ตัวประมาณค่าด้วยวิธีของซิลส์ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียว

Estimator in Single-Stage Cluster Sampling: Searls Approach

ชัชวาลย์ กองน้ำ* จิราวัลย์ จิตรถเวช และ วิชิต หล่อจีระชุณห์กุล

Chatchawan Kongnam*, Jirawan Jitthavech and Vichit Lorchirachoonkul

คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

School of Applied Statistics, National Institute of Development Administration

Received : 7 November 2016

Accepted : 28 December 2016

Published online : 31 January 2017

บทคัดย่อ

งานวิจัยนี้ได้นำเสนอตัวประมาณค่าเฉลี่ยประชากรโดยใช้วิธีการของซิลส์ สำหรับการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่าย ไม่คืนที่ ตัวประมาณค่าตัวแรก พัฒนาจากค่าเฉลี่ยของกลุ่มเมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และตัวประมาณค่าที่สอง พัฒนาจากค่าเฉลี่ยแบบอัตราส่วน เมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม ข้อมูลที่ใช้เปรียบเทียบประสิทธิภาพของตัวประมาณค่าทั้งสอง เป็นข้อมูลที่มีการเผยแพร่แล้วที่ประชากรมีการแบ่งออกเป็น 20 กลุ่ม ซึ่งจะสุ่มเลือกตัวอย่างขนาดเท่ากับ 5, 10 และ 15 กลุ่ม แต่ละขนาดตัวอย่างจะสุ่มเลือกทั้งหมด 30 ครั้ง พบว่าตัวประมาณทั้งสองตัวที่พัฒนาด้วยวิธีของซิลส์มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยประชากรแบบเดิมทุกกรณี ภายใต้เงื่อนไขที่หามาได้

คำสำคัญ : การเลือกตัวอย่างแบบกลุ่ม สัมประสิทธิ์การแปรผัน สัมประสิทธิ์สหสัมพันธ์ ค่าคลาดเคลื่อนกำลังสองเฉลี่ย ประสิทธิภาพสัมพัทธ์

Abstract

This paper presents two population mean estimators using Searls approach in single-stage cluster sampling with simple random sampling without replacement. The first estimator is developed based on the cluster mean when the coefficient of variation of cluster mean is known and the second based on the ratio of the cluster total mean and cluster size mean when the coefficients of variations of cluster total and cluster size are known. The population for efficiency comparisons of the two proposed estimators is a published data and consists of 20 clusters. Three sample sizes, 5, 10 and 15, are randomly selected from the population and 30 replications are performed for each sample size. It is found that the efficiencies of both estimators using Searls approach are higher than the traditional estimator in all cases under the conditions consistent with the derived conditions.

Keywords : cluster sampling, coefficient of variation, correlation coefficient, MSE, relative efficiency

*Corresponding author. E-mail : ch.kongnam@gmail.com

บทนำ

การเลือกตัวอย่างแบบกลุ่ม (Cluster Sampling) เป็นวิธีการเลือกตัวอย่างที่ถือว่าหน่วยตัวอย่างที่จะถูกเลือกมารวมอยู่ในตัวอย่างนั้น คือกลุ่ม (Cluster) ของหน่วยเล็กหลายหน่วยของประชากร โดยที่ลักษณะของประชากรภายในกลุ่มมีความแตกต่างกัน แต่ระหว่างกลุ่มมีลักษณะของประชากรภายในกลุ่มใกล้เคียงกัน เมื่อเลือกได้กลุ่มใดจะเก็บรวบรวมข้อมูลจากทุกหน่วยเล็กภายในกลุ่มนั้น ซึ่งเป็นการเลือกตัวอย่างขั้นเดียว (Single-Stage Sampling) ข้อดีของการเลือกตัวอย่างแบบกลุ่มคือการเตรียมกรอบตัวอย่าง (Sampling Frame) ในขอบเขตที่ศึกษาทำได้ง่าย ส่วนข้อเสียคือการประมาณค่าพารามิเตอร์ (Parameter) ในการเลือกตัวอย่างแบบกลุ่มนี้ พบว่ามักมีค่าคลาดเคลื่อนค่อนข้างสูง (Suwutthee, 2009) เนื่องจากในการเลือกตัวอย่างแบบกลุ่มจะสุ่มเลือกเพียงบางกลุ่มมาเป็นตัวแทนเพื่อประมาณค่าเฉลี่ยประชากรซึ่งแต่ละกลุ่มนั้นมีขนาดไม่เท่ากัน และระหว่างกลุ่มนั้นจะประกอบด้วยค่าสังเกตภายในกลุ่มที่เหมือนๆ กัน เพราะระหว่างกลุ่มมีลักษณะของประชากรภายในกลุ่มใกล้เคียงกัน ซึ่งอาจลดการเป็นตัวแทนที่ดีของตัวอย่าง และอาจทำให้ค่าคลาดเคลื่อนจากการเลือกตัวอย่างเพิ่มขึ้น ดังนั้นผู้วิจัยจึงสนใจศึกษาและพัฒนาตัวประมาณค่าเฉลี่ยประชากรเพื่อต้องการลดค่าคลาดเคลื่อนดังกล่าว

การพัฒนาตัวประมาณค่าเฉลี่ยประชากรในการศึกษานี้ใช้วิธีการของซิลส์ (Searls, 1964) ซึ่งอาศัยแนวคิดที่ต้องทราบค่าสัมประสิทธิ์การแปรผัน (Coefficient of Variation) ของประชากร ช่วยทำให้ตัวประมาณค่าเฉลี่ยประชากรแบบใหม่มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยประชากรแบบดั้งเดิมซึ่งเป็นตัวประมาณที่ไม่เอนเอียง (Unbiased Estimator) แต่ตัวประมาณค่าเฉลี่ยประชากรด้วยวิธีของซิลส์เป็นตัวประมาณที่เอนเอียง (Biased Estimator) วิธีการประมาณค่าของซิลส์ยอมให้ตัวประมาณค่าเฉลี่ยประชากรเป็นตัวประมาณที่เอนเอียง เพื่อให้ตัวประมาณที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำลงทำให้ตัวประมาณค่ามีประสิทธิภาพดีขึ้น

Jitthavech และ Lorchirachoonkul (2013) ได้นำวิธีของซิลส์มาใช้พัฒนาตัวประมาณค่าเฉลี่ยประชากรในการเลือกตัวอย่างสุ่มแบบง่าย (Simple Random Sampling) โดยการนำค่าสัมประสิทธิ์การแปรผันของประชากรของตัวแปรที่ศึกษาและค่าสัมประสิทธิ์การแปรผันของประชากรของตัวแปรช่วย ซึ่งเป็นตัวที่ทราบค่า มาพัฒนาตัวประมาณค่าเฉลี่ยประชากร จากนั้นทำการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณค่าเฉลี่ยประชากรแบบใหม่กับตัวประมาณค่าเฉลี่ยประชากรแบบดั้งเดิม และจากผลการศึกษาในงานวิจัยพบว่าตัวประมาณแบบถดถอยเชิงเส้น (Linear Regression Estimators) ที่ถูกพัฒนาขึ้นด้วยวิธีของซิลส์มีประสิทธิภาพมากที่สุด ในการเลือกตัวอย่างสุ่มแบบง่าย

