบนของประเทศ



ภาพที่ 3-44 คะแนน O-NET ของโรงเรียนในเมืองหลวงและปริมณฑล

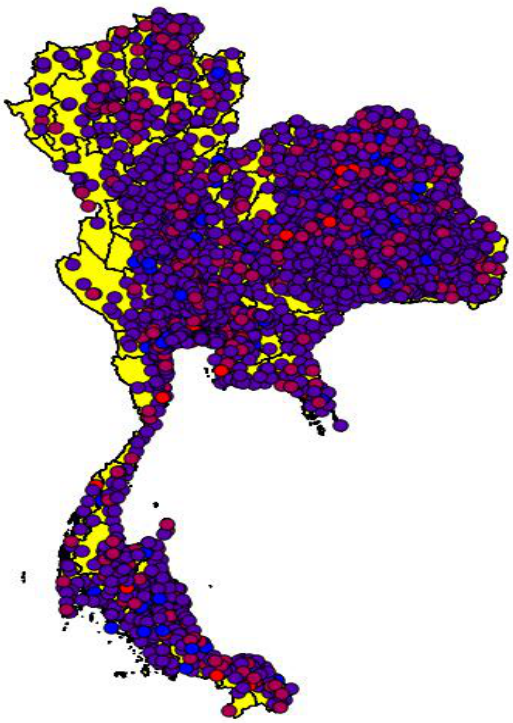
จากภาพที่ 3-43 และ 3-44 แสดงให้เห็นว่านักเรียนในภาคตะวันออกเฉียงเหนือมีคะแนน O-NET น้อยกว่านักเรียนในเมืองหลวงและปริมณฑล ซึ่งเป็นหลักฐานอย่างหนึ่งที่บ่งชี้ถึงความแตกต่างของคุณภาพการศึกษาระหว่างนักเรียนในเขตเมืองหลวงกับนักเรียนในเขตชนบท

แผนที่ประเทศไทยแสดงค่าสัมประสิทธิ์สหสัมพันธ์

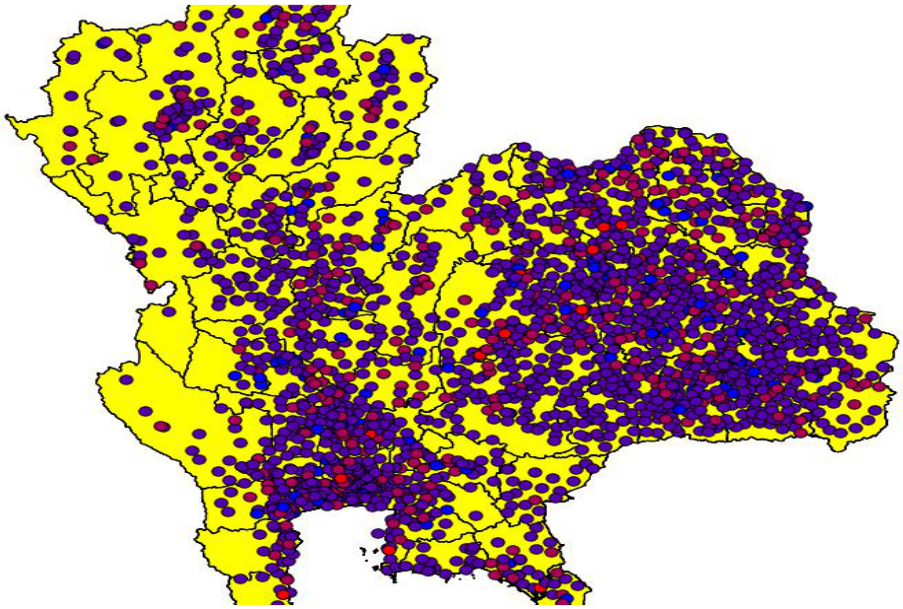
หลังจากพิจารณาในด้านเกรดเฉลี่ย Gpax และ คะแนน O-NET ของโรงเรียนทั้งประเทศแล้ว ในขั้นตอนต่อไป คือ การเปรียบเทียบตัวแปรทั้ง 2 ตัวนี้ ว่ามีความเกี่ยวข้องหรือสัมพันธ์กันอย่างไร โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ เมื่อนำมาแสดงผลในลักษณะแผนที่เชิงข้อมูลโปรแกรม QGIS ต้องกำหนดความหมายของสัญลักษณ์ในภาพ ดังนี้

ตารางที่ 3-3 ความหมายของสีตามช่วงของค่าสัมประสิทธิ์สหสัมพันธ์

ช่วงของค่าสัมประสิทธิ์สหสัมพันธ์	สี	ความหมาย
0.81-1.00	สีน้ำเงิน	มีความสัมพันธ์กันสูง
0.50-0.80	สีน้ำเงินปนม่วง	มีความสัมพันธ์กันระดับปานกลาง
0.00-0.49	สีแดงปนม่วง	มีความสัมพันธ์กันต่ำ
-1.00-0.01	สีแดง	มีความสัมพันธ์กันในทิศตรงข้าม



ภาพที่ 3-45 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง เกรดเฉลี่ย Gpax และ คะแนน O-NET



ภาพที่ 3-46 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง เกรดเฉลี่ย Gpaxและ คะแนน O-NET ในเขตตอนบนของประเทศ

จากภาพที่ 3-45 แสดงให้เห็นว่าโรงเรียนและครูอาจารย์โดยส่วนใหญ่ของประเทศให้ผลการศึกษากับนักเรียนตรงกันความรู้ความสามารถของนักเรียนอย่างแท้จริง แต่เมื่อทำการขยายภาพในภาคตะวันออกเฉียงเหนือ (ตามภาพที่ 3-46) พบว่ามีโรงเรียนที่ปล่อยเกรดเป็นจำนวนมาก ซึ่งการวิเคราะห์เกรดเฉลี่ย Gpax ไม่สอดคล้องกับคะแนน O-NET เนื่องจากโรงเรียนในภาคตะวันออกเฉียงเหนือให้เกรดเฉลี่ย Gpax แก่นักเรียนสูง แต่นักเรียนทำคะแนนสอบO-NET ได้ต่ำ ดังนั้นผลการสอบนี้จึงเหมือนกับกระจกที่สะท้อนความจริง

การสุ่มตัวอย่างแบบหลายขั้นตอน

วิธีการสุ่มตัวอย่างที่เหมาะสมที่สุดในกรณีการสุ่มตัวอย่างจากโรงเรียนทั่วประเทศซึ่งมีจำนวนประมาณสองพันกว่าโรงเรียนนั้น คือวิธีการสุ่มแบบหลายขั้นตอน เนื่องจากการสุ่มตัวอย่างแบบหลายขั้นตอนใช้ได้ดีกับประชากรขนาดใหญ่ที่แบ่งเป็นลำดับชั้นลดหลั่น (ศิริชัย กาญจนวาสี, 2545) ดังนั้นในการวิจัยครั้งนี้จึงเลือกใช้วิธีการสุ่มตัวอย่างแบบหลายขั้นตอนเพื่อนำค่าเฉลี่ยและค่าความแปรปรวนมาใช้เปรียบเทียบประสิทธิภาพกับวิธีการสุ่มแบบจุดกริดที่สร้างขึ้น และในการวางแผนการสุ่มตัวอย่างแบบหลายขั้นของงานวิจัยนี้ เริ่มต้นจากการหาขนาดกลุ่มตัวอย่าง (จำนวนโรงเรียน) ด้วยวิธีการแตกต่างกัน 5 วิธี จนได้ผลสรุปคือวิธีการใช้ตาราง Krejcie และ Morgan (รายละเอียดได้กล่าวไว้ในหัวข้อประชากรและกลุ่มตัวอย่าง) จากนั้นหา

จำนวนโรงเรียนที่ต้องการสุ่มในแต่ละภูมิภาค โดยใช้การจัดสรรตามสัดส่วนของประชากร และในขั้นตอนสุดท้ายคือการคำนวณจำนวนจังหวัดในแต่ละภูมิภาค

จากข้อมูลการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548 มีโรงเรียนต่าง ๆ ทั่วประเทศ เป็นจำนวน 2,584 โรง โดยแบ่งเป็นตามแต่ละภาคดังนี้

ตารางที่ 3-4 ร้อยละของจำนวนโรงเรียนตามภูมิภาค

ภาค	จำนวนโรงเรียน	ร้อยละของโรงเรียนตามแต่ละภูมิภาค
เหนือ	228	8.82
กลาง	615	23.8
ตะวันออก	167	6.46
ตะวันออกเฉียงเหนือ	1136	43.96
ตะวันตก	120	4.64
ใต้	318	12.31

ซึ่งการประมาณขนาดตัวอย่างโดยใช้ตาราง Krejcie และ Morgan นั้น ถ้าจำนวนประชากรเท่ากับ 2400 หน่วย ต้องใช้กลุ่มตัวอย่าง 331 หน่วย แต่ถ้าจำนวนประชากรเท่ากับ 2,600 หน่วย ต้องใช้กลุ่มตัวอย่าง 335 หน่วย ดังนั้นต้องใช้โรงเรียนจำนวน 335 โรงเรียน เป็นกลุ่มตัวอย่าง ซึ่งกลุ่มตัวอย่างขนาด 335 โรงเรียน สามารถแจกแจงตามสัดส่วน ในแต่ละภูมิภาค ดังนี้

ตารางที่ 3-5 จำนวนโรงเรียนเรียนที่ต้องถูกสุ่มออกมา

ภาค	ร้อยละของโรงเรียนตามแต่ละภูมิภาค	จำนวนโรงเรียน
เหนือ	8.82	30
กลาง	23.8	80
ตะวันออก	6.46	22
ตะวันออกเฉียงเหนือ	43.96	147
ตะวันตก	4.64	15
ใต้	12.31	41

สำหรับการหาขนาดกลุ่มตัวอย่างจังหวัดในแต่ละภูมิภาคสามารถกระทำได้โดยใช้ตาราง Krejcie และ Morgan โดยแบ่งเป็นตามแต่ละภาค ดังนี้

ตารางที่ 3-6 จำนวนจังหวัดที่ต้องถูกสุ่มออกมา

ภาค	จำนวนจังหวัด	ขนาดตัวอย่าง
เหนือ	9	9
กลาง	22	21
ตะวันออก	7	7
ตะวันออกเฉียงเหนือ	19	18
ตะวันตก	5	5
ใต้	14	14
รวม	76	74

หมายเหตุ จำนวนจังหวัดที่อ้างถึงเป็นข้อมูลในอดีตปี พ.ศ. 2548 ณ เวลานั้นประเทศไทยมีจังหวัด 76 จังหวัด

แต่ผลลัพธ์ที่แสดงตามตารางด้านบนนั้น ไม่สามารถยอมรับได้เหตุผล 2 ประการ คือ เนื่องจากในบางภูมิภาคมีจังหวัดน้อยกว่า 15 จังหวัด ซึ่งการหาขนาดตัวอย่างตามตาราง Krejcie และ Morgan ต้องใช้จำนวนทั้ง 15 จังหวัด เป็นกลุ่มตัวอย่าง ซึ่งนั้นหมายความว่ากลุ่มตัวอย่าง คือ ประชากร หรืออาจกล่าวได้ว่า ไม่เกิดกระบวนการสุ่มในบางภาค ส่วนเหตุผลอีกข้อหนึ่ง ข้อจำกัดในด้านของงบประมาณและกำลังคน ถึงแม้ว่าการทำงานวิจัยในระดับประเทศ หรือโครงการขนาดใหญ่ที่มีทุนและบุคลากรจำนวนมาก แต่ขนาดกลุ่มตัวอย่าง 74 จังหวัดนั้น ถือว่าเป็นจำนวนที่มากเกินไป ซึ่งผิดจุดมุ่งหมายการสุ่มตัวอย่าง เพราะการสุ่มตัวอย่างนั้นมีจุดมุ่งหมายในการลดความสิ้นเปลืองทรัพยากรต่าง ๆ ได้แก่ งบประมาณ บุคลากร(ผู้ช่วยนักวิจัย) ระยะเวลา ดังนั้น เพื่อให้งานวิจัยนี้มีความใกล้เคียงความเป็นจริงและสมเหตุผล จึงควรปรับขนาดกลุ่มตัวอย่างที่ใช้ลดลง ซึ่งเป็นไปตามกฎแห่งความชัดเจน (Rule of tumb) ซึ่งเป็นการกำหนดขนาดหนาดกลุ่มตัวอย่างแบบร้อยละ ในกรณีที่ประชากรน้อยกว่า 1,000 หน่วย ต้องใช้กลุ่มตัวอย่างขนาด 30% (Neuman, 1991, p. 221) ดังนั้นจำนวนจังหวัดร้อยละ 30 ของ 76 จังหวัด เท่ากับ 23 จังหวัด โดยคิดคำนวณเทียบจากร้อยละของจำนวนโรงเรียนตามแต่ละภูมิภาค ซึ่งมีรายละเอียด ดังนี้

ตารางที่ 3-7 สรุปการสุ่มโรงเรียนแยกตามรายภาค

ภาค	จำนวนจังหวัด	จำนวนโรงเรียน
เหนือ	2	30
กลาง	7	80
ตะวันออก	1	22
ตะวันออกเฉียงเหนือ	9	147
ตะวันตก	1	15
ใต้	3	41

จากรายละเอียดที่กล่าวมาสามารถสรุปการสุ่มแบบหลายขั้นตอน ดังนี้

ขั้นแรก ใช้วิธีสุ่มแบบแบ่งกลุ่ม โดยแบ่งกลุ่มจังหวัดออกเป็น 6 ภาค สุ่มจังหวัดออกมาจำนวน 23 จังหวัด ได้แก่ ภาคเหนือ 2 จังหวัด, ภาคกลาง 7 จังหวัด, ภาคตะวันออก 1 จังหวัด, ภาคตะวันออกเฉียงเหนือ 9 จังหวัด, ภาคตะวันตก 1 จังหวัด และ ภาคใต้ 3 จังหวัด

ขั้นที่ 2 สุ่มตัวอย่างโรงเรียนจากกลุ่มตัวอย่างจังหวัด โดยสุ่มโรงเรียนออกมา 335 โรงเรียน ได้แก่ ภาคเหนือ 30 โรงเรียน, ภาคกลาง 80 โรงเรียน, ภาคตะวันออก 22 โรงเรียน, ภาคตะวันออกเฉียงเหนือ 147 โรงเรียน, ภาคตะวันตก 15 โรงเรียน และ ภาคใต้ 41 โรงเรียน ซึ่งการสุ่มจังหวัดในแต่ละภูมิภาคและการสุ่มโรงเรียนในแต่ละจังหวัดใช้วิธีการสุ่มด้วยการเขียนโปรแกรม Microsoft excel

การสุ่มตัวอย่างแบบกริด

ในการเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองการสุ่มแบบจุดกริดและแบบจำลองการสุ่มแบบหลายขั้นตอนนั้น เพื่อให้ผลลัพธ์ที่ได้มีความน่าเชื่อถือจึงต้องควบคุมตัวแปรต่าง ๆ โดยการกำหนดค่าเริ่มต้นต่าง ๆ ให้เท่ากันหรือใกล้เคียงกันที่สุด ซึ่งแบบจำลองการสุ่มแบบหลายขั้นตอนใช้กลุ่มตัวอย่าง 23 จังหวัด ดังนั้นในการสร้างแบบจำลองการสุ่มแบบจุดกริดก็ต้องใช้จำนวนจุดกริด 23 จุดกริดด้วย ซึ่งมีขั้นตอนในการสุ่มตัวอย่าง ดังนี้

การหาตำแหน่งจุดกริดโดยใช้โปรแกรม QGIS

การหาตำแหน่งจุดกริด คือ การหาจุดที่ใช้ในการอ้างอิงเพื่อใช้ในการสุ่ม โดยใช้ฟังก์ชั่น (คำสั่ง Vector คำสั่ง Research tools และคำสั่ง Regular points) ของโปรแกรม QGIS เพื่อสร้างตารางกริดให้ครอบคลุมแผนที่ประเทศไทย และให้คัดเลือกจุดกริดที่อยู่บนแผนที่เท่านั้น ถ้าโปรแกรมสุ่มตำแหน่งจุดกริดได้จำนวนมากกว่าที่กำหนดให้พิจารณาตัดบางจุดที่ไม่เหมาะสมทิ้ง ถ้าโปรแกรมสุ่มตำแหน่งได้จำนวนน้อยกว่าที่วางแผน ให้ป้อนจำนวนจุดกริดมากกว่าเดิมและสร้างตารางกริดให้ถี่ขึ้น

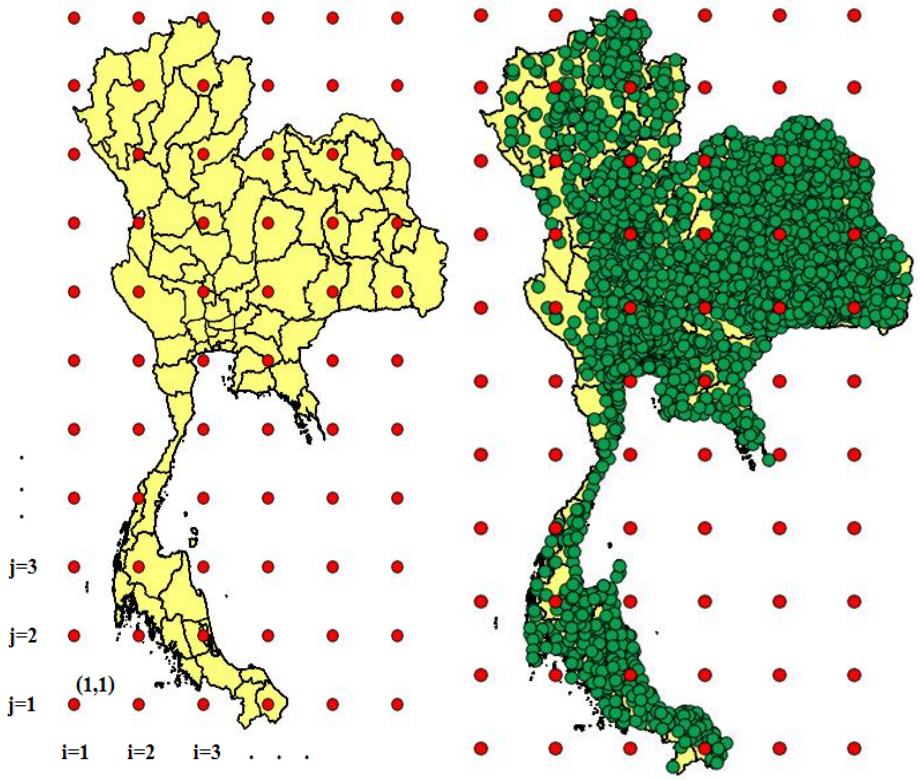
ในส่วนแผนการสุ่มแบบกริดมีลักษณะคล้ายกับแผนการสุ่มแบบหลายขั้นตอน เนื่องจากได้นำแผนการสุ่มแบบหลายขั้นตอนมาใช้เป็นต้นแบบ กล่าวคือ จำนวนจุดกริด (ของแผนการสุ่มแบบบกริด) ในแต่ละภูมิภาคต้องเท่ากับจำนวนจังหวัด (ของแผนการสุ่มแบบหลายขั้นตอน) ในแต่ละภูมิภาค ดังนั้นแผนการสุ่มแบบกริดมีรายละเอียด ดังนี้ ภาคเหนือ 2 จุดกริด, ภาคกลาง 7 จุดกริด, ภาคตะวันออก 1 จุดกริด, ภาคตะวันออกเฉียงเหนือ 9 จุดกริด, ภาคตะวันตก 1 จุดกริด และ ภาคใต้ 3 จุดกริด รวมทั้งสิ้น 23 จุดกริด

การกำหนดขนาดตัวอย่างในแต่ละจุดกริด

เมื่อได้ตำแหน่งจุดกริดตามจำนวนที่ต้องการแล้ว ให้กำหนดขนาดตัวในแต่ละจุดกริดอย่างตามสัดส่วน โดยแบ่งจำนวนกลุ่มตัวอย่างทั้งหมดออกเป็นส่วน ๆ ตามจำนวนจุดกริด ถ้าจุดกริดใดที่อยู่ในตำแหน่งที่มีความหนาแน่นประชากรสูง ที่จุดกริดนั้นต้องถูกสุ่มจำนวนตัวอย่างมากกว่าจุดกริดที่อยู่ในตำแหน่งประชากรเบาบาง

การสุ่มตัวอย่างโดยใช้วงกลมการสุ่ม

กำหนดให้จุดกริดเป็นศูนย์กลางวงกลมและขยายรัศมีวงกลมให้ครอบคลุมกลุ่มตัวอย่าง ซึ่งโรงเรียนที่อยู่ภายในรัศมีวงกลมคือกลุ่มตัวอย่างที่ใช้ในการวิจัย ซึ่งขนาดรัศมีวงกลมที่มากที่สุดเท่ากับครึ่งหนึ่งของระยะทางระหว่าง 2 จุดกริด ในกรณีที่มีจำนวนกลุ่มตัวอย่างภายในวงกลมมากกว่าที่ต้องการ/ ที่วางแผนไว้ ขอให้พิจารณาเลือกโรงเรียนที่อยู่ใกล้กับจุดกริดที่สุด โดยผลรวมของจำนวนโรงเรียนทุกจุดกริดคือ 335 โรงเรียน ซึ่งเป็นไปตามแผนการสุ่ม

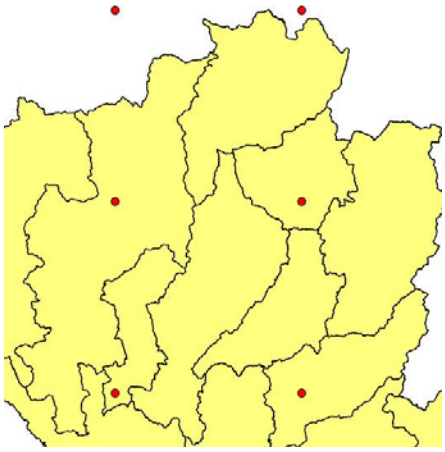


(ก) (ข)

ภาพที่ 3-47 ตำแหน่งจุดกริดที่สุ่มด้วยโปรแกรม QGIS

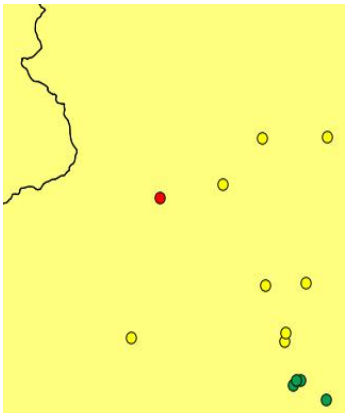
จากภาพที่ 3-47 (ก) แสดงให้เห็นว่ามีจุดกริดบางส่วนอยู่ในทะเลและอยู่ในประเทศเพื่อนบ้าน เมื่อนับจำนวน จุดกริดที่อยู่บนผืนแผ่นดินไทย มีจำนวนเท่ากับ 20 จุดกริด ซึ่งใกล้เคียงกับแผนการสุ่มที่วางไว้ ในเบื้องต้น เมื่อนำข้อมูลตำแหน่งของโรงเรียนมาพล็อตรวมกับจุดกริดที่สุ่มได้ ผลที่ได้แสดงตามภาพที่ 3-47 (ข) ดังนั้นโรงเรียน (จุดสีเขียว) ที่อยู่ใกล้กับจุดกริดสีแดงจะถูกสุ่มออกมา เพื่อคำนวณหาค่าเฉลี่ยของกลุ่มตัวอย่าง เมื่อพิจารณาจำนวนจุดกริดที่หามาได้พบว่า ยังขาดอีก 3 จุด (ตามแผนการสุ่มต้องมีจุดกริด 23 จุด) ดังนั้นต้องหาจุดกริดมาเพิ่มให้ครบ

โดยพิจารณาจุดกริดสีแดงบางส่วนที่อยู่ใกล้กับผืนแผ่นดิน ซึ่งจุดกริดเหล่านั้นจำเป็นต้องพิจารณาอย่างรอบครอบว่าจะนำมาคิดคำนวณหรือไม่ เพื่อให้งานวิจัยที่มีความสมเหตุผล โดยพิจารณาแยกเป็นรายภาคดังนี้

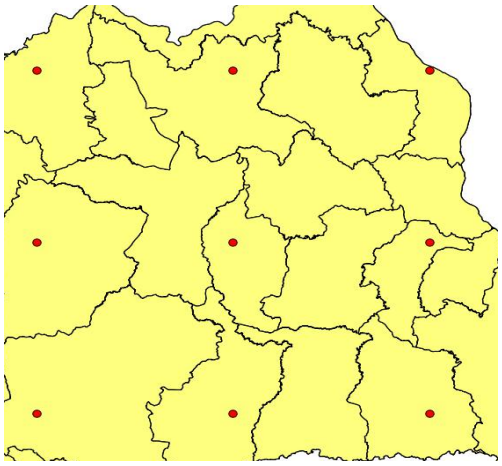


ภาพที่ 3-48 ตำแหน่งจุดกริดที่อยู่ในภาคเหนือ

จากภาพที่ 3-48 แสดงให้เห็นภาคเหนือของประเทศไทยถูกสุ่มได้ 4 จุดกริด คือ จุดกริด I = 2, j = 10 อยู่ในจังหวัด เชียงใหม่ จุดกริด i =3, j = 10 อยู่ในจังหวัดพะเยา จุดกริด i = 2, j = 9 อยู่ในจังหวัดลำพูน จุดกริด i = 3, j = 9 อยู่ในจังหวัดอุตรดิตถ์ แต่มีจุดกริดที่ควรพิจารณาอีกหนึ่งจุด คือ จุดกริด i = 3, j = 11 ที่อยู่ใกล้ชายแดนประเทศไทยด้านบนและมีระยะห่างจากจังหวัดเชียงรายประมาณ 9.59 กิโลเมตร และเมื่อเปรียบเทียบกับแบบจำลองการสุ่มแบบหลายขั้นตอนแล้ว พบว่าภาคเหนือต้องสุ่มออกมาเพียงแค่ 2 จังหวัด และโรงเรียน 30 แห่ง แต่ในแบบจำลองการสุ่มแบบจุดกริดมีโปรแกรม QGIS สุ่มออกมามีจำนวน 4 จุดกริด ดังนั้นไม่จำเป็นที่ต้องใช้ จุดกริด i = 3, j = 11 ในการสุ่มตัวอย่าง ดังนั้นในแต่ละจุดกริดต้องสุ่มโรงเรียนที่อยู่ใกล้ ๆ นั้น เป็นจำนวนประมาณ 30/4 = 7.5 โรงเรียน ยกตัวอย่าง เช่น จุดกริด i = 2, j = 10 อยู่ในจังหวัดเชียงใหม่ ซึ่งมีพิกัด ละติจูด 19 องศา 1 ลิปดา 50.2 ฟิลิปดา เหนือ ลองจิจูด 98 องศา 46 ลิปดา 40.1 ฟิลิปดา ตะวันออก ซึ่งอยู่ใกล้กับโรงเรียน จำนวน 8 โรงเรียน ได้แก่ แม่หอพระวิทยาคม แม่แตง สันป่ายางวิทยาคม สันทรายวิทยาคม แม่ริมวิทยาคม สะเมิงพิทยาคม ศึกษาสงเคราะห์เชียงใหม่ และนวมินทราชูทิศ พายัพ โรงเรียนทั้ง 8 โรงเรียน คือ จุดกริดสีเหลือง ตามภาพ 3-49

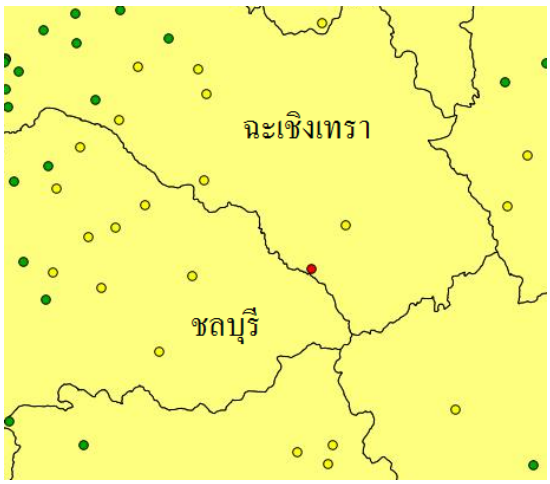


ภาพที่ 3-49 ตำแหน่งจุดกริด i = 2, j = 10 อยู่ในจังหวัดเชียงใหม่



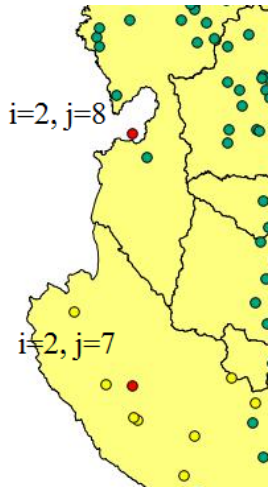
ภาพที่ 3-50 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันออกเฉียงเหนือ

จากภาพที่ 3-50 แสดงให้เห็นภาคตะวันออกเฉียงเหนือของประเทศไทยถูกสุ่มได้ 9 จุดกริด คือ จุดกริด i = 4, j = 9 อยู่ในจังหวัดเลย จุดกริด i = 5, j = 9 อยู่ในจังหวัดอุดรธานี จุดกริด i = 6, j = 9 อยู่ในจังหวัดนครพนม จุดกริด i = 4, j = 8 อยู่ในจังหวัดชัยภูมิ จุดกริด i = 5, j = 8 อยู่ในจังหวัดมหาสารคราม จุดกริด i = 6, j = 8 อยู่ในจังหวัดยโสธร จุดกริด i = 4, j = 7 อยู่ในจังหวัดนครราชสีมา จุดกริด i = 5, j = 7 อยู่ในจังหวัดบุรีรัมย์ จุดกริด i = 6, j = 7 อยู่ในจังหวัดศรีสะเกษ ซึ่งจำนวนจุดพิกัดที่สุ่มได้เท่ากันกับแบบจำลองการสุ่มแบบหลายขั้นตอนพอดี



ภาพที่ 3-51 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันออก

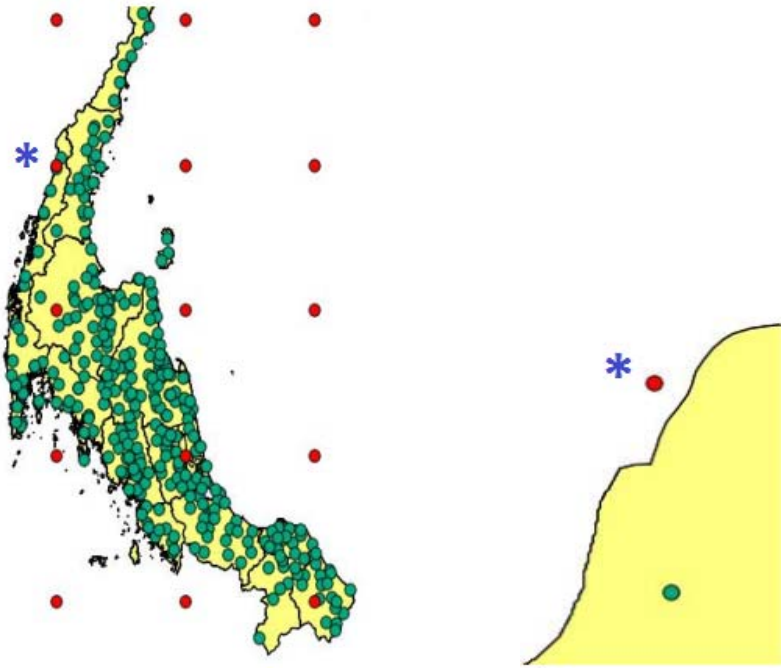
จากภาพที่ 3-51 แสดงให้เห็นภาคตะวันออกของประเทศไทยถูกสุ่มได้ 1 จุดกริด คือ จุดกริด i = 4, j = 6 อยู่ในจังหวัดฉะเชิงเทราและอยู่ใกล้กับจังหวัดชลบุรี ทำให้โรงเรียนในจังหวัดชลบุรี 9 โรงเรียน ถูกสุ่มออกมา แต่โรงเรียนในจังหวัดฉะเชิงเทราถูกสุ่มออกมาเพียง 6 โรงเรียน ซึ่งปกติจุดกริดอยู่ในจังหวัดใดก็ต้องใช้โรงเรียนที่อยู่ในจังหวัดนั้น แต่ในกรณีนี้โรงเรียนในจังหวัดฉะเชิงเทราอยู่ห่างไกลจากจุดกริดนี้ ดังนั้นถ้าต้องการสุ่มโรงเรียนให้ขนาดกลุ่มตัวอย่างเท่ากับแบบจำลองการสุ่มแบบหลายขั้นตอน (จำนวน 22 โรงเรียน) ดังนั้นวงกลมการสุ่มจึงขยายออกไปนอกจังหวัด ซึ่งเหตุการณ์แบบนี้เป็นลักษณะพิเศษของการสุ่มแบบจุดกริด ซึ่งการสุ่มแบบจุดกริดไม่ให้ความสำคัญกับเส้นเขตแดนที่มนุษย์สร้างขึ้น



ภาพที่ 3-52 ตำแหน่งจุดกริดที่อยู่ในภาคตะวันตก

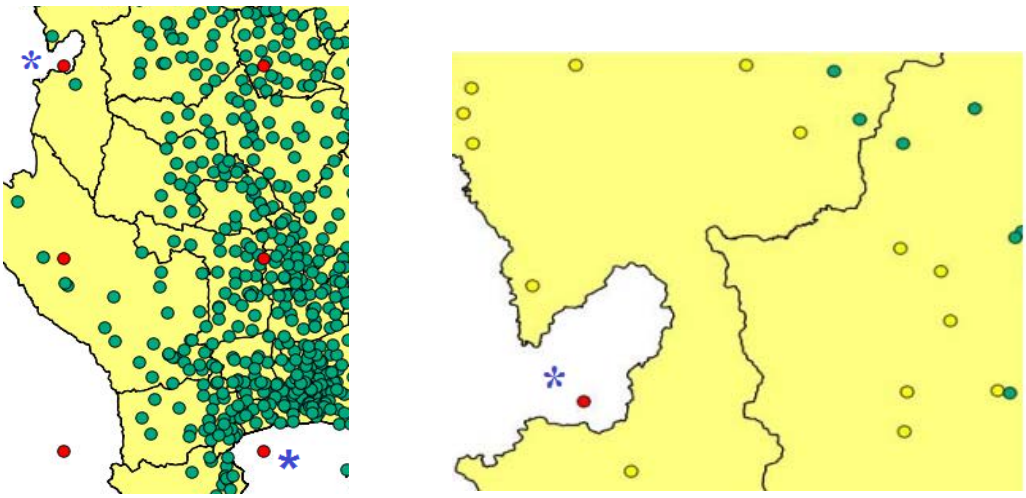
ตามแผนการสุ่มในภาคตะวันตก กำหนดให้สุ่มออกมา 1 จุดกริด และสุ่มโรงเรียน 15 โรง ซึ่งโปรแกรม QGIS สุ่มได้จังหวัดกาญจนบุรี คือ จุดกริด i = 2, j=7 ที่มีพิกัด ละติจูด 14 องศา 43 ลิปดา 58 ฟิลิปดา เหนือ ลองจิจูด 98 องศา 46 ลิปดา 40.1 ฟิลิปดา ตะวันออก (ตามภาพที่ 3-52) แต่เมื่อพิจารณาลักษณะพื้นที่กับการกระจายตัวของกลุ่มตัวอย่างพบว่าระยะห่างระหว่างจุดกริดสีแดงสองจุดที่อยู่ติดกัน ตามแนวละติจูดหรือลองจิจูดเท่ากับ 154.269 กิโลเมตร ทำให้รัศมีวงกลมการสุ่มสามารถขยายออกไปได้มากที่สุด เท่ากับ 77.1345 กิโลเมตร (ซึ่งเป็นไปตามหลักการการจัดสรรตัวอย่างตามสัดส่วนในบทที่ 2) และภายในรัศมีการสุ่ม 77.1345 กม. มีโรงเรียนรอบจุดศูนย์กลางเพียง 8 โรงเรียนเท่านั้น ดังนั้นในตำแหน่งจุดกริดนี้มีขนาดกลุ่มตัวอย่างเท่ากับ 8 โรงเรียน ดังนั้นจึงจำเป็นต้องใช้จุดกริด i=2, j=8 ซึ่งอยู่ใกล้จังหวัดตาก แต่อยู่ในประเทศพม่า โดยใช้จุดนี้ในการสุ่มตัวอย่างโรงเรียนจำนวน 7 โรงเรียน เพื่อทำให้จำนวนโรงเรียนในภาคตะวันตกครบ 15 โรงเรียน

ในภาคใต้ต้องถูกสุ่มออกมา 3 จุดกริด ซึ่งโปรแกรม QGIS ได้สุ่มจุดกริดออกมา 4 จุด คือ จุดกริด i = 2, j = 4 อยู่ในทะเล จุดกริด i = 2, j= 3 อยู่ในจังหวัดสุราษฎร์ธานี จุดกริด i = 3, j = 2 จังหวัดพัทลุง จุดกริด i = 4, j = 1 อยู่ในจังหวัดนราธิวาส ตามรูปภาพ 3-53 (ก) และจากภาพที่ 3-53 (ข) แสดงให้เห็นว่า จุดกริด i = 2, j = 4 อยู่ในทะเล (จุดกริดที่อยู่ใกล้เครื่องหมายดอกจันสีน้ำเงิน) โดยห่างจากฝั่งเพียงแค่ 458 เมตร แต่แผนการสุ่มกำหนดให้สุ่มจุดกริดในภาคใต้ 3 จุด ซึ่งโปรแกรม QGIS ได้สุ่มตำแหน่งจุดกริด 3 จุด ที่อยู่บนแผ่นดิน ดังนั้น จุดกริด i = 2, j = 4 (ที่อยู่ในทะเล) ไม่จำเป็นต้องนำใช้งาน



