

Research Article

ความสัมพันธ์ระหว่างตัวแปรเสริมและตัวแปรหลักต่อ ประสิทธิภาพของตัวประมาณค่าเฉลี่ยประชากรจากการเลือก ตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

The correlation between auxiliary variable and primary variable on the population mean estimators of ranked set sampling

ภาณุมาศ ทัสละมัย^{1*}, วิจิตรา พลเยี่ยม¹ และ อรไท พลเสน¹

Panumas Tuslamai^{1*}, Wichitra Phonyiem¹ and Orathai Polsen¹

¹ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ บางซื่อ กรุงเทพฯ 10800

¹Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangsue, Bangkok 10800

*E-mail: puui.panumas@gmail.com

Received: 15/06/2018; Accepted: 31/10/2018

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ 4 วิธี คือ วิธีการเลือกตัวอย่างแบบกลุ่มลำดับพื้นฐาน วิธีการเลือกตัวอย่างแบบกลุ่มลำดับสุ่ม จิต วิธีการเลือกตัวอย่างแบบกลุ่มลำดับสมดุล และวิธีการเลือกตัวอย่างแบบกลุ่มลำดับสองชั้นเทียบกับวิธีการเลือกตัวอย่างแบบอย่างง่าย เมื่อมีการให้ลำดับจากตัวแปรเสริมโดยที่ตัวแปรหลักและตัวแปรเสริมมีระดับความสัมพันธ์ 0.1, 0.3, 0.5, 0.7 และ 0.9 และมีความคลาดเคลื่อนจากการให้ลำดับกลุ่ม โดยศึกษาวิธีการจำลองข้อมูลสองตัวแปร คือ การแจกแจงปกติสองตัวแปร และการแจกแจงลอการิธึมสองตัวแปร ประสิทธิภาพของตัวประมาณค่าจากวิธีการเลือกตัวอย่างพิจารณาจากค่าประสิทธิภาพสัมพัทธ์ ผลการวิจัยพบว่าตัวประมาณค่าที่มีประสิทธิภาพดีที่สุดในการ

การประมาณค่าเฉลี่ยนั้นขึ้นอยู่กับระดับความสัมพันธ์ระหว่างตัวแปรเสริมและตัวแปรหลัก และการแจกแจงของข้อมูล

คำสำคัญ: ตัวแปรเสริม, ตัวแปรหลัก, วิธีการเลือกตัวอย่างแบบกลุ่มลำดับ

Abstract

The purpose of this paper was to compare the efficiency of the population mean estimator of the 4 Ranked Set Samplings which were Ranked Set Sampling (RSS), Extreme Ranked Set Sampling (ERSS), Balanced Group Ranked Set Sampling (BGRSS) and Two-Stage Ranked Set Sampling (TSRSS) compared with the population mean estimator from Simple Random Sampling (SRS) when respectively sequencing from the auxiliary variables and the primary variable 0.1, 0.3, 0.5, 0.7 and 0.9 and Error in Ranking from sequenced groups study. Bivariate data simulation was Bivariate Normal Distribution and Bivariate Lognormal Distribution. The efficiency of the population mean estimators from there samplings technique was considered from Relative efficiency. The result revealed that the efficiency of the population for mean estimator depended on the correlation between auxiliary variable and primary variable and distributions.

Keywords: auxiliary variable, primary variable, Ranked Set Sampling

บทนำ

การสำรวจตัวอย่าง (sample survey) เป็นวิธีการทางสถิติที่สำคัญในการดำเนินงานวิจัย โดยที่ผู้วิจัยทำการเก็บรวบรวมข้อมูลในเรื่องที่สนใจศึกษาจากตัวอย่าง (sample) ซึ่งเป็นส่วนหนึ่งของประชากร วิธีการเลือกตัวอย่าง (sampling method) แบ่งออกเป็น 2 ประเภท คือ การเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็น (non-probability sampling) และการเลือกตัวอย่างแบบใช้ความน่าจะเป็น (probability sampling) แต่จะใช้วิธีการเลือกตัวอย่าง แบบใดนั้นต้องพิจารณาถึงลักษณะของประชากรเป็นหลัก ตัวอย่างที่ดีควรมีลักษณะคล้ายคลึงกับประชากรที่ศึกษา นอกจากวิธีการให้ได้มาซึ่งตัวอย่างแล้ว ข้อจำกัดที่สำคัญของตัวอย่างอีกประการหนึ่ง คือ ขนาดตัวอย่าง (sample size) ควรมีขนาดใหญ่เพียงพอ เพื่อให้ตัวอย่างนั้นเป็นตัวแทนที่ดีของประชากร ซึ่งในทางปฏิบัติก็มีปัจจัยภายนอกอื่น ๆ เข้ามาเกี่ยวข้อง ได้แก่ งบประมาณค่าใช้จ่ายจำกัด หน่วยประชากรอยู่กระจัดกระจาย หรือการเก็บข้อมูลลักษณะที่ต้องการศึกษาทำได้ยาก ทำให้ขนาดตัวอย่างไม่เป็นไปตามที่ต้องการ ในขณะที่ผู้วิจัยต้องการผลการวิจัยที่น่าเชื่อถือ และถูกต้องตามหลักวิชาสถิติ ด้วยเหตุนี้ จึงมีการศึกษาถึงวิธีการเลือกตัวอย่างเพื่อช่วยแก้ไขปัญหาดังกล่าวข้างต้น