การพัฒนาตัวประมาณค่าในงานวิจัยนี้พัฒนาจากวิธีของซิลส์ของการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่าย ไม่คืนที่ เนื่องจากวิธีของซิลส์นั้นอาศัยการทราบค่าสัมประสิทธิ์การแปรผันของประชากร เมื่อนำมาใช้แล้วจะทำให้ตัวประมาณที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำลงจึงทำให้ตัวประมาณค่ามีประสิทธิภาพดีขึ้น โดยจะพัฒนาตัวประมาณค่า 2 ตัว โดยใช้หลักการการประมาณค่าแบบเดิมและตัวประมาณค่าแบบอัตราส่วนของค่าเฉลี่ยซึ่งจะนำเสนอทฤษฎีการประมาณค่าของการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่าย ไม่คืนที่ รวมทั้งการประมาณค่าเฉลี่ยประชากรโดยวิธีดั้งเดิมและวิธีอัตราส่วน หัวข้อถัดไปจะเป็นการพัฒนาตัวประมาณค่าแบบซิลส์ที่ปรับจากวิธีดั้งเดิมและวิธีอัตราส่วน ภายใต้ข้อกำหนดที่ทราบค่าสัมประสิทธิ์การแปรผันของประชากรของตัวแปรที่ศึกษา และค่าสัมประสิทธิ์การแปรผันของประชากรของตัวแปรช่วย พร้อมทั้งทฤษฎีบทในการพิสูจน์คุณสมบัติและเงื่อนไขของตัวประมาณค่าที่พัฒนาขึ้นมา หัวข้อสุดท้ายจะเป็นการใช้ตัวประมาณค่าที่ได้พัฒนาขึ้นมาใช้กับข้อมูลจากหนังสือของคอแครน (Cochran, 1977) เพื่อศึกษาถึงความสอดคล้องกับทฤษฎีที่พัฒนาขึ้น

ทบทวนทฤษฎีที่เกี่ยวข้อง

ในการดำเนินการวิจัย ผู้วิจัยได้ศึกษาทฤษฎีที่เกี่ยวข้องกับการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่าย ไม่คืนที่ และนำข้อกำหนดของซิลส์มาพัฒนาตัวประมาณค่าเฉลี่ยประชากร และศึกษาเงื่อนไขที่สอดคล้องกับตัวประมาณค่า

ตัวประมาณค่าเฉลี่ยประชากรที่เลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายไม่คืนที่

พิจารณาแบ่งประชากรออกเป็น N กลุ่มที่มีลักษณะของค่าสังเกต y_{ij} ภายในกลุ่มแตกต่างกันและลักษณะของค่าสังเกตระหว่างกลุ่มมีลักษณะคล้ายคลึงกัน และกลุ่มที่ i มีขนาดของกลุ่ม M_i หน่วย โดยที่ $i = 1, 2, \cdots, N$ กำหนดให้ y_{ij} เป็นค่าสังเกตที่สนใจศึกษา จากหน่วยสังเกตที่ j ในกลุ่มที่ i ของประชากร โดยที่ $j = 1, 2, \cdots, M_i$, $i = 1, 2, \cdots, N$

และกำหนดให้ $Y_i = \sum_{j=1}^{M_i} y_{ij}$, $\overline{Y}_i = \frac{1}{M_i}\sum_{j=1}^{M_i} y_{ij}$ เป็นค่ายอดรวม และค่าเฉลี่ยของค่าสังเกต y_{ij} ในกลุ่มที่ i ตามลำดับ

$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{\sum_{i=1}^{N} M_i}$ เป็นค่าเฉลี่ยประชากร

$\overline{M} = \frac{1}{N}\sum_{i=1}^{N} M_i$, $\overline{\overline{Y}} = \frac{1}{N}\sum_{i=1}^{N} \overline{Y}_i$ เป็นขนาดเฉลี่ยของกลุ่ม และค่าเฉลี่ยของค่าเฉลี่ยของกลุ่มในประชากร ตามลำดับ

$C_{\overline{Y}}$, C_Y และ C_M เป็นค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม ของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม

ดังนั้น ตัวประมาณค่าเฉลี่ยประชากรในการเลือกตัวอย่างแบบกลุ่ม ได้แก่

1. ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายขนาด n กลุ่ม จากประชากรขนาด N กลุ่ม โดยไม่คืนที่ เมื่อทราบค่า M_i, $i = 1, 2, \cdots, n$ จะได้ตัวประมาณค่าเฉลี่ยของกลุ่ม ที่เป็นตัวประมาณที่เอนเอียงของค่าเฉลี่ยประชากร $\overline{Y}$ ดังนี้ (Gupta and Kabe, 2011)

$$\overline{y}_1 = \frac{1}{n}\sum_{i=1}^{n} \overline{Y}_i \tag{1}$$

ค่าเอนเอียงของตัวประมาณ $\overline{y}_1$ เท่ากับ

$$\text{Bias}(\overline{y}_1) = -\frac{N-1}{N}\rho_{\overline{Y}M} C_{\overline{Y}} C_M \overline{\overline{Y}} \tag{2}$$

โดยที่ $\rho_{\overline{Y}M}$ เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าเฉลี่ยของกลุ่มกับขนาดของกลุ่ม

ความแปรปรวนของตัวประมาณ $\overline{y}_1$ เท่ากับ

$$V(\overline{y}_1) = \gamma C_{\overline{Y}}^2 \overline{\overline{Y}}^2 \tag{3}$$

โดยที่ $\gamma = \frac{1}{n} - \frac{1}{N}$ เป็นปัจจัยปรับค่ากรณีที่ประชากรมีจำนวนที่จำกัด

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\overline{y}_1$ เท่ากับ

$$\mathrm{MSE}(\bar{y}_1) = \gamma C_Y^2 \bar{\bar{Y}}^2 + \left(\frac{N-1}{N}\right)^2 \rho_{YM}^2 C_Y^2 C_M^2 \bar{\bar{Y}}^2 \quad (4)$$

2. ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายขนาด n กลุ่ม จากประชากรขนาด N กลุ่ม โดยไม่คืนที่ เมื่อทราบค่า M_i, $i = 1, 2, \cdots, n$ จะได้ตัวประมาณค่าเฉลี่ยแบบอัตราส่วน ที่เป็นตัวประมาณที่เอนเอียงของค่าเฉลี่ยประชากร $\bar{Y}$ ดังนี้ (Gupta and Kabe, 2011)

$$\bar{y}_2 = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} M_i} \quad (5)$$

ค่าเอนเอียงของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่ง (First-Order Taylor Series Approximation) เท่ากับ

$$\mathrm{Bias}(\bar{y}_2) \approx \bar{Y}\left(\gamma C_M^2 - \gamma\rho_{YM} C_Y C_M\right) \quad (6)$$

โดยที่ ρ_{YM} เป็นค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่ายอดรวมของกลุ่มกับขนาดของกลุ่ม

ความแปรปรวนของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่ง เท่ากับ

$$\mathrm{V}(\bar{y}_2) \approx \bar{Y}^2\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM} C_Y C_M\right) \quad (7)$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่ง เท่ากับ

$$\mathrm{MSE}(\bar{y}_2) \approx \bar{Y}^2\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM} C_Y C_M\right) \quad (8)$$