(ก) (ข)

ภาพที่ 3-53 ตำแหน่งจุดกริดที่อยู่ในภาคใต้



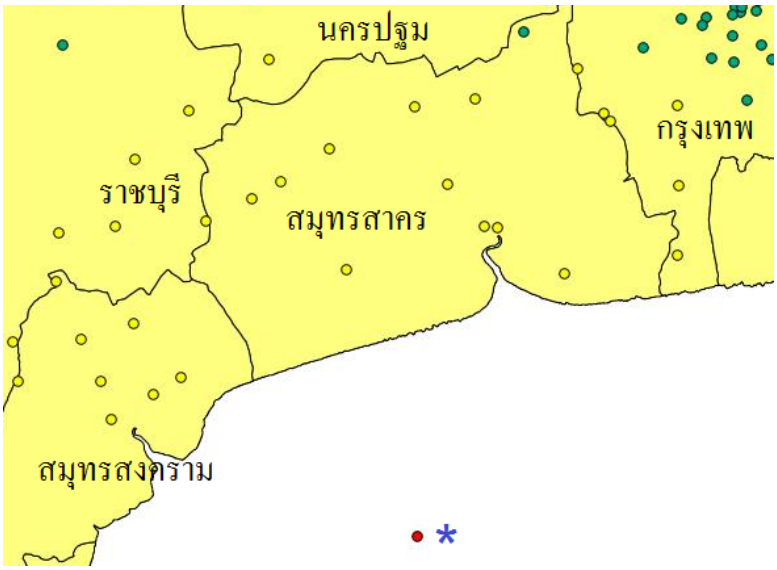
(ก) (ข)

ภาพที่ 3-54 ตำแหน่งจุดกริดที่อยู่ในภาคกลาง

ตามแผนการสุ่มในภาคกลางของประเทศไทยต้องถูกสุ่มออกมา 7 จุดกริด แต่โปรแกรม QGIS สุ่มจุดกริดได้เพียงแค่ 2 จุด คือ จุดกริด i = 3, j = 8 อยู่ในจังหวัดพิจิตร จุดกริด i = 3, j = 7 อยู่ในจังหวัดสุพรรณบุรี ใกล้แนวเขตจังหวัดอ่างทองตามภาพ 3-54 (ก) ทำให้จุดกริดที่ได้น้อยกว่า

แผนการสุ่มเป็นจำนวน 5 จุด ทั้งนี้เนื่องจากพื้นที่กรุงเทพและปริมณฑลเป็นศูนย์กลางการทำธุรกิจการลงทุน รวมทั้งเป็นที่ตั้งของศูนย์ราชการ ทำให้มีจำนวนนักเรียนและโรงเรียนเป็นจำนวนมาก และอีกเหตุผลหนึ่งคือจังหวัดของภาคกลางมีจำนวนมาก แต่มีพื้นที่เล็กกว่าเมื่อเทียบกับภาคตะวันออกเฉียงเหนือ

เพราะฉะนั้นต้องหาแนวทางในการแก้ปัญหาจำนวนจุดกริด จากภาพที่ 3-54 (ข) แสดงให้เห็นว่า จุดกริด i = 2, j = 8 อยู่ในประเทศเพื่อนบ้านและ จุดกริด i = 3, j = 6 อยู่ในทะเล (จุดกริดทั้งสองจุดอยู่ใกล้เครื่องหมายดอกจันสีน้ำเงิน) ในแนวเขตภาคกลาง สำหรับ จุดกริด i = 2, j = 8 อยู่ในเขตพม่าติดกับแนวแผ่นดินจังหวัดตาก เมื่อใช้จุดนี้เป็นจุดอ้างอิงสามารถสุ่มโรงเรียนได้ 6 โรงเรียน (นับเฉพาะโรงเรียนที่อยู่ในจังหวัดกำแพงเพชร) เนื่องจากทั้ง 6 โรงเรียนอยู่ในรัศมีวงกลมการสุ่มขนาด 77.13 กม. (เหตุผลคล้ายกับจุดกริด i = 2, j=7) ดังนั้นอีก 3 จุดกริดที่เหลือต้องสุ่มจุดกริดให้มากขึ้น เพื่อเป็นการชดเชย ยกตัวอย่างเช่น จุดกริด i = 3, j = 6 ต้องสุ่มโรงเรียนออกมา 25 โรงเรียน ได้แก่ สมุทรสงคราม 8 โรงเรียน สมุทรสาคร 10 โรงเรียน กรุงเทพฯ 6 โรงเรียน นครปฐม 1 โรงเรียน ตามภาพที่ 3-55 (ในส่วนจุดกริดสีเหลืองที่อยู่ในจังหวัดราชบุรีจะไม่นำมาคิดคำนวณ เนื่องจากเป็นจังหวัดในเขตภาคตะวันตก)

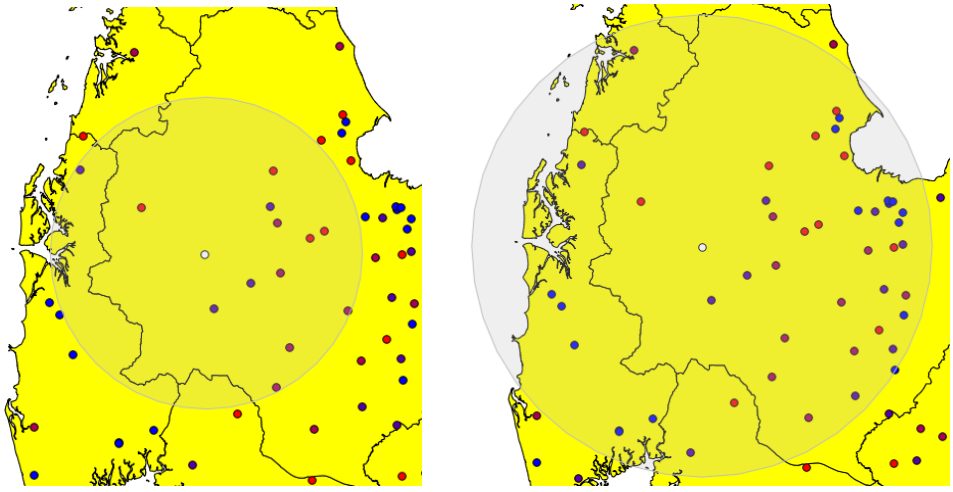


ภาพที่ 3-55 ตำแหน่งจุดกริด i = 3, j = 6

การสุ่มตัวอย่างแบบกริด-ชั้นภูมิ

เมื่อพิจารณาแผนที่คะแนนเฉลี่ย ONET ของประเทศไทย (ตามภาพที่ 3-42) พบว่า พื้นที่ส่วนใหญ่ของประเทศเป็นแบบวิวิธพันธ์ โดยโรงเรียนที่มีคะแนนเฉลี่ย ONET สูง (สีน้ำเงิน) อยู่รวมปะปนกับโรงเรียนที่มีคะแนนเฉลี่ย ONET ต่ำ (สีแดง) ดังนั้นแผนที่ประเทศไทยบางส่วนจึงมีลักษณะเป็นแบบวิวิธพันธ์ ซึ่งหลักการของการสุ่มแบบกริดใช้ระยะห่างระหว่างตำแหน่งของจุดกริดและที่ตั้งของกลุ่มตัวอย่างเป็นเกณฑ์ ซึ่งแนวคิดนี้เป็นไปเพื่อการประหยัดงบประมาณ ระยะเวลา และเจ้าหน้าที่ภาคสนาม ทำให้กลุ่มตัวอย่างที่ได้คือกลุ่มตัวอย่างที่อยู่ใกล้ ๆ กับจุดกริดนั้น แต่ในกรณีที่การกระจายตัวของประชากรเป็นแบบวิวิธพันธ์ เมื่อสุ่มตำแหน่งจุดกริดด้วยโปรแกรม QGIS อาจทำให้จุดกริดนั้นอยู่ในตำแหน่งกลุ่มประชากรย่อย (Subpopulations) และเมื่อสุ่มกลุ่มตัวอย่างบริเวณโดยรอบจุดกริดนั้นจะทำให้ได้ค่าสถิติที่ได้จากกลุ่มตัวอย่างมีความคลาดเคลื่อนจากค่าพารามิเตอร์

จากหัวข้อ “การสุ่มแบบบกริด 2 ขั้นตอน” ในบทที่ 2 ได้กล่าวถึงรายละเอียดเกี่ยวกับการสุ่มแบบกริดในกรณีที่พื้นที่เป็นแบบวิวิธพันธ์ ในการวิจัยครั้งนี้ได้นำหลักการวิธีการสุ่มแบบชั้นภูมิเข้ามาประยุกต์ใช้กับวิธีการสุ่มแบบกริด ซึ่งเรียกรวมกันว่า “วิธีการสุ่มแบบกริด-ชั้นภูมิ” โดยมุ่งหวังให้วิธีการนี้สุ่มได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร เพื่อพิสูจน์แนวคิดดังกล่าวจึงต้องทดสอบวิธีการนี้ด้วยการเปรียบเทียบประสิทธิภาพกับวิธีการสุ่มแบบหลายขั้นตอนและวิธีการสุ่มแบบกริด และเพื่อศึกษาลงลึกในรายละเอียดจึงต้องเลือกจุดกริดจุดหนึ่งออกมาเป็นตัวอย่างในการพิจารณา กำหนดให้เป็นจุดกริด i = 2, j = 3 และขยายขนาดแผนที่ที่จุดกริดดังกล่าว



(ก) (ข)

ภาพที่ 3-56 (ก) และ (ข) พื้นที่บริเวณจุดกริด i=2, j=3 ที่ประชากรมีลักษณะเป็นวิวิธพันธ์

จากภาพที่ 3-56 (ก) แสดงให้เห็นว่า ตำแหน่งจุดกริด i = 2, j = 3 คือพื้นที่ที่ประชากรมีลักษณะเป็นวิวิธพันธ์ และจุดสีขาว คือ จุดกริดที่ถูกโปรแกรม QGIS สุ่มตำแหน่ง ตามแผนการสุ่มที่ตำแหน่งนี้ต้องสุ่มตัวอย่างจำนวน 13 โรงเรียน โดยใช้จุดกริดเป็นตำแหน่งอ้างอิง เพื่อขยายรัศมีวงกลมของการสุ่มจนกระทั่งภายในมีสมาชิกครบ 13 โรงเรียน โดยโรงเรียนส่วนใหญ่เป็นกลุ่มโรงเรียนสีแดงและกลุ่มโรงเรียนสีแดงปนน้ำเงิน แต่กลุ่มโรงเรียนสีน้ำเงินและกลุ่มโรงเรียนสีน้ำเงินปนแดง ซึ่งอยู่ไม่ห่างไกลจากจุดกริดไม่ได้รับการถูกเลือก เนื่องจากสมาชิกภายในวงกลมการสุ่มมีจำนวนครบแล้ว ซึ่งค่าเฉลี่ยกลุ่มตัวอย่างนี้เท่ากับ 157.50 คะแนน แต่เมื่อขยายรัศมีวงกลมของการสุ่มให้มากที่สุด (77.13 กิโลเมตร) ทำให้ขนาดกลุ่มตัวอย่างภายในวงกลมเท่ากับ 47 โรงเรียน (ภาพที่ 3-56 ข) ซึ่งการขยายขนาดวงกลมนี้เป็นการเพิ่มกลุ่มโรงเรียนสีน้ำเงินและกลุ่มโรงเรียนน้ำเงินปนแดง ด้วยเหตุผลนี้ทำให้คะแนนเฉลี่ย ONET เพิ่มเป็น 162.82 คะแนน จากการศึกษาพื้นที่บริเวณจุดกริด i = 2, j = 3 แสดงให้เห็นว่าพื้นที่นี้มีประชากรมีลักษณะเป็นวิวิธพันธ์ ถ้าผู้วิจัยเลือกใช้กลุ่มตัวอย่างบริเวณใกล้กับจุดกริดจะทำให้ผลการทดลองมีความคลาดเคลื่อนเท่ากับ 162.82-151.69 = 11.13 คะแนน

ในโลกความเป็นจริงผู้วิจัยไม่สามารถรู้ค่าตัวแปรก่อนทำการทดลองได้ ยกตัวอย่าง เช่น ในกรณีศึกษานี้ ผู้วิจัยต้องทดสอบนักเรียนชั้นมัธยมศึกษาปีที่ 6 ทั่วประเทศไทย จำนวน 258,396 คน ซึ่งการเก็บรวบรวมข้อมูลจากประชากรจำนวนประมาณสองแสนหกหมื่นคน เป็นเรื่องที่ยุ่งยากและสิ้นเปลืองทรัพยากรอย่างมาก ซึ่งเหตุผลนี้เป็นอุปสรรคข้อใหญ่ในการใช้วิธีการสุ่มแบบชั้นภูมิ เพราะวิธีการนี้ต้องแบ่งกลุ่มโรงเรียนออกเป็นชั้นภูมิตามคะแนน ONET ดังนั้นในการวิจัยต้องอาศัยตัวแปรทางการศึกษาบางอย่างที่สามารถทราบค่าก่อนการทดลองและมีสหสัมพันธ์เชิงบวกกับตัวแปรคะแนน ONET ซึ่งตัวแปรนั้น คือ ตัวแปรขนาดโรงเรียน เพราะโรงเรียนขนาดใหญ่หรือโรงเรียนยอดนิยม (ที่มีการสอบแข่งขันกันอย่างสูง) มีคะแนนเฉลี่ยสูง เช่น โรงเรียนสาธิต เป็นต้น แต่โรงเรียนขนาดเล็กที่มีจำนวนครูและนักเรียนน้อยจะมีคะแนนเฉลี่ยต่ำเช่น โรงเรียนเทศบาล เป็นต้น และเมื่อโรงเรียนทั้งสองประเภทอยู่ใกล้กันจะเกิดปัญหาวิวิธพันธ์

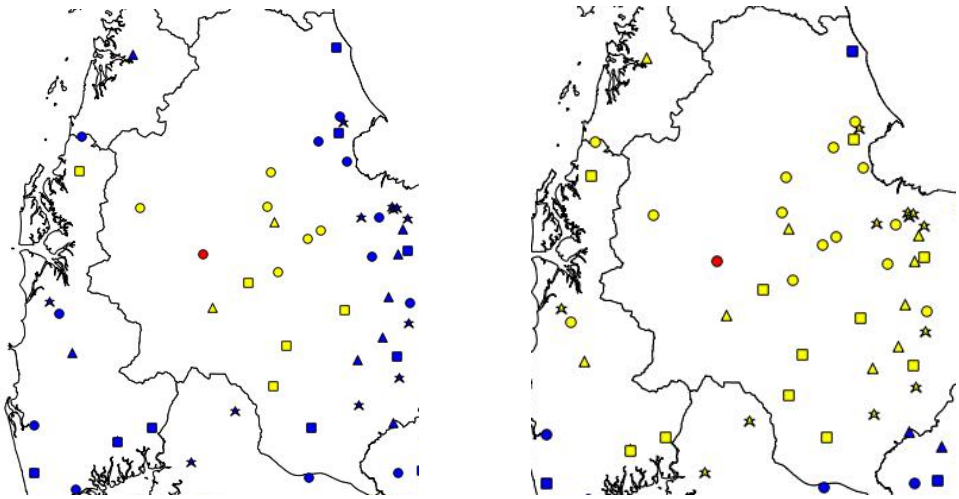
เมื่อทำการแบ่งขนาดโรงเรียนทั้งหมด 2,584 โรงเรียน ออกเป็น 4 ชั้นภูมิ (แบบควอไทล์) ตามจำนวนนักเรียน ดังนี้

ชั้นภูมิที่ 1 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ ONET อยู่ในช่วง 1-28 คน ใช้สัญลักษณ์ วงกลม

ชั้นภูมิที่ 2 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ ONET อยู่ในช่วง 28-53 คน ใช้สัญลักษณ์ สามเหลี่ยม

ชั้นภูมิที่ 3 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ ONET อยู่ในช่วง 53-111 คน ใช้สัญลักษณ์ สี่เหลี่ยม

ชั้นภูมิที่ 4 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ ONET อยู่ในช่วง 111-1615 คน ใช้สัญลักษณ์ ห้าเหลี่ยม



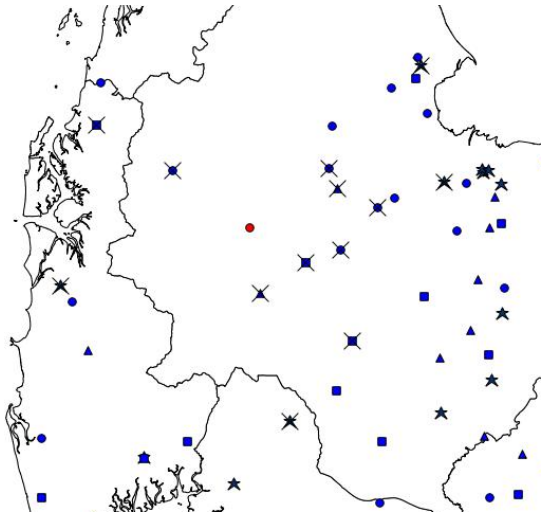
(ก) (ข)

ภาพที่ 3-57 (ก) และ (ข) การแบ่งชั้นภูมิด้วยขนาดโรงเรียน

เมื่อใช้โปรแกรม QGIS แสดงภาพโรงเรียนทั่วประเทศตาม 4 ชั้นภูมิ ดังภาพที่ 3-57 (ก) โดยจุดสีแดง คือ จุดกริด และโรงเรียนสีเหลือง คือ กลุ่มตัวอย่างจำนวน 13 โรงเรียน และ ภาพที่ 3-57 (ข) คือ รูปภาพแสดงกลุ่มตัวอย่างภายในวงกลม เมื่อขยายรัศมีวงกลมของการสุ่มให้ไกลที่สุด เป็นจำนวน 47 โรงเรียน แต่แผนการสุ่มกำหนดให้สุ่มกลุ่มตัวอย่างเพียงแค่ 13 โรงเรียน เท่านั้น ดังนั้นต้องนำโรงเรียนทั้ง 47 โรงเรียน มาแบ่งเป็นชั้นภูมิ 4 ชั้นภูมิ ได้แก่ โรงเรียนวงกลม 14 โรงเรียน, โรงเรียนสามเหลี่ยม 9 โรงเรียน, โรงเรียนสี่เหลี่ยม 11 โรงเรียน, และโรงเรียน ห้าเหลี่ยม 13 โรง และเมื่อทำการจัดสรรแบบสัดส่วน สามารถหาขนาดตัวอย่างในแต่ละชั้นภูมิ ดังนี้

ตารางที่ 3-8 จำนวนกลุ่มตัวอย่างในแต่ละชั้นภูมิ

ชั้นภูมิโรงเรียน	จำนวนโรงเรียน	จำนวนที่ถูกสุ่ม
วงกลม	14	4
สามเหลี่ยม	9	2
สี่เหลี่ยม	11	3
ห้าเหลี่ยม	13	4



ภาพที่ 3-58 กลุ่มตัวอย่างที่ได้จากวิธีการสุ่มสองขั้นตอน

จากภาพที่ 3-58 คือ ภาพแสดงผลลัพธ์สุดท้ายที่ได้จากการสุ่มสองขั้นตอน โดยขั้นตอนแรกเป็นการสุ่มแบบกริด และขั้นตอนที่สองเป็นการสุ่มแบบชั้นภูมิ ซึ่งโรงเรียนที่ถูกกากบาท คือ โรงเรียนแต่ละประเภทชั้นภูมิที่อยู่ใกล้จุดกริด (สีแดง) มากที่สุด โดยคะแนนเฉลี่ย ONET ของกลุ่มตัวอย่างเท่ากับ 159.27 คะแนน ซึ่งมีค่าใกล้เคียงกับคะแนนเฉลี่ย ONET 162.82 คะแนน และมีความคลาดเคลื่อนเท่ากับ 162.82-159.27 = 3.55 คะแนน

จากการทดสอบประสิทธิภาพการสุ่มแบบสองขั้นตอนแสดงให้เห็นว่าวิธีดังกล่าวสามารถลดความคลาดเคลื่อนที่จุดกริด i = 2, j = 3 ได้ เพื่อให้การสุ่มมีประสิทธิสูงขึ้นจึงต้องทำการทดลองต่อไปจนครบทุกจุด และในขั้นตอนสุดท้ายคือการเปรียบเทียบค่าเฉลี่ยที่คำนวนจากทุกจุดกริดกับค่าพารามิเตอร์ของโรงเรียนทั้งหมด 2,584 โรงเรียน

การวิเคราะห์ข้อมูล

การเปรียบเทียบประสิทธิภาพของวิธีการสุ่มต้องใช้ค่าสถิติเป็นตัวตัดสิน เนื่องจากค่าสถิติเป็นเครื่องชี้วัดความเป็นตัวแทนที่ดีของประชากร ถ้าวิธีการสุ่มใด ๆ ได้กลุ่มตัวอย่างที่ให้ค่าสถิติได้ใกล้เคียงกับค่าพารามิเตอร์แสดงว่าวิธีการสุ่มนั้นมีประสิทธิภาพสูง โดยค่าสถิติที่ได้จากกลุ่มตัวอย่างมี 2 แบบ คือ ค่าเฉลี่ยและความแปรปรวน ในกรณีค่าเฉลี่ยที่ได้จากแบบจำลองการสุ่ม เมื่อนำมาหาคำนวณหาค่าต่างระหว่างค่าเฉลี่ยของกลุ่มตัวอย่างกับค่าเฉลี่ยของประชากร ผลต่างระหว่างค่าทั้งสอง คือ ค่าความคลาดเคลื่อน ถ้าคิดถัวเฉลี่ยว่าให้เรียกว่า “ค่าเฉลี่ยของความคลาดเคลื่อน” ถ้าวิธีการสุ่มใดให้ “ค่าเฉลี่ยของความคลาดเคลื่อน” สูงแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่ไม่เป็นตัวแทนของประชากร ถ้าวิธีการสุ่มใดให้ “ค่าเฉลี่ยของความคลาดเคลื่อน” ต่ำแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร

$$\bar{e} = \frac{\sum_{k=1}^{m} e_k}{m} \quad \text{โดยที่ } e = \left|\sum e_{i,j}\right| = |\mu - \bar{z}| \approx \left|\mu - \hat{\bar{z}}\right|$$

$\bar{e}$ คือ ค่าเฉลี่ยของความคลาดเคลื่อน

e_k คือ ค่าความคลาดเคลื่อน เมื่อทำการการสุ่มตัวอย่างครั้งที่ k

m คือ จำนวนครั้งของการสุ่มทั้งหมด

ในกรณีค่าความแปรปรวนต้องยึดถือ “ผลรวมของค่าความแตกต่างความแปรปรวน” เป็นหลัก ถ้าวิธีการสุ่มใดมี “ผลรวมของค่าความแตกต่างความแปรปรวน” สูงแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่ไม่เป็นตัวแทนของประชากร ถ้าวิธีการสุ่มใดมี “ผลรวมของค่าความแตกต่างความแปรปรวน” ต่ำ แสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร

$$\Sigma|\Delta v| = |\Delta v|_1 + |\Delta v|_2 + |\Delta v|_3 + \ldots + |\Delta v|_m \quad \text{โดยที่ } |\Delta v| = \left|\sigma^2 - s^2\right|$$

$\Sigma|\Delta v|$ คือ ผลรวมของค่าความแตกต่างความแปรปรวน

m คือ จำนวนครั้งของการสุ่มทั้งหมด

บทที่ 4
ผลการวิเคราะห์ข้อมูล

วิธีการการสุ่มแบบกริดเป็นวิธีการสุ่มแบบหนึ่งที่ใช้กันอย่างกว้างขวางในการวิจัยเชิงวิทยาศาสตร์กายภาพ ยกตัวอย่าง เช่น การวิจัยเกี่ยวกับสัตว์และพืช การวิจัยเกี่ยวกับสภาพอากาศ การวิจัยเกี่ยวกับการแพร่ระบาดของโรคติดต่อ เป็นต้น ในการวิจัยครั้งนี้มุ่งทดสอบประสิทธิภาพวิธีการสุ่มแบบกริด เมื่อนำวิธีการดังกล่าวมาประยุกต์ใช้กับการวิจัยเชิงพฤติกรรมศาสตร์และสังคมศาสตร์ โดยนำวิธีการสุ่มตัวอย่าง 5 แบบ ได้แก่ การสุ่มอย่างง่าย การสุ่มแบบมีระบบ การสุ่มแบบชั้นภูมิ การสุ่มแบบแบ่งกลุ่ม และการสุ่มแบบหลายขั้นตอน มาเปรียบเทียบกับการสุ่มแบบกริด เพื่อตรวจสอบความแม่นยำในการประมาณค่า ซึ่งแบ่งเป็น 2 ขั้นตอนหลัก ดังนี้

1. ตอนที่ 1 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยแบบจำลองที่สร้างขึ้น

1.1 ตอนที่ 1.1 ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ

1.2 ตอนที่ 1.2 ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ

1.3 ตอนที่ 1.3 ผลการวิเคราะห์แบบจำลองการประมาณค่าเชิงพื้นที่และค่าคลาดเคลื่อนร้อยละ

2. ตอนที่ 2 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยฐานข้อมูลคะแนน O-NET

2.1 ตอนที่ 2.1 ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบหลายขั้นตอน

2.2 ตอนที่ 2.2 ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบกริด

2.3 ตอนที่ 2.3 ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบกริด-ชั้นภูมิ

2.4 ตอนที่ 2.4 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มทั้ง 3 วิธี เมื่อเทียบกับพารามิเตอร์

2.5 ตอนที่ 2.5 ผลการเปรียบเทียบแผนที่สารสนเทศที่เกิดจากวิธีการสุ่มแบบหลายขั้นตอนและวิธีการสุ่มแบบกริด/ กริด-ชั้นภูมิ

2.6 ตอนที่ 2.6 เปรียบเทียบข้อดีข้อเสียระหว่างวิธีการสุ่มแบบกริดและวิธีการสุ่มแบบกริด-ชั้นภูมิ

สัญลักษณ์ที่ใช้ในการวิเคราะห์

เพื่อให้เกิดความเข้าใจตรงกันในการแปลความหมายจากการวิเคราะห์ข้อมูล ผู้วิจัยจึงกำหนดความหมายของสัญลักษณ์ดังนี้

$\overline{X}$ หมายถึง คะแนนเฉลี่ยหรือค่าเฉลี่ย (Mean)

SD หมายถึง ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

r หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient)

* หมายถึง ความมีนัยสำคัญทางสถิติที่ระดับ .05

** หมายถึง ความมีนัยสำคัญทางสถิติที่ระดับ .01

Sig. หมายถึง ค่าความน่าจะเป็นอย่างมีนัยสำคัญทางสถิติที่คำนวณได้

H_0 หมายถึง สมมติฐานหลัก (Null hypothesis)

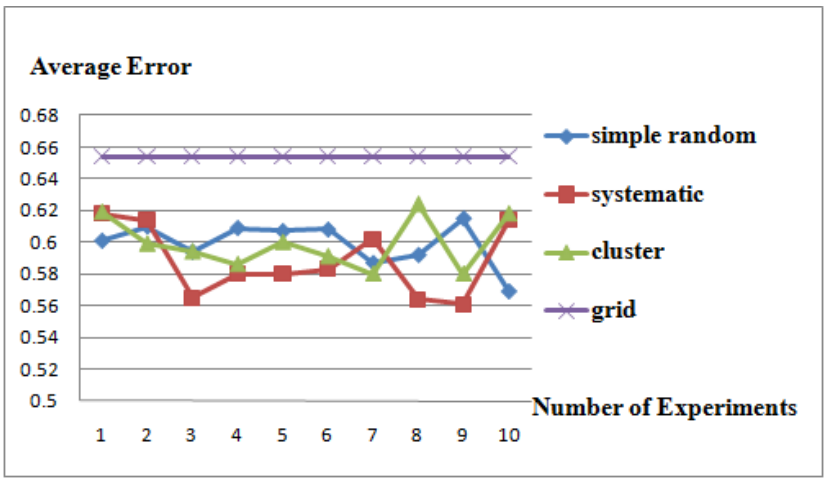
H_1 หมายถึง สมมติฐานรอง (Alternative hypothesis)

ตอนที่ 1 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยแบบจำลองที่สร้างขึ้น

1. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ

ในการวิจัยครั้งนี้ ทำการเขียนโปรแกรมคอมพิวเตอร์ เพื่อสร้างแบบจำลองต่าง ๆ ได้แก่ S0, S00, S1, S2, S3 และ S4 โดยใช้ภาษา C++ ในการคำนวณหาค่าเฉลี่ยความคลาดเคลื่อน ($\bar{e}$) และผลรวมของค่าความแตกต่างความแปรปรวน ($\Sigma|\Delta v|$) จากนั้นใช้โปรแกรม Microsoft office excel แสดงภาพ $\bar{e}$ และ $\Sigma|\Delta v|$ ในรูปแบบกราฟเส้น เพื่อให้เห็นภาพได้อย่างชัดเจนและง่ายต่อการเปรียบเทียบประสิทธิภาพระหว่างการสุ่มแบบกริดและการสุ่มแบบอื่น ๆ

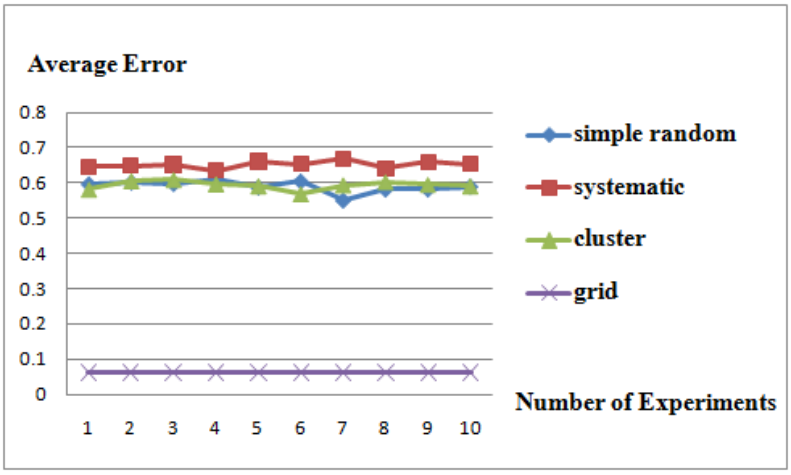
1. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S0



ภาพที่ 4-1 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S0

จากภาพที่ 4-1 แสดงให้เห็นว่า การสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนมากที่สุด คือ 0.654 ในส่วนวิธีการสุ่มแบบอื่น ๆ มีค่าเฉลี่ยความคลาดเคลื่อนใกล้เคียงกันดังนี้ การสุ่มอย่างง่าย มีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.569-0.610, การสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อน ระหว่าง 0.561-0.618 และการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.580-0.624 ซึ่งสามารถสรุปได้ว่าการสุ่มเกือบทุกรูปแบบมีประสิทธิภาพใกล้เคียงกัน ยกเว้นการสุ่มแบบกริด มีประสิทธิภาพต่ำที่สุด จากผลการวิเคราะห์แสดงให้เห็นว่าการสุ่มแบบกริดไม่เหมาะสมกับ ประชากรที่มีสหสัมพันธ์เชิงพื้นที่เป็น 0 หรือใกล้เคียง 0

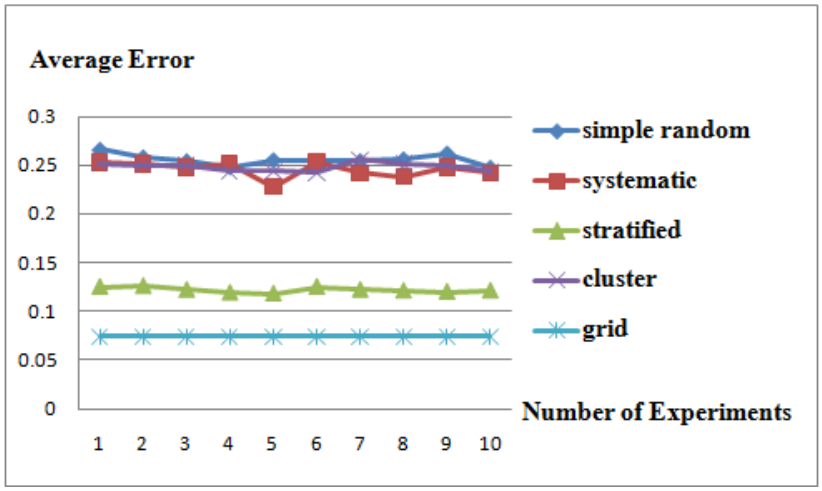
2. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S00



ภาพที่ 4-2 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S00

จากภาพที่ 4-2 แสดงให้เห็นว่า การสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนน้อยที่สุด คือ 0.062 ในส่วนวิธีการสุ่มแบบอื่น ๆ มีค่าเฉลี่ยความคลาดเคลื่อนใกล้เคียงกันดังนี้ การสุ่มอย่างง่าย มีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.550-0.612, การสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อน ระหว่าง 0.634-0.668 และการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.568-0.609 ซึ่งสามารถสรุปได้ว่าประสิทธิภาพของการสุ่มแบบกริดเป็นอันดับที่ 1 ในส่วนการสุ่มอย่างง่าย การสุ่มแบบมีระบบและการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพใกล้เคียงกัน แต่การสุ่มแบบกริด มีประสิทธิภาพมากกว่าอย่างชัดเจน ซึ่งความแตกต่างระหว่างแบบจำลอง S0 กับ S00 คือ ตำแหน่ง แถวที่ 2 คอลัมน์ที่ 8 เปลี่ยนจาก 2 เป็น 8 เท่านั้น ในส่วนตำแหน่งอื่น ๆ มีค่าเหมือนกัน แต่ผล การวิเคราะห์นั้นแตกต่างกันมาก

3. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S1

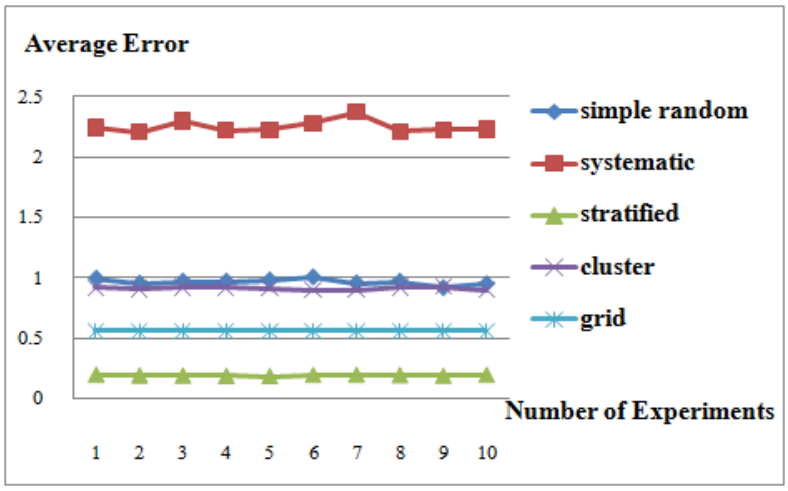


ภาพที่ 4-3 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S1

จากภาพที่ 4-3 แสดงให้เห็นว่า การสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนน้อยที่สุด คือ 0.074 การสุ่มแบบชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนน้อยเป็นลำดับที่ 2 และมีค่าระหว่าง 0.118-0.126 ในส่วนวิธีการสุ่มแบบอื่น ๆ มีค่าเฉลี่ยความคลาดเคลื่อนใกล้เคียงกันดังนี้ การสุ่มอย่างง่ายมีค่าเฉลี่ย ความคลาดเคลื่อนระหว่าง 0.247-0.266, การสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.228-0.253 และการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนระหว่าง 0.242-0.256 ซึ่งสามารถ สรุปได้ว่าประสิทธิภาพของการสุ่มแบบกริดเป็นอันดับที่ 1 การสุ่มแบบชั้นภูมิเป็นอับดับที่ 2 ในส่วนการสุ่มแบบมีระบบ, การสุ่มแบบแบ่งกลุ่มและการสุ่มอย่างง่ายเป็นอันดับสุดท้าย

จากผลการวิเคราะห์แสดงให้เห็นว่าการสุ่มแบบกริดสามารถสุ่มกลุ่มตัวอย่างที่เป็นตัวแทนประชากรได้ดี เนื่องจากจุดกริดที่เลือกไว้นั้นกระจัดกระจายครอบคลุมทุก ๆ ภาคส่วนของประชากรไว้ ในส่วนการสุ่มแบบชั้นภูมิ ซึ่งถูกคาดหวังว่าจะเป็นวิธีการที่มีประสิทธิภาพสูง แต่ผลลัพธ์ที่ได้ คือ ค่าเฉลี่ยความคลาดเคลื่อนมีค่าสูงนั้น เกิดจากชั้นภูมิแต่ละชั้นไม่เป็นเอกพันธ์อย่างแท้จริง ดังนั้น กลุ่มตัวอย่างที่ได้ในการสุ่มบางครั้งจะไม่ใช่ตัวแทนที่ดีของประชากร

4. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S2



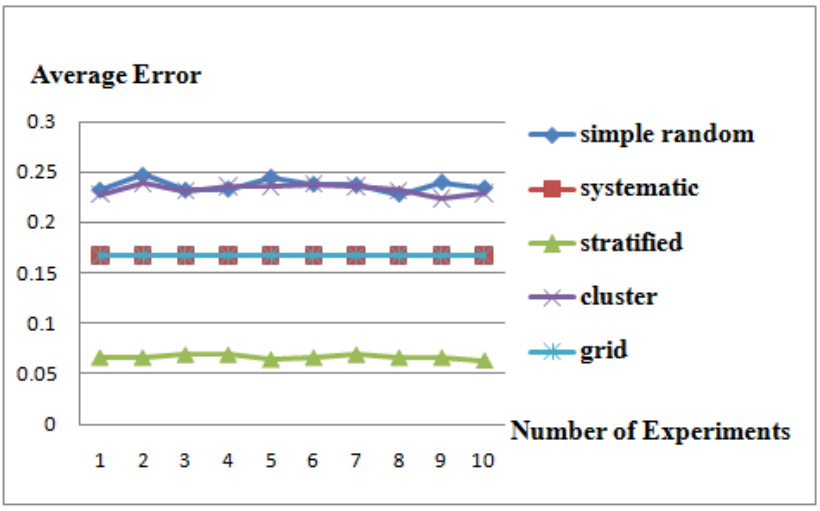
ภาพที่ 4-4 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S2

จากภาพที่ 4-4 แสดงให้เห็นว่า การสุ่มแบบชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนน้อยที่สุด และมีค่าระหว่าง 0.183-0.196 การสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนน้อยเป็นลำดับที่ 2 คือ 0.556 ในส่วนวิธีการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนใกล้เคียงกัน ซึ่งอยู่ระหว่าง 0.916-1.002 และ 0.897-0.924 สำหรับการสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อนมากที่สุด ซึ่งอยู่ระหว่าง 2.210-2.373 ซึ่งสามารถสรุปได้ว่าการสุ่มแบบมีระบบมีประสิทธิภาพต่ำที่สุด เมื่อธรรมชาติของประชากรมีการจัดเรียงตัวของค่าสังเกตกับระยะห่างของการสุ่มตรงกันพอดี จะทำให้ค่าเฉลี่ยที่ได้นั้นสูงหรือต่ำเกินความเป็นจริง เช่น

- ถ้าสุ่มได้คอลัมน์ที่ 2 จะได้สมาชิกทั้ง 9 ช่อง ที่มีค่าสังเกตดังนี้ 10 10 10 10 10 10 10 1 1
- ถ้าสุ่มได้คอลัมน์ที่ 8 จะได้สมาชิกทั้ง 9 ช่อง ที่มีค่าสังเกตดังนี้ 2 2 2 2 2 2 2 1 1

จากตัวอย่างดังกล่าวเมื่อคำนวณหาค่าเฉลี่ยจะพบว่า กลุ่มตัวอย่างทั้ง 2 กลุ่ม มีค่าค่าเฉลี่ยที่แตกต่างกันมาก

5. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S3



ภาพที่ 4-5 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S3

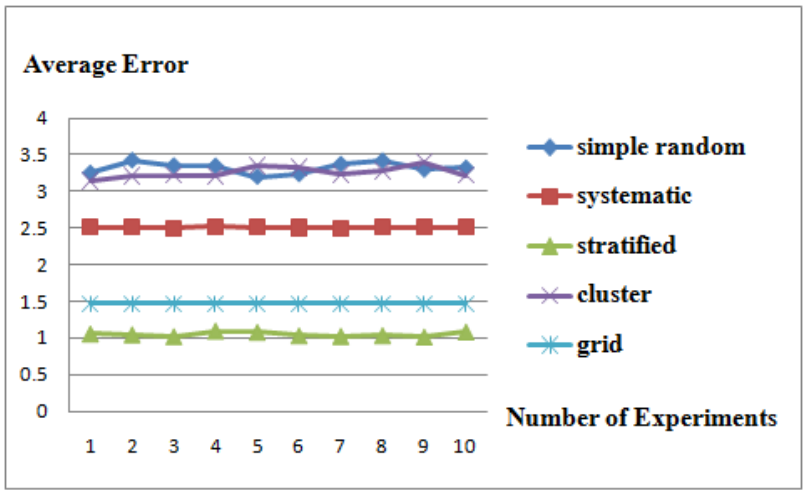
จากภาพที่ 4-5 แสดงให้เห็นว่าการสุ่มแบบชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนน้อยที่สุด และมีค่าระหว่าง 0.063-0.069 การสุ่มแบบกริดและการสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อนน้อยเป็นลำดับที่ 2 คือ 0.167 ในส่วนวิธีการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนมากที่สุดและใกล้เคียงกัน ซึ่งอยู่ระหว่าง 0.228-0.247 และ 0.224-0.239 ซึ่งสามารถสรุปได้ว่าประสิทธิภาพของการสุ่มแบบชั้นภูมิเป็นอับดับที่ 1 การสุ่มแบบกริดและการสุ่มแบบมีระบบเป็นอันดับที่ 2 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มเป็นอันดับที่ 3

จากผลการวิเคราะห์แสดงให้เห็นว่าการสุ่มแบบมีระบบและการสุ่มแบบกริดมีประสิทธิภาพใกล้เคียงกันมาก (เนื่องจากเส้นกราฟทั้ง 2 เส้น ซ้อนทับกันสนิท) ซึ่งการสุ่มแบบกริดมีประสิทธิภาพมากกว่าการสุ่มแบบมีระบบที่ทศนิยมตัวที่ 7 แต่เนื่องจากค่าเจตคติหรือทัศนคติเป็นหลักหน่วย จึงไม่จำเป็นต้องใช้ทศนิยมมากขนาดนั้นและให้ถือว่าในการวิเคราะห์กรณีนี้วิธีการทั้งสองมีประสิทธิภาพเท่ากัน

จากผลการวิเคราะห์แสดงให้เห็นว่าการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพต่ำที่สุด เพราะในบางครั้งค่าเฉลี่ยที่ได้อาจสูงเกินจริงหรือต่ำเกินจริง เนื่องจากการสุ่ม

ทั้ง 2 วิธีนี้มีหลักการว่าสมาชิกทุกตัวในประชากรมีโอกาสในการถูกเลือกเท่ากัน ดังนั้นค่าสถิติที่ได้ในแต่ละครั้งจะมีความไม่แน่นอนเมื่อเทียบกับค่าพารามิเตอร์

6. ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S4



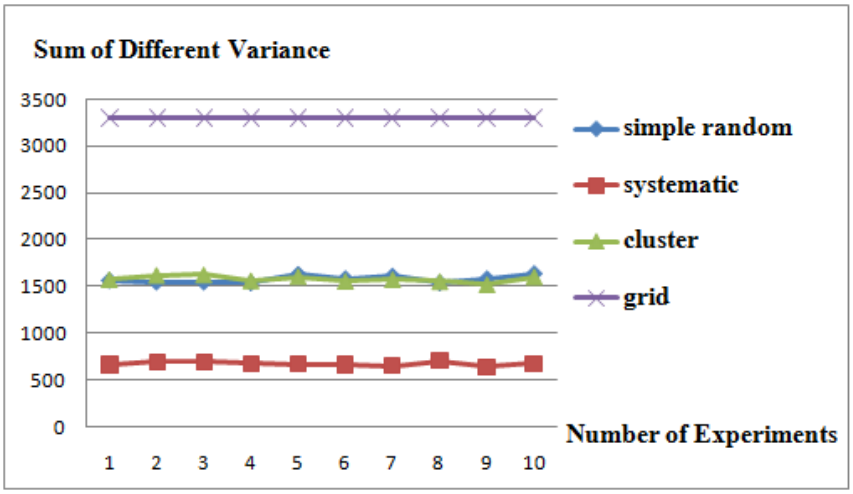
ภาพที่ 4-6 กราฟค่าเฉลี่ยความคลาดเคลื่อนของแบบจำลอง S4

จากภาพที่ 4-6 แสดงให้เห็นว่า การสุ่มแบบชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนน้อยที่สุดและมีค่าระหว่าง 1.013-1.095 การสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนน้อยเป็นลำดับที่ 2 คือ 1.465 การสุ่มแบบมีระบบมีค่าเฉลี่ยความคลาดเคลื่อนเป็นลำดับที่ 3 ซึ่งอยู่ระหว่าง 2.496-2.518 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีค่าเฉลี่ยความคลาดเคลื่อนมากที่สุดและใกล้เคียงกัน โดยมีค่าอยู่ในช่วง 3.198-3.420 และ 3.141-3.400 ซึ่งสามารถสรุปได้ว่าประสิทธิภาพของการสุ่มแบบชั้นภูมิเป็นอับดับที่ 1 การสุ่มแบบกริดเป็นอันดับที่ 2 การสุ่มแบบมีระบบเป็นอันดับที่ 3 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มเป็นอันดับสุดท้าย

จากผลการวิเคราะห์แสดงให้เห็นว่าการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพต่ำที่สุด ซึ่งผลการวิเคราะห์นี้สอดคล้องกับแบบจำลอง S3

2. การวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ

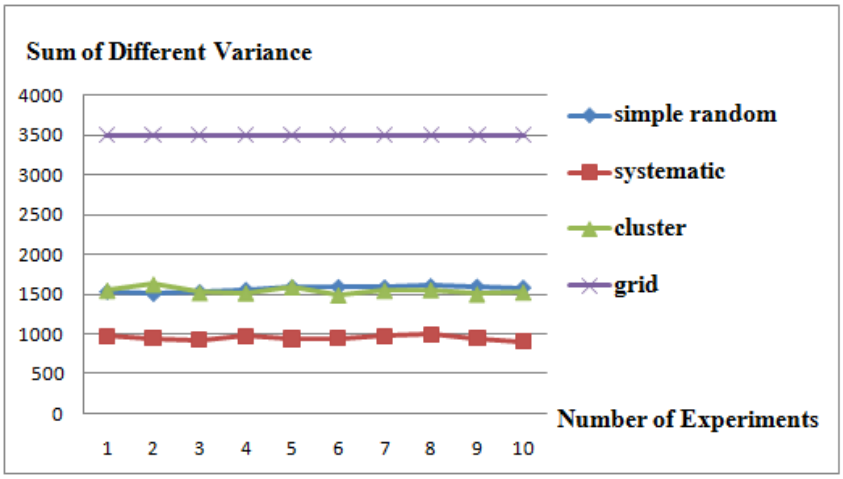
1. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S0



ภาพที่ 4-7 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S0

จากภาพที่ 4-7 แสดงให้เห็นว่า การสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยที่สุดและมีค่าระหว่าง 637.45-701.66 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 โดยมีค่าอยู่ในช่วง 1,536.37-1,634.20 และ 1,516.39-1,622.97 ในส่วนวิธีการสุ่มแบบกริดมีค่าผลรวมของค่าความแตกต่างความแปรปรวนมากที่สุด เท่ากับ 3,297.13 ซึ่งสามารถสรุปได้ว่าแบบจำลอง S0 มีสหสัมพันธ์เชิงพื้นที่เป็นศูนย์หรือใกล้เคียงศูนย์ทำให้ค่าสถิติที่ได้จากวิธีการสุ่มแบบกริดมีค่าแตกต่างจากค่าพารามิเตอร์อย่างมาก โดยมีความแปรปรวนของประชากรเท่ากับ 4.98 แต่ความแปรปรวนของกลุ่มตัวอย่างที่สุ่มได้จากวิธีการสุ่มแบบกริดมีค่าเท่ากับ 8.28 และกลุ่มตัวอย่างที่สุ่มจากวิธีการสุ่มแบบกริดในครั้งนี้ไม่เป็นตัวแทนที่ดีของประชากร

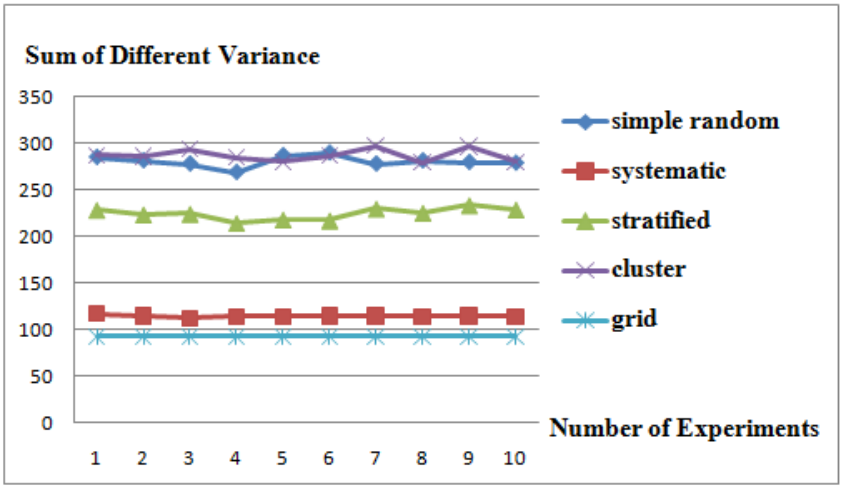
2. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S00



ภาพที่ 4-8 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S00

จากภาพที่ 4-8 แสดงให้เห็นว่า การสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยที่สุดและมีค่าระหว่าง 907.06-993.91 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 ซึ่งอยู่ระหว่าง 1,500.99-1,602.60 และ 1,491.64-1,628.23 ในส่วนวิธีการสุ่มแบบกริดมีค่าผลรวมของค่าความแตกต่างความแปรปรวนมากที่สุด เท่ากับ 3,500.38 ซึ่งสามารถสรุปได้ว่าถึงแม้ว่ากลุ่มตัวอย่างที่สุ่มได้มีค่าเฉลี่ยใกล้เคียงค่าเฉลี่ยประชากร แต่เมื่อธรรมชาติของประชากรไม่เป็นไปตามข้อตกลงเบื้องต้น ทำให้ความแปรปรวนของกลุ่มตัวอย่างมีความแตกต่างจากความแปรปรวนของประชากรมาก ดังนั้นกลุ่มตัวอย่างที่สุ่มได้ไม่เป็นตัวแทนที่ดีของประชากร จากผลลัพธ์ที่เกิดขึ้นเป็นสิ่งยืนยันว่าการสุ่มแบบกริดเหมาะสมกับประชากรที่มีการกระจายตัวของค่าตัวแปรบนแผนที่เป็นไปตามกฎข้อที่ 1 ของภูมิศาสตร์หรือสหสัมพันธ์เชิงพื้นที่ทางบวก

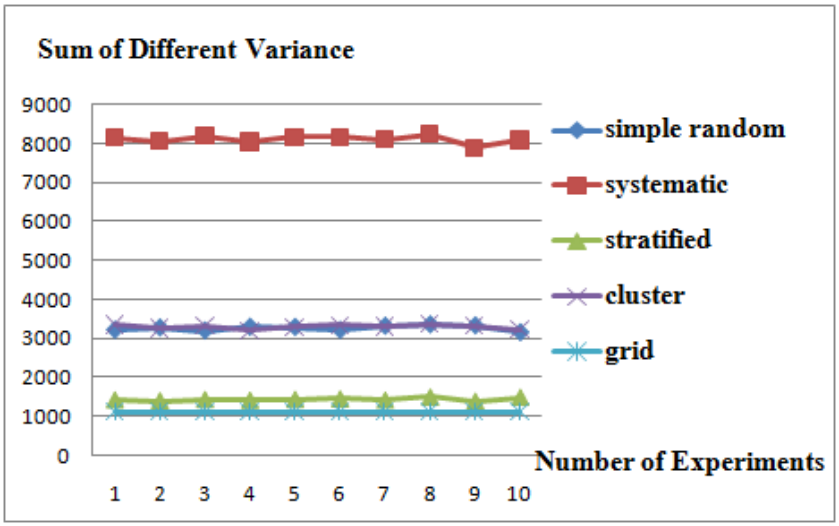
3. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S1



ภาพที่ 4-9 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S1

จากภาพที่ 4-9 แสดงให้เห็นว่า การสุ่มแบบกริดมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยที่สุด คือ 91.91 การสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 ซึ่งอยู่ระหว่าง 112.09-116.92 การสุ่มแบบชั้นภูมิมีค่าผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 3 และมีค่าระหว่าง 214.74-233.70 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวนใกล้เคียงกันและมากที่สุด โดยมีค่าอยู่ในช่วง 268.50-289.46 และ 279.58-297.19 ซึ่งสามารถสรุปได้ว่า วิธีการสุ่มแบบกริดมีผลรวมของค่าความแตกต่างระหว่างความแปรปรวนของประชากรและกลุ่มตัวอย่างน้อยกว่าวิธีการแบบอื่น ๆ เนื่องจากตำแหน่งของจุดกริดอยู่คงที่ โดยไม่เปลี่ยนแปลงไปตามจำนวนครั้งของการสุ่ม ทำให้กลุ่มตัวอย่างที่สุ่มได้เป็นกลุ่มเดิม แต่กลุ่มตัวอย่างที่สุ่มได้จากวิธีอื่น ๆ จะได้สมาชิกใหม่ ๆ เสมอ ทำให้ความแปรปรวนที่ได้มีค่ามากกว่าการสุ่มแบบกริด

4. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S2

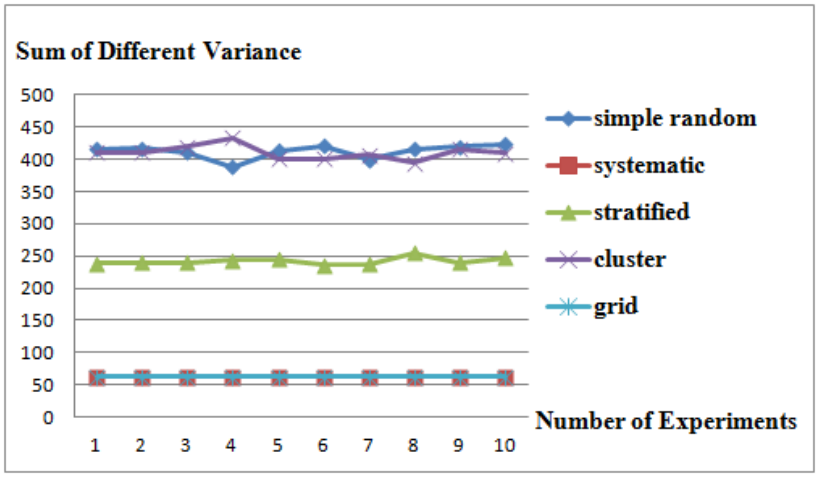


ภาพที่ 4-10 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S2

จากภาพที่ 4-10 แสดงให้เห็นว่า การสุ่มแบบกริดมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยที่สุด คือ 1,114.20 การสุ่มแบบชั้นภูมิมีค่าผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 และมีค่าระหว่าง 1383.25-1490.52 ในส่วนการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวน ใกล้เคียงกันและมีค่าน้อยเป็นลำดับที่ 3 โดยมีค่าอยู่ในช่วง 3,147.95-3,345.41 และ 3,206.39-3,361.25 สำหรับการสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนมากที่สุด ซึ่งอยู่ระหว่าง 7,900.29-8,252.66 ซึ่งสามารถสรุปได้ว่า ผลรวมของค่าความแตกต่างระหว่างความแปรปรวนของประชากรและกลุ่มตัวอย่างของการสุ่มแบบกริดน้อยกว่าวิธีการแบบอื่น ๆ การสุ่มแบบชั้นภูมิเป็นอันดับที่ 2 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มเป็นดับที่ 3 การสุ่มแบบมีระบบเป็นอันดับสุดท้าย

การสุ่มแบบมีระบบมีประสิทธิภาพต่ำที่สุดสำหรับแบบจำลอง S2 เนื่องจากการสุ่มแบบมีระบบไม่เหมาะสมที่จะนำมาใช้กับประชากรที่มีธรรมชาติของประชากรมีการจัดเรียงตัวของค่าสังเกตกับระยะห่างของการสุ่มตรงกันพอดี อาจทำให้ความแปรปรวนของกลุ่มตัวอย่างที่ได้นั้นต่ำเกินความเป็นจริง เช่น ถ้าสุ่มได้คอลัมน์ที่ 8 จะได้สมาชิกทั้ง 9 ช่อง ที่มีค่าสังเกตดังนี้ 2 2 2 2 2 2 2 1 1 ซึ่งแตกต่างจากความแปรปรวนของประชากรเป็นอย่างมาก เนื่องจากประชากรมีลักษณะวิวิธพันธ์ แต่กลุ่มตัวอย่างนี้มีลักษณะเป็นเอกพันธ์

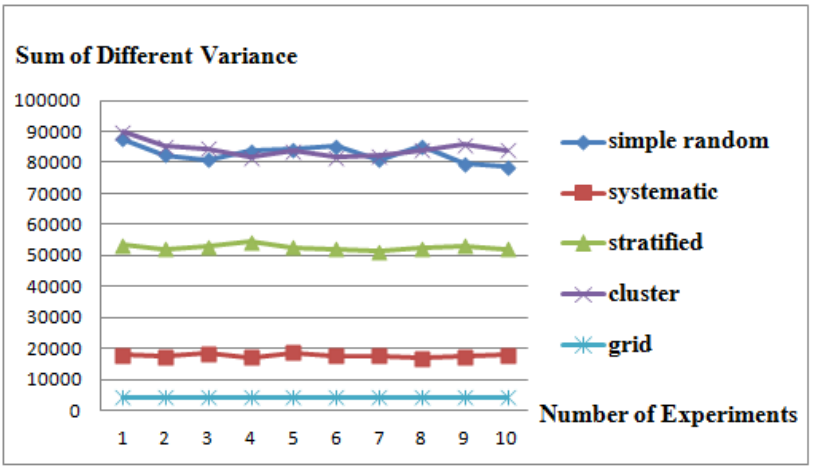
5. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S3



ภาพที่ 4-11 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S3

จากภาพที่ 4-11 แสดงให้เห็นว่า การสุ่มแบบกริดและการสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนใกล้เคียงกันและน้อยกว่าวิธีการสุ่มแบบอื่น ๆ ซึ่งวิธีการสุ่มแบบมีระบบมีค่าผลรวมของค่าความแตกต่างความแปรปรวน เท่ากับ 61.90 และวิธีการสุ่มแบบกริดมีค่าผลรวมของค่าความแตกต่างความแปรปรวน เท่ากับ 61.91 การสุ่มแบบชั้นภูมิมีค่าผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 และมีค่าระหว่าง 235.90-254.77 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวนใกล้เคียงกันและมีค่ามากที่สุด ซึ่งมีค่าระหว่าง 388.27-423.54 และ 394.44-432.89 ซึ่งสามารถสรุปได้ว่าวิธีการสุ่มอย่างง่ายและวิธีการสุ่มแบบแบ่งกลุ่มเป็นวิธีการที่มีประสิทธิภาพต่ำที่สุด เนื่องจากในบางครั้งกลุ่มตัวอย่างที่สุ่มได้มีความแปรปรวนแตกต่างจากความแปรปรวนของประชากร (พารามิเตอร์) มาก ดังนั้นกลุ่มตัวอย่างที่ได้จากทั้ง 2 วิธี จึงไม่เป็นตัวแทนที่ดีของประชากร

6. ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ โดยใช้แบบจำลอง S4



ภาพที่ 4-12 กราฟผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลอง S4

จากภาพที่ 4-12 แสดงให้เห็นว่า การสุ่มแบบกริดมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยที่สุด คือ 4,360.05 การสุ่มแบบมีระบบมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 2 และมีค่าระหว่าง 16,873.06-18,587.40 การสุ่มแบบชั้นภูมิมีผลรวมของค่าความแตกต่างความแปรปรวนน้อยเป็นลำดับที่ 3 และมีค่าระหว่าง 51,333.96-54387.65 ในส่วนการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีผลรวมของค่าความแตกต่างความแปรปรวนใกล้เคียงกันและมีค่ามากที่สุด ซึ่งมีค่าระหว่าง 78,542.26-87,439.04 และ 81,684.83-89,891.91 ซึ่งสามารถสรุปได้ว่า วิธีการสุ่มแบบกริดมีผลรวมของค่าความแตกต่างระหว่างความแปรปรวนของประชากรและกลุ่มตัวอย่างน้อยกว่าวิธีการแบบอื่น ๆ ซึ่งผลการวิเคราะห์นี้สอดคล้องกับแบบจำลอง S1

จากหัวข้อ “ผลการวิเคราะห์ค่าเฉลี่ยความคลาดเคลื่อนของวิธีการสุ่มแบบต่าง ๆ ” และ “ผลการวิเคราะห์ผลรวมของค่าความแตกต่างความแปรปรวนของวิธีการสุ่มแบบต่าง ๆ ” สามารถสรุปผลการทดลองในรูปแบบตารางได้ ดังนี้

ตารางที่ 4-1 เปรียบเทียบความแม่นยำในแต่ละวิธีการสุ่ม

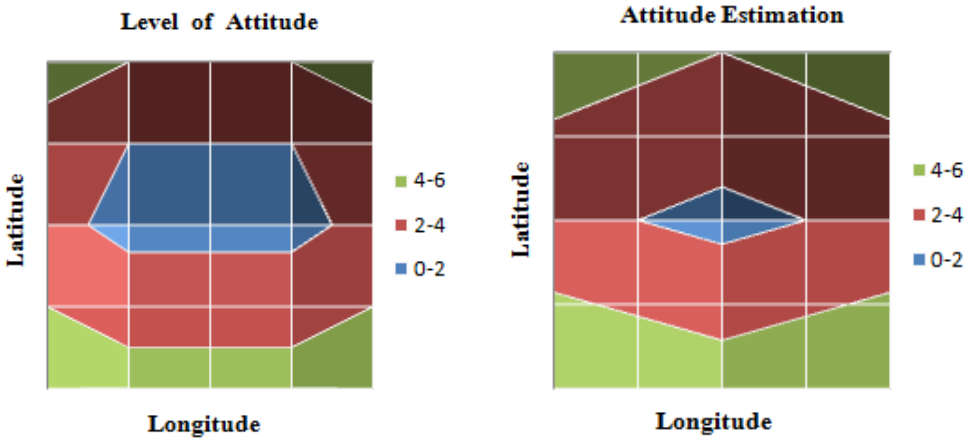
แบบจำลองการสุ่ม	วิธีการสุ่ม	ค่าเฉลี่ยความคลาดเคลื่อน	ผลรวมของความแตกต่างความแปรปรวน
S0	การสุ่มอย่างง่าย	0.569-0.610	1536.37-1634.20
	การสุ่มแบบมีระบบ	0.561-0.618	637.45-701.66
	การสุ่มแบบแบ่งกลุ่ม	0.580-0.624	1516.39-1622.97
	การสุ่มแบบกริด	0.654	3297.13
S00	การสุ่มอย่างง่าย	0.550-0.612	1500.99-1602.60
	การสุ่มแบบมีระบบ	0.634-0.668	907.06-993.91
	การสุ่มแบบแบ่งกลุ่ม	0.568-0.609	1491.64-1628.23
	การสุ่มแบบกริด	0.062	3500.38
S1	การสุ่มอย่างง่าย	0.247-0.266	268.50-289.46
	การสุ่มแบบมีระบบ	0.228-0.253	112.09-116.92
	การสุ่มแบบชั้นภูมิ	0.118-0.126	214.74-233.70
	การสุ่มแบบแบ่งกลุ่ม	0.242-0.256	279.58-297.19
	การสุ่มแบบกริด	0.074	91.91
S2	การสุ่มอย่างง่าย	0.916-1.002	3147.95-3345.41
	การสุ่มแบบมีระบบ	2.210-2.373	7900.29-8252.66
	การสุ่มแบบชั้นภูมิ	0.183-0.196	1383.25-1490.52
	การสุ่มแบบแบ่งกลุ่ม	0.897-0.924	3206.39-3361.25
	การสุ่มแบบกริด	0.556	1114.2
S3	การสุ่มอย่างง่าย	0.228-0.247	388.27-423.54
	การสุ่มแบบมีระบบ	0.167	61.9
	การสุ่มแบบชั้นภูมิ	0.063-0.069	235.90-254.77
	การสุ่มแบบแบ่งกลุ่ม	0.224-0.239	394.44-432.89
	การสุ่มแบบกริด	0.167	61.91
S4	การสุ่มอย่างง่าย	3.198-3.420	78542.26-87439.04
	การสุ่มแบบมีระบบ	2.496-2.518	16873.06-18587.40
	การสุ่มแบบชั้นภูมิ	1.013-1.095	51333.96-54387.65
	การสุ่มแบบแบ่งกลุ่ม	3.141-3.400	81684.83-89891.91
	การสุ่มแบบกริด	1.465	4360.05

3. ผลการวิเคราะห์แบบจำลองการประมาณค่าเชิงพื้นที่และค่าคลาดเคลื่อนร้อยละ

ในการวิจัยครั้งนี้ ทำการเขียนโปรแกรมคอมพิวเตอร์ เพื่อสร้างแบบจำลองต่าง ๆ เช่น SI1, SI2 และ SI3 โดยใช้โปรแกรม Microsoft office excel ทำการประมาณค่าเชิงพื้นที่และค่าคลาดเคลื่อนร้อยละ (Percent error) รวมทั้งแสดงภาพแบบสามมิติ เพื่อให้เห็นภาพได้อย่างชัดเจนและง่ายต่อการการวิเคราะห์ความคลาดเคลื่อน

ความคลาดเคลื่อนที่เกิดขึ้นในการทดสอบเป็นตัวบ่งชี้ถึงประสิทธิภาพของการสุ่มแบบกริดและความเหมาะสมในการนำระเบียบวิธีเชิงตัวเลขมาประยุกต์ ซึ่งค่าเฉลี่ยความคลาดเคลื่อนหรือค่าคลาดเคลื่อนร้อยละที่หามาได้นั้นสามารถยอมรับได้หรือไม่นั้น ให้พิจารณาที่การแปลความหมายของค่าประมาณว่าผิดไปจากความเป็นจริงมากน้อยเพียงใด

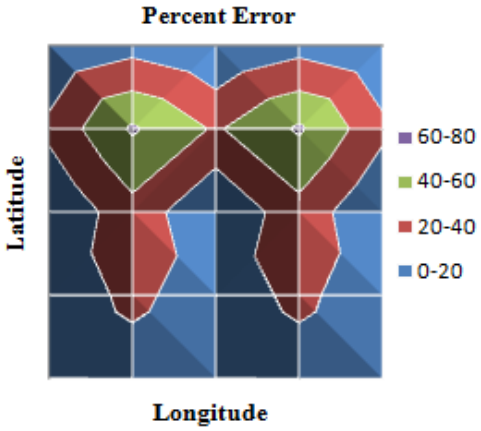
1. ผลการวิเคราะห์แบบจำลอง SI1



(ก) (ข)

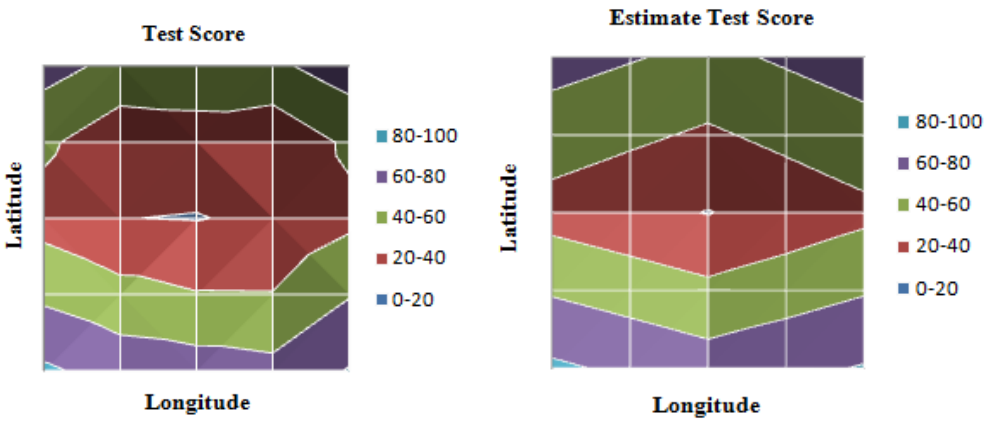
ภาพที่ 4-13 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงระดับทัศนคติ

จากภาพที่ 4-13 (ก) และ (ข) แสดงให้เห็นว่าค่าประมาณมีค่าใกล้เคียงกับค่าจริง แต่ในบริเวณพื้นที่ตอนกลางนั้นค่าประมาณมีความแตกต่างจากค่าจริงค่อนข้างมาก ซึ่งตำแหน่งที่มีคลาดเคลื่อนมากที่สุด คือ ค่าจริงเท่ากับ 2 และค่าประมาณเท่ากับ 3.25 ดังนั้นค่าความคลาดเคลื่อนร้อยละเท่ากับ 62.5% ตามภาพที่ 4-14 ซึ่งเกณฑ์การแปลความหมายค่าเฉลี่ยของเบสต์ (Best, 1981, pp. 204-208) ให้ความหมายว่า 2 หมายถึง คุณภาพพอใช้ แต่ 3.25 หมายถึง คุณภาพปานกลาง ดังนั้น การประมาณค่าทำให้เกิดความผิดพลาดไป 1 ระดับ



ภาพที่ 4-14 แผนที่ค่าคลาดเคลื่อนร้อยละ

2. ผลการวิเคราะห์แบบจำลอง SI2



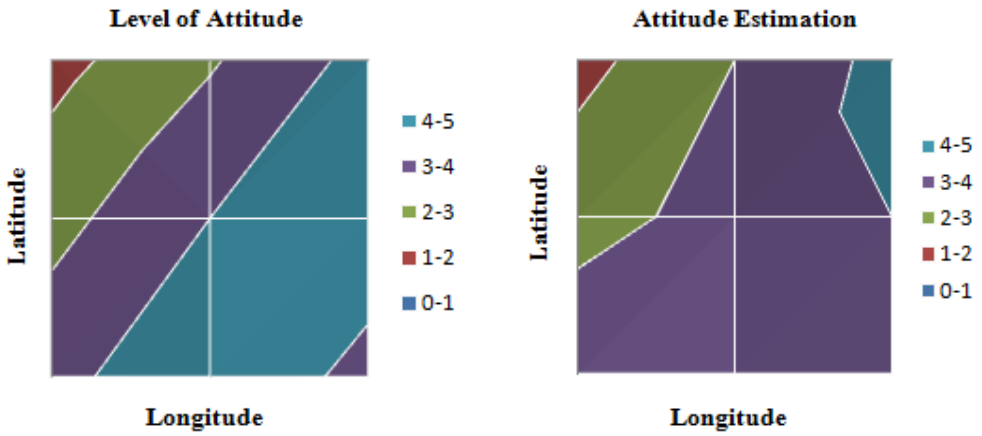
(ก) (ข)

ภาพที่ 4-15 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงคะแนนการทดสอบ

จากภาพที่ 4-15 (ก) และ (ข) แสดงให้เห็นว่าค่าประมาณมีค่าใกล้เคียงกับค่าจริง แต่ในบริเวณพื้นที่ตอนกลางนั้นค่าประมาณมีความแตกต่างจากค่าจริงค่อนข้างมาก ซึ่งตำแหน่งที่มีคลาดเคลื่อนมากที่สุด คือ ค่าจริงเท่ากับ 24.25 และค่าประมาณเท่ากับ 44.94 ดังนั้น ค่าความคลาดเคลื่อนร้อยละเท่ากับ 85.31% ตามภาพที่ 4-16 ถ้าสมมติให้คะแนนการทดสอบนี้เป็นคะแนน TOEFL IBT ซึ่งค่าจริง 24.25 จะอยู่ในช่วงคะแนน 19-29 แต่ค่าประมาณ 44.94 จะอยู่ในช่วงคะแนน 41-52 ดังนั้น ค่าการประมาณค่าทำให้เกิดความผิดพลาดไป 2 ระดับ

ภาพที่ 4-16 แผนที่ค่าคลาดเคลื่อนร้อยละ

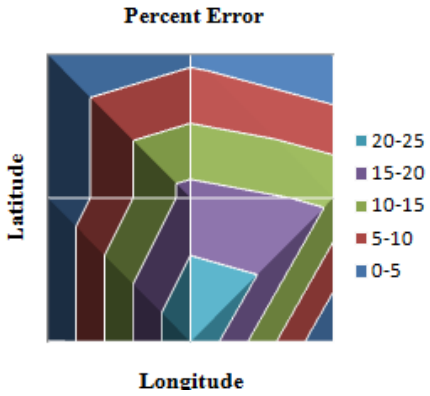
3. ผลการวิเคราะห์แบบจำลอง SI3



(ก) (ข)

ภาพที่ 4-17 (ก) แผนที่การประมาณค่าเชิงพื้นที่และ (ข) แผนที่แสดงระดับทัศนคติ

จากภาพที่ 4-17 (ก) และ (ข) แสดงให้เห็นว่าค่าประมาณมีค่าใกล้เคียงกับค่าจริง แต่ในบริเวณพื้นที่ตอนกลางจนถึงด้านตะวันออกเฉียงใต้นั้นค่าประมาณมีความแตกต่างจากค่าจริงค่อนข้างมาก ซึ่งตำแหน่งที่มีคลาดเคลื่อนมากที่สุดคือ ค่าจริงเท่ากับ 4.889 และค่าประมาณเท่ากับ 3.667 ดังนั้นค่าความคลาดเคลื่อนร้อยละเท่ากับ 25% ตามภาพที่ 4-18 ซึ่งเกณฑ์การแปลความหมายค่าเฉลี่ยของเบสต์ (Best, 1981, pp. 204-208) ให้ความหมายว่า 4.889 หมายถึง คุณภาพดีมาก แต่ 3.667 หมายถึง คุณภาพดี ดังนั้นการประมาณค่าทำให้เกิดความผิดพลาดไป 1 ระดับ



ภาพที่ 4-18 แผนที่ค่าคลาดเคลื่อนร้อยละ

จากหัวข้อ ผลการวิเคราะห์แบบจำลอง SI1 SI2 และ SI3 สามารถสรุปผลได้ตามตารางที่ 4-2 ดังนี้