ปี ค.ศ. 1952 McIntyre เสนอวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ (Rank Set Sampling: RSS) และทำการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ RSS กับตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบง่าย (SRS) และพบว่าตัวประมาณ RSS มีประสิทธิภาพดีกว่า ต่อมา Dell & Clutter (1972) เปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยประชากรจากวิธีการเลือกตัวอย่างแบบ SRS และ RSS โดยคำนึงถึงค่าใช้จ่ายในการเลือกตัวอย่างและมีความคลาดเคลื่อนในการให้ลำดับกลุ่ม (error in ranking) และพบว่า ตัวประมาณจากการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพดีกว่าตัวประมาณจากการเลือกตัวอย่างแบบ SRS ทั้งในกรณีที่ไม่มีและมีความคลาดเคลื่อนจากการให้ลำดับกลุ่ม

นอกจากวิธีการเลือกตัวอย่างแบบ RSS แล้ว ต่อ ๆ มา มีการพัฒนาวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ และนำเสนออย่างต่อเนื่อง อาทิเช่น การเลือกตัวอย่างแบบกลุ่มลำดับสุดขีด (Extreme Rank Set Sampling: ERSS) การเลือกตัวอย่างแบบกลุ่มลำดับสมดุล (Balance Group Rank Set Sampling: BGRSS) การเลือกตัวอย่างแบบกลุ่มลำดับสองขั้น (Two-Stage Ranked Set Sampling: TSRSS) ซึ่ง Al-Ommari & Al-Saleh (2010) ได้ศึกษาวิธีการเลือกตัวอย่างแบบ TSRSS และเปรียบเทียบตัวประมาณค่าเฉลี่ยประชากรจากการเลือกตัวอย่างแบบ TSRSS, RSS, ERSS และ BGRSS เทียบกับ SRS พบว่าเมื่อประชากรมีการแจกแจงแบบสมมาตร ตัวประมาณจากวิธีการเลือกตัวอย่างแบบ TSRSS มีประสิทธิภาพดีที่สุด และ Laongkaew (2015) พบว่า ถึงแม้ว่าจะมีความคลาดเคลื่อนจากการให้ลำดับกลุ่ม ตัวประมาณค่าเฉลี่ยจากการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพดีกว่าตัวประมาณจากวิธีการเลือกแบบ SRS

จากงานวิจัยที่ผ่านมาการให้ลำดับหน่วยในชุดตัวอย่างที่สุ่มจะเป็นการให้ลำดับตามลักษณะของตัวแปรที่ศึกษา เมื่อไม่มีและมีความผิดพลาดจากการให้ลำดับ แต่ในกรณีที่การวัดค่าลักษณะที่ต้องการศึกษาซึ่งถือเป็นตัวแปรหลักทำได้ยาก และการให้ลำดับค่าของตัวแปรหลักโดยการสังเกตด้วยสายตาอาจจะไม่ชัดเจน การให้ลำดับหน่วยในชุดตัวอย่างที่ถูกสุ่มโดยใช้ค่าวัดของตัวแปรเสริมจึงเป็นทางเลือกหนึ่ง แต่ทั้งนี้ความแม่นยำของการให้ลำดับหน่วยในชุดตัวอย่างน่าจะขึ้นอยู่กับความสัมพันธ์ระหว่างตัวแปรหลักและตัวแปรเสริมดังกล่าว ผู้วิจัยจึงจะทำการศึกษาถึงประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับโดยให้ลำดับหน่วยในชุดตัวอย่างที่สุ่มจากค่าของตัวแปรเสริม เมื่อตัวแปรหลักและตัวแปรเสริมมีความสัมพันธ์กันในระดับต่าง ๆ เปรียบเทียบกับตัวประมาณจากการเลือกตัวอย่างแบบง่าย