ค่าคลาดเคลื่อนกำลังสองเฉลี่ยและความแปรปรวนของตัวประมาณ $\bar{y}_2$ ในสมการที่ (7) และ (8) มีค่าเท่ากัน เนื่องจากใช้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่งเท่านั้น

ผลการวิจัยและวิจารณ์ผล

ผลการวิจัยเชิงทฤษฎี

ตัวประมาณค่าเฉลี่ยประชากรที่พัฒนามาจากวิธีของซิลส์

ตัวประมาณค่าเฉลี่ยประชากรโดยใช้วิธีการของซิลส์ สำหรับการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่าย ไม่คืนที่ ตัวประมาณค่าตัวแรก พัฒนาจากค่าเฉลี่ยของกลุ่ม เมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และตัวประมาณค่าที่สอง พัฒนาจากค่าเฉลี่ยแบบอัตราส่วน เมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม มีทฤษฎีที่เกี่ยวข้อง 4 ทฤษฎีดังนี้

ทฤษฎีบทที่ 1 ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายขนาด n กลุ่ม จากประชากรขนาด N กลุ่ม โดยไม่คืนที่ เมื่อกำหนดให้ $\bar{Y}_i$ เป็นค่าเฉลี่ยของกลุ่มที่ $i\,;\ i = 1, 2, \ldots, n$ และทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่มในประชากร ($C_{\bar{Y}}$) และค่าสัมประสิทธิ์การแปรผันของขนาดของกลุ่มในประชากร (C_M) เป็นตัวที่ทราบค่า และ $\gamma C_{\bar{Y}}^2$ มีค่าน้อยกว่า 1 จะได้ตัวประมาณค่าเฉลี่ยประชากรที่เอนเอียงของค่าเฉลี่ยประชากร $\bar{Y}$ ด้วยวิธีของซิลส์ที่พัฒนาจากค่าเฉลี่ยของกลุ่ม ดังนี้

$$\bar{y}_{1_s} = \left(\frac{1}{1 + \gamma C_{\bar{Y}}^2} \right) \frac{\sum_{i=1}^{n} \bar{Y}_i}{n} \tag{9}$$

ค่าเอนเอียงของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ $\mathrm{Bias}\left(\bar{y}_{1_s}\right) = \mathrm{Bias}\left(\bar{y}_1\right) - \dfrac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}}{1 + \gamma C_{\bar{Y}}^2}$ (10)

ความแปรปรวนของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ $\mathrm{V}\left(\bar{y}_{1_s}\right) = \dfrac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}^2}{\left(1 + \gamma C_{\bar{Y}}^2\right)^2}$ (11)

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ

$$\mathrm{MSE}\left(\bar{y}_{1_s}\right) = \mathrm{MSE}\left(\bar{y}_1\right) - \left[\frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1 + \gamma C_{\bar{Y}}^2\right)} - 2\left(\frac{N-1}{N}\right) \rho_{\bar{Y}M} \frac{\gamma C_M C_{\bar{Y}}^3 \bar{\bar{Y}}^2}{\left(1 + \gamma C_{\bar{Y}}^2\right)} \right] \tag{12}$$

พิสูจน์ ตัวประมาณค่าด้วยวิธีของซิลส์ที่นำเสนอเมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่มเป็นดังนี้

$$\bar{y}_{1_s} = \left(\frac{1}{1 + \gamma C_{\bar{Y}}^2} \right) \frac{\sum_{i=1}^{n} \bar{Y}_i}{n}$$

ค่าเอนเอียงของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ

$$\begin{aligned}
\mathrm{Bias}\left(\bar{y}_{1_s}\right) &= \frac{\bar{\bar{Y}}}{1 + \gamma C_{\bar{Y}}^2} - \bar{Y} \\
&= \mathrm{Bias}\left(\bar{y}_1\right) + \frac{\bar{\bar{Y}}}{1 + \gamma C_{\bar{Y}}^2} - \bar{\bar{Y}} \\
&= \mathrm{Bias}\left(\bar{y}_1\right) - \frac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}}{1 + \gamma C_{\bar{Y}}^2}
\end{aligned} \tag{13}$$

ความแปรปรวนของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ

$$\mathrm{V}\left(\bar{y}_{1_s}\right) = \frac{1}{\left(1+\gamma C_{\bar{Y}}^2\right)^2}\mathrm{V}\left(\bar{y}_1\right) = \frac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)^2} \tag{14}$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_{1_s}$ เท่ากับ

$$\begin{aligned}
\mathrm{MSE}\left(\bar{y}_{1_s}\right) &= \frac{\mathrm{V}\left(\bar{y}_1\right)}{\left(1+\gamma C_{\bar{Y}}^2\right)^2} + \mathrm{Bias}^2\left(\bar{y}_1\right) - 2\,\mathrm{Bias}\left(\bar{y}_1\right)\frac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}}{\left(1+\gamma C_{\bar{Y}}^2\right)} + \frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)^2} \\
&= \mathrm{MSE}\left(\bar{y}_1\right) + \frac{\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)^2} + \frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)^2} - \gamma C_{\bar{Y}}^2 \bar{\bar{Y}}^2 - \frac{2\,\mathrm{Bias}\left(\bar{y}_1\right)\gamma C_{\bar{Y}}^2 \bar{\bar{Y}}}{\left(1+\gamma C_{\bar{Y}}^2\right)} \\
&= \mathrm{MSE}\left(\bar{y}_1\right) - \left[\frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)} - 2\left(\frac{N-1}{N}\right)\rho_{\bar{Y}M}\,\frac{\gamma C_M C_{\bar{Y}}^3 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)}\right]
\end{aligned} \tag{15}$$

ทฤษฎีบทที่ 2 ตัวประมาณ $\bar{y}_{1_s}$ จะมีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_1$ ภายใต้เงื่อนไข

$$\rho_{\bar{Y}M} < \frac{\gamma C_{\bar{Y}}}{2\left(\frac{N-1}{N}\right)C_M} \tag{16}$$

เมื่อ $\rho_{\bar{Y}M}$ เป็นค่าสัมประสิทธิ์สหสัมพันธ์ของค่าเฉลี่ยของกลุ่มและขนาดของกลุ่มในประชากร γ เป็นปัจจัยปรับค่ากรณีประชากรมีจำนวนที่จำกัด $C_{\bar{Y}}$, C_M เป็นค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และของขนาดของกลุ่มตามลำดับ และ N เป็นจำนวนกลุ่มของประชากร

พิสูจน์ ตัวประมาณ $\bar{y}_{1_s}$ จะมีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_1$ เมื่อ $\mathrm{MSE}(\bar{y}_1) > \mathrm{MSE}(\bar{y}_{1_s})$ ซึ่งเขียนได้ดังนี้

$$\mathrm{MSE}\left(\bar{y}_1\right) > \mathrm{MSE}\left(\bar{y}_1\right) - \left[\frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)} - 2\left(\frac{N-1}{N}\right)\rho_{\bar{Y}M}\,\frac{\gamma C_M C_{\bar{Y}}^3 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)}\right]$$

$$\frac{\gamma^2 C_{\bar{Y}}^4 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)} > 2\left(\frac{N-1}{N}\right)\rho_{\bar{Y}M}\,\frac{\gamma C_M C_{\bar{Y}}^3 \bar{\bar{Y}}^2}{\left(1+\gamma C_{\bar{Y}}^2\right)}$$

$$\rho_{\bar{Y}M} < \frac{\gamma C_{\bar{Y}}}{2\left(\frac{N-1}{N}\right)C_M} \tag{17}$$