ตารางที่ 4-2 รายละเอียดการประมาณค่าเชิงพื้นที่และค่าคลาดเคลื่อนร้อยละ

แบบจำลอง	กริดเซลล์ที่มีค่า error มากที่สุด	ค่าคลาดเคลื่อนร้อยละ	ระดับ	สรุป
SI1	ค่าจริง = 2 ค่าประมาณ = 3.25	62.50%	คุณภาพพอใช้ คุณภาพปานกลาง	ผิดพลาด 1 ระดับ
SI2	ค่าจริง = 24.25 ค่าประมาณ = 44.94	85.31%	ช่วงคะแนน 19-29 ช่วงคะแนน 41-52	ผิดพลาด 2 ระดับ
SI3	ค่าจริง = 4.89 ค่าประมาณ = 3.67	25%	คุณภาพดีมาก คุณภาพดี	ผิดพลาด 1 ระดับ

ตอนที่ 2 ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบหลายขั้นตอน โดยอาศัยฐานข้อมูลคะแนน O-NET

1. ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบหลายขั้นตอน

ในการสุ่มตัวอย่างนั้นสามารถกระทำได้หลายวิธี เช่น การจับฉลาก หรือ การใช้ตารางการสุ่ม แต่ในการวิจัยครั้งนี้ ผู้วิจัยได้เขียนโปรแกรมสุ่มจังหวัดและโรงเรียน เพื่อให้วิธีการสุ่ม

ตัวอย่างมีความเป็นรูปธรรม มีหลักฐานในการพิสูจน์และปราศจากอคติ ซึ่งโปรแกรมนี้เขียนขึ้นจากโปรแกรม Microsoft excel โดยการเรียกโปรแกรมทำงาน (Run program) จำนวน 30 ครั้ง และตามแผนการสุ่มต้องสุ่มโรงเรียนทางภาคเหนือ จำนวน 30 โรงเรียน จาก 2 จังหวัด, โรงเรียนทางภาคตะวันออกเฉียงเหนือ จำนวน 147 โรงเรียน จาก 9 จังหวัด, โรงเรียนทางภาคกลาง จำนวน 80 โรงเรียน จาก 7 จังหวัด, โรงเรียนทางภาคตะวันออก จำนวน 22 โรงเรียน จาก 1 จังหวัด, โรงเรียนทางภาคตะวันตก จำนวน 15 โรงเรียน จาก 1 จังหวัด, และโรงเรียนทางภาคใต้ จำนวน 41 โรงเรียน จาก 3 จังหวัด โดยมีรายละเอียดผลการสุ่มตัวอย่างตามผนวก 1 ซึ่งสามารถสรุปได้ ดังนี้

ตารางที่ 4-3 ผลลัพธ์ที่ได้จากวิธีการสุ่มแบบหลายขั้นตอน

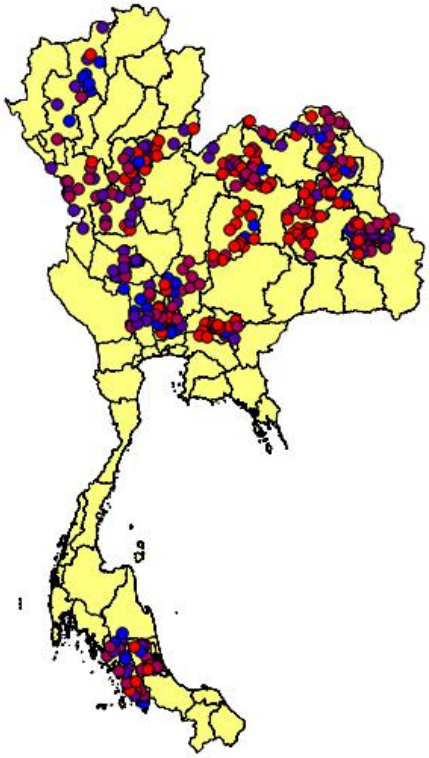
ภูมิภาค	รายชื่อจังหวัด	จำนวนโรงเรียน
เหนือ	เชียงใหม่	15
เหนือ	อุตรดิตถ์	15
ตะวันออกเฉียงเหนือ	อำนาจเจริญ	16
ตะวันออกเฉียงเหนือ	ชัยภูมิ	17
ตะวันออกเฉียงเหนือ	กาฬสินธุ์	16
ตะวันออกเฉียงเหนือ	เลย	16
ตะวันออกเฉียงเหนือ	มหาสารคาม	16
ตะวันออกเฉียงเหนือ	หนองบัวลำภู	16
ตะวันออกเฉียงเหนือ	หนองคาย	16
ตะวันออกเฉียงเหนือ	สกลนคร	17
ตะวันออกเฉียงเหนือ	ยโสธร	17
ภาคกลาง	กำแพงเพชร	11
ภาคกลาง	ลพบุรี	11
ภาคกลาง	ปทุมธานี	11
ภาคกลาง	อยุธยา	12
ภาคกลาง	สุโขทัย	12

ตารางที่ 4-3 (ต่อ)

ภูมิภาค	รายชื่อจังหวัด	จำนวนโรงเรียน
ภาคกลาง	สุพรรณบุรี	12
ภาคกลาง	อุทัยธานี	11
ภาคตะวันออก	ปราจีนบุรี	22
ภาคตะวันตก	ตาก	15
ภาคใต้	พัทลุง	14
ภาคใต้	สตูล	12
ภาคใต้	ตรัง	15
รวมทุกภูมิภาค	จำนวน 23 จังหวัด	335

หมายเหตุ การแบ่งส่วนภูมิภาคเป็นไปตามการจัดทำของราชบัณฑิตยสภา ตามแนวทางคณะกรรมการภูมิศาสตร์แห่งชาติ

เมื่อนำกลุ่มตัวอย่างที่สุ่มได้มาลงตำแหน่งในโปรแกรม QGIS จะได้ผลลัพธ์ตามภาพที่ 4-19 ซึ่งภาพนี้ได้แสดงให้เห็นตำแหน่งที่ตั้งของกลุ่มตัวอย่าง และการกระจายตัวของคะแนนเฉลี่ย O-NET ในภาพรวมของประเทศ (ซึ่งคะแนนเฉลี่ย O-NET นั้น จะอยู่ในลักษณะของเฉดสี โดยที่สีแดงหมายถึงคะแนนต่ำและสีน้ำเงินหมายถึงคะแนนสูง)



ภาพที่ 4-19 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบหลายขั้นตอน

จากผลลัพธ์ของการสุ่มที่กล่าวมาข้างต้นเป็นผลลัพธ์เพียงแค่ครั้งเดียว แต่ในการวิจัยที่ดีไม่ควรนำผลลัพธ์ที่ได้เพียงแค่ครั้งเดียวมาทำการสรุปผล ซึ่งอาจทำให้เกิดข้อผิดพลาดและปราศจากความน่าเชื่อถือ ดังนั้นจึงต้อง Run program เป็นจำนวน 30 ครั้ง โดยผลการวิเคราะห์จะนำเสนอต่อไปในหัวข้อ “ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มทั้ง 3 วิธี เมื่อเทียบกับพารามิเตอร์”

2. ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบกริด

จากผลการสุ่มตำแหน่งจุดกริดด้วยโปรแกรม QGIS ได้จำนวนจุดกริด 22 จุดกริด ซึ่งใช้ฃจุดกริดได้น้อยกว่าแผนการสุ่มที่วางไว้ (ตามแผนการสุ่มตัวอย่างแบบหลายขั้นตอนกำหนดจำนวนจังหวัดเท่ากับ 23 จังหวัด) โดยแต่ละจุดกริดต้องสุ่มโรงเรียนที่อยู่ใกล้ ๆ กับจุดกริดนั้น ๆ และจำนวนโรงเรียนที่ต้องสุ่มออกมาในแต่ละจุดกริดมีจำนวนแตกต่างกันขึ้นอยู่กับสัดส่วนจำนวนประชากรโรงเรียนในแต่ละภูมิภาค ถ้าภูมิภาคใดมีจำนวนประชากรโรงเรียนมาก กลุ่มตัวอย่างในแต่ละจุดกริดต้องถูกสุ่มออกมามากขึ้นด้วย

ตามแผนการสุ่มตัวอย่าง ภาคเหนือของประเทศไทยต้องสุ่มจุดกริดให้ได้ 2 จุดกริด และสุ่มโรงเรียนให้ได้ 30 โรงเรียน แต่โปรแกรม QGIS สุ่มจุดกริดได้ 4 จุด (ซึ่งเกินกว่าความต้องการ)

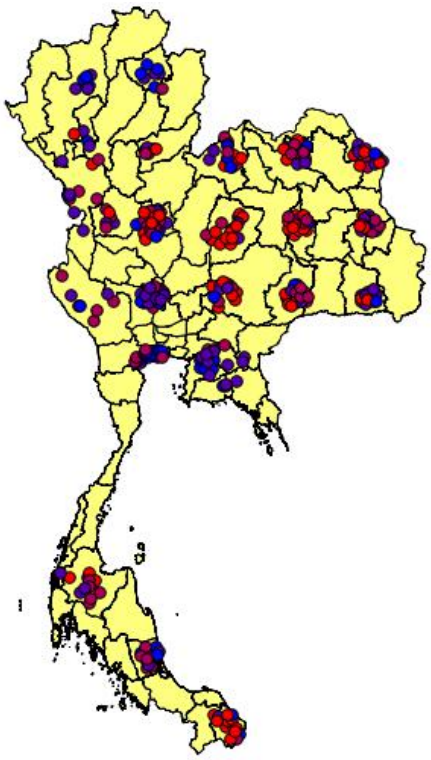
คือ จุดกริด i = 2, j = 10 อยู่ในจังหวัด เชียงใหม่ จุดกริด i = 3, j = 10 อยู่ในจังหวัดพะเยา จุดกริด i = 2, j = 9 อยู่ในจังหวัดลำพูน จุดกริด i = 3, j = 9 อยู่ในจังหวัดอุตรดิตถ์ ดังนั้นแต่ละจุดกริดต้องสุ่มโรงเรียนเป็นจำนวนประมาณ 30/ 4 = 7.5 โรงเรียน ยกตัวอย่างเช่น จุดกริด i = 2, j = 10 อยู่ในจังหวัดเชียงใหม่ ซึ่งอยู่ใกล้กับโรงเรียนจำนวน 8 โรงเรียน ได้แก่ แม่หอพระวิทยาคม แม่แตงสันป่ายางวิทยาคม สันทรายวิทยาคม แม่ริมวิทยาคม สะเมิงพิทยาคม ศึกษาสงเคราะห์เชียงใหม่ และนวมินทราชูทิศ พายัพ ส่วนจุดกริดอีก 3 จุดที่เหลือ มีรายละเอียดตามตารางที่ 4-3

ตามแผนการสุ่มตัวอย่าง ภาคกลางของประเทศไทยที่ต้องถูกสุ่มออกมา 7 จุดกริด และสุ่มโรงเรียนให้ได้ 80 โรงเรียน แต่โปรแกรม QGIS สุ่มจุดกริดได้เพียงแค่ 2 จุด คือ จุดกริด i = 3, j = 8 อยู่ในจังหวัดพิจิตร จุดกริด i = 3, j = 7 อยู่ในจังหวัดสุพรรณบุรี ใกล้แนวเขตจังหวัดอ่างทอง ทำให้จุดกริดที่ได้น้อยกว่าแผนการสุ่มที่วาง 5 จุด ทั้งนี้เนื่องจากจำนวนโรงเรียนในภาคกลางมีมากเมื่อเทียบกับขนาดของพื้นที่ ดังนั้นเพื่อให้เป็นไปตามแผนการสุ่มจึงจำเป็นต้องเพิ่มจุดกริด โดยนับรวมกับ จุดกริด i = 2, j = 8 ที่อยู่ในประเทศเพื่อนบ้าน และ จุดกริด i = 3, j = 6 ที่อยู่ในทะเล รวมเป็นจำนวนจุดกริดทั้งหมด 4 จุดกริด (ซึ่งจำนวนจุดกริดรวมทั้งหมดยังคงน้อยกว่าแผนการสุ่มที่คาดการณ์) ดังนั้นทั้ง 4 จุดกริด ต้องเพิ่มจำนวนโรงเรียนต่อจุดกริดให้มากขึ้น เพื่อให้ครบ 80 โรงเรียน แต่ในจุดกริด i = 2, j = 8 อยู่ในประเทศพม่าติดกับชายแดนจังหวัดตาก เมื่อใช้จุดนี้เป็นจุดอ้างอิงสามารถสุ่มโรงเรียนได้ 6 โรงเรียนเท่านั้น เพราะโรงเรียนอื่น ๆ นั้น อยู่ห่างไกลจุดกริด i = 2, j = 8 มากเกินไป ดังนั้นจุดกริดที่เหลือทั้ง 3 จุดกริดต้องสุ่มโรงเรียนในแต่ละจุดกริดเท่ากับ 24 หรือ 25 โรงเรียน ยกตัวอย่าง เช่น จุดกริด i = 3, j = 6 ต้องสุ่มโรงเรียนออกมา 25 โรงเรียน ได้แก่ สมุทรสงคราม 8 โรงเรียน สมุทรสาคร 10 โรงเรียน กรุงเทพฯ 6 โรงเรียน นครปฐม 1 โรงเรียน ในส่วนของภาคตะวันออก ภาคตะวันออกเฉียงเหนือ ภาคตะวันตก ภาคใต้ มีผลลัพธ์ที่ได้จากการสุ่มตัวอย่างตามรายละเอียดในผนวก 2 ซึ่งสามารถสรุปได้ ดังนี้

ตารางที่ 4-4 ผลลัพธ์ที่ได้จากวิธีการสุ่มแบบกริด

ตำแหน่ง	จำนวนโรงเรียน	จังหวัด	ภูมิภาค	ค่าเฉลี่ย (คะแนน)	ส่วนเบี่ยงเบนมาตรฐาน (คะแนน)
i = 2, j = 9	7	ลำพูน	เหนือ	158.04	10.03
i = 2, j = 10	8	เชียงใหม่	เหนือ	171.80	15.41
i = 3, j = 9	7	อุตรดิตถ์	เหนือ	168.21	31.35
i = 3, j = 10	8	พะเยา	เหนือ	165.13	8.13
i = 2, j = 8	13	อยู่ในประเทศพม่า	กลางและตะวันตก	156.49	9.87
i = 3, j = 6	25	อยู่ในทะเล	กลาง	175.80	22.17
i = 3, j = 7	25	สุพรรณบุรี	กลาง	159.98	7.76
i = 3, j = 8	24	พิจิตร	กลาง	155.92	16.88
i = 4, j = 6	22	ฉะเชิงเทรา	ตะวันออก	168.12	19.32
i = 4, j = 9	16	เลย	ตะวันออกเฉียงเหนือ	157.96	20.09
i = 5, j = 9	16	อุดรธานี	ตะวันออกเฉียงเหนือ	154.55	8.93
i = 6, j = 9	16	นครพนม	ตะวันออกเฉียงเหนือ	153.02	9.53
i = 4, j = 8	16	ชัยภูมิ	ตะวันออกเฉียงเหนือ	147.2	7.22
i = 5, j = 8	19	มหาสารคาม	ตะวันออกเฉียงเหนือ	150.88	17.45
i = 6, j = 8	16	ยโสธร	ตะวันออกเฉียงเหนือ	152.29	9.89
i = 4, j = 7	16	นครราชสีมา	ตะวันออกเฉียงเหนือ	152.54	13.48
i = 5, j = 7	16	บุรีรัมย์	ตะวันออกเฉียงเหนือ	154.98	9.95
i = 6, j = 7	16	ศรีสะเกษ	ตะวันออกเฉียงเหนือ	158.82	10.81
i = 2, j = 7	8	กาญจนบุรี	ตะวันตก	161.39	25.24
i = 2, j = 3	13	สุราษฎร์ธานี	ภาคใต้	151.69	6.12
i = 3, j = 2	14	พัทลุง	ภาคใต้	166.2	17.46
i = 4, j = 1	14	นราธิวาส	ภาคใต้	149.06	13.85
จำนวน 22 จุดกริด	335	-	-	158.49	16.54

เมื่อหาโรงเรียนได้ครบในแต่ละจุดกริดได้ตามตารางที่ 4-4 จากนั้นนำกลุ่มตัวอย่างที่สุ่มได้มาลงตำแหน่งในโปรแกรม QGIS จะได้ผลลัพธ์ตามภาพที่ 4-20 ซึ่งภาพนี้ได้แสดงให้เห็นตำแหน่งที่ตั้งของกลุ่มตัวอย่างและการกระจายตัวของคะแนนเฉลี่ย O-NET ในภาพรวมของประเทศ โดยโรงเรียนที่ถูกสุ่มออกมานั้นจะอยู่ใกล้ ๆ กัน บริเวณจุดกริดที่กำหนดไว้



ภาพที่ 4-20 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบกริด

3. ผลการวิเคราะห์แบบจำลองการสุ่มตัวอย่างแบบกริด-ชั้นภูมิ

เนื่องจากวิธีการสุ่มแบบกริด-ชั้นภูมิมีขั้นตอนการปฏิบัติเช่นเดียวกับวิธีการสุ่มแบบกริด ดังนั้นสามารถนำผลลัพธ์ที่หาได้จากวิธีการสุ่มแบบกริดมาใช้ได้ ยกตัวอย่างเช่น จำนวนจุดกริด ตำแหน่งจุดกริด จำนวนกลุ่มตัวอย่างในแต่ละจุดกริด เป็นต้น ในส่วนความแตกต่างระหว่างการสุ่มแบบกริดและการสุ่มแบบกริด-ชั้นภูมิ คือ การเลือกกลุ่มตัวอย่างกล่าวคือกลุ่มตัวอย่างที่ได้จากวิธีการสุ่มแบบกริดเป็นโรงเรียนที่อยู่บริเวณใกล้เคียงกับจุดกริด แต่กลุ่มตัวอย่างที่ได้จากวิธีการสุ่มแบบกริด-ชั้นภูมิ อาจเป็นโรงเรียนที่อยู่ไกลกว่า ทั้งนี้ขึ้นอยู่การแบ่งชั้นภูมิ และตัวแปรที่ถูกนำมาใช้ในการแบ่งชั้นภูมิ คือ ตัวแปรขนาดโรงเรียน โดยพิจารณาข้อมูลจากจำนวนรักเรียนที่เข้าสอบ O-NET และก่อนที่จะแบ่งประชากรในแต่ละจุดกริดเป็นชั้นภูมิต้องทดสอบความสัมพันธ์ของ

ตัวแปร โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายของเพียร์สัน (Pearson's product-moment correlation coefficient) เพื่อพิสูจน์ความสัมพันธ์ระหว่างขนาดโรงเรียน (จำนวนนักเรียน) และคะแนน O-NET

H_0: ขนาดโรงเรียนไม่มีความสัมพันธ์กับคะแนน O-NET

H_1: ขนาดโรงเรียนมีความสัมพันธ์กับคะแนน O-NET

ตารางที่ 4-5 ผลการทดสอบความสัมพันธ์ระหว่างขนาดโรงเรียนและคะแนน O-NET

Pearson correlation	จำนวนโรงเรียน	Sig
0.780**	2584	0.00

หมายเหตุ ** หมายถึง ความมีนัยสำคัญทางสถิติที่ระดับ 0.01 (2-tailed)

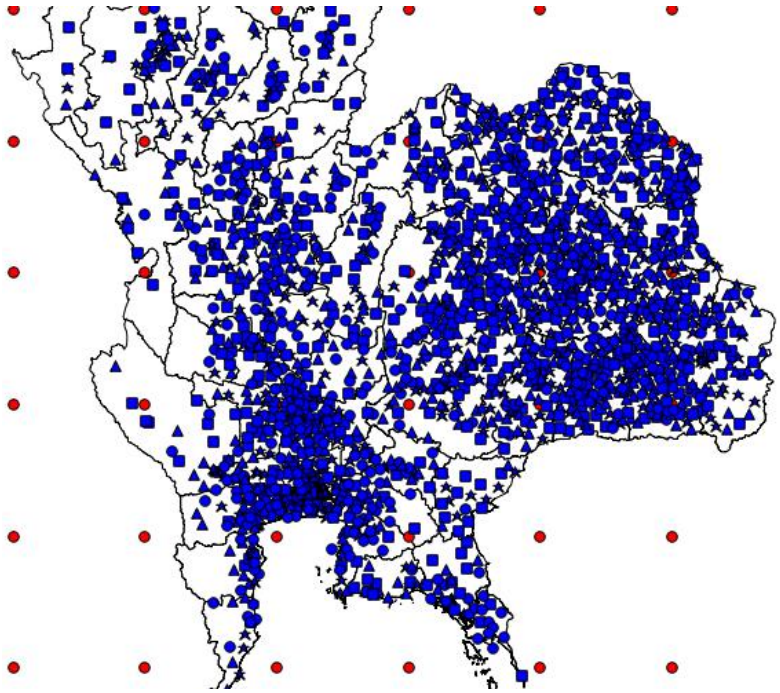
จากตารางที่ 4-5 ผลการทดสอบความสันพันธ์ระหว่างระหว่างขนาดโรงเรียนและคะแนน O-NET โดยใช้สถิติการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์อย่างง่ายของเพียร์สัน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 สามารถวิเคราะห์ได้ว่า ขนาดโรงเรียนมีความสัมพันธ์ทางบวกกับคะแนน O-NET ที่สัมประสิทธิ์สหสัมพันธ์ (**r**) เท่ากับ 0.78 ซึ่งมีความสัมพันธ์ในระดับสูง (0.71-0.90 มีความสัมพันธ์กันระดับสูง) ดังนั้นประเทศไทยมีโรงเรียนที่มีนักเรียนเข้าสอบ O-NET ในปีการศึกษา 2548 ทั้งหมด 2,584 โรงเรียน เมื่อแบ่งชั้นภูมิโรงเรียนตามจำนวนนักเรียน เป็น 4 ชั้นภูมิ (แบบควอไทล์) จะแบ่งชั้นภูมิได้ดังนี้

ชั้นภูมิที่ 1 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 1-28 คน ใช้สัญลักษณ์ วงกลม (Q_1)

ชั้นภูมิที่ 2 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 28-53 คน ใช้สัญลักษณ์ สามเหลี่ยม (Q_2)

ชั้นภูมิที่ 3 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 53-111 คน ใช้สัญลักษณ์ สี่เหลี่ยม (Q_3)

ชั้นภูมิที่ 4 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 111-1615 คน ใช้สัญลักษณ์ ห้าเหลี่ยม (Q_4)



ภาพที่ 4-21 การแบ่งชั้นภูมิของโรงเรียนทั่วประเทศ

ภาพที่ 4-21 แสดงให้เห็นการแบ่งชั้นภูมิของโรงเรียนทั่วประเทศตาม 4 ชั้นภูมิ และถ้าขยายรัศมีวงกลมของการสุ่มในแต่ละจุดกริดและทำการจัดสรรแบบสัดส่วน โดยมีรายละเอียดผลการสุ่มตัวอย่างตามผนวก 3 ซึ่งสามารถสรุปได้ ดังนี้

ตารางที่ 4-6 ค่าสถิติและจำนวนกลุ่มตัวอย่างที่ได้จากการจัดสรรแบบสัดส่วน

ตำแหน่ง	จำนวนสมาชิก	จำนวนโรงเรียนในแต่ละประเภท				ค่าเฉลี่ย (คะแนน)	ส่วนเบี่ยงเบนมาตรฐาน (คะแนน)
		⌂	□	Δ	o		
i = 2, j = 9	ขนาดกลุ่มตัวอย่าง	2	3	2	0	159.5	15.85
	จำนวนทั้งหมด	5	8	8	1		
i = 2, j = 10	ขนาดกลุ่มตัวอย่าง	3	2	2	1	169.66	17.03
	จำนวนทั้งหมด	18	12	9	4		
i = 3, j = 9	ขนาดกลุ่มตัวอย่าง	2	2	1	2	161.69	22.12
	จำนวนทั้งหมด	17	12	12	15		

ตารางที่ 4-6 (ต่อ)

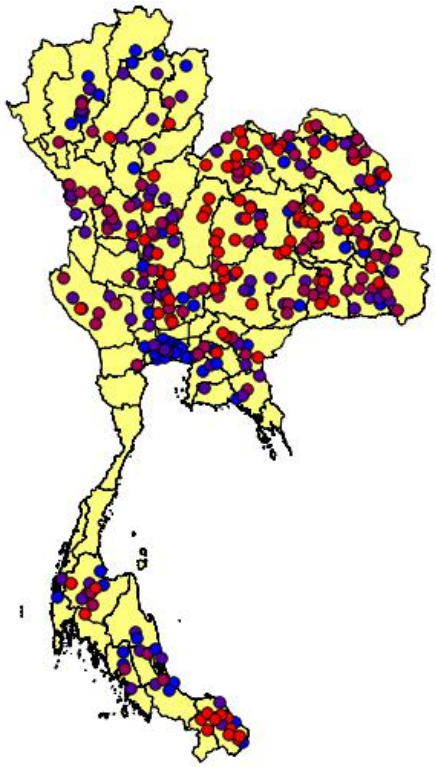
ตำแหน่ง	จำนวนสมาชิก	จำนวนโรงเรียนในแต่ละประเภท				ค่าเฉลี่ย (คะแนน)	ส่วนเบี่ยงเบนมาตรฐาน (คะแนน)
		⌂	□	Δ	o		
i = 3, j = 10	ขนาดกลุ่มตัวอย่าง	4	2	1	1	168.04	10.46
	จำนวนทั้งหมด	23	10	9	7		
i = 2, j = 8	ขนาดกลุ่มตัวอย่าง	0	6	4	3	154.02	3.79
	จำนวนทั้งหมด	1	6	4	3		
i = 3, j = 6	ขนาดกลุ่มตัวอย่าง	15	5	2	3	193.55	27.05
	จำนวนทั้งหมด	108	34	14	26		
i = 3, j = 7	ขนาดกลุ่มตัวอย่าง	5	7	5	8	160.6	20.1
	จำนวนทั้งหมด	30	36	30	45		
i = 3, j = 8	ขนาดกลุ่มตัวอย่าง	5	6	7	6	159.75	21.59
	จำนวนทั้งหมด	18	25	30	25		
i = 4, j = 6	ขนาดกลุ่มตัวอย่าง	4	6	6	6	164.29	17.04
	จำนวนทั้งหมด	16	23	21	23		
i = 4, j = 9	ขนาดกลุ่มตัวอย่าง	3	6	4	3	148.8	7.47
	จำนวนทั้งหมด	9	17	11	8		
i = 5, j = 9	ขนาดกลุ่มตัวอย่าง	3	3	6	4	155.44	18.15
	จำนวนทั้งหมด	21	23	39	31		
i = 6, j = 9	ขนาดกลุ่มตัวอย่าง	3	4	4	5	153.55	7.84
	จำนวนทั้งหมด	11	19	20	22		
i = 4, j = 8	ขนาดกลุ่มตัวอย่าง	4	4	5	3	149.48	11.35
	จำนวนทั้งหมด	19	18	27	16		
i = 5, j = 8	ขนาดกลุ่มตัวอย่าง	4	4	6	5	151.08	13.93
	จำนวนทั้งหมด	37	38	61	52		
i = 6, j = 8	ขนาดกลุ่มตัวอย่าง	3	4	6	3	153.77	9.05
	จำนวนทั้งหมด	24	30	45	22		
i = 4, j = 7	ขนาดกลุ่มตัวอย่าง	4	3	5	4	154.12	12.29
	จำนวนทั้งหมด	22	15	27	22		
i = 5, j = 7	ขนาดกลุ่มตัวอย่าง	4	4	5	3	154	9.59

	จำนวนทั้งหมด	30	30	39	28		

ตารางที่ 4-6 (ต่อ)

ตำแหน่ง	จำนวนสมาชิก	จำนวนโรงเรียนในแต่ละประเภท				ค่าเฉลี่ย (คะแนน)	ส่วนเบี่ยงเบนมาตรฐาน (คะแนน)
		⌂	□	Δ	o		
i = 6, j = 7	ขนาดกลุ่มตัวอย่าง	4	4	5	3	157.17	9.84
	จำนวนทั้งหมด	28	33	40	27		
i = 2, j = 7	ขนาดกลุ่มตัวอย่าง	0	4	3	1	161.39	25.24
	จำนวนทั้งหมด	0	4	3	1		
i = 2, j = 3	ขนาดกลุ่มตัวอย่าง	4	3	2	4	159.27	13.81
	จำนวนทั้งหมด	13	11	9	14		
i = 3, j = 2	ขนาดกลุ่มตัวอย่าง	3	4	3	4	170.74	19.21
	จำนวนทั้งหมด	17	20	20	24		
i = 4, j = 1	ขนาดกลุ่มตัวอย่าง	1	3	3	7	149.38	14.3
	จำนวนทั้งหมด	3	6	8	17		
จำนวน 22 จุดกริด	ขนาดกลุ่มตัวอย่าง	80	89	87	79	159.97	19.41

เมื่อใช้โปรแกรม QGIS แสดงตำแหน่งของกลุ่มตัวอย่างในแต่ละจุดกริด โดยแสดงค่าเป็นเฉดสีตามคะแนนเฉลี่ย O-NET ของโรงเรียน จะได้ผลลัพธ์ตามภาพที่ 4-22



ภาพที่ 4-22 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบกริด-ชั้นภูมิ

ภาพที่ 4-22 ตำแหน่งของกลุ่มตัวอย่างที่ได้จากการสุ่มแบบกริด-ชั้นภูมิจากผลลัพธ์ของการสุ่มที่กล่าวมาข้างต้นเป็นผลลัพธ์เพียงแค่ครั้งเดียว แต่ในการวิจัยที่ดีไม่ควรนำผลลัพธ์ที่ได้เพียงแค่ครั้งเดียวมาทำการสรุปผล ซึ่งอาจทำให้เกิดข้อผิดพลาดและปราศจากความน่าเชื่อถือ ดังนั้นจึงต้อง Run program เป็นจำนวน 30 ครั้ง โดยผลการวิเคราะห์จะนำเสนอต่อไปในหัวข้อ “ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มทั้ง 3 วิธี เมื่อเทียบกับพารามิเตอร์”

4. ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างทั้ง 3 วิธี เมื่อเทียบกับพารามิเตอร์

การที่จะพิสูจน์ว่าค่าสถิติที่ได้จากการวัดคุณลักษณะกลุ่มตัวอย่างนั้น สามารถบรรยายลักษณะที่สำคัญของกลุ่มประชากรที่ต้องการศึกษาได้ถูกต้องหรือไม่ ต้องนำค่าสถิติไปเปรียบเทียบกับค่าพารามิเตอร์ที่ ถ้าค่าสถิติใกล้เคียงกับค่าพารามิเตอร์แสดงว่าวิธีการสุ่มนั้นมีประสิทธิภาพสูง หรือกล่าวได้ว่าค่าสถิติที่ได้เป็นตัวแทนที่ดีของประชากร ดังนั้นในการเปรียบเทียบประสิทธิภาพระหว่างแบบจำลองการสุ่มแบบกริด-ชั้นภูมิ แบบจำลองการสุ่มแบบกริดและแบบจำลองการสุ่มแบบหลายขั้นตอนนั้น ผู้วิจัยใช้ค่าสถิติที่ได้จากกลุ่มตัวอย่าง คือ ค่าเฉลี่ยกลุ่มตัวอย่าง เพื่อเปรียบเทียบกับค่าพารามิเตอร์ คือ ค่าเฉลี่ยประชากร และจากข้อมูลคะแนนสอบ O-NET ของนักเรียน

ชั้นมัธยมศึกษาปีที่ 6 ของโรงเรียนต่าง ๆ ทั่วประเทศ เป็นจำนวน 2,584 โรงเรียน มีค่าเฉลี่ยของประชากรเท่ากับ 161.67 คะแนน และความแปรปรวนของประชากรเท่ากับ 472.62 หรือส่วนเบี่ยงเบนมาตรฐานเท่ากับ 21.74 คะแนน ผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างทั้ง 3 วิธี เมื่อเทียบกับพารามิเตอร์เป็นไปตามตารางที่ 4-7

ตารางที่ 4-7 ค่าเฉลี่ยความคลาดเคลื่อนของแต่ละวิธีการสุ่ม

การสุ่มครั้งที่	การสุ่มแบบหลายขั้นตอน			การสุ่มแบบกริด			การสุ่มแบบกริด-ชั้นภูมิ		
	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror
1	161.67	156.56	5.11	161.67	158.49	3.18	161.67	159.97	1.7
2	161.67	162.96	1.29	161.67	158.49	3.18	161.67	159.89	1.78
3	161.67	157.44	4.23	161.67	158.49	3.18	161.67	160.69	0.98
4	161.67	158.81	2.86	161.67	158.49	3.18	161.67	159.59	2.08
5	161.67	163.92	2.25	161.67	158.49	3.18	161.67	159.67	2
6	161.67	157.86	3.81	161.67	158.49	3.18	161.67	159.51	2.16
7	161.67	158.72	2.95	161.67	158.49	3.18	161.67	158.97	2.7
8	161.67	159.34	2.33	161.67	158.49	3.18	161.67	159.93	1.74
9	161.67	163.97	2.3	161.67	158.49	3.18	161.67	159.49	2.18
10	161.67	161.32	0.35	161.67	158.49	3.18	161.67	159.45	2.22
11	161.67	158.95	2.72	161.67	158.49	3.18	161.67	159.24	2.43
12	161.67	158.65	3.02	161.67	158.49	3.18	161.67	159.72	1.95
13	161.67	158.55	3.12	161.67	158.49	3.18	161.67	158.74	2.93
14	161.67	159.64	2.03	161.67	158.49	3.18	161.67	159.69	1.98
15	161.67	158.34	3.33	161.67	158.49	3.18	161.67	159.2	2.47
16	161.67	157.26	4.41	161.67	158.49	3.18	161.67	159.84	1.83
17	161.67	162.63	0.96	161.67	158.49	3.18	161.67	159.91	1.76
18	161.67	159.79	1.88	161.67	158.49	3.18	161.67	160.02	1.65
19	161.67	159.89	1.78	161.67	158.49	3.18	161.67	158.88	2.79
20	161.67	157.53	4.14	161.67	158.49	3.18	161.67	158.74	2.93
21	161.67	163.6	1.93	161.67	158.49	3.18	161.67	159.45	2.22
22	161.67	158.07	3.6	161.67	158.49	3.18	161.67	159.7	1.97

ตารางที่ 4-7 ค่าเฉลี่ยความคลาดเคลื่อนของแต่ละวิธีการสุ่ม

การสุ่มครั้งที่	การสุ่มแบบหลายขั้นตอน			การสุ่มแบบกริด			การสุ่มแบบกริด-ชั้นภูมิ		
	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror	ค่าพารามิเตอร์	ค่าสถิติ	ค่าerror
23	161.67	157.37	4.3	161.67	158.49	3.18	161.67	159.48	2.19
24	161.67	161.88	0.21	161.67	158.49	3.18	161.67	159.72	1.95
25	161.67	157.13	4.54	161.67	158.49	3.18	161.67	159.79	1.88
26	161.67	161.49	0.18	161.67	158.49	3.18	161.67	159.55	2.12
27	161.67	161.85	0.18	161.67	158.49	3.18	161.67	160.25	1.42
28	161.67	158.32	3.35	161.67	158.49	3.18	161.67	160.04	1.63
29	161.67	163.65	1.98	161.67	158.49	3.18	161.67	160.33	1.34
30	161.67	158.68	2.99	161.67	158.49	3.18	161.67	159.41	2.26
	ค่าความแปรปรวน	5.323		ค่าความแปรปรวน	0		ค่าความแปรปรวน	0.202	
	ค่าเฉลี่ยความคลาดเคลื่อน		2.52	ค่าเฉลี่ยความคลาดเคลื่อน		3.18	ค่าเฉลี่ยความคลาดเคลื่อน		2.05

จากตารางที่ 4-7 แสดงให้เห็นความแม่นยำของวิธีการสุ่มทั้ง 3 วิธี โดยใช้ค่าเฉลี่ยความคลาดเคลื่อนเป็นตัวเปรียบเทียบ ซึ่งวิธีการสุ่มทั้งหมดมีความแม่นยำที่สามารถยอมรับได้ เนื่องจากวิธีที่มีค่าเฉลี่ยความคลาดเคลื่อนมากที่สุด (การสุ่มแบบกริด) มีค่าเท่ากับ 3.18 คะแนน หรืออาจกล่าวได้ว่าโดยปกติเมื่อใช้วิธีการสุ่มแบบกริดจะมีค่าความคลาดเคลื่อนไม่เกิน 3.18 คะแนน ซึ่งถือว่าน้อยมากจากค่าจริง (161.67 คะแนน) ในส่วนของประสิทธิภาพการสุ่มนั้นพิจารณาจากทรัพยากรที่ใช้ในการวิจัย ถึงแม้ว่าในการทำวิจัยครั้งนี้จะไม่มีการลงพื้นที่ เพื่อเก็บรวบรวมข้อมูล แต่ในความเป็นจริงมีงานวิจัยเป็นจำนวนมากที่มีข้อจำกัดด้านทรัพยากร และเนื่องจากวิธีการสุ่มทั้ง 3 วิธี ใช้กลุ่มตัวอย่างเท่ากับ 335 โรงเรียนจากประชากร 2584 โรงเรียน จึงสามารถลดความสิ้นเปลืองในการวิจัยได้อย่างมาก เช่น ค่าใช้จ่าย ระยะเวลา เจ้าหน้าที่ที่เก็บรวบรวมข้อมูล เป็นต้น ดังนั้นวิธีการสุ่มทั้ง 3 วิธี จึงอยู่ในเกณฑ์ที่ยอมรับได้