วิธีการทดลอง

งานวิจัยนี้ศึกษาโดยใช้วิธีการจำลองข้อมูลสองตัวแปรด้วยโปรแกรม R คือ การแจกแจงปกติสองตัวแปรใช้ฟังก์ชัน `rmultnorm` ในแพ็คเกจ (package) `MSBVAR` โดยที่ตัวแปรหลัก Y มีค่าเฉลี่ยเท่ากับ 2 และความแปรปรวนเท่ากับ 4 ตัวแปรเสริม X มีค่าเฉลี่ยเท่ากับ 2 และความแปรปรวนเท่ากับ 1 และการแจกแจงลอการิทึมสองตัวแปรใช้ฟังก์ชัน `rlnorm.rplus` ในแพ็คเกจ (package) `compositions` โดยตัวแปรหลัก Y และตัวแปรเสริม X มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 สร้างประชากรขนาด 10,000 โดยกำหนดความสัมพันธ์

ระหว่างตัวแปรเสริม X และตัวแปรหลัก Y เท่ากับ 0.1, 0.3, 0.5, 0.7, และ 0.9 เลือกตัวอย่างแบบอย่างง่ายขนาด mr โดยกำหนด $m = 3$ และ 6 และ $r = 2$ เลือกตัวอย่างแบบกลุ่มลำดับ โดยสุ่มตัวอย่างขนาด m จากประชากร (X, Y) สร้างความคลาดเคลื่อนจากการให้ลำดับกลุ่ม (ε) ขนาด m ที่มีการแจกแจงปกติมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวน (σ_ε^2) เท่ากับ 0.05 และ 0.30 โดยนำค่า ε ที่ได้บวกเข้ากับค่าของตัวแปรเสริม X ของหน่วยตัวอย่างที่สุ่มได้ ซึ่งใช้ในขั้นตอนการจัดลำดับ ดำเนินการตามแผนการเลือกตัวอย่างแบบกลุ่มลำดับทั้ง 4 วิธี ครอบรอบจะได้ตัวอย่างขนาด mr หน่วย กำหนดทำซ้ำเท่ากับ 100,000 รอบในแต่ละสถานการณ์ เหน้ที่ใช้ในการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าพิจรณาจากค่าประสิทธิภาพสัมพัทธ์ (R.E.) ระหว่างวิธีการเลือกตัวอย่างแบบกลุ่มลำดับและวิธีการเลือกตัวอย่างแบบอย่างง่าย วิธีการที่ให้ค่า R.E. สูงที่สุดจะเป็นวิธีการที่ดีที่สุด ในงานวิจัยครั้งนี้ได้ใช้วิธีการเลือกตัวอย่าง 4 วิธีดังนี้

1. การเลือกตัวอย่างแบบกลุ่มลำดับพื้นฐาน (Ranked Set Sampling: RSS)

McIntyre (1952) ได้เสนอการเลือกตัวอย่างแบบกลุ่มลำดับพื้นฐาน (RSS) มีกระบวนการสุ่มตัวอย่างดังนี้
ขั้นที่ 1 สุ่มตัวอย่างขนาด m หน่วย จำนวน m ชุด จากประชากร ให้ลำดับหน่วยตัวอย่างในแต่ละชุดจากน้อยไปมาก

ขั้นที่ 2 วัดค่าลักษณะที่สนใจของประชากรโดยที่

ชุดที่ 1 วัดค่าจากหน่วยตัวอย่าง ณ ลำดับที่ 1

ชุดที่ 2 วัดค่าจากหน่วยตัวอย่าง ณ ลำดับที่ 2

...

ชุดที่ m วัดค่าจากหน่วยตัวอย่าง ณ ลำดับที่ m

จะได้ค่าวัดลักษณะที่สนใจ จำนวน m ค่า ถ้าทำซ้ำขั้นที่ 1 และขั้นที่ 2 จำนวน r รอบจะได้ข้อมูลตัวอย่างขนาด mr หน่วย