ในกรณีไม่ทราบค่า $C_{\bar{Y}}$ และ C_M อาจประมาณได้จากตัวอย่างด้วย $\hat{C}_{\bar{Y}} = \dfrac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\bar{Y}_i - \bar{\bar{y}}\right)^2}}{\bar{\bar{y}}}$ และ $\hat{C}_M = \dfrac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(M_i - \bar{m}\right)^2}}{\bar{m}}$ โดยที่ $\bar{\bar{y}} = \dfrac{1}{n}\sum_{i=1}^{n}\bar{Y}_i$ และ $\bar{m} = \dfrac{1}{n}\sum_{i=1}^{n}M_i$

ในการพัฒนาตัวประมาณค่าแบบซิลส์จากค่าเฉลี่ยแบบอัตราส่วน จะใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง (Second-Order Taylor Series Approximation) ซึ่งจะทำการพิสูจน์ตามทฤษฎีบทที่ 3 เพื่อให้มีความแม่นยำในการหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยมากขึ้น ซึ่งทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าแตกต่างกับความแปรปรวนของตัวประมาณ $\bar{y}_2$

ทฤษฎีบทที่ 3 ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายขนาด n กลุ่ม จากประชากรขนาด N กลุ่ม โดยไม่คืนที่ เมื่อทราบค่า $M_i,\ i = 1, 2, \cdots, n$ จะได้ความแปรปรวนของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง (Second-Order Taylor Series Approximation) เท่ากับ

$$\mathrm{V}\left(\bar{y}_2\right) \approx \bar{Y}^2\left[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_Y C_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM}C_Y C_M\right)^2 + O\left(n^{-2}\right)\right] \qquad (18)$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\mathrm{MSE}\left(\bar{y}_2\right) \approx \bar{Y}^2\left[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_Y C_M\right) + O\left(n^{-2}\right)\right] \qquad (19)$$

เมื่อ $O\left(n^{-2}\right) = \dfrac{\theta_1}{n^2}\left(4C_{12} - 2C_{03} - 2C_{21}\right) + \dfrac{\theta_2}{n^2}\left(3C_{02}C_{20} + 6C_{11}^2 + 9C_{02}^2 - 18C_{02}C_{11}\right)$ โดยที่

$\theta_1 = \dfrac{(N-n)(N-2n)}{(N-1)(N-2)},\quad \theta_2 = \dfrac{N(N-n)(N-n-1)}{(N-1)(N-2)(N-3)}$ (Sharma, Verma, Sanaullah and Singh, 2013)

และ $C_{rs} = \dfrac{E\left[(Y - E(Y))^r (M - E(M))^s\right]}{E^r(Y)E^s(M)}$ (Tin, 1965)

พิสูจน์ ตัวประมาณ $\bar{y}_2$ เขียนได้ดังนี้ $\bar{y}_2 = \dfrac{\sum_{i=1}^{n}Y_i}{\sum_{i=1}^{n}M_i} = \dfrac{\bar{Y}_t(1 + \delta_{\bar{y}_t})}{\bar{M}(1 + \delta_{\bar{m}})}$

กำหนดให้ $\delta_{\bar{y}_t} = \dfrac{\left(\bar{y}_t - \bar{Y}_t\right)}{\bar{Y}_t}$, $\delta_{\bar{m}} = \dfrac{\left(\bar{m} - \bar{M}\right)}{\bar{M}}$ จะได้ $\mathrm{E}\left(\delta_{\bar{y}_t}\right) = 0$, $\mathrm{E}\left(\delta_{\bar{m}}\right) = 0$, $\mathrm{E}\left(\delta^2_{\bar{y}_t}\right) = \gamma C_Y^2$, $\mathrm{E}\left(\delta^2_{\bar{m}}\right) = \gamma C_M^2$, $\mathrm{E}\left(\delta_{\bar{y}_t}\delta_{\bar{m}}\right) = \gamma\rho_{YM} C_Y C_M$ เมื่อ $\bar{y}_t = \dfrac{1}{n}\sum_{i=1}^{n} Y_i$, $\bar{m} = \dfrac{1}{n}\sum_{i=1}^{n} M_i$, $\bar{Y}_t = \dfrac{1}{N}\sum_{i=1}^{N} Y_i$ และ $\bar{M} = \dfrac{1}{N}\sum_{i=1}^{N} M_i$ และจะได้ค่าคาดหมายต่อไปนี้ (Sharma, Verma, Sanaullah and Singh, 2013)

$$\mathrm{E}\left(\delta^2_{\bar{y}_t}\delta_{\bar{m}}\right) = \frac{\theta_1}{n^2}C_{21}, \qquad \mathrm{E}\left(\delta_{\bar{y}_t}\delta^2_{\bar{m}}\right) = \frac{\theta_1}{n^2}C_{12},$$

$$\mathrm{E}\left(\delta^3_{\bar{m}}\right) = \frac{\theta_1}{n^2}C_{03}, \qquad \mathrm{E}\left(\delta_{\bar{y}_t}\delta^3_{\bar{m}}\right) = 3\frac{\theta_2}{n^2}C_{02}C_{11},$$

$$\mathrm{E}\left(\delta^4_{\bar{m}}\right) = 3\frac{\theta_2}{n^2}C_{02}^2 \text{ และ } \mathrm{E}\left(\delta^2_{\bar{y}_t}\delta^2_{\bar{m}}\right) = \frac{\theta_2}{n^2}\left(C_{02}C_{20} + 2C_{11}^2\right)$$

ความแปรปรวนของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\begin{aligned}
\mathrm{V}\left(\bar{y}_2\right) &= \mathrm{E}\left(\bar{y}_2^2\right) - \left[\mathrm{E}\left(\bar{y}_2\right)\right]^2 \\
&\approx \Big[\bar{Y}^2\,\mathrm{E}(1 + \delta^2_{\bar{y}_t} + 3\delta^2_{\bar{m}} + 3\delta^2_{\bar{y}_t}\delta^2_{\bar{m}} + 5\delta^4_{\bar{m}} + 2\delta_{\bar{y}_t} - 2\delta_{\bar{m}} - 4\delta_{\bar{y}_t}\delta_{\bar{m}} + 6\delta_{\bar{y}_t}\delta^2_{\bar{m}} - 4\delta^3_{\bar{m}} \\
&\quad -8\delta_{\bar{y}_t}\delta^3_{\bar{m}} - 2\delta^2_{\bar{y}_t}\delta_{\bar{m}})\Big] - \Big[\bar{Y}\,\mathrm{E}(1 + \delta_{\bar{y}_t} - \delta_{\bar{m}} - \delta_{\bar{y}_t}\delta_{\bar{m}} + \delta^2_{\bar{m}} + \delta_{\bar{y}_t}\delta^2_{\bar{m}} - \delta^3_{\bar{m}} - \delta_{\bar{y}_t}\delta^3_{\bar{m}} + \delta^4_{\bar{m}})\Big]^2 \\
&\approx \bar{Y}^2\Big[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_YC_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM}C_YC_M\right)^2 \\
&\quad + \frac{\theta_1}{n^2}\left(4C_{12} - 2C_{03} - 2C_{21}\right) + \frac{\theta_2}{n^2}\left(3C_{02}C_{20} + 6C_{11}^2 + 9C_{02}^2 - 18C_{02}C_{11}\right)\Big] \\
&\approx \bar{Y}^2\Big[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_YC_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM}C_YC_M\right)^2 + O\left(n^{-2}\right)\Big]
\end{aligned} \tag{20}$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_2$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\begin{aligned}
\mathrm{MSE}\left(\bar{y}_2\right) &= \mathrm{V}\left(\bar{y}_2\right) + \left[\mathrm{Bias}\left(\bar{y}_2\right)\right]^2 \\
&\approx \bar{Y}^2\Big[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_YC_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM}C_YC_M\right)^2 + O\left(n^{-2}\right)\Big] \\
&\quad + \Big[\bar{Y}\left(\gamma C_M^2 - \gamma\rho_{YM}C_YC_M\right)\Big]^2 \\
&\approx \bar{Y}^2\Big[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_YC_M\right) + O\left(n^{-2}\right)\Big]
\end{aligned} \tag{21}$$