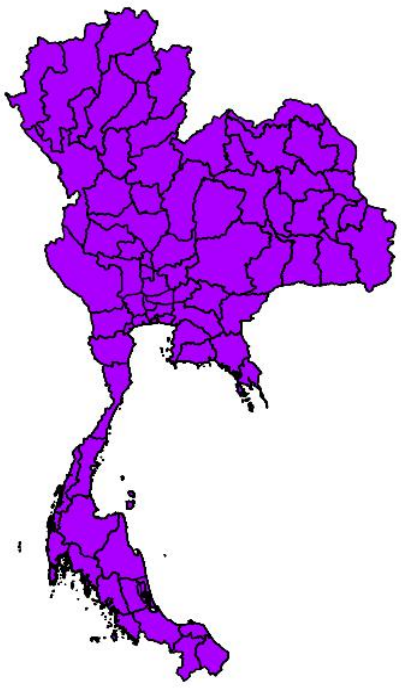
จากผลการทดลอง พบว่า วิธีการสุ่มแบบกริดเป็นวิธีที่มีความแม่นยำต่ำกว่าวิธีการสุ่มแบบหลายขั้นตอน และวิธีการสุ่มแบบกริดชั้นภูมิ เนื่องจากในบางพื้นที่มีสหสัมพันธ์เชิงพื้นที่ต่ำ (ใกล้เคียง 0) ดังนั้นกลุ่มตัวอย่างที่ถูกสุ่มออกมาจึงเป็นกลุ่มตัวอย่างที่อยู่ใกล้กับตำแหน่งจุดกริด ซึ่งเป็นประชากรกลุ่มย่อยที่ไม่ใช่ตัวแทนประชากรที่ดี ในส่วนวิธีการสุ่มแบบหลายขั้นตอนเป็น

วิธีการสุ่มที่ให้โอกาสกับสมาชิกทุกตัวในประชากรได้รับการสุ่มออกมา ดังนั้นในบางครั้งอาจสุ่มได้โรงเรียนที่มีคะแนนเฉลี่ย O-NET สูงเป็นจำนวนมาก แต่ในบางครั้งอาจสุ่มได้โรงเรียนที่คะแนนเฉลี่ย O-NET ต่ำเป็นจำนวนมาก ทำให้เกิดความคลาดเคลื่อนและมีความแปรปรวนสูง (5.323) ซึ่งแตกต่างจากวิธีการสุ่มแบบกริด-ชั้นภูมิ ที่มีสมาชิกภายในกลุ่มตัวอย่างมาจากต่างชั้นภูมิ ทำให้มีความหลากหลายและมาจากทุกกลุ่มก้อนของประชากร หรือกลุ่มตัวอย่างเป็นตัวแทนที่ดีของประชากร ทำให้ค่าเฉลี่ยความคลาดเคลื่อนต่ำที่สุด (2.05 คะแนน) และมีความแปรปรวนต่ำ

5. ผลการเปรียบเทียบแผนที่สารสนเทศที่เกิดจากวิธีการสุ่มแบบหลายขั้นตอนและวิธีการสุ่มแบบกริด/ กริด-ชั้นภูมิ

วิธีการสุ่มตัวอย่างแบบหลายขั้นตอนเป็นวิธีการที่ถูกสร้างขึ้นด้วยหลักวิชาสถิติ โดยมีจุดมุ่งหมายในการวัดแนวโน้มเข้าสู่ส่วนกลาง เพื่อหาค่าที่เป็นตัวแทนสำหรับอ้างอิงชุดข้อมูล ดังนั้นสารสนเทศที่ได้จากวิธีการสุ่มตัวแบบหลายขั้นตอนจะมีลักษณะเป็นบทสรุปในภาพรวม

เมื่อใช้วิธีการสุ่มตัวอย่างแบบหลายขั้นตอนกับฐานข้อมูลของคะแนน O-NET ปีการศึกษา 2548 โดยใช้ขนาดกลุ่มตัวอย่างเท่ากับ 335 โรงเรียน จากจังหวัดทั้งหมด 23 จังหวัด ซึ่งกลุ่มตัวอย่างที่สุ่มได้มีค่าเฉลี่ยเท่ากับ 156.56 คะแนน เมื่อนำมาแสดงเป็นแผนที่ตามเฉดสีในโปรแกรม QGIS พบว่าค่า 156.56 คะแนน อยู่ในช่วงคะแนนสอบของควอร์ไทล์ที่ 3 ซึ่งแทนด้วยสีน้ำเงินปนม่วง ตามภาพที่ 4-23



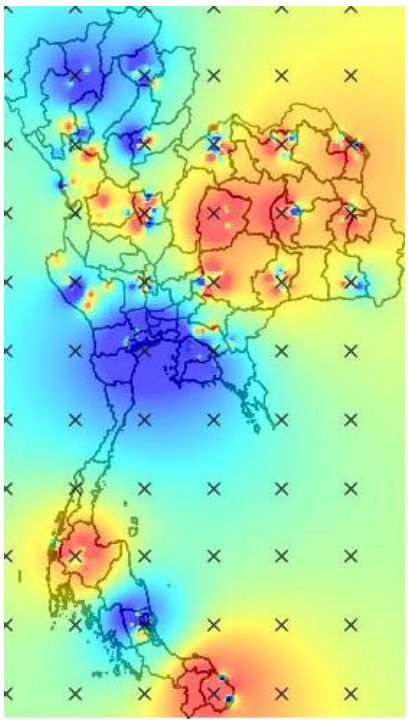
ภาพที่ 4-23 คะแนนเฉลี่ย O-NET ของประเทศไทยที่ได้จากวิธีการสุ่มแบบหลายขั้นตอน

ซึ่งสารสนเทศที่ได้นี้สามารถนำไปใช้อ้างอิงกับปัจจัยอื่น ๆ ในกระบวนการการศึกษาของประเทศไทย เช่น มาตรฐานและความเหมาะสมของข้อสอบ O-NET, แรงจูงใจของนักเรียนต่อการสอบ O-NET, คุณภาพและปริมาณของครู/ ครูอัตราจ้าง, งบประมาณ, นโยบายการศึกษาของกระทรวงศึกษา, เทคโนโลยีทางการศึกษา, ภัยพิบัติทางธรรมชาติ ฯลฯ ซึ่งสารสนเทศที่ได้จากการวิเคราะห์นี้จะถูกนำไปใช้ใช้ในการวางแผนและแก้ปัญหาที่เกิดขึ้นในระดับประเทศ

แต่ในการสร้างแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดจากวิธีการสุ่มแบบกริดหรือแบบกริด-ชั้นภูมินั้น ได้นำวิธีการประมาณค่าเชิงพื้นที่ (Spatial interpolation) มาประยุกต์ใช้โดยใช้จุดที่ทราบค่ามาทำการประมาณค่าตำแหน่งอื่น ๆ ในแผนที่ที่ไม่ทราบค่า/ ไม่ได้เก็บข้อมูล ดังนั้นแผนที่ฯ ที่เกิดขึ้นด้วยวิธีการประมาณค่าฯ จะแสดงข้อมูลคะแนนเฉลี่ย O-NET เป็นแบบค่าตัวแปรเชิงพื้นที่

การประมาณค่าเชิงพื้นที่สำหรับโปรแกรม QGIS มีฟังก์ชั่นประมาณค่าตัวแปรของแบบจำลองสามมิติอยู่ 2 วิธี คือ Inverse distance weighted (IDW) และ Triangulated irregular network (TIN) ซึ่ง TIN เป็นวิธีการประมาณค่าโดยใช้รูปทรงสามเหลี่ยม แต่การเก็บข้อมูลของงานวิจัยนี้ได้เก็บรวบรวมข้อมูลในลักษณะรูปสี่เหลี่ยมขนาดเท่ากันครอบคลุมทั้งแผนที่ ดังนั้นต้องใช้วิธี IDW ในการประมาณค่าตัวแปร ซึ่งวิธี IDW มีแนวคิดแบบค่าเฉลี่ยถ่วงน้ำหนักกล่าว คือ จุดกริดที่อยู่ใกล้กันจะมีความสำคัญหรืออิทธิพลของค่าตัวแปรมากกว่าจุดกริดที่อยู่ไกลออกไป

เมื่อใช้วิธีการสุ่มตัวอย่างแบบแบบกริดกับฐานข้อมูลของคะแนน O-NET ปีการศึกษา 2548 โดยใช้ขนาดกลุ่มตัวอย่างเท่ากับ 335 โรงเรียน จากจุดกริดทั้งหมด 22 จุดกริด ซึ่งกลุ่มตัวอย่างที่สุ่มได้มีค่าเฉลี่ยเท่ากับ 158.49 คะแนน เมื่อนำมาแสดงเป็นแผนที่ด้วยวิธี IDW ได้ผลการวิเคราะห์ตามภาพที่ 4-24



ภาพที่ 4-24 คะแนนเฉลี่ย O-NET ของประเทศไทยที่ได้จากวิธีการสุ่มแบบกริด

ซึ่งสารสนเทศที่ได้จากวิธีการสุ่มตัวอย่างแบบกริด (หรือแบบกริด-ชั้นภูมิ) นั้น มีความแตกต่างจากวิธีการสุ่มแบบหลายขั้นตอนอย่างชัดเจน เพราะวิธีการสุ่มแบบหลายขั้นตอนเป็นการประมาณค่าแบบภาพรวม (ประเทศ) แต่ในส่วนการสุ่มแบบกริดเป็นการประมาณค่าทั้งในระดับภาพรวมและในรายละเอียดที่ตำแหน่งต่าง ๆ (จังหวัด อำเภอ ตำบล) ดังนั้นเมื่อนำไปใช้อ้างอิงกับปัจจัยอื่น ๆ ในกระบวนการการศึกษาของประเทศไทยจะทำให้ทราบรายละเอียดในระดับท้องถิ่นและเข้าใจภาพรวมของประเทศ เช่น สถานการณ์ความไม่สงบในจังหวัดชายแดนใต้ เช่น การวางเพลิงโรงเรียน การลอบสังหารครู ฯลฯ ความเหลื่อมล้ำของโอกาสทางการศึกษาระหว่างนักเรียนในชนบทกับนักเรียนในเขตกรุงเทพและปริมณฑล เช่น การกวดวิชา ปัญหาโรงเรียนขนาดเล็ก ระดับการศึกษาและฐานะของผู้ปกครอง ฯลฯ

6. เปรียบเทียบข้อดีข้อเสียระหว่างวิธีการสุ่มแบบกริดและวิธีการสุ่มแบบกริด-ชั้นภูมิ

1. ความคลาดเคลื่อนในการวิจัย

วิธีการสุ่มแบบกริด-ชั้นภูมิเป็นวิธีการสุ่มแบบหนึ่งที่มีความถูกต้องเที่ยงตรงสูง เนื่องจากวิธีการนี้ได้นำสมาชิกทั้งหมดที่อยู่ภายในวงกลมการสุ่มมาแบ่งชั้นภูมิ ดังนั้นทุก ๆ สมาชิกมีโอกาสที่จะได้รับการสุ่ม เพื่อเป็นกลุ่มตัวอย่าง และการแบ่งชั้นภูมินี้ยังเป็นการบังคับให้ผู้วิจัยต้องเลือกสมาชิกในทุก ๆ ประชากรย่อย ดังนั้นกลุ่มตัวอย่างที่ถูกสุ่มออกมาจึงเป็นตัวแทนที่ดีของประชากร

แต่วิธีการสุ่มกลุ่มตัวอย่างแบบกริดจะเลือกกลุ่มตัวอย่างที่อยู่ใกล้ ๆ กับจุดกริดนั้น และถ้าจุดกริดนั้นอยู่ในพื้นที่ที่ประชากรมีการกระจายตัวเป็นแบบวิวิธพันธ์จะทำให้ข้อมูลที่สำรวจจากกลุ่มตัวอย่างมีความคลาดเคลื่อนสูง

2. ความสิ้นเปลืองทรัพยากร

การเก็บรวบรวมข้อมูลในการทำวิจัย พนักงานสำรวจจะต้องลงพื้นที่จริงเพื่อสัมภาษณ์ สังเกตหรือสอบถามกับกลุ่มตัวอย่างทำให้เกิดค่าใช้จ่ายต่าง ๆ เช่นค่าเดินทางค่าอาหารค่าที่พัก ฯลฯ นอกจากนี้ยังสิ้นเปลืองเวลาในการเดินทางสำรวจข้อมูลหรือต้องใช้พนักงานสำรวจหลายคนเพื่อแยกย้ายกันไปเก็บข้อมูลตามแต่ละพื้นที่ และภาพที่ 4-24 แสดงให้ว่าการสุ่มแบบกริด-ชั้นภูมิทำให้กลุ่มตัวอย่างกระจายตัวครอบคลุมแผนที่ประเทศไทย ซึ่งจะส่งผลในการเก็บข้อมูลเพราะการกระจายตัวแบบนี้จะเป็นการเพิ่มระยะทางและเวลาในการรวบรวมข้อมูล ทำให้สิ้นเปลืองทรัพยากร ซึ่งแตกต่างจากภาพที่ 4-20 ที่เป็นการสุ่มจุดกริดนั้น กลุ่มตัวอย่างที่ได้จะอยู่ในบริเวณเดียวกับจุดกริด ดังนั้นการสุ่มแบบกริดจะสูญเสียทรัพยากรน้อยกว่า

หมายเหตุ ในงานวิจัยได้ใช้ฐานข้อมูลฐานข้อมูลของคะแนน O-NET ปีการศึกษา 2548 (ศูนย์ปฏิบัติการ GPA-กระทรวงศึกษาธิการ) ซึ่งการเก็บข้อมูลจากแบบทดสอบจะใช้วิธีให้อาจารย์คุมสอบที่ศูนย์สอบ ดังนั้นค่าใช้จ่ายที่เกิดขึ้นคือค่าเบี้ยเลี้ยงในการคุมสอบ และค่าใช้จ่ายในเรื่องการจัดทำข้อสอบ (ค่าพิมพ์/ โรเนียว) กระดาษทด กระดาษคำตอบก็จะเท่ากัน นอกจากนี้ค่าใช้จ่ายในการเดินทางไปศูนย์สอบนั้น นักเรียนและผู้ปกครองต้องเป็นผู้ออกค่าใช้จ่ายเอง ดังนั้นวิธีการสุ่มทั้งสองวิธีจึงมีค่าใช้จ่ายพอ ๆ กัน

แต่ค่าใช้จ่ายในงานวิจัยนี้ไม่สามารถนำมาใช้เป็นบทสรุปได้ เนื่องจากยังมีสำรวจข้อมูลอีกหลายประเภทที่ผู้วิจัยจะต้องเสียค่าใช้จ่ายเอง เช่น การสำรวจสถานการณ์ระดับสติปัญญาเด็กนักเรียนไทย พ.ศ. 2554 โดยสถาบันราชานุกูล กรมสุขภาพจิต ซึ่งต้องสำรวจกลุ่มตัวอย่างเด็กนักเรียนจำนวน 72,780 คน จากทุกจังหวัดทั่วประเทศ และในการทดสอบนี้ไม่สามารถส่งเอกสารทางไปรษณีย์หรือจดหมายอิเล็กทรอนิกส์ไปให้กับกลุ่มตัวอย่างได้ เนื่องจากแบบทดสอบระดับ IQ ต้องใช้เจ้าหน้าที่ที่มีความเชี่ยวชาญในการกำกับดูแล (สถาบันราชานุกูล กรมสุขภาพจิต, 2554) ทำให้เกิดความสิ้นเปลืองทรัพยากรในการทำวิจัย (เช่น งบประมาณ ระยะเวลา และจำนวนเจ้าหน้าที่) ซึ่งงานวิจัยเชิงสำรวจแบบนี้จะขึ้นอยู่กับการกระจายตัวของกลุ่มตัวอย่างเป็นสำคัญ เนื่องจากพนักงานสำรวจข้อมูลต้องลงไปเก็บข้อมูลในพื้นที่จริง

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

วัตถุประสงค์ของการวิจัย

การวิจัย เรื่อง “แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด” มีวัตถุประสงค์เพื่อตรวจสอบความแม่นยำในการประมาณค่าและเปรียบเทียบประสิทธิภาพของการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ ซึ่งวิธีการสุ่มมีทั้งหมด 6 วิธี ได้แก่ 1) วิธีการสุ่มอย่างง่าย 2) วิธีการสุ่มแบบมีระบบ 3) วิธีการสุ่มแบบแบ่งชั้นภูมิ 4) วิธีการสุ่มแบบกลุ่ม 5) วิธีการสุ่มแบบหลายขั้นตอน และ 6) วิธีการสุ่มแบบกริด ซึ่งวัตถุประสงค์การวิจัยสามารถสามารถแบ่งตามขั้นตอนการวิจัยได้ 2 ส่วน คือ ส่วนที่ 1 เป็นการทดลองการสุ่มแบบกริดโดยอาศัยแบบจำลองที่สร้างขึ้น ซึ่งเป็นการทดสอบประสิทธิระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ (วิธีการสุ่มอย่างง่าย วิธีการสุ่มแบบมีระบบ วิธีการสุ่มแบบแบ่งชั้นภูมิ และวิธีการสุ่มแบบกลุ่ม) ในส่วน 2 เป็นกรณีศึกษาของการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548 เพื่อทดสอบว่าวิธีการสุ่มแบบกริดสามารถใช้งานกับฐานข้อมูลจริงได้ โดยทำการเปรียบเทียบประสิทธิและความแม่นยำในการประมาณค่าระหว่างการสุ่มแบบหลายขั้นตอน การสุ่มแบบกริดและการสุ่มแบบกริด-ชั้นภูมิ และเปรียบเทียบประสิทธิภาพของแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดจากวิธีการสุ่มทั้ง 3 วิธี

ประชากรและกลุ่มตัวอย่าง

ประชากรของการวิจัยมี 2 รูปแบบ คือประชากรของแบบจำลองที่สร้างขึ้น และประชากรจากข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ประชากรที่ใช้ในการศึกษามีดังนี้

1. ประชากรสำหรับกรณีแบบจำลองการสุ่ม คือ ค่าตัวแปรทุก ๆ ค่าที่อยู่ในตารางแบบจำลองนั้น ๆ ได้แก่ ประชากรของแบบจำลอง S0 เท่ากับ 81 ช่อง, ประชากรของแบบจำลอง S00 เท่ากับ 81 ช่อง, ประชากรของแบบจำลอง S1 เท่ากับ 81 ช่อง, ประชากรของแบบจำลอง S2 เท่ากับ 81 ช่อง, ประชากรของแบบจำลอง S3 เท่ากับ 144 ช่อง, ประชากรของแบบจำลอง S4 เท่ากับ 144 ช่อง, ประชากรของแบบจำลอง SI1 เท่ากับ 144 ช่อง, ประชากรของแบบจำลอง SI2 เท่ากับ 144 ช่อง และประชากรของแบบจำลอง SI3 เท่ากับ 195 ช่อง

2. ประชากรสำหรับกรณีศึกษา คือโรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 ทั้งหมด 2,584 โรงเรียน

เนื่องจากประชากรของการวิจัยมี 2 รูปแบบ ดังนั้นกลุ่มตัวอย่างที่ถูกสุ่มออกมาเพื่อศึกษาวิจัยต้องมี 2 รูปแบบด้วย คือกลุ่มตัวอย่างของแบบจำลองที่สร้างขึ้น และกลุ่มตัวอย่างจากฐานข้อมูลคะแนน O-NET ดังนี้

3. กลุ่มตัวอย่างสำหรับกรณีแบบจำลองการสุ่ม คือ กลุ่มตัวอย่างที่ถูกสุ่มมาจากแบบจำลองแผนที่ขนาดต่าง ๆ ได้แก่ กลุ่มตัวอย่างของแบบจำลอง S0 เท่ากับ 9 ช่อง, กลุ่มตัวอย่างของแบบจำลอง S00 เท่ากับ 9 ช่อง, กลุ่มตัวอย่างของแบบจำลอง S1 เท่ากับ 9 ช่อง, กลุ่มตัวอย่างของแบบจำลอง S2 เท่ากับ 9 ช่อง, กลุ่มตัวอย่างของแบบจำลอง S3 เท่ากับ 36 ช่อง, กลุ่มตัวอย่างของแบบจำลอง S4 เท่ากับ 36 ช่อง, กลุ่มตัวอย่างของแบบจำลอง SI1 เท่ากับ 36 ช่อง, กลุ่มตัวอย่างของแบบจำลอง SI2 เท่ากับ 36 ช่อง และกลุ่มตัวอย่างของแบบจำลอง SI3 เท่ากับ 36 ช่อง

4. กลุ่มตัวอย่างสำหรับกรณีศึกษาคือโรงเรียนที่มีนักเรียนชั้นมัธยมศึกษาปีที่ 6 เข้ารับการทดสอบทางการศึกษาแห่งชาติขั้นพื้นฐาน (O-NET) ปีการศึกษา 2548 จำนวน 335 โรงเรียน

ตัวแปรที่ใช้ในการวิจัย

การวิจัย เรื่อง “แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด” เป็นการวิจัยที่มุ่งทดสอบประสิทธิภาพของการสุ่มแบบกริด เมื่อเทียบกับวิธีการสุ่มรูปแบบอื่น ๆ โดยพิจารณาจากค่าสถิติที่กับค่าพารามิเตอร์ ถ้าค่าสถิติที่ได้มีความใกล้เคียงกับค่าพารามิเตอร์แสดงว่ากลุ่มตัวอย่างเป็นตัวแทนที่ดีของประชากร ดังนั้นตัวแปรอิสระของการวิจัย คือ วิธีการสุ่มแบบต่าง ๆ และตัวแปรตามของการวิจัย คือ ค่าสถิติ ซึ่งการวิจัยครั้งนี้แบ่งการทดลองได้เป็น 2 ขั้นตอน ดังนั้นตัวแปรการวิจัยจึงมี 2 รูปแบบ เช่นกัน ดังนี้

ขั้นตอนที่ 1: การทดลองการสุ่มแบบกริดโดยอาศัยแบบจำลองที่สร้างขึ้น ตัวแปรอิสระ คือ วิธีการสุ่มตัวอย่าง ซึ่งมี 5 วิธี ได้แก่ การสุ่มตัวอย่างแบบง่าย การสุ่มตัวอย่างแบบมีระบบ การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ การสุ่มตัวอย่างแบบกลุ่ม และการสุ่มตัวอย่างแบบกริด ตัวแปรตาม คือ ค่าสถิติที่ได้จากวิธีการสุ่ม ได้แก่ ค่าเฉลี่ยของความคลาดเคลื่อนและผลรวมของค่าความแตกต่างความแปรปรวน

ขั้นตอนที่ 2: การทดลองการสุ่มแบบกริดโดยอาศัยข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น ตัวแปรอิสระคือวิธีการสุ่มตัวอย่าง ซึ่งมี 3 วิธี ได้แก่ การสุ่มแบบกริด การสุ่มแบบกริด-ชั้นภูมิ การสุ่มแบบหลายขั้นตอน ตัวแปรตาม คือ ค่าสถิติที่ได้จากวิธีการสุ่ม ได้แก่ ค่าเฉลี่ยของความคลาดเคลื่อนและความแปรปรวนของกลุ่มตัวอย่าง

เครื่องมือที่ใช้ในการทดลอง

การวิจัยนี้เป็นการวิจัยเชิงทดลอง โดยการสร้างแบบจำลองและใช้ฐานข้อมูลทุติยภูมิในการตรวจสอบความแม่นยำและประสิทธิภาพของการสุ่มแบบกริด ดังนั้นเครื่องมือที่ใช้ในการวิจัยจึงมี 3 ประเภท คือ เครื่องคอมพิวเตอร์ โปรแกรมคอมพิวเตอร์ และข้อมูลที่ใช้ในการวิจัย

1. เครื่องคอมพิวเตอร์ CPU: Intel Core i5-6200U RAM: 4 GB DDR3L HDD: 1 TB 5400 RPM

2. ข้อมูลที่ใช้ในการวิจัย ประกอบด้วยฐานข้อมูลคะแนน O-NET จากศูนย์ปฏิบัติการ GPA สพฐ. โดยมีรายละเอียดที่เกี่ยวข้องกับโรงเรียน เช่น รายชื่อโรงเรียน รหัสโรงเรียน จังหวัด คะแนนเฉลี่ย คะแนนสูงสุด คะแนนต่ำสุด ฯลฯ และข้อมูลสารสนเทศภูมิศาสตร์ของประเทศไทย คือ ข้อมูลเชิงพื้นที่ (Spatial data) และข้อมูลเชิงคุณลักษณะ (Attribute data) โดยข้อมูลเชิงพื้นที่ประกอบด้วย ข้อมูลรูปแบบปิด (Polygon) คือ แผนที่ประเทศไทย ชั้นข้อมูลเชิงตำแหน่ง (Point) คือ พิกัดภูมิศาสตร์ของโรงเรียน ในส่วนข้อมูลเชิงคุณลักษณะเป็นข้อมูลที่ใช้อธิบายประกอบข้อมูลเชิงพื้นที่นั้น เช่น ค่าละติจูด ค่าลองจิจูด คะแนนเฉลี่ย ฯลฯ

3. โปรแกรมคอมพิวเตอร์ได้แก่ โปรแกรม Microsoft excel , โปรแกรม QGIS และโปรแกรมภาษา C ซึ่งโปรแกรม QGIS และ โปรแกรม Microsoft excel เป็นโปรแกรมที่ถูกนำมาใช้ในการแสดงผลค่าสถิติและค่าพารามิเตอร์ในรูปแบบแผนภาพ 2 มิติ นอกจากนี้โปรแกรม QGIS มีฟังก์ชั่นในการสุ่มตำแหน่งของจุดกริดและฟังก์ชั่นในการสร้างวงกลมการสุ่ม ในส่วนโปรแกรมที่ใช้ในสุ่มตัวอย่าง คือ โปรแกรมภาษา C (ขั้นตอนการวิจัยที่ 1) และ โปรแกรม Microsoft excel (ขั้นตอนการวิจัยที่ 2) โดยคำนวณหาค่าพารามิเตอร์จากประชากรและคำนวณหาค่าสถิติจากกลุ่มตัวอย่างในแต่ละวิธีการสุ่ม เพื่อเปรียบเทียบประสิทธิภาพของการสุ่มแบบต่าง ๆ

แบบจำลองการสุ่ม

แบบจำลองการสุ่มเป็นเครื่องมือสำคัญในการเปรียบเทียบประสิทธิภาพวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ ซึ่งแบบจำลองการสุ่ม คือ แผนที่แสดงค่าตัวแปรเชิงพื้นที่ โดยมีเส้นแนวนอนเป็นเส้นละติจูดและมีเส้นแนวตั้งเป็นเส้นลองจิจูด ดังนั้นพื้นที่ทั้งหมดจึงถูกแบ่งด้วยเส้นแนวตั้งและแนวนอนออกเป็นกริดเซลล์ ซึ่งแต่ละกริดเซลล์มีประชาชนอาศัยอยู่รวมกันเป็นสังคม ประชากรในแต่ละพื้นที่มีความเชื่อ ทัศนคติและความคิดเห็นต่อเรื่องใดเรื่องหนึ่งคล้ายกัน โดยกำหนดให้ค่าตัวแปรที่วัดได้จากสมาชิก (Element) ภายในแต่กริดเซลล์มีค่าเท่ากัน ซึ่งแบบจำลองการสุ่มมีทั้งหมด 9 แบบ ได้แก่ S0, S00, S1, S2, S3, S4, SI1, SI2 และ SI3 ซึ่งแต่ละแบบมีทิศทางและการจัดเรียงรูปแบบตัวแปรเชิงพื้นที่ที่ต่างกัน เนื่องจากถูกออกแบบและพัฒนาด้วยแนวคิดที่ต่างกัน

วิธีการสุ่มตัวอย่างแบบหลายขั้นตอน

จากข้อมูลการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548 มีโรงเรียนต่าง ๆ ทั่วประเทศ เป็นจำนวน 2,584 โรง โดยแบ่งเป็นตามแต่ละภาค ดังนี้ ภาคเหนือ 228 โรงเรียน คิดเป็นร้อยละ 8.82, ภาคกลาง 615 โรงเรียน คิดเป็นร้อยละ 23.8, ภาคตะวันออก 167 โรงเรียน คิดเป็นร้อยละ 6.46, ภาคตะวันออกเฉียงเหนือ 1,136 โรงเรียน คิดเป็นร้อยละ 43.96, ภาคตะวันตก 120 โรงเรียน คิดเป็นร้อยละ 4.64 และ ภาคใต้ 318 โรงเรียน คิดเป็นร้อยละ 12.31 ซึ่งการประมาณขนาดตัวอย่างโดยใช้ตาราง Krejcie และ Morgan นั้น ถ้าจำนวนประชากรเท่ากับ 2400 หน่วย ต้องใช้กลุ่มตัวอย่าง 331 หน่วย แต่ถ้าจำนวนประชากรเท่ากับ 2,600 หน่วย ต้องใช้กลุ่มตัวอย่าง 335 หน่วย ดังนั้นต้องใช้โรงเรียนจำนวน 335 โรงเรียน เป็นกลุ่มตัวอย่าง ซึ่งกลุ่มตัวอย่างขนาด 335 โรงเรียน สามารถแจกแจงตามสัดส่วน ในแต่ละภูมิภาค ดังนี้ ภาคเหนือ 30 โรงเรียน, ภาคกลาง 80 โรงเรียน, ภาคตะวันออก 22 โรงเรียน, ภาคตะวันออกเฉียงเหนือ 147 โรงเรียน, ภาคตะวันตก 15 โรงเรียน และ ภาคใต้ 41 โรงเรียน

หมายเหตุ การแบ่งส่วนภูมิภาคเป็นไปตามแนวทางคณะกรรมการภูมิศาสตร์แห่งชาติ (จัดทำโดยราชบัณฑิตยสภา)

การกำหนดจำนวนจังหวัดในแต่ละภาค คิดคำนวณตามกฎแห่งความชัดเจน (Rule of tumb) ซึ่งเป็นการกำหนดขนาดหนาดกลุ่มตัวอย่างแบบร้อยละ ในกรณีที่ประชากรน้อยกว่า 1,000 หน่วย ต้องใช้กลุ่มตัวอย่างขนาด 30% (Neuman, 1991, p. 221) ดังนั้นจำนวนจังหวัดร้อยละ 30 ของ 76 จังหวัด เท่ากับ 23 จังหวัด โดยคิดคำนวณเทียบจากร้อยละของจำนวน โรงเรียนตามแต่ละภูมิภาค ซึ่งมีรายละเอียดดังนี้ ภาคเหนือ 2 จังหวัด, ภาคกลาง 7 จังหวัด, ภาคตะวันออก 1 จังหวัด, ภาคตะวันออกเฉียงเหนือ 9 จังหวัด, ภาคตะวันตก 1 จังหวัด และ ภาคใต้ 3 จังหวัด

หมายเหตุ จำนวนจังหวัดที่อ้างถึงเป็นข้อมูลในอดีตปี พ.ศ. 2548 ณ เวลานั้นประเทศไทยมีจังหวัด 76 จังหวัด

เมื่อทำการสุ่มตัวอย่างจังหวัดจำนวน 23 จังหวัด ตามแต่ละภูมิภาคดังนี้ ภาคเหนือ ได้แก่ เชียงใหม่และอุตรดิตถ์, ภาคตะวันออกเฉียงเหนือ ได้แก่ อำนาจเจริญ ชัยภูมิ กาฬสินธุ์ เลยมหาสารคาม หนองบัวลำภู หนองคาย สกลนคร และยโสธร, ภาคกลาง ได้แก่ กำแพงเพชร ลพบุรี ปทุมธานี อยุธยา สุโขทัย สุพรรณบุรี และอุทัยธานี, ภาคตะวันออก คือ ปราจีนบุรี, ภาคตะวันตก คือ ตาก, และภาคใต้ ได้แก่ พัทลุง สตูล และตรัง แต่ผลการทดลองนี้เป็นผลลัพธ์ที่เกิดจากการสุ่มเพียงแค่ครั้งเดียว ซึ่งต้องทำการสุ่มอีกจนครบ 30 ครั้ง

วิธีการสุ่มตัวอย่างแบบกริด

วิธีการสุ่มแบบกริดมีขั้นตอนและวิธีการบางส่วนคล้ายกับวิธีการสุ่มแบบหลายขั้นตอน เพราะใช้วิธีการสุ่มแบบหลายขั้นตอนเป็นต้นแบบกล่าวคือวิธีการสุ่มแบบกริดจะต้องใช้กลุ่มตัวอย่างขนาด 335 โรงเรียน และใช้จุดกริดเท่ากับจำนวนจังหวัดที่ได้จากวิธีการสุ่มแบบหลายขั้นตอน (23 จังหวัด) ดังนั้นแผนการสุ่มตัวอย่างของวิธีการสุ่มแบบกริดต้องสุ่มตำแหน่งจุดกริดให้ได้ 23 จุดกริด โดยใช้โปรแกรม QGIS ในการสุ่มตำแหน่งบนแผนที่ประเทศไทย และเมื่อใช้โปรแกรม QGIS สุ่มตำแหน่งแล้ว พบว่ามีจุดกริดที่อยู่บนผืนแผ่นดินไทย มีจำนวนเท่ากับ 20 จุดกริด ซึ่งใกล้เคียงกับแผนการสุ่มที่วางไว้ ในเบื้องต้น แต่ยังขาดอีก 3 จุด ทำให้ต้องพิจารณาจุดกริดที่อยู่ติดกับชายแดนหรือในทะเล เพื่อนำมาเพิ่มให้ครบตามแผน ซึ่งได้ผลลัพธ์รวมเป็นจำนวนจุดกริดทั้งหมด 22 จุดกริด และโรงเรียนทั้งหมด 335 โรงเรียน

วิธีการสุ่มตัวอย่างแบบกริด-ชั้นภูมิ

วิธีการสุ่มแบบกริด-ชั้นภูมิมีขั้นตอนและวิธีการบางส่วนคล้ายกับวิธีการสุ่มแบบกริด เพราะใช้วิธีการสุ่มแบบกริดเป็นต้นแบบกล่าวคือใช้ตำแหน่งจุดกริดและขนาดกลุ่มตัวอย่างในแต่ละจุดกริดเช่นเดียวกัน แต่กลุ่มตัวอย่างที่ถูกสุ่มออกมาต่างกัน ซึ่งมีขั้นตอนในการสุ่ม คือ แบ่งโรงเรียนทั้งหมด 2,584 โรงเรียน ออกเป็น 4 ชั้นภูมิ (แบบควอไทล์) โดยใช้ตัวแปรขนาดโรงเรียนเป็นตัวเกณฑ์ในการแบ่งชั้นภูมิ ซึ่งได้ผลลัพธ์ คือ ชั้นภูมิที่ 1 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 1-28 คน ชั้นภูมิที่ 2 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 28-53 คน ชั้นภูมิที่ 3 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 53-111 คน และชั้นภูมิที่ 4 คือ กลุ่มโรงเรียนที่มีจำนวนนักเรียนเข้าสอบ O-NET อยู่ในช่วง 111-1615 คน ในขั้นตอนต่อไปคือการขยายรัศมีวงกลมของการสุ่มในแต่ละจุดกริดให้ครอบคลุมกลุ่มตัวอย่างทั้งหมด และทำการจัดสรรแบบสัดส่วน เพื่อทำให้ทราบว่าในแต่ละจุดกริดต้องใช้สุ่มโรงเรียนในแต่ละชั้นภูมิเท่าไหร่

การวิเคราะห์ข้อมูล

การศึกษาวิจัย เรื่อง “แบบจำลองการประมาณค่าเชิงพื้นที่ของการสุ่มแบบกริด” เป็นการทดลองเพื่อศึกษาตัวแปรค่าสถิติที่เปลี่ยนแปลงไปตามวิธีการสุ่มแบบต่าง ๆ และทำการเปรียบเทียบค่าสถิติที่ได้กับค่าพารามิเตอร์ เพื่อศึกษาเปรียบเทียบประสิทธิภาพของการสุ่มในแต่ละวิธี ซึ่งการวิจัยครั้งนี้แบ่งการทดลองได้เป็น 2 ขั้นตอน ดังนั้นการวิเคราะห์ข้อมูลจึงมี 2 รูปแบบ เช่นกัน ดังนี้

ขั้นตอนที่ 1

การเปรียบเทียบประสิทธิภาพของวิธีการสุ่มต้องใช้ค่าสถิติเป็นตัวตัดสิน เนื่องจากค่าสถิติเป็นเครื่องชี้วัดความเป็นตัวแทนที่ดีของประชากร ถ้าวิธีการสุ่มใด ๆ ได้กลุ่มตัวอย่างที่ให้ค่าสถิติได้ใกล้เคียงกับค่าพารามิเตอร์แสดงว่าวิธีการสุ่มนั้นมีประสิทธิภาพสูง โดยค่าสถิติที่ได้จากกลุ่มตัวอย่างมี 2 แบบ คือ ค่าเฉลี่ยและความแปรปรวน ในกรณีค่าเฉลี่ยต้องยึดถือ “ค่าเฉลี่ยของความคลาดเคลื่อน” เป็นหลัก ถ้าวิธีการสุ่มใดให้ “ค่าเฉลี่ยของความคลาดเคลื่อน” สูงแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่ไม่เป็นตัวแทนของประชากร ถ้าวิธีการสุ่มใดให้ “ค่าเฉลี่ยของความคลาดเคลื่อน” ต่ำแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร ซึ่งมีสูตรดังนี้

$$\bar{e} = \frac{\sum_{k=1}^{m} e_k}{m} \quad \text{โดยที่ } e = \left|\sum e_{i,j}\right| = |\mu - \bar{z}| \approx \left|\mu - \hat{\bar{z}}\right|$$