ให้ $Y_{[a]b}; a = 1, 2, \dots, m; b = 1, 2, \dots, r$ เป็นตัวอย่างแบบกลุ่มลำดับที่มีขนาดชุดตัวอย่าง m หน่วย ทำการสุ่ม r รอบ ตัวประมาณค่าเฉลี่ย (\hat{Y}_{RSS}) ของประชากรจากการเลือกตัวอย่างแบบ RSS จำนวนดังนี้

$$\hat{Y}_{RSS} = \frac{1}{mr} \sum_{b=1}^r \sum_{a=1}^m Y_{[a]b} \quad (1)$$

2. การเลือกตัวอย่างแบบกลุ่มลำดับสุดขีด (Extream Ranked Set Sampling: ERSS)

Samawi et al. (1996) ได้เสนอการเลือกตัวอย่างแบบกลุ่มลำดับสุดขีด (ERSS) มีกระบวนการสุ่ม ดังนี้
ขั้นที่ 1 สุ่มตัวอย่างขนาด m หน่วย จำนวน m ชุดจากประชากร ให้ลำดับหน่วยตัวอย่างในแต่ละชุดจากน้อยไปมาก

ขั้นที่ 2 วัดค่าลักษณะที่สนใจของประชากร โดยที่

- ถ้า m เป็นจำนวนคู่

วัดค่าจากหน่วยตัวอย่าง ณ ตำแหน่งลำดับที่ 1 ของตัวอย่าง $\frac{m}{2}$ ชุดแรก

วัดค่าจากหน่วยตัวอย่าง ณ ตำแหน่งลำดับสูงสุดของตัวอย่าง $\frac{m}{2}$ ชุดหลัง (ชุดที่ $\frac{m}{2} + 1$ ถึง m)

- ถ้า m เป็นจำนวนคี่

วัดค่าจากหน่วยตัวอย่าง ณ ตำแหน่งลำดับที่ 1 ของตัวอย่าง $\frac{m-1}{2}$ ชุดแรก (ชุดที่ 1 ถึง $\frac{m-1}{2}$)

วัดค่าจากหน่วยตัวอย่าง ณ ตำแหน่งลำดับมัธยฐาน จากชุดตัวอย่างที่ $\frac{m+1}{2}$

วัดค่าจากหน่วยตัวอย่าง ณ ตำแหน่งลำดับสูงสุดของตัวอย่าง $\frac{m-1}{2}$ ชุดหลัง (ชุดที่ $\frac{m+3}{2}$ ถึง m)

จะได้ค่าวัดลักษณะที่สนใจ จำนวน m ค่า ถ้าทำซ้ำขั้นที่ 1 และขั้นที่ 2 จำนวน r รอบ จะได้ข้อมูลตัวอย่างขนาด mr หน่วย

ให้ $Y_{[a]b}; a = 1, 2, \dots, m; b = 1, 2, \dots, r$ เป็นตัวอย่างแบบกลุ่มลำดับที่มีขนาดชุดตัวอย่าง m หน่วย ทำการสุ่ม r รอบ ตัวประมาณค่าเฉลี่ย $\left(\hat{Y}_{ERSS}\right)$ ของประชากรจากการเลือกตัวอย่างแบบ ERSS คำนวณดังนี้

$$\hat{Y}_{ERSS} = \frac{1}{mr} \sum_{b=1}^r \sum_{a=1}^m Y_{[a]b} \quad (2)$$

3. วิธีการเลือกตัวอย่างแบบกลุ่มลำดับสมดุล (Balance Group Ranked Set Sampling: BGRSS)

Jemain et al. (2009) ได้เสนอการเลือกตัวอย่างแบบกลุ่มลำดับสมดุล (BGRSS) มีกระบวนการสุ่มดังนี้

ขั้นที่ 1 สุ่มตัวอย่างขนาด $m = 3k; k = 1, 2, \dots$ ชุดตัวอย่าง ขนาด m หน่วย จากประชากร

ขั้นที่ 2 ให้ลำดับขนาดหน่วยตัวอย่างในแต่ละชุดตัวอย่างจากน้อยไปมาก

ขั้นที่ 3 จัดแบ่ง $3k$ ชุดตัวอย่าง ออกเป็น 3 กลุ่มๆ ละ k ชุด โดยสุ่ม

กลุ่มที่ 1 เลือกหน่วยตัวอย่าง ณ ลำดับขนาดที่ 1 จากทุกชุด

กลุ่มที่ 2 เลือกหน่วยตัวอย่าง ณ ลำดับมัธยฐาน จากทุกชุด

กลุ่มที่ 3 เลือกหน่วยตัวอย่าง ณ ลำดับสุดท้าย จากทุกชุด

ครบ 1 รอบ จะได้ตัวอย่างขนาด $m = 3k$ หน่วย ถ้าทำซ้ำจำนวน r รอบ จะได้ตัวอย่างขนาด $3kr$ หน่วย