ทฤษฎีบทที่ 4 ในการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายขนาด n กลุ่ม จากประชากรขนาด N กลุ่ม โดยไม่คืนที่ เมื่อกำหนดให้ Y_i เป็นค่ายอดรวมของกลุ่มที่ $i\,;\; i = 1, 2, \ldots, n$ และทราบค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่มในประชากร (C_Y) และค่าสัมประสิทธิ์การแปรผันของขนาดของกลุ่มในประชากร (C_M) เป็นตัวที่ทราบค่า และ γC_Y^2 มีค่าน้อยกว่า 1 จะได้ตัวประมาณค่าเฉลี่ยประชากรที่เอนเอียงของค่าเฉลี่ยประชากร $\overline{Y}$ ด้วยวิธีของซิลส์ที่พัฒนาจากค่าเฉลี่ยแบบอัตราส่วน ดังนี้

$$\overline{y}_{2_s} = \left(\frac{1}{1+\gamma C_Y^2}\right)\frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} M_i} \tag{22}$$

ค่าเอนเอียงของตัวประมาณ $\overline{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่ง เท่ากับ

$$\mathrm{Bias}\left(\overline{y}_{2_s}\right) \approx \mathrm{Bias}\left(\overline{y}_2\right)\left[\frac{C_M^2 - C_Y^2 - \rho_{YM} C_Y C_M}{\left(1+\gamma C_Y^2\right)\left(C_M^2 - \rho_{YM} C_Y C_M\right)}\right] \tag{23}$$

ความแปรปรวนของตัวประมาณ $\overline{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\mathrm{V}\left(\overline{y}_{2_s}\right) \approx \frac{\overline{Y}^2}{\left(1+\gamma C_Y^2\right)^2}\left[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM} C_Y C_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM} C_Y C_M\right)^2 + O\left(n^{-2}\right)\right] \tag{24}$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\overline{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\mathrm{MSE}\left(\overline{y}_{2_s}\right) \approx \frac{\mathrm{MSE}\left(\overline{y}_2\right)}{\left(1+\gamma C_Y^2\right)^2} - \frac{\gamma^2 C_Y^2 \overline{Y}^2}{\left(1+\gamma C_Y^2\right)^2}\left[2\left(C_M^2 - \rho_{YM} C_Y C_M\right) - C_y^2\right] \tag{25}$$

พิสูจน์ ตัวประมาณค่าด้วยวิธีของซิลส์ที่นำเสนอเมื่อทราบค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่มเป็นดังนี้

$$\overline{y}_{2_s} = \left(\frac{1}{1+\gamma C_Y^2}\right)\frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} M_i}$$

ค่าเอนเอียงของตัวประมาณ $\overline{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่หนึ่ง เท่ากับ

$$\mathrm{Bias}\left(\overline{y}_{2_s}\right) \approx \frac{1}{1+\gamma C_Y^2}\left[\overline{Y}\left(1+\gamma C_M^2 - \gamma\rho_{YM} C_Y C_M\right)\right] - \overline{Y}$$

$$\approx \frac{\mathrm{Bias}(\bar{y}_2)}{1+\gamma C_Y^2} - \frac{\gamma C_Y^2}{1+\gamma C_Y^2}\left[\frac{\mathrm{Bias}(\bar{y}_2)}{\gamma\left(C_M^2 - \rho_{YM}C_Y C_M\right)}\right]$$

$$\approx \mathrm{Bias}(\bar{y}_2)\left[\frac{C_M^2 - C_Y^2 - \rho_{YM}C_Y C_M}{\left(1+\gamma C_Y^2\right)\left(C_M^2 - \rho_{YM}C_Y C_M\right)}\right] \tag{26}$$

ความแปรปรวนของตัวประมาณ $\bar{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\mathrm{V}\left(\bar{y}_{2_s}\right) = \frac{1}{\left(1+\gamma C_Y^2\right)^2}\mathrm{V}(\bar{y}_2)$$

$$\approx \frac{\bar{Y}^2}{\left(1+\gamma C_Y^2\right)^2}\left[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM}C_Y C_M\right) - \left(\gamma C_M^2 - \gamma\rho_{YM}C_Y C_M\right)^2 + O\left(n^{-2}\right)\right] \tag{27}$$

และ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\bar{y}_{2_s}$ ภายใต้การกระจายอนุกรมเทย์เลอร์ถึงอันดับที่สอง เท่ากับ

$$\mathrm{MSE}\left(\bar{y}_{2_s}\right) \approx \frac{\mathrm{V}(\bar{y}_2)}{\left(1+\gamma C_Y^2\right)^2} + \mathrm{Bias}^2(\bar{y}_2)\frac{\left(C_M^2 - C_Y^2 - \rho_{YM}C_Y C_M\right)^2}{\left(1+\gamma C_Y^2\right)^2\left(C_M^2 - \rho_{YM}C_Y C_M\right)^2}$$

$$\approx \frac{\mathrm{MSE}(\bar{y}_2)}{\left(1+\gamma C_Y^2\right)^2} - \frac{\mathrm{Bias}^2(\bar{y}_2)}{\left(1+\gamma C_Y^2\right)^2}\left[1 - \frac{\left(C_M^2 - C_Y^2 - \rho_{YM}C_Y C_M\right)^2}{\left(C_M^2 - \rho_{YM}C_Y C_M\right)^2}\right]$$

$$\approx \frac{\mathrm{MSE}(\bar{y}_2)}{\left(1+\gamma C_Y^2\right)^2} - \frac{\gamma^2 C_Y^2 \bar{Y}^2}{\left(1+\gamma C_Y^2\right)^2}\left[2\left(C_M^2 - \rho_{YM}C_Y C_M\right) - C_Y^2\right] \tag{28}$$

ทฤษฎีบทที่ 5 ตัวประมาณ $\bar{y}_{2_s}$ จะมีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_2$ ภายใต้เงื่อนไข

$$\rho_{YM} < \frac{\left(1+\gamma C_Y^2\right)\gamma C_Y^2 + \left(4+\gamma C_Y^2\right)\gamma C_M^2 + \left(2+\gamma C_Y^2\right)O\left(n^{-2}\right)}{\left(6+2\gamma C_Y^2\right)\gamma C_Y C_M} \tag{29}$$