$\bar{e}$ คือ ค่าเฉลี่ยของความคลาดเคลื่อน

e_k คือ ค่าความคลาดเคลื่อน เมื่อทำการการสุ่มตัวอย่างครั้งที่ k

m คือ จำนวนครั้งของการสุ่มทั้งหมด

ในกรณีค่าความแปรปรวนต้องยึดถือ “ผลรวมของค่าความแตกต่างความแปรปรวน” เป็นหลัก ถ้าวิธีการสุ่มใดมี “ผลรวมของค่าความแตกต่างความแปรปรวน” สูงแสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่ไม่เป็นตัวแทนของประชากร ถ้าวิธีการสุ่มใดมี “ผลรวมของค่าความแตกต่างความแปรปรวน” ต่ำ แสดงว่าวิธีการสุ่มนั้นสุ่มได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร ซึ่งมีสูตรดังนี้

$$\Sigma|\Delta v| = |\Delta v|_1 + |\Delta v|_2 + |\Delta v|_3 + \ldots + |\Delta v|_m \quad \text{โดยที่ } |\Delta v| = \left|\sigma^2 - s^2\right|$$

$\Sigma|\Delta v|$ คือ ผลรวมของค่าความแตกต่างความแปรปรวน

m คือ จำนวนครั้งของการสุ่มทั้งหมด

ขั้นตอนที่ 2

การทดลองการสุ่มแบบกริดโดยอาศัยข้อมูลทุติยภูมิ O-NET ที่ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. จัดทำขึ้น โดยตัวแปรตามคือค่าสถิติที่ได้จากการสุ่มตัวอย่าง ได้แก่ ค่าเฉลี่ยของความคลาดเคลื่อนและความแปรปรวน

การสรุปผล

ตอนที่ 1 สรุปผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยแบบจำลองที่สร้างขึ้น

เมื่อใช้วิธีการสุ่มแบบกริดกับแบบจำลองที่ลักษณะประชากรมีสหสัมพันธ์เชิงพื้นที่เป็นศูนย์ (แบบจำลอง S0) ค่าเฉลี่ยของกลุ่มตัวอย่างที่ได้จะความแตกต่างกับค่าเฉลี่ยประชากรสูง (มีความคลาดเคลื่อนสูง) ถึงแม้ผู้วิจัยได้ปรับค่าสังเกต (ตำแหน่งแถวที่ 2 คอลัมน์ที่ 8 เปลี่ยนจาก 2 เป็น 8) ในแบบจำลองและเรียกชื่อแบบจำลองนี้ว่า “แบบจำลอง S00” เพื่อให้ค่าเฉลี่ยของกลุ่มตัวอย่างใกล้เคียงกับค่าเฉลี่ยของประชากร แต่ค่าความแตกต่างความแปรปรวนระหว่างกลุ่มตัวอย่างและประชากรยังมีค่ามาก จากผลการทดลองในแบบจำลอง S0 และ S00 แสดงให้เห็นว่าการสุ่มตัวอย่างแบบกริดเหมาะสมกับประชากรที่มีสหสัมพันธ์เชิงพื้นที่ทางบวก

ในส่วนแบบจำลอง S1, S2, S3 และ S4 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อใช้ในการเปรียบเทียบวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ สำหรับกรณีที่ประชากรมีสหสัมพันธ์เชิงพื้นที่ทางบวก จากผลการทดลอง พบว่า วิธีการสุ่มแบบแบ่งชั้นภูมิ วิธีการสุ่มแบบมีระบบและวิธีการสุ่มแบบกริด เป็นวิธีการที่มีประสิทธิภาพสูง แต่วิธีการสุ่มอย่างง่ายและวิธีการสุ่มแบบแบ่งกลุ่มมีประสิทธิต่ำในทุกแบบจำลอง นอกจากนี้วิธีการสุ่มแบบกริดมีค่าผลรวมของค่าความแตกต่างความแปรปรวนของแบบจำลองต่ำที่สุดในทุกแบบจำลอง เนื่องจากกลุ่มตัวอย่างที่ได้จากวิธีการสุ่มแบบกริดเป็นกลุ่มตัวอย่างเดิมในทุก ๆ ครั้ง ของการสุ่ม เพราะตำแหน่งของวิธีการสุ่มนี้เป็นตำแหน่งเดิม จึงทำให้ค่าความแตกต่างความแปรปรวนมีค่าคงที่ทุกครั้ง ซึ่งแตกต่างจากวิธีการสุ่มแบบอื่น ๆ ที่กลุ่มตัวอย่างมีการเปลี่ยนแปลงทุกครั้ง ทำให้ค่าความแตกต่างความแปรปรวนมีค่าสูง ๆ ต่ำ ๆ จนกระทั่งผลรวมของค่าความแตกต่างความแปรปรวนมีค่าสูง

แบบจำลอง SI1, SI2 และ SI3 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อทดสอบความแม่นยำในการประมาณค่าเชิงพื้นที่ จากผลการทดลอง พบว่า เมื่อสหสัมพันธ์เชิงพื้นที่มีค่าเป็นบวกจะสามารถทำการประมาณค่าเชิงพื้นที่ได้อย่างถูกต้องและแผนที่ระบบสารสนเทศภูมิศาสตร์ที่เกิดขึ้นมีความน่าเชื่อถือและสามารถนำไปใช้งานได้ โดยส่วนใหญ่ค่าจริงและค่าประมาณในแต่ละจุดกริดมีความใกล้เคียงกัน ถึงแม้ในบางตำแหน่งจะมีค่าความคลาดเคลื่อนร้อยละที่แตกต่างกันมาก แต่เมื่อแปลความหมายตามเกณฑ์พบว่ามีความผิดพลาดเพียงแค่ 1-2 ระดับ โดยแบบจำลอง SI1 มีค่าความคลาดเคลื่อนร้อยละมากที่สุดเท่ากับ 62.5% แต่มีการแปลความหมายความผิดพลาดไป 1 ระดับ สำหรับแบบจำลอง SI2 มีค่าความคลาดเคลื่อนร้อยละมากที่สุดเท่ากับ 85.31% ทำให้มีการแปลความหมายความผิดพลาดไป 2 ระดับ และแบบจำลอง SI3 มีคลาดเคลื่อนร้อยละมากที่สุดเท่ากับ 25% ซึ่งมีการแปลความหมายความผิดพลาดไปเพียงแค่ 1 ระดับ

ตอนที่ 2 สรุปผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างวิธีการสุ่มแบบกริด วิธีการสุ่มแบบกริด-ชั้นภูมิ และวิธีการสุ่มแบบหลายขั้นตอน โดยอาศัยฐานข้อมูลคะแนน O-NET

จากการศึกษาในหัวข้อ “การเปรียบเทียบประสิทธิภาพของวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ” เป็นวิธีการเปรียบเทียบประสิทธิภาพ โดยอาศัยแบบจำลองที่ตนเองสร้างขึ้น แต่ในการศึกษาในหัวข้อ “การเปรียบเทียบประสิทธิวิธีการสุ่มแบบกริด วิธีการสุ่มแบบกริด-ชั้นภูมิ และวิธีการสุ่มแบบหลายขั้นตอน” ผู้วิจัยนำฐานข้อมูลจะคะแนนการทดสอบการศึกษาระดับชาติขั้นพื้นฐาน ปีการศึกษา 2548 มาประยุกต์ใช้เป็นประชากรและกลุ่มตัวอย่าง เพื่อเป็นการทดสอบความตรงภายนอกของงานวิจัย

เพื่อเป็นการควบคุมตัวแปรแทรกซ้อนระหว่างแบบจำลองการสุ่มแบบหลายขั้นตอน แบบจำลองการสุ่มแบบกริด และแบบจำลองการสุ่มแบบกริด-ชั้นภูมิ จึงกำหนดขนาดกลุ่มตัวอย่างโรงเรียนเท่ากัน คือ 335 โรงเรียน และจำนวนจังหวัดที่ใช้ในแบบจำลอง ฯ มีจำนวนใกล้เคียงกันมาก โดยแบบจำลองการสุ่มแบบหลายขั้นตอนใช้จำนวนจังหวัด 23 จังหวัด และแบบจำลองการสุ่มแบบกริด (หรือการสุ่มแบบกริด-ชั้นภูมิ) ใช้จำนวนจุดกริด 22 จุด ซึ่งสาเหตุที่โมเดลทั้ง 2 มีจำนวนจังหวัดไม่เท่ากัน เนื่องจากภาคกลางของประเทศไทยเป็นภาคที่มีโรงเรียนจำนวนมากเมื่อเทียบกับพื้นที่ หรือมีความหนาแน่นของประชากรโรงเรียนมาก ซึ่งตามแผนการสุ่มต้องสุ่มจุดกริดออกมา 7 จุด และสุ่มโรงเรียนให้ได้ 80 โรงเรียน แต่โปรแกรม QGIS สุ่มจุดกริดได้เพียงแค่ 2 จุด อย่างไรก็ตามขนาดตัวอย่างที่ 23 จังหวัด และ 22 จุดกริด ถือว่ามีความใกล้เคียงกัน

การสุ่มตัวอย่างโรงเรียนต่าง ๆ ทั่วประเทศจำนวน 335 โรงเรียน เพื่อศึกษาคะแนนสอบ O-NET ของนักเรียนชั้นมัธยมศึกษาปีที่ 6 พบว่า วิธีการสุ่มแบบหลายขั้นตอนมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.52 คะแนนและมีความแปรปรวนเท่ากับ 5.323 วิธีการสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 3.18 คะแนน และไม่มีความแปรปรวน และวิธีการสุ่มแบบกริด-ชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.05 คะแนน และความแปรปรวนเท่ากับ 0.202

จากผลการทดลอง พบว่า วิธีการสุ่มทั้ง 3 วิธี มีความแม่นยำที่สามารถยอมรับได้ เนื่องจากวิธีการที่มีความแม่นยำต่ำที่สุดมีค่าค่าเฉลี่ยความคลาดเคลื่อนเพียงแค่ 3.18 คะแนน จากค่าพารามิเตอร์ 161.67 คะแนน ซึ่งวิธีการสุ่มแบบกริดเป็นวิธีที่มีความแม่นยำต่ำกว่าวิธีการสุ่มแบบหลายขั้นตอนและวิธีการสุ่มแบบกริดชั้นภูมิ เนื่องจากในบางพื้นที่มีสหสัมพันธ์เชิงพื้นที่ต่ำ (ใกล้เคียง 0) ดังนั้นกลุ่มตัวอย่างที่ถูกสุ่มออกมาจึงเป็นกลุ่มตัวอย่างที่อยู่ใกล้กับตำแหน่งจุดกริด ซึ่งเป็นประชากรกลุ่มย่อยที่ไม่ใช่ตัวแทนประชากรที่ดี ในส่วนวิธีการสุ่มแบบหลายขั้นตอนเป็นวิธีการสุ่มที่ให้โอกาสกับสมาชิกทุกตัวในประชากรได้รับการสุ่มออกมา ดังนั้นในบางครั้งอาจสุ่มได้โรงเรียนที่มีคะแนนเฉลี่ย O-NET สูงเป็นจำนวนมาก แต่ในบางครั้งอาจสุ่มได้โรงเรียนที่คะแนนเฉลี่ย O-NET

ต่ำเป็นจำนวนมาก ทำให้เกิดความคลาดเคลื่อนและมีความแปรปรวนสูง (5.323) ซึ่งแตกต่างจากวิธีการสุ่มแบบกริด-ชั้นภูมิ ที่มีสมาชิกภายในกลุ่มตัวอย่างมาจากต่างชั้นภูมิ ทำให้มีความหลากหลายและมาจากทุกกลุ่มก้อนของประชากร หรือกลุ่มตัวอย่างเป็นตัวแทนที่ดีของประชากร ทำให้วิธีการสุ่มแบบกริด-ชั้นภูมิเป็นวิธีการสุ่มที่มีความแม่นยำที่สุด

การอภิปรายผล

ตอนที่ 1 อภิปรายผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริดกับการสุ่มแบบอื่น ๆ โดยอาศัยแบบจำลองที่สร้างขึ้น

แบบจำลอง S0 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อทดสอบประสิทธิภาพวิธีการสุ่มแบบกริดในกรณีที่สหสัมพันธ์เชิงพื้นที่เป็น 0 จากผลการทดลองพบว่าวิธีการสุ่มแบบกริดมีค่าเฉลี่ยฯ มากที่สุด คือ 0.654 และมีค่าผลรวมฯ มากที่สุด คือ 3,297.13 ดังนั้นวิธีการสุ่มแบบกริดมีประสิทธิภาพต่ำที่สุด ซึ่งผลการทดลองนี้เป็นไปตามข้อตกลงเบื้องต้นของวิธีการสุ่มแบบกริดในข้อที่ 2

แบบจำลอง S00 แบบจำลองพัฒนามาจากแบบจำลอง S0 โดยปรับค่าสังเกตที่ตำแหน่งแถวที่ 2 คอลัมน์ที่ 8 เปลี่ยนจาก 2 เป็น 8 เพื่อทำให้ค่าเฉลี่ยของกลุ่มตัวอย่างที่สุ่มจากแบบจำลอง S00 ใกล้เคียงกับค่าเฉลี่ยของประชากร แต่เมื่อพิจารณาความแปรปรวนของการสุ่มแบบกริด พบว่าค่าผลรวมฯ ของวิธีการสุ่มแบบกริดยังคงมีค่ามากกว่าวิธีการสุ่มแบบอื่น ๆ คือ 3,297.13 ซึ่งผลการทดลองของแบบจำลอง S0 และ S00 แสดงให้เห็นว่าการสุ่มตัวอย่างแบบกริดเหมาะสมกับประชากรที่มีสหสัมพันธ์เชิงพื้นที่ทางบวกเท่านั้น

แบบจำลอง S1 เป็นแบบจำลองที่ถูกสร้างขึ้น เพื่อทดสอบการสุ่มแบบชั้นภูมิ ในกรณีที่มีบางชั้นภูมิที่มีความแปรปรวนภายในสูง โดยชั้นภูมิที่ 2 และ 3 มีความแปรปรวนภายในเท่ากับ 0.225 และมีค่ามากกว่าความแปรปรวนภายในของชั้นภูมิที่ 1 (0.194) และชั้นภูมิที่ 4 (0.183) แต่ตามหลักการและแนวคิดของการแบ่งชั้นภูมิ คือ การแบ่งชั้นภูมิต้องมีความแปรปรวนภายในต่ำ ดังนั้นเมื่อใช้วิธีการสุ่มกับแบบจำลอง S1 ทำให้ประสิทธิภาพของวิธีการสุ่มแบบชั้นภูมิลดลง

จากผลการทดลองแบบจำลอง S1 วิธีการสุ่มแบบชั้นภูมิมีประสิทธิภาพเป็นอันดับที่ 2 (การสุ่มแบบกริดมีประสิทธิภาพเป็นอันดับที่ 1) ซึ่งมีสาเหตุมาจากการแบ่งประชากรเป็นประชากรย่อมปฏิบัติได้ยาก เนื่องจากขาดขอบเขตที่ชัดเจน (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131)

แบบจำลอง S2 เป็นแบบจำลองที่ใช้ทดสอบการสุ่มแบบมีระบบ ในกรณีที่ธรรมชาติของประชากรมีการจัดเรียงตัวของค่าสังเกตกับระยะห่างของการสุ่มตรงกันพอดี จากผลการทดลองพบว่า วิธีการสุ่มแบบมีระบบมีค่าเฉลี่ยฯ มากที่สุด โดยมีค่าอยู่ในช่วง 2.210-2.373 และมีค่าผลรวมฯ

มากที่สุด โดยมีค่าอยู่ในช่วง 7,900.29-8,252.66 ดังนั้นวิธีการสุ่มแบบมีระบบจึงมีประสิทธิภาพต่ำที่สุดสำหรับการทดลองในแบบจำลอง S2 ซึ่งผลการทดลองนี้สอดคล้องกับ ศิริชัย กาญจนวาสี (2545, หน้า 123-131) ที่กล่าวถึงจุดด้อยของวิธีการสุ่มแบบมีระบบว่า “ถ้าหน่วย ถ้าบัญชีรายชื่อของประชากรจัดเรียงอย่างเป็นระบบ หรือสมาชิกของประชากรมีการเปลี่ยนแปลงขึ้นลงเป็นวงจรหรือช่วง ทำให้เกิดความลำเอียงในการสุ่ม”

แบบจำลอง S3 เป็นแบบเป็นแบบจำลองที่ถูกออกแบบให้แผนที่มีขนาดใหญ่ (144 ช่อง) กว่าแบบจำลองอื่น ๆ เพื่อศึกษาผลกระทบที่เกิดขึ้นในกรณีที่มีการเปลี่ยนแปลงขนาดของแผนที่ จากผลการทดลองพบว่า การสุ่มแบบชั้นภูมิมีประสิทธิภาพเป็นอันดับ 1, การสุ่มแบบกริดและการสุ่มแบบมีระบบมีประสิทธิภาพเป็นอันดับ 2 และการสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพเป็นอันดับ 3

จากผลการทดลองทั้งหมดของแบบจำลอง S3 มีความสอดคล้องตามทฤษฎีทุกประการดังนี้ 1) เหตุผลที่วิธีการสุ่มแบบชั้นภูมิเป็นวิธีที่มีประสิทธิภาพสูงที่สุดสอดคล้องกับแนวคิดของศิริชัย กาญจนวาสี (2545, หน้า 123-131) พบว่า วิธีการสุ่มแบบชั้นภูมิมีประสิทธิภาพสูงในเชิงการวิเคราะห์ทางสถิติ และสอดคล้องกับแนวคิดของสมบัติ ท้ายเรือคำ (2551, หน้า 55) พบว่าวิธีการสุ่มแบบชั้นภูมิช่วยลดความคลาดเคลื่อนในการประมาณค่าพารามิเตอร์ของกลุ่มประชากรได้มากกว่าใช้กลุ่มตัวอย่างที่ได้จากการสุ่มแบบอย่างง่าย 2) เหตุผลที่วิธีการสุ่มแบบมีระบบมีประสิทธิภาพเป็นอันดับ 2 เพราะวิธีการสุ่มแบบมีระบบเป็นวิธีที่มีประสิทธิภาพสูงกว่าวิธีการสุ่มอย่างง่าย ซึ่งสอดคล้องกับแนวคิดของศิริชัย กาญจนวาสี (2545, หน้า 123-131) พบว่า ถ้าประชากรจัดเรียงไว้อย่างสุ่มวิธีการสุ่มแบบมีระบบจะมีประสิทธิภาพสูงกว่าวิธีสุ่มแบบง่าย 3) เหตุผลที่วิธีการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพเป็นอันดับสุดท้าย เพราะวิธีนี้มีประสิทธิภาพต่ำกว่าวิธีเลือกกลุ่มตัวอย่างแบบใช้ความน่าจะเป็นวิธีอื่น ๆ (สุชาดา บวรกิติวงศ์, 2548, หน้า 119; สมบัติ ท้ายเรือคำ, 2551, หน้า 55) 4) ขนาดของแบบจำลอง S3 ที่ใหญ่กว่าแบบจำลองอื่น ๆ ไม่มีผลกระทบต่อการทดลอง ผลการทดลองของแบบจำลองนี้เป็นการยืนยันผลการทดลองแบบจำลองขนาดเล็ก (81 ช่อง) กล่าวคือ ค่าสถิติที่ได้จากกลุ่มตัวอย่างจะใกล้เคียงกับค่าพารามิเตอร์มากน้อยเพียงใดขึ้นอยู่กับวิธีการสุ่มและการกระจายตัวของค่าตัวแปรบนแผนที่/ สหสัมพันธ์เชิงพื้นที่

แบบจำลอง S4 เป็นแบบจำลองที่ถูกพัฒนาจากแบบจำลอง S3 โดยออกแบบให้แผนที่มีค่าตัวแปรที่หลากหลาย จากผลการทดลอง พบว่า การสุ่มแบบชั้นภูมิมีประสิทธิภาพเป็นอับดับที่ 1 การสุ่มแบบกริดมีประสิทธิภาพเป็นอันดับที่ 2 การสุ่มแบบมีระบบมีประสิทธิภาพเป็นอันดับที่ 3 การสุ่มอย่างง่ายและการสุ่มแบบแบ่งกลุ่มมีประสิทธิภาพเป็นอันดับสุดท้าย จากผลการทดลองของแบบจำลอง S4 ใกล้เคียงกับแบบจำลอง S3 เกือบทุกประการ ยกเว้นกรณีที่การสุ่มแบบกริด

เป็นอันดับที่ 2 และการสุ่มแบบมีระบบเป็นอันดับที่ 3 แต่ในแบบจำลอง S3 นั้น วิธีการสุ่มทั้ง 2 วิธี มีประสิทธิภาพเท่ากัน ซึ่งสาเหตุที่ทำให้ผลการทดลองแตกต่างกัน เนื่องจากค่าตัวแปรของแบบจำลองของแบบจำลอง S4 มีความหลากหลายมากกว่า โดยมีค่าอยู่ในช่วง 0-100 คะแนน แต่ค่าตัวแปรของแบบจำลอง S3 มีค่าระหว่าง 1-7 ระดับ ดังนั้นผลลัพธ์ที่ได้จากแบบจำลอง S4 จึงแสดงให้เห็นถึงความแตกต่างด้านประสิทธิภาพระหว่างการสุ่มแบบกริดและการสุ่มแบบมีระบบ

จากผลการทดลองของแบบจำลอง S4 เป็นการยืนยันแนวคิดและทฤษฎีของข้อเสียของวิธีการสุ่มแบบมีระบบ เพราะกลุ่มตัวอย่างที่ได้จากวิธีการสุ่มแบบมีระบบ ในแต่ละครั้งจะมีความแตกต่างกันและไม่ได้เป็นตัวแทนที่ดีของกลุ่มประชากร (สมบัติ ท้ายเรือคำ, 2551, หน้า 54)

แบบจำลอง SI1, SI2 และ SI3 เป็นแบบจำลองมีวัตถุประสงค์ ในการสร้างขึ้นมาแตกต่างจากแบบจำลอง S0, S00, S1, S2, S3 และ S4 เพราะแบบจำลอง S0-S4 เป็นแบบจำลองที่ถูกสร้างเพื่อใช้เปรียบเทียบวิธีการสุ่มแบบกริดกับวิธีการสุ่มแบบอื่น ๆ แต่แบบจำลอง SI1 – SI3 เป็นแบบจำลองที่ใช้ในการทดสอบทฤษฎีการประมาณค่าเชิงพื้นที่ของ Donald Shepard โดยแบบจำลองทั้งหมดมีค่าสหสัมพันธ์เชิงพื้นที่ทางบวก จากผลการทดลองพบว่า เมื่อสหสัมพันธ์เชิงพื้นที่มีค่าเป็นบวกนั้น วิธีประมาณค่าเชิงพื้นที่สามารถทำการประมาณค่าได้ใกล้เคียงกับความเป็นจริง ดังนั้นผลการทดลองที่เกิดขึ้นเป็นไปตามทฤษฎีของ Shepard

ตอนที่ 2 อภิปรายผลการวิเคราะห์ความแม่นยำและประสิทธิภาพระหว่างการสุ่มแบบกริด การสุ่มแบบกริด-ชั้นภูมิ และการสุ่มแบบหลายขั้นตอน โดยอาศัยฐานข้อมูลคะแนน O-NET

ผลลัพธ์ที่ได้จากวิธีการสุ่มแบบหลายขั้นตอนมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.52 คะแนนและมีความแปรปรวนเท่ากับ 5.323 วิธีการสุ่มแบบกริดมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 3.18 คะแนน และไม่มีความแปรปรวน และวิธีการสุ่มแบบกริด-ชั้นภูมิมีค่าเฉลี่ยความคลาดเคลื่อนเท่ากับ 2.05 คะแนน และความแปรปรวนเท่ากับ 0.202 จากผลการทดลองพบว่าวิธีการสุ่มแบบกริด-ชั้นภูมิให้ค่าสถิติใกล้เคียงค่าพารามิเตอร์มากกว่าวิธีการสุ่มแบบกริดและวิธีการสุ่มแบบหลายขั้นตอน สรุปได้ว่าวิธีการสุ่มแบบกริด-ชั้นภูมิเป็นวิธีการที่มีประสิทธิภาพสูงที่สุด

สาเหตุที่ค่าสถิติวิธีการสุ่มแบบกริดมีค่าแตกต่างจากค่าพารามิเตอร์มาก เนื่องจากความแม่นยำของวิธีการสุ่มแบบกริดขึ้นอยู่กับความหนาแน่นของการสุ่ม (Sampling density) แต่ในการวิจัยครั้งนี้ได้กำหนดระยะห่างระหว่างจุดกริดมากเกินไป คือ 154.27 กิโลเมตร (เนื่องจากเป็นการควบคุมตัวแปรระหว่างจำนวนจังหวัดและจำนวนจุดกริดให้ใกล้เคียงกัน) นอกจากนั้นพื้นที่ส่วนใหญ่ของประเทศเป็นแบบวิวิธพันธ์ โดยโรงเรียนที่มีคะแนนเฉลี่ย ONET สูง อยู่รวมปะปนกับ

โรงเรียนที่มีคะแนนเฉลี่ย ONET ต่ำ เมื่อใช้จุดกริดเป็นศูนย์กลางวงกลมการสุ่มตัวอย่าง และเลือกกลุ่มตัวอย่างที่อยู่ใกล้ ๆ กับจุดกริด อาจทำให้กลุ่มตัวอย่างที่ได้จุดกริดนั้นอยู่ในตำแหน่งประชากรกลุ่มย่อย (Subpopulations) ซึ่งไม่ได้เป็นตัวแทนที่ดีของประชากร ทำให้ได้ค่าสถิติที่ได้จากกลุ่มตัวอย่างมีความคลาดเคลื่อนจากค่าพารามิเตอร์ ซึ่งผลการทดลองนี้สอดคล้องกับการวิจัยของ Dong-Sheng et al. (2011) และ Majasalmi et al. (2012) โดยงานวิจัยทั้ง 2 เรื่อง เป็นการศึกษาความสัมพันธ์ระหว่างความหนาแน่นของการสุ่ม กับค่าความคลาดเคลื่อนมาตรฐานสัมพัทธ์ (Coefficient of variation) และให้ผลสรุปตรงกันว่า เมื่อใช้จำนวนจุดกริดน้อยลงจะทำให้ค่าความคลาดเคลื่อนมีค่ามากขึ้น ดังนั้นผลการทดลองที่ได้มีความสอดคล้องกับหลักการและเหตุผล

สาเหตุที่วิธีการสุ่มแบบหลายขั้นตอนมีประสิทธิภาพต่ำ เนื่องจากวิธีการสุ่มแบบหลายขั้นตอนในการวิจัยครั้งนี้เป็นการสุ่มแบบแบ่งกลุ่ม 2 ขั้นตอน กล่าวคือ ขั้นตอนที่ 1 คือ ใช้วิธีการสุ่มแบบแบ่งกลุ่ม โดยแบ่งประเทศไทยออกเป็น 6 ภาค และในแต่ละภาคให้สุ่มจังหวัดออกมา และขั้นตอนที่ 2 คือ สุ่มกลุ่มตัวอย่างโรงเรียนจากโรงเรียนทั้งหมดในแต่ละจังหวัด ดังนั้นวิธีการสุ่มแบบหลายขั้นตอนในการวิจัยครั้งนี้จึงมีประสิทธิภาพเทียบเท่ากับวิธีการสุ่มแบบแบ่งกลุ่ม 2 ขั้นตอน และตามหลักการ/ ทฤษฎีของวิธีการสุ่มแบบแบ่งกลุ่มได้อธิบายไว้ว่า “วิธีนี้มีประสิทธิภาพต่ำกว่าวิธีเลือกกลุ่มตัวอย่างแบบใช้ความน่าจะเป็นวิธีอื่น ๆ” (สุชาดา บวรกิติวงศ์, 2548, หน้า 119; สมบัติ ท้ายเรือคำ, 2551, หน้า 55) ดังนั้นผลการทดลองที่ได้มีความสอดคล้องกับหลักการและเหตุผล

สาเหตุที่วิธีการสุ่มแบบกริด-ชั้นภูมิมีประสิทธิภาพสูง เนื่องจากพื้นที่ของประเทศไทยมีการกระจายตัวของโรงเรียนที่มีคะแนน O-NET แบบวิวิธพันธ์ ดังนั้นการสุ่มให้ได้กลุ่มตัวอย่างที่มาจากทุกกลุ่มก้อนทุกภาคส่วนของประชากรต้องนำวิธีการสุ่มแบบชั้นภูมิมาประยุกต์ใช้ โดยการแบ่งประชากรเป็นชั้นภูมิตามค่าตัวแปรขนาดโรงเรียน จากนั้นสุ่มกลุ่มตัวอย่างในแต่ละชั้นภูมิด้วยวิธีจัดสรรตัวอย่างแบบสัดส่วน ซึ่งจุดเด่นของวิธีการสุ่มแบบชั้นภูมิมีหลายด้าน ได้แก่ เป็นวิธีการที่สามารถควบคุมตัวแปรแทรกซ้อนได้ดี, กลุ่มตัวอย่างที่ได้มีความเป็นตัวแทนประชากรย่อย และมีประสิทธิภาพสูงในเชิงการวิเคราะห์ทางสถิติ (ศิริชัย กาญจนวาสี, 2545, หน้า 123-131) ดังนั้นผลการทดลองที่ได้มีความสอดคล้องกับหลักการและเหตุผล

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

เมื่อวันที่ 6 พฤศจิกายน พ.ศ. 2555 นายชินภัทร ภูมิรัตน เลขาธิการคณะกรรมการการศึกษาขั้นพื้นฐาน (กพฐ.) กล่าวว่า สพฐ.ตัดสินใจเลื่อนแผนการจัดสอบ National Test หรือ NT ทุกระดับชั้น

ตามนโยบายของนายสุชาติ ธาดาธำรงเวช อดีต รมว.ศึกษาธิการ ออกไปก่อน เนื่องจากการจัดสอบ NT ทุกระดับชั้นนั้น เป็นงานในสเกลที่ใหญ่มาก ใช้งบประมาณรองรับถึงประมาณ 280 ล้านบาท (ศูนย์ข้อมูล & ข่าวสืบสวนเพื่อสิทธิพลเมือง (TCIJ), 2555) ด้วยปัญหาดังกล่าวทำให้ ณ ปัจจุบัน สพฐ. มีการจัดสอบ NT เฉพาะนักเรียนระดับ ป.3 และ ป.6 เท่านั้น

ตามความเห็นของผู้วิจัย ในการจัดสอบ NT ควรจัดสอบทุกระดับชั้น เนื่องจากการทดสอบในระดับ ป.3 และ ป.6 เป็นการประเมินผลผลิตหรือผลลัพธ์ตามหลักการประเมินโครงการ ถ้าผลการประเมินได้บทสรุปว่าการเรียนการสอนในระดับช่วงชั้น ป.1-ป.3 ไม่ดี ดังนั้น สพฐ. และโรงเรียนจะแก้ไขผลผลิตนี้ (นักเรียนที่สำเร็จการศึกษาชั้นปีที่ 3) ไม่ได้ แต่ถ้ามีการประเมินทุกระดับชั้นจะเปรียบได้กับการประเมินกระบวนการ ทำให้มีเวลามากพอในการแก้ไขและปรับปรุงก่อนที่นักเรียนจะสำเร็จการศึกษา แต่การจัดสอบทุกระดับชั้นย่อมต้องใช้งบประมาณที่สูง ดังนั้น วิธีการที่ดีที่สุด คือ การสุ่มตัวอย่างให้ได้ตัวแทนที่ดีของประชากร ซึ่งสถาบันการทดสอบการศึกษาระดับนานาชาติได้ใช้วิธีการสุ่มตัวอย่างนักเรียนในการทดสอบความรู้ ยกตัวอย่างเช่น โครงการการประเมินผลนักเรียนนานาชาติ (PISA) และโครงการศึกษาแนวโน้มการจัดการศึกษาคณิตศาสตร์และวิทยาศาสตร์ (TIMSS) เป็นต้น โดยใช้ขนาดกลุ่มตัวอย่างไม่เกิน 10,000 คน (ขนาดกลุ่มตัวอย่างของ PISA 2015 เท่ากับ 8,249 และขนาดกลุ่มตัวอย่างของ TIMSS 2015 เท่ากับ6,482 คน) ซึ่งจะทำให้ สพฐ. และประเทศชาติประหยัดงบประมาณได้มากกว่าการจัดนักเรียนทุกคนในระดับ ป.3 และ ป.6 เข้าสอบ

การที่ สพฐ. กำหนดให้นักเรียนในระดับ ป.3 และ ป.6 เข้าสอบ NT เนื่องจากไม่ต้องการสุ่มตัวอย่าง โดยคาดหวังว่าผลลัพธ์ที่ได้คือค่าพารามิเตอร์จากประชากรที่ไม่มีความคลาดเคลื่อน แต่ปัญหาใหม่ที่เกิดขึ้นตามมาคือความคลาดเคลื่อนที่ไม่เกี่ยวข้องกับการสุ่มตัวอย่าง (Non-sampling errors) และปัญหาการสุ่มตัวอย่างที่ล้นเกินไปหรือมากเกินไป (Oversampling) ซึ่งผู้วิจัยได้อธิบายไว้แล้วในหัวข้อ “ขนาดตัวอย่างมีความสำคัญน้อยกว่าความเป็นตัวแทนที่ดีของประชากร” ยกตัวอย่าง เช่น นักเรียนไม่ตั้งใจทำข้อสอบเพราะเป็นการสอบที่ไม่มีผลต่ออนาคตและชีวิต การเฉลยคำตอบในห้องสอบเพื่อให้คะแนนโรงเรียนออกมาดี เป็นต้น ดังนั้นวิธีการที่ดีสุด คือ กลุ่มตัวอย่างที่เป็นอาสาสมัครในการสอบ โดยไม่ต้องบังคับ และกลุ่มตัวอย่างนี้ให้คัดเลือกจากผู้มีระดับสติปัญญาในระดับเกณฑ์เฉลี่ยของโรงเรียนนั้น ๆ นอกจากนั้น สพฐ. ยังเคยจัดสอบ LAS (Local assessment system) (ปัจจุบันยกเลิกแล้ว) ซึ่งเป็นการประเมินคุณภาพการศึกษาขั้นพื้นฐาน ระดับเขตพื้นที่การศึกษา สำหรับนักเรียนระดับชั้น ป.2, ป.5, ม.2 และ ม.5 ของทุกโรงเรียนในเขตพื้นที่การศึกษา เพื่อให้ทราบถึงคุณภาพการศึกษาในระดับท้องถิ่น ซึ่งมีแนวคิดเช่นเดียวกับการสุ่มแบบกริด/ การสุ่มแบบกริด-ชั้นภูมิ ในเรื่องการให้สารสนเทศในระดับท้องถิ่น ดังนั้นการสุ่มแบบกริด/ การสุ่มแบบกริด-ชั้นภูมิจึงมีความเหมาะสมที่จะนำมาใช้กับการสอบ NT

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

ควรมีการศึกษาเปรียบเทียบประสิทธิภาพวิธีการสุ่มแบบหลายขั้นตอน วิธีการการสุ่มแบบกริด และวิธีการสุ่มแบบกริด-ชั้นภูมิ ภายใต้ผลกระทบจากเงื่อนไขอื่น ๆ ดังนี้

1. เงื่อนไขด้านเนื้อหาของข้อมูล เนื่องจากการวิจัยครั้งนี้เป็นการวิจัยที่นำฐานข้อมูลคะแนน O-NET ปี พ.ศ. 2548 มาใช้ในการทดสอบ เพื่อเปรียบเทียบประสิทธิภาพ ดังนั้นในการทำวิจัยครั้งต่อไปควรใช้ฐานข้อมูลประเภทอื่น ๆ เช่น ฐานข้อมูลผู้สูงอายุ ฐานข้อมูลแรงงานต่างด้าว เป็นต้น เพื่อเป็นการทดสอบประสิทธิภาพของการสุ่มแบบกริด เมื่อค่าตัวแปรเชิงพื้นที่มีการกระจายตัวในรูปแบบอื่น ๆ

2. เงื่อนไขด้านความละเอียดของตารางกริด เนื่องจากการวิจัยในครั้งนี้เป็นการทดสอบการสุ่มตัวอย่างด้วยแผนที่ประเทศไทย โดยใช้จำนวนจุดกริดเท่ากับ 22 จุด ดังนั้นในการทำวิจัยครั้งต่อไป ควรใช้แผนที่ประเภทอื่น ๆ อาจเป็นแผนที่ที่มีขนาดใหญ่มากขึ้น เช่น แผนที่อาเซียน แผนที่เอเชีย เป็นต้น หรือแผนที่ที่มีขนาดเล็กลง เช่น แผนที่ภูมิภาคของประเทศไทย แผนที่จังหวัด เป็นต้น แต่ถ้าใช้แผนที่ประเทศไทยควรกำหนดจำนวนจุดกริดให้มากขึ้น (> 22 จุดกริด) หรือน้อยลง (< 22 จุดกริด) เพื่อศึกษาประสิทธิภาพการประมาณค่าเชิงพื้นที่ของวิธีการสุ่มแบบกริด นอกจากนั้นจากผลประเมินของประเทศไทยตามโครงการประเมินผลนักเรียนร่วมกับนานาชาติประจำปี 2558 (PISA 2015) โดยการเก็บข้อมูลจากนักเรียนกลุ่มตัวอย่างอายุ 15 ปี จำนวน 8,249 คน ใน 273 โรงเรียนทุกสังกัดการศึกษา เมื่อช่วงเดือนสิงหาคม พ.ศ.2558 โดยใช้แบบทดสอบประเมินความสามารถในการใช้ความรู้และทักษะด้านวิทยาศาสตร์ การอ่าน และคณิตศาสตร์ พร้อมทั้งแบบสอบถามนักเรียนและผู้บริหารโรงเรียน ซึ่งผลการประเมิน PISA 2015 ของประเทศไทย พบว่าคะแนนวิชาวิทยาศาสตร์ได้อันดับที่ 54 วิชาคณิตศาสตร์ได้อันดับ 54 และทักษะการอ่านได้อันดับที่ 57 จากทั้งหมด 70 ประเทศ ซึ่งคะแนนทั้งหมดต่ำกว่าเกณฑ์ค่าเฉลี่ยรวม และเมื่อเปรียบเทียบกับการทดสอบครั้งก่อน (PISA 2012) ถือว่าประเทศไทยมีอันดับและคะแนนลดลงในทุกวิชา นอกจากนี้ประเทศไทยยังได้เข้าร่วมวิจัยกับสมาคม IEA เพื่อประเมินผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์และวิทยาศาสตร์ของนักเรียนชั้น ม.2 ตามโครงการ TIMSS โดยเก็บข้อมูลเมื่อเดือนกุมภาพันธ์ พ.ศ. 2558 จากนักเรียนกลุ่มตัวอย่างชั้น ม.2 จำนวน 6,482 คน ใน 204 โรงเรียนของทุกสังกัด ผลการทดสอบ TIMSS 2015 ของนักเรียนไทยระดับ ม.2 พบว่า มีคะแนนเฉลี่ยสูงขึ้น โดยมีคะแนนเฉลี่ยคณิตศาสตร์ 431 คะแนน และ วิทยาศาสตร์ 456 คะแนน ซึ่งสูงกว่าคะแนนเฉลี่ย TIMSS 2011 จากผลการประเมินทั้ง PISA 2015 และ TIMSS 2015 แสดงผลการทดสอบที่มีความขัดแย้งกัน ทั้งที่เป็นเหตุการณ์ที่เกิดใกล้เคียงกัน จึงควรนำวิธีการสุ่มแบบกริด และ/ หรือ กริด-ชั้นภูมิ มาทำการเปรียบกับวิธีการสุ่มตัวอย่างทั้ง 2 โครงการ พร้อมทั้งจำลองสถาการณ์ในการเปรียบเทียบประสิทธิภาพ

บรรณานุกรม

กรมแผนที่ทหาร. (2555). **การใช้โปรแกรมภูมิสารสนเทศในการสร้างแบบจำลองความสูงสามมิติของภูมิประเทศ**. เข้าถึงได้จาก http://www.rtsd.mi.th/ index.php? option=com_ content&view=article&id=374:2012-10-10-05-42-47&catid=12:knownledge-management&Itemid=28&lang=th

กรมสรรพสามิต. (2555). **โครงการสำรวจความพึงพอใจของประชาชน ต่อการปฏิบัติงานของข้าราชการกรมสรรพสามิต ประจำปีงบประมาณ พ.ศ. 2555**. เข้าถึงได้จาก http:// prweb.excise.go.th/survey.html

กรรณิการ์ มิ่งสอน. (2547). **การสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสุ่มสองขั้น สำหรับการประมาณอัตราส่วนของนักเรียนต่อครูในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา.** วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, บัณฑิตวิทยาลัย, มหาวิทยาลัยเกษตรศาสตร์.