ให้ $Y_{[a]b}; a = 1, 2, \dots, m; b = 1, 2, \dots, r$ เป็นตัวอย่างแบบกลุ่มลำดับที่มีขนาดชุดตัวอย่าง m หน่วย ทำการสุ่ม r รอบ ตัวประมาณค่าเฉลี่ย $\left(\hat{Y}_{BGRSS}\right)$ ของประชากรจากการเลือกตัวอย่างแบบ BGRSS คำนวณดังนี้

$$\hat{\bar{Y}}_{BGRSS} = \frac{1}{mr} \sum_{b=1}^r \sum_{a=1}^m Y_{[a]b} \quad (3)$$

4. การเลือกตัวอย่างแบบกลุ่มลำดับสองขั้น (Two-Stage Ranked Set Sampling: TSRSS)

Al-Ommari & Al-Saleh (2010) ได้เสนอการเลือกตัวอย่างแบบกลุ่มลำดับสองขั้น (TSRSS) มีกระบวนการเลือก 4 ขั้นตอน ดังนี้

ขั้นที่ 1 สุ่มตัวอย่างขนาด $m^3 = 27k^3$; $k = 1, 2, \dots$ หน่วยจากประชากร แบ่ง $27k^3$ หน่วยตัวอย่างออกเป็น $m^2 = 9k^2$ ชุดตัวอย่าง ขนาด $3k$ หน่วย

ขั้นที่ 2 จัดแบ่ง $9k^2$ ชุดตัวอย่าง ออกเป็น 3 กลุ่มๆ ละ $3k^2$ ชุด

ขั้นที่ 3 ให้ลำดับขนาดหน่วยตัวอย่างในแต่ละชุดตัวอย่างจากน้อยไปมาก

กลุ่มที่ 1 เลือกหน่วยตัวอย่างลำดับต่ำที่สุดจากแต่ละชุดตัวอย่างในกลุ่ม

กลุ่มที่ 2 เลือกหน่วยตัวอย่าง ณ ลำดับมัธยฐานจากแต่ละชุดตัวอย่างในกลุ่ม

กลุ่มที่ 3 เลือกหน่วยตัวอย่างลำดับสูงสุดจากแต่ละชุดตัวอย่างในกลุ่ม

ในขั้นที่ 3 ในแต่ละกลุ่มจะประกอบด้วย k ชุดตัวอย่าง

ขั้นที่ 4 ให้ลำดับขนาดหน่วยตัวอย่างในแต่ละชุดตัวอย่างในขั้นที่ 3 จากน้อยไปมาก

กลุ่มที่ 1 เลือกหน่วยตัวอย่างลำดับต่ำที่สุดจากแต่ละชุดตัวอย่างในกลุ่ม

กลุ่มที่ 2 เลือกหน่วยตัวอย่าง ณ ลำดับมัธยฐานจากแต่ละชุดตัวอย่างในกลุ่ม

กลุ่มที่ 3 เลือกหน่วยตัวอย่างลำดับสูงสุดจากแต่ละชุดตัวอย่างในกลุ่ม

ทำครบ 1 รอบ จะได้ตัวอย่างขนาด $m = 3k$ หน่วยตัวอย่างเมื่อทำการเลือกครบ r รอบ จะได้หน่วยตัวอย่างทั้งหมด mr หน่วยตัวอย่าง

ให้ $Y_{[a]b}$; $a = 1, 2, \dots, m$; $b = 1, 2, \dots, r$ เป็นตัวอย่างแบบกลุ่มลำดับที่มีขนาดชุดตัวอย่าง m หน่วย

ทำการสุ่ม r รอบ ตัวประมาณค่าเฉลี่ย $\left(\hat{\bar{Y}}_{TSRSS}\right)$ ของประชากรจากการเลือกตัวอย่างแบบ TSRSS คำนวณดังนี้

$$\hat{\bar{Y}}_{TSRSS} = \frac{1}{mr} \sum_{b=1}^r \sum_{a=1}^m Y_{[a]b} \quad (4)$$

เกณฑ์ในการพิจารณาประสิทธิภาพของตัวประมาณค่า

เกณฑ์ที่ใช้ในการพิจารณาการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับทั้ง 4 วิธีกับวิธีการเลือกตัวอย่างแบบอย่างง่าย คือ ค่าประสิทธิภาพสัมพัทธ์ (Relative efficiency: R.E.)