เมื่อ ρ_{YM} เป็นค่าสัมประสิทธิ์สหสัมพันธ์ของค่ายอดรวมของกลุ่มและขนาดของกลุ่มในประชากร γ เป็นปัจจัยปรับค่ากรณีประชากรมีจำนวนที่จำกัด และ C_Y , C_M เป็นค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม ตามลำดับ โดยที่ $O\left(n^{-2}\right) = \frac{\theta_1}{n^2}\left(4C_{12} - 2C_{03} - 2C_{21}\right) + \frac{\theta_2}{n^2}\left(3C_{02}C_{20} + 6C_{11}^2 + 9C_{02}^2 - 18C_{02}C_{11}\right)$

พิสูจน์ ตัวประมาณ $\bar{y}_{2_s}$ จะมีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_2$ เมื่อ $\mathrm{MSE}(\bar{y}_2) > \mathrm{MSE}(\bar{y}_{2_s})$ ซึ่งเขียนได้ดังนี้

$$\mathrm{MSE}(\bar{y}_2) > \frac{\mathrm{MSE}(\bar{y}_2)}{\left(1+\gamma C_Y^2\right)^2} - \frac{\gamma^2 C_Y^2 \bar{Y}^2}{\left(1+\gamma C_Y^2\right)^2}\left[2\left(C_M^2 - \rho_{YM} C_Y C_M\right) - C_Y^2\right]$$

$$\left(2+\gamma C_Y^2\right)\left[\left(\gamma C_Y^2 + \gamma C_M^2 - 2\gamma\rho_{YM} C_Y C_M\right) + O\left(n^{-2}\right)\right] > \gamma\left[C_Y^2 - 2\left(C_M^2 - \rho_{YM} C_Y C_M\right)\right]$$

$$\rho_{YM} < \frac{\left(1+\gamma C_Y^2\right)\gamma C_Y^2 + \left(4+\gamma C_Y^2\right)\gamma C_M^2 + \left(2+\gamma C_Y^2\right)O\left(n^{-2}\right)}{\left(6+2\gamma C_Y^2\right)\gamma C_Y C_M} \tag{30}$$

ในกรณีไม่ทราบค่า C_Y, C_M และ $O(n^{-2})$ อาจประมาณได้จากตัวอย่างด้วย $\hat{C}_Y = \dfrac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{y}_t\right)^2}}{\bar{y}_t}$ และ $\hat{C}_M = \dfrac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(M_i - \bar{m}\right)^2}}{\bar{m}}$ และ $\hat{O}(n^{-2})$ เมื่อ $\hat{C}_{rs} = \dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\left(Y_i - \bar{y}_t\right)^r\left(M_i - \bar{m}\right)^s}{\bar{y}_t^r\,\bar{m}^s}$ (Sharma, Verma, Sanaullah and Singh, 2013) โดยที่ $\bar{y}_t = \dfrac{1}{n}\sum_{i=1}^{n} Y_i$ และ $\bar{m} = \dfrac{1}{n}\sum_{i=1}^{n} M_i$

ผลการวิจัยเชิงตัวเลขและการวิจารณ์ผลการวิจัย

จากผลการศึกษา พบว่าเมื่อนำตัวประมาณค่าเฉลี่ยประชากรที่พัฒนาด้วยวิธีของซิลส์มาใช้กับข้อมูลจากหนังสือของคอแครน (Cochran, 1977) ดังแสดงไว้ในตารางที่ 1 เพื่อตรวจสอบตัวประมาณค่าที่พัฒนาขึ้น โดยศึกษาทั้งหมด 2 กรณี ได้แก่ กรณีทราบค่า $C_{\bar{Y}}$, C_Y และ C_M จากประชากร และ กรณีใช้ค่าประมาณด้วย $\hat{C}_{\bar{Y}}$, $\hat{C}_Y$ และ $\hat{C}_M$ จากตัวอย่าง และทำการเปรียบเทียบตัวประมาณค่าเฉลี่ยประชากร ด้วยข้อมูลที่ประชากรมีการแบ่งออกเป็น 20 กลุ่ม ขนาดตัวอย่างที่ทำการศึกษาเท่ากับ 5, 10 และ 15 กลุ่ม ตามลำดับ แต่ละขนาดตัวอย่างจะสุ่มเลือกทั้งหมด 30 ครั้ง โดยแต่ละขนาดตัวอย่างจะทำการสุ่มเลือกกลุ่มด้วยวิธีการเลือกตัวอย่างสุ่มแบบง่าย และแต่ละขนาดตัวอย่างจะตรวจสอบเงื่อนไขที่หามาได้ของตัวประมาณค่าเฉลี่ยประชากรที่พัฒนามาจากวิธีของซิลส์ จากนั้นทำการประมาณค่าเฉลี่ยประชากรด้วยวิธีของซิลส์และวิธีแบบเดิมทั้ง 30 ครั้ง หาค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่าด้วยสมการที่ (31) และนำค่าเฉลี่ยของทั้ง 30 ครั้ง เป็นผลการวิจัยเชิงตัวเลข จะได้ผลลัพธ์ของตัวประมาณค่าเฉลี่ยของกลุ่มวิธีแบบเดิม ($\bar{y}_1$) และตัวประมาณค่าเฉลี่ยประชากรแบบซิลส์ที่ปรับจากค่าเฉลี่ยของกลุ่มวิธีแบบเดิม ($\bar{y}_{1_s}$) ซึ่งแสดงดังตารางที่ 2 และได้ผลลัพธ์ของตัวประมาณค่าเฉลี่ยแบบอัตราส่วนแบบเดิม ($\bar{y}_2$) และตัวประมาณค่าเฉลี่ยประชากรแบบซิลส์ที่ปรับจากวิธีอัตราส่วนแบบเดิม ($\bar{y}_{2_s}$) ซึ่งแสดงดังตารางที่ 3

$$\text{ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า} = \left| \frac{\text{ค่าประมาณค่าเฉลี่ยประชากรที่ได้} - \text{ค่าเฉลี่ยประชากรจริง}}{\text{ค่าเฉลี่ยประชากรจริง}} \right| \quad (31)$$

ตารางที่ 1 ขนาดของกลุ่ม ค่ายอดรวมของกลุ่ม และค่าเฉลี่ยของกลุ่ม ของประชากรที่มีการแบ่งเป็น 20 กลุ่ม

กลุ่มที่	**ขนาดของกลุ่ม** (M_i)	**ค่ายอดรวมของกลุ่ม** (Y_i)	**ค่าเฉลี่ยของกลุ่ม** ($\overline{Y_i}$)
1	15	577.5	38.5
2	19	845.5	44.5
3	19	836.0	44.0
4	16	768.0	48.0
5	16	832.0	52.0
6	19	731.5	38.5
7	18	774.0	43.0
8	18	900.0	50.0
9	18	720.0	40.0
10	18	945.0	52.5
11	20	1,350.0	67.5
12	18	729.0	40.5
13	19	1,092.5	57.5
14	19	1,254.0	66.0
15	18	621.0	34.5
16	16	824.0	51.5
17	16	880.0	55.0
18	19	655.5	34.5
19	19	950.0	50.0
20	19	836.0	44.0
รวม	359	17,121.5	952.0