แก้วใจ คำสุข. (2543). **ความเที่ยงสัมพัทธ์ของแผนการสุ่มตัวอย่างแบบแบ่งชั้นภูมิเทียบกับแผนการสุ่มตัวอย่างแบบธรรมดาในการประมาณค่าสัดส่วนในเชิงประจักษ์.** วิทยานิพนธ์สาธารณสุขศาสตรมหาบัณฑิต, สาขาวิชาชีวสถิติ, บัณฑิตวิทยาลัย, มหาวิทยาลัยขอนแก่น.

ดวงใจ ปวีณอภิชาต. (2535). **การเปรียบเทียบค่าประมาณพารามิเตอร์ของแผนการสุ่มตัวอย่างแบบแบ่งชั้นที่มีตัวแปรจำแนกชั้นภูมิ และวิธีการกำหนดขนาดกลุ่มตัวอย่างย่อย ที่แตกต่างกัน: กรณีศึกษาผลสัมฤทธิ์ทางการเรียน.** วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาสถิติการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ธวัช สุขแซว. (2539). **การเปรียบเทียบแผนแบบการสุ่มตัวอย่างแบบแบ่งชั้นภูมิชนิดสองทางและชนิดทางเดียวศึกษากรณีการประมาณจำนวนวันลาของข้าราชการครูในโรงเรียนเขตกรุงเทพมหานคร สังกัดกรมสามัญศึกษา.** วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, บัณฑิตวิทยาลัย, มหาวิทยาลัยเกษตรศาสตร์.

ธัญลักษณ์ เหลืองวิสุทธิ์. (2540). **โพล: การเมืองของการวิจัยเชิงสำรวจ.** วิทยานิพนธ์รัฐศาสตรมหาบัณฑิต, สาขาวิชาการปกครอง, บัณฑิตวิทยาลัย, มหาวิทยาลัยธรรมศาสตร์.

ธีรวัฒน์ ประกอบผล. (2552). **คู่มือการเขียนโปรแกรมภาษา C.** กรุงเทพฯ: รีไวว่า.

ธีรวุฒิ เอกะกุล. (2543). **ระเบียบวิธีวิจัยทางพฤติกรรมศาสตร์และสังคมศาสตร์.** อุบลราชธานี: สถาบันราชภัฏอุบลราชธานี.

นพดล กรรณิกา. (2551, 28 มิถุนายน). อย่าเชื่อโพลล์**. มติชนรายวัน,** 31(11067).

นิภาพร โพธิ์ชัย. (2544). **การเปรียบเทียบการสุ่มตัวอย่างแบบสองชั้นและการสุ่มตัวอย่างแบบง่าย สำหรับข้อมูลทวินาม: ศึกษากรณีการหาปริมาตรไม้ยูคาลิปตัส.** กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์.

ประชา สุวัฒนพันธุ์กุล. (2534). **การเปรียบเทียบประสิทธิภาพ ของการเลือกตัวอย่างแบบโควต้ากับการสุ่มตัวอย่างแบบชั้นภูมิ โดยวิธีการจำลองค่า.** วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ประชุม สุวัตถี. (2517). **ทฤษฎีการสำรวจด้วยตัวอย่าง.** กรุงเทพฯ: สถาบันบัณฑิตพัฒนบริหารศาสตร์.

ประภาพร ช่างไม้. (2551). **คู่มือการเขียนโปรแกรมภาษา C ฉบับผู้เริ่มต้น.** นนทบุรี: ไอดีซีอินโฟดิสทริบิวเตอร์ เซ็นเตอร์.

พรพรรณ วีระปรียากูร. (2548). ทางบรรจบหรือทางคู่ขนานระหว่างโพลกับการวิจัยเชิงสำรวจ. **วารสารวิจัยและสาระสถาปัตยกรรม/ การผังเมือง**, **3**, 241-250.

มนตรี พิริยะกุล. (2530). **เทคนิคการสำรวจด้วยกลุ่มตัวอย่าง** (พิมพ์ครั้งที่ 4). กรุงเทพฯ: ป. สัมพันธ์พาณิชย์.

มยุรี ศรีชัย. (2539). **เทคนิคการสุ่มกลุ่มตัวอย่าง** (พิมพ์ครั้งที่ 2). กรุงเทพฯ: วี. เจ. พริ้นติ้ง.

มหาวิทยาลัยเชียงใหม่. (2554). **การเผยแพร่ของศาสนาอิสลามในดินแดนเอเชียตะวันออกเฉียงใต้** เข้าถึงได้จาก http://www.thaiartcmu.com/e_book/104201_muslim_in_southeast_asia

มารยาท โยทองยศ และปราณี สวัสดิสรรพ์. (2551). **การกำหนดขนาดของกลุ่มตัวอย่างเพื่อการวิจัย**. กรุงเทพฯ: ศูนย์บริการวิชาการสถาบันส่งเสริมการวิจัยและพัฒนานวัตกรรม.

มูลนิธิคริสต์ศาสนิกชนออร์โธดอกซ์ในประเทศไทย. (ม.ป.ป.). **ประวัติศาสตร์ศาสนาคริสต์** เข้าถึงได้จาก http://www.orthodox.or.th/index.php?content=history&lang=th

ยงยุทธ มัยลาภ. (2549)**. นาม (**NAM) **นั้นสำคัญไฉน**. เข้าถึงได้จาก http://www.thaiworld.org/th/thailand_monitor/answer.php?question_id=573

วัฒนา สุนทรชัย. (2549). ความแม่นยำของโพลล์ (Poll). **วารสารนักบริหาร, 26**(3), 102-105.

ศิรประภา มโนมัธย์. (2539). **ประสิทธิภาพของแผนแบบการสุ่มตัวอย่างแบบกลุ่มปรับ เมื่อตัวอย่างขั้นต้น ใช้วิธีการสุ่มแบบง่าย แบบมีชั้นภูมิและแบบมีระบบ ภายใต้แบบจำลองประชากรที่กำหนด**. กรุงเทพฯ: มหาวิทยาลัยศิลปากร.

ศิริชัย กาญจนวาสี. (2545). **หนังสือสถิติประยุกต์สำหรับการวิจัย** (พิมพ์ครั้งที่ 3). กรุงเทพฯ: ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2556). มาตรฐานการทำโพลล์. **วารสารการวิจัยสังคมศาสตร์**, 1-15.

ศิริลักษณ์ สุวรรณวงศ์. (2538). **ทฤษฎีและเทคนิคการสุ่มตัวอย่าง.** กรุงเทพฯ: โอเดียนสโตร์ .

ศูนย์ข้อมูล & ข่าวสืบสวนเพื่อสิทธิพลเมือง (TCIJ). (2555). **สพฐ. เลื่อนสอบNTไปปี 56 'ชินภัทร' ชี้ปัญหา-งบมาก.** เข้าถึงได้จาก http://www.tcijthai.com/news/2012/07/scoop/1438

ศูนย์ปฏิบัติการ GPA สำนักทดสอบทางการศึกษา สพฐ. (2549). **รายงานคะแนน O-Net นักเรียนชั้นมัธยมศึกษาปีที่ 6 สรุปเป็นภาพรวมของโรงเรียน โรงเรียนลักษณะพิเศษ จังหวัดเขตพื้นที่การศึกษาและเขตตรวจราชการ ปีการศึกษา 2548.** เข้าถึงได้จาก http://gpa.moe.go.th/gpasearch/

ศูนย์สารสนเทศชุมชน คณะเกษตรศาสตร์ มหาวิทยาลัยอุบลราชธานี. (ม.ป.ป.). **คู่มือการใช้งานโปรแกรม Quantum gis version 1.7.1**. เข้าถึงได้จาก ttp://202.28.48.140/dssgis/Qgis.pdf

สถาบันทดสอบทางการศึกษาแห่งชาติ (สทศ.). (2555). **O-NET คือ.** เข้าถึงได้จาก http://www.niets.or.th/index.php/exam_information/view_se1/1/1

สถาบันราชานุกูล กรมสุขภาพจิต. (2554). **รายงานการสำรวจสถานการณ์ระดับสติปัญญาเด็กนักเรียนไทย พ.ศ. 2554.** เข้าถึงได้จาก http://www.smartteen.net/iqeq/result.html

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2554). **ผลการประเมิน PISA 2009 การอ่าน คณิตศาสตร์ และวิทยาศาสตร์.** กรุงเทพฯ: อรุณการพิมพ์

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2556 ก). **ผลการประเมิน PISA 2012 คณิตศาสตร์ การอ่าน และวิทยาศาสตร์ บทสรุปสำหรับผู้บริหาร.** สมุทรปราการ: แอดวานซ์ พริ้นติ้ง เซอร์วิส.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2556 ข). **สรุปผลการวิจัยโครงการ TIMSS 2011 ชั้นประถมศึกษาปีที่ 4.** สมุทรปราการ: แอดวานซ์ พริ้นติ้ง เซอร์วิส.

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี. (2556 ค). **สรุปผลการวิจัยโครงการ TIMSS 2011 ชั้นมัธยมศึกษาปีที่ 2.** สมุทรปราการ: แอดวานซ์ พริ้นติ้ง เซอร์วิส.

สถาบันสารสนเทศทรัพยากรน้ำและเกษตร (องค์การมหาชน). (2555). **คู่มือการใช้โปรแกรมระบบสารสนเทศภูมิศาสตร์ (QUANTUM GIS) เบื้องต้น**. เข้าถึงได้จาก http://provinces.haii.or.th/index.php/remos.html?func=startdown&id=10

สมบัติ ท้ายเรือคำ. (2551). **ระเบียบวิธีวิจัยสำหรับมนุษยศาสตร์และสังคมศาสตร์.** กาฬสินธุ์: ประสานการพิมพ์.

สำนักสถิติแห่งชาติ. (ม.ป.ป.). **เทคนิคการสุ่มตัวอย่างและการประมาณค่า. กลุ่มระเบียบวิธีสถิติ สำนักนโยบายและวิชาการสถิติ.** เข้าถึงได้จาก http://service.nso.go.th/nso/nsopublish/Toneminute/files/55/A3-16.pdf

สำนักงานคณะกรรมการป้องกันและปราบปรามการทุจริตในภาครัฐ. (2553).**โครงการสำรวจความพึงพอใจของประชาชนต่อมาตรการป้องกัน และปราบปรามการทุจริตในภาครัฐของกระทรวงยุติธรรม ประจำปี พ.ศ. 2553.** เข้าถึงได้จาก www.chainat.go.th/sub/cgd/POC/2553.doc

สุชาดา บวรกิติวงศ์. (2548). **สถิติประยุกต์ทางพฤติกรรมศาสตร์.** กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.

สุทิน ชนะบุญ. (2544). **การเปรียบเทียบผลการสุ่มตัวอย่างแบบง่ายกับการสุ่มตัวอย่างแบบแบ่งชั้นภูมิโดยใช้ข้อมูลความจำเป็นพื้นฐานในการประยุกต์ใช้ Lot acceptance sampling เพื่อประเมินความครอบคลุมของการชั่งน้ำหนักในเด็กอายุต่ำกว่า 5 ปี.** วิทยานิพนธ์สาธารณสุขศาสตรมหาบัณฑิต, สาขาวิชาชีวสถิติ, บัณฑิตวิทยาลัย, มหาวิทยาลัยขอนแก่น.

สุระ พัฒนเกียรติ. (2554). **การประยุกต์ใช้ข้อมูลการรับรู้ระยะไกลเบื้องต้นเพื่อการจัดการสิ่งแวดล้อมและทรัพยากร.** กรุงเทพฯ: คณะสิ่งแวดล้อมและทรัพยากรศาสตร์ มหาวิทยาลัยมหิดล.

สุรินทร์ นิยมมางกูร. (2546). **เทคนิคการสุ่มตัวอย่าง.** กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์.

องอาจ นัยพัฒน์. (2551). **วิธีวิทยาการวิจัยเชิงปริมาณและเชิงคุณภาพทางพฤติกรรมศาสตร์ และสังคมศาสตร์** (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สามลดา.

Allison, F. (2008). **Recent tensions between Russia and the west.** Retrieved from http://www.csa.com/discoveryguides/russia/review.php

Anderson, J. D. (1995). **Computational fluid dynamics.** Singapore: McGraw-Hill.

Best, J. W. (1981). **Research in education.** Boston: Allyn and Bacon.

Cochran, W. G. (1977). **Sampling techniques** (3rd ed.). New York: John Wiley and Sons.

Dinkins, C. P., & Jones, C. (2008). **Soil sampling strategies.** Retrieved from http://landresources.montana.edu/soilfertility/PDFbyformat/publication%20pdfs/Soil_Sampling_Strat_MT200803AG.pdf

Dong-Sheng, Y., Zhong-Qi, Z., Hao, Y., Xue-Zheng, S., Man-Zhi, T., Wei-Xia, S., & Hong-Jie, W. (2011). Effect of soil sampling density on detected spatial variability of soil organic carbon in a red soil region of China. **Pedosphere, 21**(2), 207-213.

Ferguson, R. B., & Hergert, G.W. (2009). **Soil sampling for precision agriculture**. Retrieved from http://www.ianrpubs.unl.edu/live/ec154/build/ec154.pdf

Franklin, E., Moraes, J., Landeiro, V. L., Souza, J. L. P., Pequeno, P. A. C. L., Magnusson, W. E., & Morais, J. W. (2013). Geographic position of sample grid and removal of uncommon species affect multivariate analyses of diverse assemblages: The case of oribatid mites (Acari: oribatida). **Ecological Indicators, 34**, 172-180.

Gallego, F. J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A., & Kussul, O. (2014). Efficiency assessment of usin. **Applied Earth Observation and Geoinformation, 29,** 22-30.

Griffith, D. A. (2009). **Spatial autocorrelation.** Retrieved from http://booksite. elsevier.com/brochures/hugy/SampleContent/Spatial-Autocorrelation.pdf

Henry, G. T. (1991). **Pratical sampling.** California: Sage.

Jessen, R. J. (1978). **Statistics survey techniques.** New York: John Wiley and Sons.

Kish, L. (1965). **Survey sampling.** New York: John Wiley and Sons.

Krajewski, S. A., & Gibbs, B. L. (1994). **Understanding contouring: A practical guide to spatial estimation and contouring using a computer and basics of using variograms.** Boulder CO: Gibbs Associates.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. **Educational and Psychological Measurement, 30**(3), 607-610.

Lin, Q., Li, H., Luo, W., Lin, Z., & Li, B. (2013). Optimal soil-sampling design for rubber tree management based on fuzzy clustering. **Forest Ecology and Management, 308**, 214-222.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (1999). **Geographical information systems: Principles, techniques, management and applications** (2nd ed.). Chichester: John Wiley.

Majasalmi, T., Rautiainen, M., Stenberg, P., & Rita, H. (2012). Optimizing the sampling scheme for LAI-2000 measurements in a boreal forest. **Agricultural and Forest Meteorology, 155**, 38-43.

Mallarino, A. P., & Wittry, D. (2001). Management zones soil sampling: A better alternative to grid and soil type sampling?, In **The Integrated Crop Management Conf. Proceedings. Dec. 5-6, 2001, Ames, IA.** (pp. 159-164). Iowa State Univ: Extension.

Martínez-Coronado, A., Oyarzun, R., Esbrí, J. M., Llanos, W., & Higueras, P. (2011). Sampling high to extremely high hg concentrations at the cerco de almadenejos, almadén mining district (spain): The old metallurgical precinct (1794 to 1861 AD) and surrounding areas. **Journal of Geochemical Exploration, 109**, 70-77.

Midwest Laboratories. (2009). **Soil sampling methods.** Retrieved from http://agrienergy.net/docs/lab-information/soil-sampling.pdf

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. **Biometrika**, **37**(1), 17-23.

Mookum, T. (2004). **Finite difference methods for finding a control parameter in two-dimentional parabolic equation with Neumann boundary conditions.** Chiang Mai: Chiang Mai University.

Neuman, W. L. (1991). **Social research methods: Qualitative and quantitative approaches**. Boston: Allyn and Bacon.

Onofiok, O. E. (1993). **Determination of spatial and temporal variations in soil organic matter in tropical soil using different sampling schemes.** New York: John Wiley and Sons.

Pew Research Religion & Public Life Project. (2012). **The global religious landscape**. Retrieved from http://www.pewforum.org/2012/12/18/global-religious-landscape-exec/#

Radil, S. M. (2011). **Spatializing social networks: Making space for theory in spatial analysis.** Illinois: University of Illinois at Urbana Champaign.

Sánchez-Fernández, D., Lobo, J. M., Abellán, P., & Millán, A. (2011). How to identify future sampling areas when information is biased and scarce: An example using predictive models for species richness of Iberian water beetles**. Journal for Nature Conservation, 19**, 54-59.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. In **Proceedings of the 1968 ACM National Conference**. (pp. 517-524). New York: ACM.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. **Economic Geography**, **46**(2), 234-240.

Yamane, T. (1973). **Statistics: An introductory analysis** (3rd ed.). New York: Harper and Row.

ภาคผนวก

ภาคผนวก ก
คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบหลายขั้นตอน

คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบหลายขั้นตอน

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
1	08500102:วัฒโนทัยพายัพ	215.61
2	08500104:กาวิละวิทยาลัย	194.66
3	08500301:แม่แจ่ม	159.16
4	08500401:เชียงดาววิทยาคม	162.23
5	08500491:ศึกษาสงเคราะห์เชียงดาว	144.89
6	08500601:แม่แตง	168.31
7	08500603:สันป่ายางวิทยาคม	168.74
8	08500702:นวมินทราชูทิศ พายัพ	207.21
9	08500801:สะเมิงพิทยาคม	159.37
10	08501001:แม่อายวิทยาคม	164.41
11	08501101:พร้าววิทยาคม	169.8
12	08501202:น้ำบ่อหลวงวิทยาคม	151.78
13	08501401:สันทรายวิทยาคม	170.62
14	08501601:ฮอดพิทยาคม	166.95
15	08501801:อมก๋อยวิทยาคม	152.86
16	07530105:แสนตอวิทยา	144.7
17	07530106:วังกะพี้พิทยาคม	151.08
18	07530107:น้ำริดวิทยา	158.71
19	07530201:ตรอนตรีสินธุ์	157.01
20	07530301:ท่าปลาประชาอุทิศ	144.42
21	07530401:น้ำปาดชนูปถัมภ์	165.07
22	07530501:ฟากท่าวิทยา	156.33
23	07530601:บ้านโคกวิทยาคม	148.52
24	07530701:พิชัย	170.98
25	07530702:บ้านโคนพิทยา	142.05
26	07530703:ดาราพิทยาคม	137.89
27	07530801:ลับแลศรีวิทยา	148.8
28	07530802:ลับแลพิทยาคม	160.69
29	07530803:ด่านแม่คำมันพิทยาคม	146.19
30	07530901:ทองแสนขันวิทยา	144.56

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
31	10370106:นาจิกพิทยาคม	144.36
32	10370107:นายมวิทยาคาร	148.74
33	10370108:สร้างนกทาวิทยาคม	150.52
34	10370109:อำนาจเจริญ 2	153.49
35	10370191:ศึกษาสงเคราะห์อำนาจเจริญ	162.67
36	10370201:ชานุมานวิทยาคม	154.06
37	10370202:คำเขื่อนแก้ววิทยาคม	142.58
38	10370301:ปทุมราชวงศา	159.48
39	10370302:ลืออวิทยาคม	159.17
40	10370401:พนาศึกษา	160.24
41	10370402:ศรีคูณวิทยบัลลังก์	149.92
42	10370501:เสนางคนิคม	158.17
43	10370502:นาเวียงจุลดิศวิทยา	152.7
44	10370601:หัวตะพานวิทยาคม	163.34
45	10370602:ศรีเจริญศึกษา	149.33
46	10370603:จิกดู่วิทยา	148.65
47	11360102:สตรีชัยภูมิ	186.74
48	11360105:บ้านค่ายวิทยา	145.23
49	11360108:กาญจนาภิเษกวิทยาลัย ชัยภูมิ	183.38
50	11360111:ชีลองวิทยา	123.22
51	11360203:ลุ่มลำชีนิรมิตวิทยา	145.18
52	11360303:โนนสะอาดวิทยา	141.78
53	11360402:บ้านเป้าวิทยา	141.61
54	11360405:บ้านเดื่อวิทยาคม	136.82
55	11360503:คูเมืองวิทยา	142.78
56	11360603:หนองบัวบานวิทยา	142.51
57	11360701:บำเหน็จณรงค์วิทยาคม	154.92
58	11360704:เริงรมย์วิทยาคม	143.23
59	11360901:เทพสถิตวิทยา	150.88
60	11360904:ตรีประชาพัฒนศึกษา	140
61	11361003:กวางโจนศึกษา	147.11
62	11361006:โคกสะอาดวิทยา	134.18
63	11361201:แก้งคร้อวิทยา	182.03

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
64	10460302:ฟ้าแดดสูงยางวิทยาคาร	154.92
65	10460305:ธัญญาพัฒนวิทย์	144.3
66	10460502:กุฉินารายณ์	140.65
67	10460505:โนนคำวิทยา	138.29
68	10460601:เขาวงพิทยาคาร	169.55
69	10460701:ยางตลาดวิทยาคาร	149.36
70	10460704:เขาพระนอนวิทยาคม	135.42
71	10460707:คลองขามวิทยาคาร	135.53
72	10460803:หัวหินวัฒนาลัย	134.58
73	10460901:สหัสขันธ์ศึกษา	156.23
74	10461001:คำม่วง	149.37
75	10461102:ดงกลางพัฒนศึกษา	144.85
76	10461203:หนองชุมแสงวิทยาคม	144.42
77	10461301:สมเด็จพิทยาคม	162.51
78	10461304:ประชารัฐพัฒนศึกษา	132.42
79	10461307:ผาเสวยรังสรรค์	135.21
80	09420701:ภูเรือวิทยา	156.98
81	09420801:ท่าลี่วิทยา	156.87
82	09420901:ศรีสงครามวิทยา	167.74
83	09420902:เหมืองแบ่งวิทยาคม	154.5
84	09420903:วังทรายขาววิทยา	138.18
85	09420904:เขาหลวงวิทยา	135.35
86	09420905:เซโลวิทยาคม	145.07
87	09420906:ผาน้อยวิทยาคม	138.03
88	09420907:เลยสว่างวิทยาคม	135.73
89	09420991:ศึกษาสงเคราะห์เลย	157.76
90	09421001:ภูกระดึงวิทยาคม	154.53
91	09421101:ภูหลวงวิทยา	139.83
92	09421201:สันติวิทยาสรรพ์	156.73
93	09421301:เอราวัณวิทยาคม	155.78
94	09421302:ผาอินทร์แปลงวิทยา	151.38
95	09421303:ผาสามยอดวิทยาคม	146.69
96	10440104:โคกก่อพิทยาคม	144.09

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
97	10440201:แกดำวิทยาคาร	144.23
98	10440302:เขวาไร่ศึกษา	152.05
99	10440305:วังยาวศึกษาวิทย์	141.28
100	10440401:กันทรวิชัย	147.09
101	10440404:นาสีนวนพิทยาสรรค์	135.76
102	10440501:เชียงยืนพิทยาคม	151.82
103	10440601:บรบือ	146.03
104	10440604:โนนแดงวิทยาคม	146.9
105	10440607:โนนราษีวิทยา	141.44
106	10440702:ปอพานพิทยาคม รัชมังคลาภิเษก	137.82
107	10440802:เมืองเตาวิทยาคม	153.91
108	10440902:นาข่าวิทยาคม	155.9
109	10440905:งัวบาวิทยาคม	154.09
110	10440908:ขามป้อมพิทยาคม	148.49
111	10441003:มัธยมดงยาง	138.15
112	09390104:เฉลิมพระเกียรติสมเด็จพระศรีนครินทร์ หนองบัวลำภู	168.1
113	09390201:คำแสนวิทยาสรรค์	161.84
114	09390202:กุดดินจี่พิทยาคม	148.11
115	09390203:ฝั่งแดงวิทยาสรรค์	149.62
116	09390301:โนนสังวิทยาคาร	148.67
117	09390302:โนนเมืองวิทยาคาร	143.18
118	09390303:กุดดู่พิทยาคม	146.24
119	09390304:หนองเหมือดแอ่วิทยาคาร	144.35
120	09390401:ศรีบุญเรืองวิทยาคาร	162.05
121	09390402:ยางหล่อวิทยาคาร	154.6
122	09390403:กุงแก้วิทยาคาร	153.95
123	09390404:นากอกวิทยาคาร	144.79
124	09390405:กุดสะเทียนวิทยาคาร	153.58
125	09390501:สุวรรณคูหาพิทยาสรรค์	146.99
126	09390502:ดงมะไฟวิทยาคม	151.89
127	09390503:ภูซางใหญ่วิทยาคม	144.65
128	09430201:ท่าบ่อ	169.81

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
129	09430204:ท่าบ่อพิทยาคม	145.11
130	09430302:นาสวรรค์พิทยาคม	146.7
131	09430401:พรเจริญวิทยา	158.3
132	09430501:ชุมพลโพนพิสัย	171.37
133	09430504:นาหนังพัฒนศึกษา	148.42
134	09430591:ราชประชานุเคราะห์ 27 จังหวัดหนองคาย	159.72
135	09430603:โนนคำพิทยาคม	149.06
136	09430702:พระพุทธบาทวิทยาคม	150.87
137	09430802:วังม่วงพิทยาคม	148.57
138	09430903:โสกก่ามวิทยา	149.46
139	09430906:เจ็ดสีวิทยาคาร	154.14
140	09431003:สมสนุกพิทยาคม	140.77
141	09431201:ศรีวิไลวิทยา	153.98
142	09431401:น้ำสวยวิทยา	163.89
143	09431502:นาดีพิทยาคม	162.76
144	09470101:สกลราชวิทยานุกูล	209.05
145	09470103:สกลนครพัฒนศึกษา	165.79
146	09470105:ท่าแร่ศึกษา	158.3
147	09470107:เฉลิมพระเกียรติฯ สกลนคร	178.5
148	09470202:โพธิแสนวิทยา	144.55
149	09470302:ร่มไทรวิทยา	141.3
150	09470402:บ้านบัวราษฎร์บำรุง	148.94
151	09470404:ภูริทัตต์วิทยา	138.93
152	09470406:ช้างมิ่งพิทยานุกูล	151.2
153	09470501:พังโคนวิทยาคม	168.36
154	09470601:มัธยมวาริชภูมิ	150.82
155	09470603:วาริชวิทยา	150.39
156	09470801:มัธยมวานรนิวาส	159.49
157	09470803:ธาตุทองอำนวยวิทย์	142.65
158	09470805:หนองแวงวิทยา	142.43
159	09470901:คำตากล้าราชประชาสงเคราะห์	151.15
160	09471001:บ้านม่วงพิทยาคม	169.15
161	10350105:ศิลาทองพิทยาสรรค์	154.58

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
162	10350201:ทรายมูลวิทยา	158
163	10350202:ดงมะไฟพิทยาคม	167.74
164	10350301:กุดชุมวิทยาคม	163.42
165	10350302:กำแมดขันติธรรมวิทยาคม	155.63
166	10350303:โพนงามพิทยาคาร	149.92
167	10350304:หนองแหนพัฒนวิทยาคม	151.05
168	10350305:เมืองกลางประชานุกูล	146.03
169	10350401:คำเขื่อนแก้วชนูปถัมภ์	174.81
170	10350402:สมเด็จพระญาณสังวร	162.39
171	10350403:กู่จานวิทยาคม	155.11
172	10350404:โพนทันเจริญวิทย์	148.84
173	10350491:ราชประชานุเคราะห์ 28	160.32
174	10350501:ป่าติ้ววิทยา	160.13
175	10350502:ศรีฐานกระจายศึกษา	142.79
176	10350601:มหาชนะชัยวิทยาคม	151.92
177	10350602:ตระกูลประเทืองวิทยาคม	147.86
178	07620109:หนองกองพิทยาคม	152.13
179	07620201:ไทรงามพิทยาคม	152.59
180	07620301:คลองลานวิทยา	156.67
181	07620303:คลองลานพัฒนาจินดาศักดิ์	153.59
182	07620402:โค้งไผ่วิทยา	147.7
183	07620404:สลกบาตรวิทยา	147.92
184	07620502:วังแขมวิทยาคม	158.23
185	07620601:พรานกระต่ายพิทยาคม	156.28
186	07620603:เรืองวิทย์พิทยาคม	154.86
187	07620702:มัธยมพัชรกิติยาภา 2 กำแพงเพชร	158.9
188	07620801:ทุ่งทรายวิทยา	157.94
189	06160103:โคกตูมวิทยา	154.05
190	06160105:พระนารายณ์	182.48
191	06160201:พัฒนานิคม	153.93
192	06160203:ขุนรามวิทยา	149.16
193	06160302:จุฬาภรณราชวิทยาลัย ลพบุรี	218.69
194	06160402:ชัยบาดาลพิทยาคม	153.26

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
195	06160501:ท่าวุ้งวิทยาคาร	145.5
196	06160601:บ้านหมี่วิทยา	168.56
197	06160604:บ้านชีวิทยา	142.26
198	06160801:สระโบสถ์วิทยาคาร	151.41
199	06160902:ยางรากวิทยา	153.37
200	01130201:ธรรมศาสตร์คลองหลวงวิทยาคม	186.97
201	01130202:มัธยมวัดหัตถสารเกษตร	153.83
202	01130203:สวนกุหลาบวิทยาลัย รังสิต	218
203	01130301:ธัญบุรี	185.59
204	01130302:ธัญรัตน์	186.18
205	01130303:สายปัญญารังสิต	170.01
206	01130401:หนองเสือวิทยาคม	160.49
207	01130501:บัวแก้วเกษร	159.88
208	01130502:วรราชาทินัดดามาตุวิทยา	121.25
209	01130503:สุนทโรเมตตาประชาสรรค์	151.65
210	01130504:จุฬาภรณราชวิทยาลัย ปทุมธานี	241.27
211	06140103:อยุธยานุสรณ์	170.47
212	06140201:ท่าเรือ"นิตยานุกูล"	166.07
213	06140301:นครหลวง"อุดมรัชต์วิทยา"	149.01
214	06140401:บางไทรวิทยา	148.92
215	06140601:บางปะอิน"ราชานุเคราะห์ ๑"	164.61
216	06140603:อุดมศีลวิทยา	145.5
217	06140801:ผักไห่"สุทธาประมุข"	157.67
218	06140803:วัดโพธิ์ผักไห่ (เวชพันธุ์อนุสรณ์)	146.8
219	06141001:ลาดบัวหลวงไพโรจน์วิทยา	182.24
220	06141102:วังน้อย (พนมยงค์วิทยา)	151.29
221	06141202:สาคลีวิทยา	149.4
222	06141301:บางซ้ายวิทยา	152.21
223	07640103:บ้านสวนวิทยาคม	158.61
224	07640105:ลิไทพิทยาคม	139.54
225	07640202:ตลิ่งชันวิทยานุสรณ์	149.28
226	07640302:บ้านใหม่เจริญผลพิทยาคม	150.01
227	07640402:หนองตูมวิทยา	149.25

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
228	07640501:เมืองเชลียง	165.27
229	07640503:บ้านแก่งวิทยา	155.42
230	07640601:ศรีสำโรงชนูปถัมภ์	183.7
231	07640603:วังทองวิทยา	144.38
232	07640701:สวรรค์อนันต์วิทยา	209.99
233	07640703:หนองกลับวิทยาคม	149.6
234	07640801:ศรีนคร	169.21
235	05720203:ทุ่งคลีโคกช้างวิทยา	163.98
236	05720301:บรรหารแจ่มใสวิทยา 3	172.12
237	05720401:บางปลาม้า"สูงสุมารผดุงวิทย์"	172.32
238	05720403:บางแม่หม้ายรัฐราษฎร์รังสฤษดิ์	157.11
239	05720502:สรวงสุทธาวิทยา	155.13
240	05720601:บรรหารแจ่มใสวิทยา 1	177.26
241	05720701:สองพี่น้องวิทยา	159.68
242	05720703:บ่อสุพรรณวิทยา	151
243	05720705:หนองวัลย์เปรียงวิทยา	159.19
244	05720802:ทุ่งแฝกพิทยาคม	154.27
245	05720901:อู่ทอง	184.17
246	05720903:สระยายโสมวิทยา	164.94
247	06610302:บ่อยางวิทยา	156.38
248	06610401:หนองฉางวิทยา	198.96
249	06610402:บ้านทุ่งนาวิทยา	160.17
250	06610403:ทุ่งโพวิทยา	155.58
251	06610501:หนองขาหย่างวิทยา	177.15
252	06610502:กาญจนาภิเษกวิทยาลัย อุทัยธานี	163.63
253	06610601:บ้านไร่วิทยา	162.16
254	06610602:การุ้งวิทยาคม	149.85
255	06610603:ทองหลางวิทยาคม	160.31
256	06610604:หนองจอกประชานุสรณ์	153.49
257	06610605:วังหินวิทยาคม	158.25
258	12250101:ปราจิณราษฎรอำรุง	201.89
259	12250102:ปราจีนกัลยาณี	176.63
260	12250103:อดุลศาสนกิจศึกษา	162.88

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
261	12250104:มัธยมวัดป่ามะไฟ	146.59
262	12250105:ไทยรัฐวิทยา 7	145.96
263	12250106:ประชารัฐพัฒนา	147.26
264	12250201:กบินทร์วิทยา	159.02
265	12250202:ลาดตะเคียนราษฎร์บำรุง	148.28
266	12250203:กบินทร์บุรี	153.33
267	12250204:วังตะเคียนวิทยาคม	156.98
268	12250205:วังดาลวิทยาคม	147.39
269	12250206:นนทรีวิทยาคม	145
270	12250207:เตรียมอุดมศึกษาน้อมเกล้า กบินทร์บุรี	170.48
271	12250301:มณีเสวตรอุปถัมภ์	147.8
272	12250302:ร่มเกล้า ปราจีนบุรี	142.18
273	12250303:ทุ่งใหญ่วิทยาคาร	153.38
274	12250601:ชิตใจชื่น	141.33
275	12250602:ศรีรักษ์ราษฎร์บำรุง	146
276	12250603:กระทุ่มแพ้ววิทยา	147.54
277	12250701:ประจันตราษฎร์บำรุง	160.3
278	12250702:สุวรรณวิทยา	141.89
279	12250703:วัดพรหมประสิทธิ์	144.89
280	07630201:บ้านตาก"ประชาวิทยาคาร"	161.96
281	07630202:ทุ่งฟ้าวิทยาคม	152.01
282	07630301:สามเงาวิทยาคม	154.1
283	07630302:ยกกระบัตรวิทยาคม	147.81
284	07630401:แม่ระมาดวิทยาคม	150.58
285	07630402:แม่จะเราวิทยาคม	152.69
286	07630501:ท่าสองยางวิทยาคม	162.74
287	07630601:สรรพวิทยาคม	186.81
288	07630602:แม่ปะวิทยาคม	156.21
289	07630603:ด่านแม่ละเมาวิทยาคม	151.56
290	07630604:แม่กุวิทยาคม	150.74
291	07630701:พบพระวิทยาคม	156.69
292	07630801:อุ้มผางวิทยาคม	159.69
293	07630901:นาโบสถ์พิทยาคม	154.31