$$R.E.(\theta_{SRS}, \theta_{RSS}) = \frac{MSE(\hat{\theta}_{SRS})}{MSE(\hat{\theta}_{RSS})} \quad (5)$$

โดยที่ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) สามารถคำนวณได้ดังนี้

$$MSE(\hat{\theta}) = \frac{\sum_{i=1}^j (\hat{\theta}_i - \theta_i)^2}{j} \quad (6)$$

โดยที่ θ_i คือ ค่าพารามิเตอร์จากการทำซ้ำรอบที่ i

$\hat{\theta}_i$ คือ ค่าประมาณของพารามิเตอร์จากการทำซ้ำรอบที่ i

j คือ จำนวนรอบของการจำลองซ้ำ

ถ้า $R.E. > 1$ แสดงว่าตัวประมาณจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับมีประสิทธิภาพดีกว่าตัวประมาณจากวิธีการเลือกตัวอย่างแบบอย่างง่าย

ถ้า $R.E. = 1$ แสดงว่าตัวประมาณจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับมีประสิทธิภาพเท่ากับตัวประมาณจากวิธีการเลือกตัวอย่างแบบอย่างง่าย

ถ้า $R.E. < 1$ แสดงว่าตัวประมาณจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับมีประสิทธิภาพน้อยกว่าตัวประมาณจากวิธีการเลือกตัวอย่างแบบอย่างง่าย

ผลการทดลอง

การเปรียบเทียบประสิทธิภาพสัมพัทธ์ (R.E.) ของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับทั้ง 4 วิธีกับวิธีการเลือกตัวอย่างแบบอย่างง่ายในสถานการณ์ต่าง ๆ เมื่อประชากรมีการแจกแจงปกติสองตัวแปร และการแจกแจงลอการิธึมสองตัวแปร มีผลการทดลองดังนี้

ตารางที่ 1 และ 2 แสดงผลการเปรียบเทียบประสิทธิภาพสัมพัทธ์ (R.E.) ของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับทั้ง 4 วิธีกับวิธีการเลือกตัวอย่างแบบอย่างง่าย กรณีมีและไม่มี ความคลาดเคลื่อนจากการให้ลำดับกลุ่ม พบว่าเมื่อประชากรมีการแจกแจงปกติสองตัวแปร ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ TSRSS มีประสิทธิภาพสูงกว่าตัวประมาณค่าเฉลี่ยที่ได้จากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับวิธีอื่นๆ ในทุกระดับความคลาดเคลื่อนจากการให้ลำดับกลุ่มและทุกระดับความสัมพันธ์ที่ศึกษา และเมื่อประชากรมีการแจกแจงลอการิธึมสองตัวแปร ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพสูงกว่าตัวประมาณค่าเฉลี่ยที่ได้จากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับวิธีอื่นๆ ในทุกระดับความคลาดเคลื่อนจากการให้ลำดับกลุ่มและทุกระดับความสัมพันธ์ที่ศึกษา นอกจากนี้เมื่อระดับความสัมพันธ์ของตัวแปรเสริม X และตัวแปรหลัก Y เพิ่มขึ้น ประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างยังดีขึ้น

ตารางที่ 1. ค่าประสิทธิภาพสัมพัทธ์ (R.E.) ของตัวประมาณค่าเฉลี่ย เมื่อมีความคลาดเคลื่อนจาก การให้ลำดับกลุ่ม