หมายเหตุ ข้อมูลจากหนังสือของคอแครน (Cochran, 1977) หน้า 306

ค่าเฉลี่ยประชากร เท่ากับ

$$\overline{Y} = \frac{17,121.5}{359} = 47.6922$$

ตารางที่ *2* ค่าประมาณค่าเฉลี่ยประชากร ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า จำแนกตามตัวประมาณ $\bar{y}_1$ และ $\bar{y}_{1_s}$ ขนาดตัวอย่างและค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และของขนาดของกลุ่ม จากประชากรและจากตัวอย่าง ที่แต่ละขนาดตัวอย่างมีการสุ่มเลือกทั้งหมด 30 ครั้ง และแต่ละขนาดตัวอย่างสอดคล้องกับเงื่อนไขที่หามาได้

ขนาดตัวอย่าง (กลุ่ม)	**ค่าสัมประสิทธิ์การแปรผัน**	**ตัวประมาณค่าเฉลี่ยประชากร**	**ค่าประมาณค่าเฉลี่ยประชากร**	**ค่าเอนเอียง**	MSE	**ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า (%)**
$n = 5$	$C_{\bar{Y}} = 0.1947$	$\bar{y}_1$	49.9733	-0.0650	14.2148	5.0454
	$C_M = 0.0777$	$\bar{y}_{1_s}$	49.6909	-0.3474	14.1713	4.5647
	$\hat{C}_{\bar{Y}} = 0.2436$	$\bar{y}_1$	49.9733	-0.0943	22.2934	5.0454
	$\hat{C}_M = 0.0914$	$\bar{y}_{1_s}$	49.5306	-0.5370	22.1747	4.3471
$n = 10$	$C_{\bar{Y}} = 0.1947$	$\bar{y}_1$	48.2767	-0.0262	4.4227	3.1672
	$C_M = 0.0777$	$\bar{y}_{1_s}$	48.1854	-0.1175	4.4191	3.1532
	$\hat{C}_{\bar{Y}} = 0.2112$	$\bar{y}_1$	48.2767	-0.0309	5.2172	3.1672
	$\hat{C}_M = 0.0849$	$\bar{y}_{1_s}$	48.1690	-0.1386	5.2120	3.1555
$n = 15$	$C_{\bar{Y}} = 0.1947$	$\bar{y}_1$	47.9778	-0.0031	1.4540	0.8861
	$C_M = 0.0777$	$\bar{y}_{1_s}$	47.9475	-0.0334	1.4533	0.8645
	$\hat{C}_{\bar{Y}} = 0.2091$	$\bar{y}_1$	47.9778	-0.0032	1.6773	0.8861
	$\hat{C}_M = 0.0729$	$\bar{y}_{1_s}$	47.9428	-0.0381	1.6763	0.8617

จากตารางที่ 2 ในกรณีทราบค่า $C_{\bar{Y}}$ และ C_M ของประชากร เมื่อพิจารณาจากค่า MSE พบว่า ตัวประมาณ $\bar{y}_{1_s}$ มีค่า MSE น้อยกว่าตัวประมาณ $\bar{y}_1$ ทุกขนาดตัวอย่าง ดังนั้น ตัวประมาณ $\bar{y}_{1_s}$ มีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_1$ ภายใต้เงื่อนไขที่หามาได้จากสมการที่ (16)

ในกรณีใช้ค่าประมาณด้วย $\hat{C}_{\bar{Y}}$ และ $\hat{C}_M$ จากตัวอย่าง เมื่อพิจารณาจากค่า MSE พบว่า ตัวประมาณ $\bar{y}_{1_s}$ มีค่า MSE น้อยกว่าตัวประมาณ $\bar{y}_1$ ทุกขนาดตัวอย่าง ดังนั้น ตัวประมาณ $\bar{y}_{1_s}$ มีประสิทธิภาพมากกว่าตัวประมาณ $\bar{y}_1$ ภายใต้เงื่อนไขที่หามาได้จากสมการที่ (16)

จากขนาดตัวอย่างในตารางที่ 2 เมื่อพิจารณาขนาดตัวอย่างเปลี่ยนแปลงจากขนาดเล็กไปขนาดใหญ่ จะสังเกตเห็นว่า เมื่อตัวอย่างมีขนาดเล็ก MSE จะมีค่ามาก และลดลงเมื่อมีขนาดใหญ่

เมื่อพิจารณาจากค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า พบว่า ตัวประมาณ $\bar{y}_{1_s}$ มีเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่าน้อยกว่าตัวประมาณ $\bar{y}_1$ ทุกขนาดตัวอย่าง ทั้งในกรณีทราบค่า $C_{\bar{Y}}$ และ C_M จากประชากร และในกรณีใช้ค่าประมาณด้วย $\hat{C}_{\bar{Y}}$ และ $\hat{C}_M$ จากตัวอย่าง

***ตารางที่** 3* ค่าประมาณค่าเฉลี่ยประชากร ค่าเอนเอียง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า จำแนกตามตัวประมาณ $\overline{y}_2$ และ $\overline{y}_{2_s}$ ขนาดตัวอย่างและค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม จากประชากรและจากตัวอย่าง ที่แต่ละขนาดตัวอย่างมีการสุ่มเลือกทั้งหมด 30 ครั้ง และแต่ละขนาดตัวอย่างสอดคล้องกับเงื่อนไขที่หามาได้

ขนาดตัวอย่าง (กลุ่ม)	ค่าสัมประสิทธิ์การแปรผัน	ตัวประมาณค่าเฉลี่ยประชากร	ค่าประมาณค่าเฉลี่ยประชากร	ค่าเอนเอียง	MSE	ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า (%)
$n = 5$	$C_Y = 0.2270$	$\overline{y}_2$	50.0527	-0.0121	15.6609	5.1912
	$C_M = 0.0777$	$\overline{y}_{2_s}$	49.6687	-0.3961	15.5785	4.5375
	$\hat{C}_Y = 0.2849$	$\overline{y}_2$	50.0527	-0.0215	24.9821	5.1912
	$\hat{C}_M = 0.0914$	$\overline{y}_{2_s}$	49.4489	-0.6251	24.7679	4.2271
$n = 10$	$C_Y = 0.2270$	$\overline{y}_2$	48.3059	-0.0024	5.0869	3.1730
	$C_M = 0.0777$	$\overline{y}_{2_s}$	48.1817	-0.1266	5.0768	3.1526
	$\hat{C}_Y = 0.2417$	$\overline{y}_2$	48.3059	-0.0023	5.7800	3.1730
	$\hat{C}_M = 0.0849$	$\overline{y}_{2_s}$	48.1649	-0.1434	5.7667	3.1482
$n = 15$	$C_Y = 0.2270$	$\overline{y}_2$	47.9809	-0.0001	1.7433	0.8972
	$C_M = 0.0777$	$\overline{y}_{2_s}$	47.9397	-0.0413	1.7420	0.8678
	$\hat{C}_Y = 0.2336$	$\overline{y}_2$	47.9809	-0.0005	1.8473	0.8972
	$\hat{C}_M = 0.0729$	$\overline{y}_{2_s}$	47.9373	-0.0441	1.8459	0.8664

จากตารางที่ 3 ในกรณีทราบค่า C_Y และ C_M ของประชากร เมื่อพิจารณาจากค่า MSE พบว่า ตัวประมาณ $\overline{y}_{2_s}$ มีค่า MSE น้อยกว่าตัวประมาณ $\overline{y}_2$ ทุกขนาดตัวอย่าง ดังนั้น ตัวประมาณ $\overline{y}_{2_s}$ มีประสิทธิภาพมากกว่าตัวประมาณ $\overline{y}_2$ ภายใต้เงื่อนไขที่หามาได้จากสมการที่ (29)