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
294	07630902:วังเจ้าวิทยาคม	157.48
295	03930302:หานโพธิ์พิทยาคม	151.45
296	03930401:ตะโหมด	151.34
297	03930402:ประชาบำรุง	151.9
298	03930501:ควนขนุน	153.84
299	03930502:พนางตุง	154
300	03930503:อุดมวิทยายน	145.98
301	03930505:ดอนศาลานำวิทยา	157.25
302	03930506:ปัญญาวุธ	185.42
303	03930601:หารเทา	162.52
304	03930602:ปากพะยูนพิทยาคาร	162.82
305	03930603:ควนพระสาคริินทร์	150.31
306	03930604:มัธยมเกาะหมาก	153.99
307	03930701:ศรีบรรพตพิทยาคม	144.83
308	03930801:ป่าบอนพิทยาคม	146
309	02910101:สตูลวิทยา	163.18
310	02910102:พิมานพิทยาสรรค์	186.24
311	02910103:จุฬาภรณราชวิทยาลัย สตูล	251.51
312	02910201:ควนโดนวิทยา	154.93
313	02910301:ควนกาหลงวิทยาคม"นิคมวัฒนา"	155.64
314	02910401:ท่าแพผดุงวิทย์	146.58
315	02910402:สาครพิทยาคาร	159.22
316	02910501:กำแพงวิทยา	161.15
317	02910502:ละงูพิทยาคม	147.71
318	02910601:ทุ่งหว้าวรวิทย์	162.48
319	02910602:ท่าศิลาบำรุงราษฎร์	153.95
320	02910701:ปาล์มพัฒนวิทย์	142.95
321	04920401:ปะเหลียนผดุงศิษย์	165.58
322	04920402:คันธพิทยาคาร	156.11
323	04920403:ทุ่งยาวผดุงศิษย์	165.93
324	04920501:สิเกาประชาผดุงวิทย์	155.07
325	04920601:ห้วยยอด	190.25
326	04920602:สามัคคีศึกษา	165.7

ลำดับที่	ชื่อโรงเรียน	คะแนนเฉลี่ย O-NET
327	04920603:ห้วยนางราษฎร์บำรุง	161.87
328	04920604:ลำภูราเรืองวิทย์	150.97
329	04920606:บางดีวิทยาคม	171.46
330	04920701:วังวิเศษ	161.34
331	04920702:รัษฎานุประดิษฐ์อนุสรณ์	152.57
332	04920801:สวัสดิ์รัตนาภิมุข	171.42
333	04920802:นาโยงวิทยาคม	167.88
334	04920901:คลองปางวิทยาคม	169.88
335	04921001:หาดสำราญวิทยาคม	155.01
	คะแนน O-NET ของกลุ่มตัวอย่าง	156.56
	ความแปรปรวนกลุ่มตัวอย่าง	250.75
	ค่าเบี่ยงเบนมาตรฐานกลุ่มตัวอย่าง	15.83

ภาคผนวก ข

คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบกริด

คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบกริด

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
1	2109	9.142467	98.9661	156.73
2	2126	9.243169	98.97402	146.63
3	2128	9.049858	99.08126	146.22
4	2112	8.848778	98.80486	156
5	2129	9.137351	98.59612	142.6
6	2131	9.072698	99.12113	140.98
7	2115	9.095995	98.9851	154.12
8	2117	8.840883	99.18925	152.7
9	2119	8.950558	98.99386	152.46
10	2120	8.621054	98.98302	152.44
11	1203	9.246222	98.41841	160.35
12	2104	8.921812	98.90975	159.95
13	2123	8.737104	99.02093	150.75
14	1208	7.61698	100.0769	200.89
15	1225	7.355109	100.1203	151.34
16	1905	7.765652	100.3086	168.53
17	1209	7.61472	100.0695	200.75
18	1226	7.383562	100.2116	150.31
19	1210	7.65009	99.99872	185.42
20	1212	7.46671	100.1376	163.46
21	1215	7.609141	100.1143	161.92
22	1913	7.66489	100.3778	157.1
23	1217	7.413465	100.1586	160.17
24	1219	7.770536	100.1107	154
25	1222	7.577371	99.99682	152.17
26	1903	7.62549	100.3284	169.26
27	1224	7.509025	100.1306	151.45
28	987	6.165416	101.6996	144.06
29	988	6.405791	101.4544	142.5
30	990	6.179809	101.833	139.55

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
31	991	6.307002	101.7184	139.07
32	992	6.32537	101.7925	136.85
33	993	6.14469	101.8908	133.13
34	977	6.02333	101.9621	180.91
35	978	6.421958	101.8215	175.33
36	980	6.397147	101.5256	158.1
37	981	5.927879	101.8838	152.11
38	982	5.940774	101.772	147.93
39	983	6.305841	101.5106	146.45
40	984	6.079057	101.8765	146.4
41	986	6.403653	101.7075	144.46
42	2571	18.84677	98.73392	159.37
43	2572	19.11114	99.02785	159.31
44	2557	18.91571	98.93629	176.36
45	2560	18.91905	98.99645	170.62
46	2562	19.04956	98.87164	168.74
47	2563	19.11006	98.93013	168.31
48	2549	18.84237	98.96363	207.21
49	2566	18.85422	98.9656	164.45
50	1191	19.14806	100.3807	161.58
51	1192	19.16945	100.2788	161.33
52	1194	19.06775	100.0584	158.21
53	1180	19.33695	100.1308	176.23
54	1182	18.89522	100.2862	175.31
55	1186	19.16348	99.99692	168.91
56	1189	19.25605	100.3508	166.41
57	1020	18.84963	100.4377	153.07
58	2568	17.40476	98.45535	164.27
59	1702	17.79995	98.94888	174.96
60	1704	17.71746	98.97633	163.08
61	654	17.34657	99.05547	147.81
62	1691	17.44374	99.1574	150.83
63	1708	18.01689	98.88289	157.06

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
64	2581	17.94787	98.69446	148.29
65	2450	17.64538	100.2479	144.38
66	2434	17.63319	100.0977	215.8
67	2435	17.62788	100.0972	211.22
68	2439	17.6996	100.1215	158.71
69	2442	17.57843	100.146	151.55
70	2443	17.566	100.1116	151.08
71	2447	17.65919	100.3224	144.7
72	154	14.74186	98.63169	156.56
73	156	15.15289	98.45243	154.45
74	158	14.63095	99.47371	151.82
75	160	14.22434	99.07131	150.51
76	144	14.77899	99.33939	163.31
77	162	14.44837	99.13422	149.24
78	164	14.5495	98.78411	143.05
79	134	14.53507	98.80898	222.14
80	650	16.80442	98.76566	151.56
81	651	16.65664	98.59078	150.74
82	636	16.7132	98.57318	186.81
83	640	16.03175	98.8588	159.69
84	642	16.38414	98.69144	156.69
85	643	16.76174	98.58879	156.21
86	646	16.67666	99.14468	154.31
87	410	13.67551	101.3213	149.71
88	373	13.04099	101.9101	164.55
89	448	13.36119	101.2295	152.71
90	416	13.67201	101.433	144.56
91	1603	12.96107	101.6167	170.43
92	420	13.45052	101.1707	224.33
93	1604	12.94014	101.6748	170.42
94	422	13.29491	101.1633	219.16
95	1980	13.51112	102.0458	165.68
96	1611	12.9745	101.6839	160.95

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
97	394	13.46599	101.4435	168.9
98	1985	13.41746	102.0076	156.85
99	432	13.37696	101.2807	178.35
100	396	13.62629	101.4493	164.86
101	433	13.28916	101.4221	175.26
102	435	13.26555	101.2526	171.4
103	402	13.38142	101.7063	159.66
104	439	13.52681	101.2135	165.89
105	440	13.41961	101.3341	164.97
106	404	13.57803	101.2872	158.38
107	1145	13.75781	101.6626	149.82
108	442	13.1477	101.3619	161.85
109	257	16.45729	99.31574	152.13
110	241	16.18641	99.48118	159.3
111	260	16.31887	99.40002	150.2
112	263	16.4116	99.38374	146.42
113	250	16.18368	99.32611	156.67
114	254	16.10758	99.3213	153.59
115	1262	16.40585	100.3938	131.57
116	1232	16.44711	100.3262	202.57
117	1233	16.21373	100.4408	181.73
118	938	16.05781	100.0558	153.72
119	1234	16.03368	100.3751	175.44
120	235	16.14961	99.91778	169.05
121	1236	16.43815	100.3096	165.58
122	1238	16.17438	100.5192	159.07
123	1243	16.27261	100.1894	153.59
124	1244	16.23662	100.0817	153.56
125	1245	16.41782	100.1767	151.92
126	1246	16.13739	100.1241	151.84
127	1247	16.2282	100.3558	150.84
128	1248	16.07577	100.3013	149.72
129	952	15.98925	100.0814	142.15

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
130	1249	16.09251	100.2653	149.28
131	954	15.86951	100.1113	139.92
132	1250	16.03584	100.1801	149.1
133	1251	15.98056	100.2276	148.9
134	1252	16.31507	100.282	147.67
135	921	15.90647	100.3057	189.11
136	1255	16.33681	100.3827	142.01
137	1256	16.39568	100.0735	141.99
138	1258	16.19448	100.2257	141.66
139	2073	14.61038	100.0776	162.62
140	2332	14.59955	100.2629	158.08
141	2074	14.84923	100.0918	162.51
142	2334	14.75618	100.3009	150.95
143	2335	14.50843	100.2409	149.56
144	2337	14.68873	100.38	145.31
145	2080	14.78921	100.0071	158.43
146	2083	14.69056	100.1846	155.13
147	2085	14.76248	100.1756	154.27
148	2089	14.81651	100.1146	143.71
149	2021	14.75696	100.3896	164.17
150	2022	14.80354	100.4465	162.9
151	2023	14.86094	100.39	160.61
152	2024	14.80173	100.3407	159.56
153	2025	14.88215	100.3255	158.66
154	2062	14.62348	100.0118	177.26
155	2026	14.74747	100.4524	156.97
156	2027	14.92812	100.2692	153.86
157	2068	14.60386	100.156	166.07
158	2328	14.58709	100.3537	171.08
159	2070	14.75161	100.0921	165.28
160	2329	14.59946	100.3591	168.18
161	2330	14.70299	100.2937	165.56
162	2072	14.902	100.1963	163.98

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
163	2331	14.66074	100.4095	164.69
164	1958	13.42849	100.021	212.59
165	1959	13.4148	99.99891	206.21
166	1960	13.42454	99.95703	179.83
167	1961	13.47088	99.98345	177.75
168	1962	13.39405	99.96501	158.02
169	1963	13.45883	99.94082	157.15
170	1964	13.42504	99.89	150.25
171	1965	13.50565	99.92126	150.18
172	1973	13.51465	100.1541	167.01
173	1974	13.5116	100.3289	161.01
174	1975	13.6452	100.209	153.47
175	1976	13.5719	100.0779	153.46
176	1966	13.548	100.2752	219
177	1967	13.54953	100.265	200.85
178	1969	13.6514	100.2573	179.97
179	1970	13.58532	100.1008	178.65
180	1971	13.58325	100.2346	170.09
181	1972	13.61175	100.1405	169.39
182	105	13.58218	100.4209	170.99
183	108	13.6406	100.3608	168.91
184	77	13.67604	100.3394	187.97
185	30	13.63375	100.3654	220.41
186	65	13.64668	100.4193	195.66
187	116	13.5266	100.4194	151.31
188	690	13.68332	100.0912	154.87
189	1718	17.4558	101.715	159.52
190	1720	17.30343	101.8013	157.76
191	1721	17.45989	101.3659	156.98
192	1722	17.62493	101.4345	156.87
193	1739	17.79969	101.7288	144.93
194	1724	17.3699	101.8932	155.78
195	1741	17.54918	101.8392	140.54

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
196	1728	17.89743	101.6632	153.58
197	1745	17.38151	101.8173	138.03
198	1729	17.7508	101.7044	153.34
199	1747	17.32811	101.6428	135.35
200	1715	17.48984	101.7195	209.64
201	1716	17.72972	101.7169	199.11
202	1733	17.60513	101.7114	150.38
203	1717	17.30012	101.7905	167.74
204	1734	17.47832	101.9737	147.83
205	2398	17.7777	103.1625	151.39
206	2400	17.67143	103.1701	150.96
207	2384	17.5299	103.1219	156.04
208	2401	17.54801	103.0609	150.68
209	2419	17.45362	103.2569	147.33
210	2369	17.70168	103.2641	175.05
211	2421	17.823	103.0927	146.73
212	2389	17.47947	103.2444	155.26
213	2373	17.69519	102.8916	168.49
214	2390	17.7427	103.0393	154.45
215	2424	17.45093	102.8758	145.69
216	2427	17.64715	102.8217	141.14
217	2377	17.35886	103.1054	163.66
218	2378	17.38927	103.2386	162.76
219	2395	17.58517	103.2064	151.67
220	2396	17.57902	102.9882	151.47
221	732	17.52083	104.6871	144.32
222	699	17.6264	104.2558	164.84
223	716	17.45767	104.4713	152.46
224	717	17.6535	104.4418	152.26
225	734	17.54758	104.2103	142.95
226	702	17.7571	104.306	161.62
227	723	17.28802	104.4659	149.82
228	1879	17.33084	104.3366	146.15

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
229	707	17.40762	104.7749	159.13
230	724	17.62658	104.346	149.07
231	708	17.58793	104.3101	158.77
232	725	17.4418	104.6523	147.92
233	742	17.58669	104.43	138.14
234	710	17.35633	104.6194	158.52
235	727	17.38997	104.725	146.77
236	696	17.55404	104.6115	175.64
237	477	16.29629	101.9345	156.88
238	494	15.85459	101.4186	147.22
239	1346	15.95874	101.3062	152.69
240	481	16.08059	101.8022	154.56
241	515	16.01601	101.7654	142.78
242	500	15.76405	101.5579	146.24
243	520	16.37719	101.9702	141.61
244	504	15.75542	101.7863	144.74
245	521	16.29974	102.0227	141.34
246	488	16.18792	101.6032	150.46
247	522	15.89798	101.6733	140.33
248	472	15.88602	101.9392	158.77
249	474	16.17996	101.9275	158.36
250	525	15.88767	101.7929	138.31
251	509	16.01934	101.4882	144.07
252	527	16.11551	101.8912	136.82
253	1407	16.04204	103.1092	151.68
254	1444	16.24738	103.1905	131.08
255	1408	16.31625	102.9564	150.94
256	1416	16.32411	103.0664	147.79
257	1419	15.94563	103.0348	146.9
258	1421	16.03917	103.118	146.03
259	1422	16.06907	102.9205	146
260	1390	16.1863	103.2991	205.88
261	1428	16.03708	103.2629	144.09

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
262	1392	16.18364	103.3024	181.72
263	1431	16.22169	102.9163	141.28
264	1394	16.23668	103.0684	162.56
265	1432	16.07858	103.2241	141.11
266	1397	16.18295	103.2104	155.52
267	1434	16.15797	103.0241	139.09
268	1439	16.34004	103.2139	135.76
269	1403	16.23656	103.2718	152.84
270	1404	16.29658	102.9216	152.05
271	1441	16.09082	103.0077	134.31
272	1514	16.03239	104.5101	145.52
273	2347	16.04524	104.6697	158.17
274	1839	16.03239	104.5101	125.99
275	1466	16.3514	104.3825	146.16
276	2353	16.03457	104.5755	152.7
277	1453	16.35176	104.4963	162.3
278	2354	15.99916	104.5659	152.23
279	1505	16.13442	104.7288	151.92
280	1490	16.19854	104.5397	167.62
281	1491	16.04687	104.3475	163.42
282	1508	16.27995	104.4425	148.93
283	1458	16.24085	104.6168	155.72
284	1492	16.26674	104.4037	162.67
285	1510	16.166	104.374	148.75
286	1511	16.09368	104.4174	148.49
287	1513	16.12355	104.3241	146.03
288	818	14.96686	101.7471	148.55
289	786	14.93175	101.5181	155.72
290	837	14.80268	101.9039	144.13
291	787	14.89565	101.8172	155.63
292	838	14.64455	101.8719	143.72
293	805	14.7838	101.4775	151.38
294	843	14.50532	101.5456	142.62

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
295	2016	14.92157	101.4269	147.32
296	827	14.75053	101.8208	146.73
297	828	14.97657	101.6813	146.46
298	812	14.96585	101.5707	150.61
299	779	14.85589	101.5659	157.58
300	814	14.63851	101.5295	148.91
301	848	15.0205	101.6186	140.06
302	748	14.71601	101.4376	197.47
303	765	14.88992	101.7221	163.74
304	1055	14.68272	102.8787	155.28
305	2179	14.61839	103.2831	150.66
306	1092	14.78503	102.8479	139.32
307	1076	14.55606	102.9255	148.2
308	1044	14.95018	103.0914	159.18
309	1028	14.6117	103.076	180.45
310	1079	14.61532	103.0793	146.71
311	1080	14.53429	102.9798	146.65
312	1030	14.94493	103.0463	168.68
313	1067	14.80384	103.2403	151
314	1051	14.62857	103.1938	156.13
315	1035	14.8377	102.9715	165.85
316	1052	14.79467	103.0616	155.51
317	1069	14.72029	103.1531	150.29
318	1070	14.90321	103.1439	149.95
319	2159	14.71096	103.3014	155.77
320	1769	14.75353	104.6121	159.25
321	1753	14.62109	104.4171	172.15
322	1773	14.62559	104.4953	158.5
323	1757	14.8204	104.4609	165.32
324	1825	14.67617	104.5274	144.95
325	1776	14.74656	104.3572	157.92
326	1760	14.68787	104.4301	163.15
327	1811	14.62852	104.3521	148.79

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	คะแนนเฉลี่ย O-NET
328	1761	14.62657	104.5842	162.54
329	1762	14.57032	104.4525	162.29
330	1763	14.78217	104.6696	162.2
331	1831	14.68721	104.368	139.26
332	1783	14.75706	104.4577	156.53
333	1817	14.6565	104.4604	147.68
334	1784	14.86014	104.589	156.08
335	1751	14.63604	104.6489	184.58
คะแนน O-NET เฉลี่ยของกลุ่มตัวอย่าง				158.49
ความแปรปรวนกลุ่มตัวอย่าง				273.41
ค่าเบี่ยงเบนมาตรฐานกลุ่มตัวอย่าง				16.54

ภาคผนวก ค

คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบกริด-ชั้นภูมิ

คะแนนเฉลี่ย O-NET ของกลุ่มตัวอย่างด้วยวิธีการสุ่มแบบกริด-ชั้นภูมิ

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
1	10	13.7510919	100.502324	355	254.09
2	15	13.8143746	100.655561	513	247.53
3	21	13.771239	100.508722	293	233.97
4	27	13.7448794	100.482747	425	227.5
5	46	13.678323	100.623557	161	201.17
6	48	13.7037145	100.529297	196	200.18
7	58	13.7108316	100.43911	67	197.31
8	62	13.745985	100.757982	235	196.29
9	64	13.7194513	100.463443	107	195.8
10	71	13.7073239	100.546645	321	190.71
11	85	13.7756938	100.472377	33	184.03
12	105	13.5821836	100.420861	73	170.99
13	112	13.6953463	100.487387	40	165.48
14	113	13.7353651	100.432922	17	164.78
15	130	8.5466368	98.872799	118	147.33
16	134	14.535068	98.808975	60	222.14
17	144	14.7789861	99.3393924	103	163.31
18	148	14.3422017	99.7538047	30	160.58
19	154	14.7418561	98.6316923	54	156.56
20	156	15.1528896	98.4524341	37	154.45
21	158	14.6309454	99.47371	28	151.82
22	160	14.2243428	99.0713133	26	150.51
23	162	14.4483697	99.1342155	45	149.24
24	164	14.5494974	98.7841081	55	143.05
25	180	16.6537104	103.293455	172	154.15
26	198	16.5040537	104.182489	92	146.21
27	230	16.5487415	103.422194	21	130.07
28	237	16.0588563	99.8580408	96	165.91
29	241	16.1864066	99.4811793	34	159.3
30	247	16.1713806	99.7985449	45	158.23

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
31	248	16.5964486	99.8549118	81	158.15
32	250	16.1836777	99.3261051	99	156.67
33	254	16.1075833	99.3213004	41	153.59
34	257	16.4572893	99.3157355	21	152.13
35	258	16.5471847	99.7517738	26	151.52
36	260	16.3188662	99.4000236	27	150.2
37	261	15.9945181	99.7970836	67	147.92
38	263	16.411601	99.3837397	76	146.42
39	268	16.541204	102.113234	389	184.48
40	270	16.470496	102.853883	214	178.92
41	309	16.553229	102.948042	15	150.05
42	321	15.856035	102.89715	63	147.34
43	327	16.663699	101.911079	80	146.72
44	331	16.527234	102.995692	43	145.62
45	332	15.665013	102.78785	26	145.55
46	362	16.515639	102.115259	7	135.5
47	367	12.765954	101.851652	62	176.9
48	376	12.8234453	101.965763	3	162.67
49	379	13.2281564	102.277261	78	161.54
50	383	12.9376647	102.056713	31	155.13
51	391	13.684028	101.073395	260	184.77
52	392	13.7286037	101.377926	101	176.56
53	394	13.4659922	101.443504	17	168.9
54	395	13.604569	101.02066	45	167.98
55	405	13.7748947	101.205769	35	155.16
56	411	13.634341	101.076045	8	149.56
57	416	13.672011	101.433012	21	144.56
58	422	13.2949105	101.163335	112	219.16
59	440	13.4196103	101.334081	51	164.97
60	458	15.0490964	100.01106	57	156.32
61	462	15.262673	100.050535	70	150.7
62	463	15.2963025	100.188076	31	150.27
63	473	15.9228704	102.291186	111	158.73

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
64	477	16.2962909	101.934473	139	156.88
65	478	16.3429687	102.275698	86	156.41
66	482	15.8495085	102.067122	35	152.23
67	484	15.3891448	101.449126	78	150.88
68	494	15.8545928	101.418595	83	147.22
69	496	15.6558081	101.48484	60	147.18
70	509	16.0193437	101.488211	125	144.07
71	518	16.1800998	102.324152	19	142.36
72	521	16.2997409	102.022714	45	141.34
73	525	15.887673	101.792943	41	138.31
74	560	19.6095352	99.8462114	43	172.41
75	595	7.561168	99.60645	430	203.49
76	597	7.7894312	99.6307987	229	190.25
77	619	7.452595	99.633843	42	148.68
78	640	16.0317472	98.8587987	77	159.69
79	642	16.3841377	98.6914432	80	156.69
80	643	16.7617352	98.5887855	51	156.21
81	645	16.8057938	99.0529916	88	154.81
82	646	16.676657	99.1446811	42	154.31
83	648	16.9621	98.582082	52	152.69
84	650	16.8044244	98.765656	27	151.56
85	651	16.6566436	98.5907771	62	150.74
86	652	16.9816015	98.5249267	92	150.58
87	670	13.7889598	100.048913	279	199.41
88	678	13.8018984	100.299086	78	171.79
89	685	13.80699	100.217787	25	158.95
90	697	17.0503846	104.663872	133	168.62
91	701	16.9441772	104.509177	83	163.58
92	704	16.9506119	104.484728	151	161.15
93	705	17.48939	104.10239	85	160.85
94	711	17.2192875	104.639372	58	156.38
95	714	14.9185855	104.715236	37	152.72
96	715	17.8250663	104.150445	42	152.64

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
97	718	17.5845177	104.105936	11	151.73
98	721	17.8777455	104.209512	18	150.42
99	724	17.6265787	104.34603	17	149.07
100	726	17.20914	104.685328	44	147.68
101	729	17.477724	104.134989	25	145.07
102	730	17.1342967	104.755398	29	144.97
103	733	17.0532623	104.438691	34	144.1
104	734	17.5475782	104.21035	56	142.95
105	749	14.910331	102.069882	212	182.64
106	763	15.06691	102.12013	184	164.75
107	783	15.151327	102.510268	49	156.35
108	784	15.207011	101.786962	475	156.16
109	786	14.931749	101.518071	39	155.72
110	805	14.783799	101.477529	37	151.38
111	820	15.134788	101.601941	55	148.4
112	832	15.282198	101.55899	29	145.85
113	839	15.252102	101.842296	11	143.66
114	850	14.620651	102.142245	26	138.67
115	851	15.627472	102.686087	36	138.07
116	867	8.05881	99.892	32	171.27
117	878	8.16922	99.864443	90	162.1
118	918	15.710488	100.125238	518	247.54
119	919	15.2708536	100.361626	196	206.97
120	925	15.634865	100.129668	71	164.02
121	926	15.7661331	99.7993709	112	162.16
122	928	15.8642688	100.598832	167	159.68
123	936	15.5950223	100.085811	77	154.77
124	946	15.6314052	100.2774	8	147.56
125	948	15.1895391	100.531312	21	147.23
126	950	15.2781847	100.318593	54	146.69
127	953	15.8989422	100.030875	23	140.84
128	959	13.8811491	100.531233	340	214.73
129	968	13.933068	100.392927	151	180.79

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
130	973	14.9089711	101.980534	27	156.1
131	977	6.0233299	101.962073	103	180.91
132	978	6.4219583	101.82148	178	175.33
133	980	6.3971471	101.525647	32	158.1
134	984	6.0790574	101.876454	63	146.4
135	985	6.5117084	101.658981	33	144.59
136	987	6.1654157	101.699588	36	144.06
137	991	6.307002	101.718418	99	139.07
138	993	6.1446899	101.890808	6	133.13
139	998	19.2952426	100.861955	174	174.46
140	1005	18.712147	100.770035	75	162.38
141	1017	18.5839972	100.508576	19	153.83
142	1023	18.1811164	100.502871	28	147.45
143	1028	14.611704	103.075981	415	180.45
144	1041	14.639279	102.7639	118	161.09
145	1048	15.090136	103.176505	116	157.38
146	1064	14.6273	102.855088	44	151.92
147	1068	14.45079	102.838133	36	150.83
148	1070	14.903213	103.143924	26	149.95
149	1136	13.959614	101.692079	149	170.48
150	1140	13.837627	101.870135	55	156.98
151	1142	13.9770001	101.752449	33	153.33
152	1149	13.98677	101.623678	18	147.39
153	1157	14.069788	101.580365	29	141.89
154	1163	6.791346	101.495315	19	160.26
155	1172	6.5583749	101.536599	19	137.7
156	1179	19.5188841	100.311541	357	187.32
157	1184	19.3741761	99.78788	175	171.24
158	1192	19.1694524	100.278822	109	161.33
159	1196	8.8665612	98.3328771	157	188.22
160	1203	9.246222	98.418413	62	160.35
161	1216	7.8132182	99.9451811	70	160.47
162	1219	7.770536	100.110678	21	154

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
163	1233	16.2137253	100.440832	285	181.73
164	1234	16.0336772	100.375085	310	175.44
165	1237	16.3283751	100.571796	31	159.61
166	1247	16.2282013	100.355803	42	150.84
167	1250	16.0358403	100.180053	5	149.1
168	1251	15.9805554	100.227561	31	148.9
169	1259	16.560364	100.091921	28	141.17
170	1261	16.094336	100.585869	4	131.63
171	1273	16.6876758	100.432873	57	162.22
172	1275	16.758335	100.118473	48	160.66
173	1276	16.865229	100.194525	47	159.56
174	1286	16.4091135	100.681254	51	154.95
175	1296	17.152378	100.214555	16	147.08
176	1356	16.459096	101.227966	49	148.38
177	1361	16.68335	101.286567	12	146.52
178	1364	16.322618	101.185488	46	145.36
179	1376	17.8929135	99.5916437	150	162.97
180	1380	18.2627857	100.193528	92	155.93
181	1391	15.84928	103.37616	368	188.14
182	1394	16.2366786	103.068398	228	162.56
183	1396	15.8303712	103.207828	99	155.9
184	1401	15.9611296	103.339723	34	154.09
185	1406	15.6976514	103.126584	59	151.82
186	1409	15.827434	103.420924	79	149.72
187	1412	15.943932	103.406043	42	148.49
188	1416	16.324109	103.066429	45	147.79
189	1420	15.9455942	103.438528	40	146.59
190	1426	16.2148785	103.347911	19	144.53
191	1444	16.2473809	103.190478	25	131.08
192	1466	16.351397	104.382474	24	146.16
193	1488	15.6523364	104.310299	251	174.81
194	1497	15.8304228	104.380485	68	160.13
195	1504	15.8562101	104.281121	37	154.07

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
196	1507	16.0725466	104.196546	48	149.92
197	1515	15.7831168	104.319885	24	142.79
198	1522	6.589048	101.263717	16	147.48
199	1523	6.477415	101.135293	26	146.35
200	1524	6.266734	101.262802	17	139.13
201	1525	6.477002	101.431918	11	138.86
202	1533	15.733797	103.972877	85	166.41
203	1537	16.099753	103.844871	41	164.35
204	1547	16.387768	104.164735	118	157.39
205	1581	16.09748	103.99026	36	144.65
206	1582	16.237142	103.900572	29	143.97
207	1610	12.974366	101.220909	53	162.19
208	1651	14.855031	100.991347	89	153.93
209	1659	14.928743	101.182223	16	149.16
210	1665	14.726684	100.509313	4	132.69
211	1680	19.145404	99.609961	231	161.32
212	1693	17.911038	99.34105	155	148.22
213	1700	18.294316	98.815367	168	192.73
214	1703	18.436848	98.8071	49	163.22
215	1705	18.401995	98.82598	75	159.57
216	1712	18.016202	99.026226	37	152.47
217	1720	17.303425	101.801307	56	157.76
218	1722	17.624927	101.434504	85	156.87
219	1726	17.22556	101.863746	51	154.5
220	1728	17.897434	101.663186	157	153.58
221	1730	17.102927	101.867485	96	152.55
222	1735	17.249689	102.025223	72	146.69
223	1736	17.359652	101.250405	17	145.71
224	1737	17.803621	101.944727	82	145.69
225	1740	18.022879	101.895189	67	143.49
226	1741	17.549181	101.839238	34	140.54
227	1745	17.381509	101.81725	39	138.03
228	1747	17.328113	101.642764	35	135.35

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
229	1751	14.636036	104.648876	405	184.58
230	1769	14.753526	104.612085	30	159.25
231	1770	14.676934	104.847985	65	159.23
232	1779	14.540385	104.131709	47	156.68
233	1783	14.757064	104.45774	54	156.53
234	1789	14.941996	104.241363	98	155.22
235	1793	14.84723	104.783421	26	154.87
236	1796	14.92754	104.510874	78	152.42
237	1799	14.885113	103.989732	40	152.17
238	1812	15.139576	104.303341	14	148.71
239	1829	15.052626	104.634537	24	142.93
240	1844	17.847837	103.572611	101	169.15
241	1849	17.773467	103.95607	22	160.27
242	1854	17.046153	104.271742	145	157.29
243	1885	17.752521	103.708701	18	142.43
244	1894	7.204916	100.597046	538	206.97
245	1898	7.051961	100.528486	52	185.5
246	1900	7.842899	100.251489	16	177.91
247	1910	7.177222	100.260061	79	160.95
248	1913	7.66489	100.377789	18	157.1
249	1918	7.256857	100.396451	7	149.14
250	1929	7.073842	99.779345	67	162.48
251	1944	13.662939	100.53664	240	182.16
252	1945	13.630035	100.535948	204	181.65
253	1950	13.578829	100.857577	97	170.26
254	1964	13.425036	99.890002	12	150.25
255	1966	13.547999	100.275173	395	219
256	1969	13.651398	100.257322	209	179.97
257	1977	13.828378	102.060228	221	184.08
258	1985	13.417457	102.007606	31	156.85
259	1995	13.647321	102.00325	65	151.65
260	1997	13.592347	102.257963	12	143.5
261	1999	14.589216	100.995512	162	179.45

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
262	2001	14.735387	100.776713	83	170.34
263	2009	14.486957	100.784578	61	152.8
264	2016	14.921567	101.426881	36	147.32
265	2019	14.889087	100.417507	423	200.81
266	2022	14.803536	100.446504	22	162.9
267	2024	14.801734	100.340699	33	159.56
268	2028	15.025207	100.365171	12	148.46
269	2031	17.295986	99.838351	303	209.99
270	2047	16.954467	99.965592	44	149.68
271	2059	14.484929	100.123958	426	214.35
272	2066	14.455275	100.121793	71	167.75
273	2070	14.75161	100.092101	181	165.28
274	2076	14.199867	100.129518	124	159.68
275	2095	9.11183	99.237485	197	178.65
276	2096	9.385784	99.183661	121	177.14
277	2104	8.921812	98.909754	78	159.95
278	2109	9.142467	98.966096	28	156.73
279	2112	8.848778	98.804863	52	156
280	2115	9.095995	98.9851	44	154.12
281	2119	8.950558	98.99386	24	152.46
282	2123	8.737104	99.020929	54	150.75
283	2128	9.049858	99.08126	17	146.22
284	2129	9.137351	98.596115	18	142.6
285	2139	15.118709	103.609909	203	166.47
286	2164	14.807794	103.347147	79	154.37
287	2167	15.227838	103.561809	57	153.86
288	2169	14.961194	103.520758	23	153.08
289	2178	14.780158	103.540722	57	151
290	2188	15.075073	103.462643	43	147.35
291	2191	14.689121	103.737537	58	146.27
292	2200	14.643271	103.471145	34	144.74
293	2214	15.075905	103.539043	9	138.92
294	2241	17.780731	102.505303	46	153.26

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
295	2245	18.139222	103.384464	35	151.07
296	2250	17.940085	102.853143	22	149.81
297	2259	17.851804	102.431449	44	147.97
298	2261	18.092877	103.283404	58	146.68
299	2270	17.738621	102.702019	33	141.61
300	2275	17.660796	102.647007	27	140.31
301	2279	17.297981	102.212766	170	161.84
302	2284	17.624214	102.235401	25	151.89
303	2285	17.325734	102.066438	120	151.52
304	2294	17.560714	102.153282	15	144.65
305	2308	14.543039	100.5227	21	154.71
306	2319	14.4561581	100.319224	28	146.3
307	2320	14.2754213	100.573044	2	145.5
308	2322	14.3245343	100.555095	10	144.48
309	2336	14.6764834	100.457675	34	145.6
310	2349	15.9863039	104.747401	59	155.55
311	2360	15.5864788	104.57941	15	148.65
312	2364	17.4079288	102.777621	552	215.24
313	2373	17.6951938	102.89159	133	168.49
314	2377	17.3588578	103.105433	228	163.66
315	2396	17.5790195	102.988178	67	151.47
316	2403	17.8249999	103.327561	13	150.19
317	2412	17.2135032	103.268733	37	148.82
318	2420	17.6098221	103.376779	49	146.87
319	2439	17.6995965	100.121512	63	158.71
320	2457	15.3775076	100.01891	40	175.23
321	2470	15.1754289	99.7003875	64	149.85
322	2479	15.2766494	104.824707	207	170.81
323	2484	15.24602	104.899189	165	165.25
324	2507	14.4968682	104.997571	144	155.55
325	2514	15.6637943	104.943573	32	152.88
326	2515	15.5135294	104.732435	134	152.41
327	2534	15.3556643	104.462737	40	147.85

ลำดับที่	รหัสโรงเรียน	ละติจูด	ลองจิจูด	นักเรียนที่เข้าสอบ	คะแนนเฉลี่ย O-NET
328	2549	18.8423698	98.9636314	290	207.21
329	2554	18.737808	99.146976	176	180.69
330	2563	19.110055	98.930125	118	168.31
331	2564	18.173369	98.605554	76	166.95
332	2566	18.854222	98.965597	82	164.45
333	2574	18.607563	98.814235	44	158.09
334	2576	18.540947	98.798525	11	155.75
335	2579	17.80094	98.3569	70	152.86
คะแนน O-NET เฉลี่ยของกลุ่มตัวอย่าง					159.97
ความแปรปรวนกลุ่มตัวอย่าง					376.72
ค่าเบี่ยงเบนมาตรฐานกลุ่มตัวอย่าง					19.41

ประวัติย่อของผู้วิจัย

ชื่อ-สกุล	นาวาตรีพงศ์พัฒน์ อิศรกุล
วัน เดือน ปีเกิด	9 กุมภาพันธ์ พ.ศ. 2524
สถานที่เกิด	จังหวัดนครสวรรค์
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 204/ 33 ตำบลปากน้ำ อำเภอเมือง จังหวัดสมุทรปราการ
ตำแหน่งและประวัติการทำงาน	
พ.ศ. 2548-ปัจจุบัน	อาจารย์ฝ่ายศึกษาโรงเรียนนายเรือ
ประวัติการศึกษา	
พ.ศ. 2547	วิศวกรรมศาสตรบัณฑิต (วิศวกรรมไฟฟ้าอิเล็กทรอนิกส์) โรงเรียนนายเรือ
พ.ศ. 2551	วิทยาศาสตรมหาบัณฑิต (คณิตศาสตร์ประยุกต์) มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
พ.ศ. 2560	ปรัชญาดุษฎีบัณฑิต (วิชาวิจัย วัดผลและสถิติการศึกษา) มหาวิทยาลัยบูรพา