m	σ_ε^2	ρ	การแจกแจงปกติสองตัวแปร				การแจกแจงลอการิธึมสองตัวแปร			
			RSS	ERSS	BGRSS	TSRSS	RSS	ERSS	BGRSS	TSRSS
3	0.05	0.1	1.00153	1.00672	1.01887	1.06100	1.01611	1.00275	1.01240	0.66310
		0.3	1.03697	1.04149	1.05668	1.14119	1.05307	1.01124	1.04733	0.73226
		0.5	1.12053	1.13090	1.14792	1.28000	1.13326	1.02879	1.08558	0.81579
		0.7	1.27522	1.29135	1.31111	1.64515	1.18618	1.03947	1.14429	0.86182
		0.9	1.57509	1.59422	1.61013	2.35797	1.28464	1.10925	1.25059	0.90678
	0.3	0.1	1.01327	1.00360	1.00624	1.04970	1.00893	1.00097	1.00242	0.51651
		0.3	1.05153	1.02979	1.03575	1.06318	1.03013	1.00562	1.01421	0.57785
		0.5	1.11959	1.10042	1.10131	1.19241	1.08282	1.01463	1.06700	0.65092
		0.7	1.23766	1.21143	1.21946	1.39137	1.14516	1.02231	1.12460	0.77232
		0.9	1.44225	1.41006	1.42559	1.85980	1.25807	1.10229	1.24536	0.86031
6	0.05	0.1	1.01212	1.00711	1.23636	1.25142	1.00140	0.37609	0.70855	0.14088
		0.3	1.07331	1.01872	1.28943	1.29116	1.01164	0.52106	0.82162	0.18126
		0.5	1.21538	1.13064	1.41622	1.50397	1.11414	0.64950	0.90151	0.22085
		0.7	1.49061	1.35142	1.67714	1.94239	1.25926	0.74727	0.95178	0.47378
		0.9	2.16014	1.81602	2.22068	2.88751	1.47416	0.99394	0.96864	0.72730
	0.3	0.1	1.01884	1.00431	1.22765	1.24526	1.01284	0.35403	0.64627	0.11199
		0.3	1.06254	1.01269	1.26882	1.28122	1.03002	0.48226	0.78885	0.20248
		0.5	1.17520	1.10009	1.37004	1.39047	1.06986	0.61649	0.84106	0.30694
		0.7	1.36239	1.26231	1.56841	1.77205	1.20400	0.78848	0.88922	0.49315
		0.9	1.77585	1.56049	1.65215	2.29518	1.40366	0.97189	0.95794	0.57999

หมายเหตุ ตัวหนา คือ ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างที่มีประสิทธิภาพมากที่สุด

ตารางที่ 2. ประสิทธิภาพสัมพัทธ์ (R.E.) ของตัวประมาณค่าเฉลี่ย เมื่อไม่มีความคลาดเคลื่อนจากการให้ลำดับกลุ่ม

m	ρ	การแจกแจงปกติสองตัวแปร				การแจกแจงลอการิธึมสองตัวแปร			
		RSS	ERSS	BGRSS	TSRSS	RSS	ERSS	BGRSS	TSRSS
3	0.1	1.0100	1.00236	1.00777	1.04764	1.01024	1.00738	1.00225	0.62637
	0.3	1.05021	1.03223	1.04107	1.10860	1.08051	1.01154	1.02478	0.68098
	0.5	1.13826	1.12359	1.12959	1.19351	1.11803	1.07235	1.07123	0.79450
	0.7	1.30891	1.29658	1.29942	1.45535	1.14812	1.12174	1.13962	0.89293
	0.9	1.64303	1.62325	1.63114	2.30787	1.30846	1.29047	1.29912	0.93829
6	0.1	1.01715	1.00655	1.02494	1.35836	1.01307	0.41243	0.75711	0.10480
	0.3	1.05405	1.03532	1.08999	1.38146	1.03446	0.52208	0.86467	0.31665
	0.5	1.21111	1.15194	1.23216	1.66512	1.12641	0.66481	0.93485	0.49742
	0.7	1.50757	1.38752	1.70486	2.09945	1.23907	0.91062	0.9505	0.74229
	0.9	2.24799	1.88928	2.29852	3.22056	1.50658	0.98258	0.97033	0.87494

หมายเหตุ ตัวหนา คือ ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างที่มีประสิทธิภาพมากที่สุด

ตารางที่ 3. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณจากวิธีการเลือกตัวอย่างแบบอย่างง่าย

m	ρ	การแจกแจงปกติ สองตัวแปร	การแจกแจงลอการิธึม สองตัวแปร
3	0.1	0.67580	0.89298
	0.3	0.66108	0.82696
	0.5	0.65557	0.73987
	0.7	0.65246	0.73699
	0.9	0.63687	0.64954
6	0.1	0.34015	0.40148
	0.3	0.33221	0.39649
	0.5	0.33206	0.37802
	0.7	0.33024	0.35212
	0.9	0.32416	0.35143