ในกรณีใช้ค่าประมาณด้วย $\hat{C}_Y$ และ $\hat{C}_M$ จากตัวอย่าง เมื่อพิจารณาจากค่า MSE พบว่า ตัวประมาณ $\overline{y}_{2_s}$ มีค่า MSE น้อยกว่าตัวประมาณ $\overline{y}_2$ ทุกขนาดตัวอย่าง ดังนั้น ตัวประมาณ $\overline{y}_{2_s}$ มีประสิทธิภาพมากกว่าตัวประมาณ $\overline{y}_2$ ภายใต้เงื่อนไขที่หามาได้จากสมการที่ (29)

จากขนาดตัวอย่างในตารางที่ 3 เมื่อพิจารณาขนาดตัวอย่างเปลี่ยนแปลงจากขนาดเล็กไปขนาดใหญ่ จะสังเกตเห็นว่า เมื่อตัวอย่างมีขนาดเล็ก MSE จะมีค่ามาก และลดลงเมื่อมีขนาดใหญ่

เมื่อพิจารณาจากค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่า พบว่า ตัวประมาณ $\overline{y}_{2_s}$ มีเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมพัทธ์ของตัวประมาณค่าน้อยกว่าตัวประมาณ $\overline{y}_2$ ทุกขนาดตัวอย่าง ในกรณีทราบค่า C_Y และ C_M ของประชากร และในกรณีใช้ค่าประมาณด้วย $\hat{C}_Y$ และ $\hat{C}_M$ จากตัวอย่าง

จากผลการวิจัยเชิงตัวเลขในตารางที่ 2 และ 3 ถึงแม้ว่าค่า MSE ของตัวประมาณค่าเฉลี่ยประชากรที่นำเสนอใหม่จะให้ผลลัพธ์เชิงตัวเลขที่ดีกว่าทั้งหมด แต่ตัวเลขที่ได้นั้นไม่ได้แสดงให้เห็นว่าตัวประมาณค่าเฉลี่ยประชากรที่นำเสนอใหม่ดีกว่าตัวประมาณค่าเฉลี่ยประชากรแบบเดิมอย่างชัดเจน เพราะว่าตัวเลขที่ได้นั้นใกล้เคียงกันมาก

***ตารางที่** 4* ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่เพิ่มขึ้น (เปอร์เซ็นต์) จากกรณีใช้ค่าประมาณสัมประสิทธิ์การแปรผันจากตัวอย่างเมื่อไม่ทราบค่าสัมประสิทธิ์การแปรผันของประชากร

ขนาดตัวอย่าง (กลุ่ม)	**ตัวประมาณค่าเฉลี่ยประชากร**			
	$\overline{y}_1$	$\overline{y}_{1_s}$	$\overline{y}_2$	$\overline{y}_{2_s}$
5	36.24	36.09	37.31	37.10
10	15.23	15.21	11.99	11.96
15	13.31	13.30	5.63	5.63

ในกรณีใช้ค่าประมาณด้วย $\hat{C}_{\overline{Y}}$ และ $\hat{C}_M$ จากตัวอย่าง เมื่อไม่ทราบค่า $C_{\overline{Y}}$ และ C_M ของประชากร สำหรับตัวประมาณ $\overline{y}_1$ และ $\overline{y}_{1_s}$ พบว่าตัวประมาณค่าเฉลี่ยประชากรดังกล่าวแต่ละขนาดตัวอย่างจะมีค่า MSE เพิ่มขึ้นตามที่แสดงในตารางที่ 4 และพบว่าจะมีค่าลดลงเมื่อขนาดตัวอย่างเปลี่ยนแปลงจากขนาดเล็กไปขนาดใหญ่

ในกรณีใช้ค่าประมาณด้วย $\hat{C}_Y$ และ $\hat{C}_M$ จากตัวอย่าง เมื่อไม่ทราบค่า C_Y และ C_M ของประชากร สำหรับตัวประมาณ $\overline{y}_2$ และ $\overline{y}_{2_s}$ พบว่าตัวประมาณค่าเฉลี่ยประชากรดังกล่าวแต่ละขนาดตัวอย่างจะมีค่า MSE เพิ่มขึ้นตามที่แสดงในตารางที่ 4 และพบว่าจะมีค่าลดลงเมื่อขนาดตัวอย่างเปลี่ยนแปลงจากขนาดเล็กไปขนาดใหญ่

สรุปผลการวิจัย

จากการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยประชากร พบว่า ตัวประมาณค่าเฉลี่ยประชากรแบบซิลส์ที่พัฒนามาจากค่าเฉลี่ยของกลุ่ม มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยประชากรแบบเดิม ทั้งในกรณีที่ทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และของขนาดของกลุ่ม จากประชากร และในกรณีประมาณค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม และของขนาดของกลุ่ม จากตัวอย่าง และตัวประมาณค่าเฉลี่ยประชากรแบบซิลส์ที่พัฒนามาจากค่าเฉลี่ยแบบอัตราส่วน มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยประชากรแบบเดิม ทั้งในกรณีที่ทราบค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม จากประชากร และในกรณีประมาณค่าสัมประสิทธิ์การแปรผันของค่ายอดรวมของกลุ่ม และของขนาดของกลุ่ม จากตัวอย่าง เมื่อตัวประมาณค่าเฉลี่ยประชากรแบบซิลส์ทั้ง 2 ตัว ดังกล่าว มีข้อกำหนดสอดคล้องกับเงื่อนไขที่หามาได้

การประมาณค่าเฉลี่ยประชากรที่พัฒนาจากวิธีของซิลส์ สำหรับการเลือกตัวอย่างแบบกลุ่มขั้นเดียวแบบง่ายไม่คืนที่ควรทราบค่าสัมประสิทธิ์การแปรผันของค่าเฉลี่ยของกลุ่ม ของค่ายอดรวมของกลุ่ม และขนาดของกลุ่มของประชากร แต่หากไม่ทราบค่าสัมประสิทธิ์การแปรผันดังกล่าวของประชากร อาจประมาณค่าสัมประสิทธิ์การแปรผันดังกล่าวได้จากตัวอย่าง

เอกสารอ้างอิง

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

Gupta, A.K. and Kabe, D.G. (2011). *Theory of Sample Surveys.* Singapore: World Scientific Publishing.

Jitthavech, J. and Lorchirachoonkul, V. (2013). Estimators in Simple Random Sampling: Searls Approach. *Songklanakarin Journal of Science and Technology*, *35*,(6), 749-760.

Searls, D.T. (1964). The Utilization of a Known Coefficient of Variation in the Estimation Procedure. *American Statistical Association Journal, 56*, 1225-1226.

Sharma, P., Verma, H.K., Sanaullah, A. and Singh, R. (2013). Some Exponential Ratio-Product Type Estimators Using Information on Auxiliary Attributes under Second Order Approximation. *International Journal of Statistics and Economics, 12*, 58-66.

Suwutthee, P. (2009). *Sample Surveys: Sampling Designs and Analysis.* Bangkok: WVO Officer of Printing Mill. (in Thai)

Tin, M. (1965). Comparisons of Some Ratio Estimators. *Journal of American Statistics Association*, *60,* 294-307.