วิจารณ์และสรุปผลการทดลอง

จากงานวิจัยที่ผ่านมาการให้ลำดับหน่วยในชุดตัวอย่างที่สุ่มจะเป็นการให้ลำดับตามลักษณะของตัวแปรที่ศึกษา แต่ในกรณีที่การวัดค่าลักษณะที่ต้องการศึกษาซึ่งถือเป็นตัวแปรหลักทำได้ยาก การให้ลำดับหน่วยโดยใช้ค่าวัดของตัวแปรเสริมจึงเป็นทางเลือกหนึ่ง แต่ทั้งนี้การให้ลำดับหน่วยในชุดตัวอย่างน่าจะขึ้นอยู่กับความสัมพันธ์ระหว่างตัวแปรหลักและตัวแปรเสริมดังกล่าว McIntyre (1952) ได้ทำการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ RSS กับตัวประมาณจากวิธีการเลือกตัวอย่างแบบง่าย (SRS) และพบว่าตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพดีกว่า Dell & Clutter (1972) เปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยประชากรจากวิธีการเลือกตัวอย่างแบบ SRS และ RSS โดยคำนึงถึงค่าใช้จ่ายในการเลือกตัวอย่างและมีความคลาดเคลื่อนในการให้ลำดับกลุ่ม และพบว่าตัวประมาณจากการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพดีกว่าตัวประมาณจากการเลือกตัวอย่างแบบ SRS ทั้งในกรณีที่ไม่มีและมีความคลาดเคลื่อนจากการให้ลำดับกลุ่ม Al-Ommari & Al-Saleh (2010) ได้ศึกษาวิธีการเลือกตัวอย่างแบบ TSRSS และเปรียบเทียบตัวประมาณค่าเฉลี่ยประชากรจากการเลือกตัวอย่างแบบ TSRSS, RSS, ERSS และ BGRSS เทียบกับตัวประมาณค่าเฉลี่ยจาก SRS พบว่าเมื่อประชากรมีการแจกแจงแบบสมมาตรตัวประมาณจากวิธีการเลือกตัวอย่างแบบ TSRSS มีประสิทธิภาพดีที่สุด ผู้วิจัยจึงทำการศึกษาถึงประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ โดยให้ลำดับหน่วยจากชุดตัวอย่างที่สุ่มได้จากค่าของตัวแปรเสริม เมื่อตัวแปรหลักและตัวแปรเสริมมีความสัมพันธ์กันในระดับต่าง ๆ

ผลการวิจัยเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ 4 วิธี กับตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบอย่างง่าย เมื่อตัวแปรหลักและตัวแปรเสริมมีความสัมพันธ์กัน โดยที่มีและไม่มี ความคลาดเคลื่อนจากการให้ลำดับกลุ่ม พบว่าเมื่อประชากรมีการแจกแจงปกติสองตัวแปร ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ TSRSS มีประสิทธิภาพดีกว่าตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับวิธีอื่นๆ ในทุกระดับความสัมพันธ์และทุกระดับความคลาดเคลื่อนจากการให้ลำดับกลุ่มในกรณีที่ศึกษา แต่เมื่อประชากรมีการแจกแจงลอการิธึมสองตัวแปร ตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบ RSS มีประสิทธิภาพดีกว่าตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับวิธีอื่น ๆ ในทุกระดับความสัมพันธ์และทุกระดับความคลาดเคลื่อนจากการให้ลำดับกลุ่มในกรณีที่ศึกษา นอกจากนั้นประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างขึ้นอยู่กับระดับความสัมพันธ์ระหว่างตัวแปรหลักและตัวแปรเสริม กล่าวคือ เมื่อระดับความสัมพันธ์ของตัวแปรหลักและตัวแปรเสริมเพิ่มขึ้น ประสิทธิภาพของตัวประมาณค่าเฉลี่ยจากวิธีการเลือกตัวอย่างแบบกลุ่มลำดับ 4 วิธีก็มีแนวโน้มเพิ่มขึ้น

เอกสารอ้างอิง

- Al-Omari, A. I. & Al-Saleh, M. F. (2010). Improvement in estimating the population mean using two-stage balanced groups Ranked Set Sampling. *METRON - International Journal of Statistics*, 68(2), 185-196.
- Dell, T. R. & Clutter, J. L. (1972) Ranked set sampling theory with order statistics background. *Biometric*, 28(2), 545-555.
- Jemain, A. A., Al-Omari, A., & Ibrahim, K. (2009). Balanced groups ranked set samples for estimating the population median. *Journal of Applied Statistical Science*, 17(1), 39-46.
- Laongkaew, M. (2015) *A comparison on parameter estimates from a simple random sample and a ranked set sample* (M.S. Thesis) Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok. (in Thai)
- McIntyre, G. A. (1952). A method of unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research*, 3(4), 385-390. <https://doi.org/10.1071/AR9520385>
- Samawi, H. M., Ahmed, M. S. & Abu-Dayyeh, W. (1996). Estimating the population mean using Extreme Ranked Set Sampling. *Biometrical Journal*, 38(5), 577-586.