ตัวประมาณอัตราส่วนสำหรับประมาณค่าเฉลี่ยประชากรที่ปรับใหม่ด้วยการถดถอยที่มีความแกร่ง

ศศินันท์ เชยชิต[1*] และ นวลพรรณ ลอว์สัน[1]

[1]ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพ ประเทศไทย
[*]c.sasinun94@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเสนอตัวประมาณอัตราส่วนสำหรับประมาณค่าเฉลี่ยประชากรในกรณีที่ข้อมูลมีค่าผิดปกติขึ้นใหม่โดยวิธีการประมาณด้วยการถดถอยที่มีความแกร่ง โดยปรับจากตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรของ Nangsue (2009) ซึ่งในการวิจัยครั้งนี้จะทำการศึกษาวิธีการประมาณการถดถอยที่มีความแกร่งโดยวิธี Huber M พร้อมทั้งศึกษาถึงค่าความเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณที่เสนอใหม่ และทำการเปรียบเทียบประสิทธิภาพของตัวประมาณที่เสนอใหม่กับวิธีเดิมโดยใช้เกณฑ์ความคลาดเคลื่อนกำลังสองเฉลี่ย ผลการศึกษาจากข้อมูลที่ได้จากการจำลองพบว่า ตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรที่นำเสนอขึ้นใหม่มีประสิทธิภาพสูงกว่าตัวประมาณดั้งเดิมภายใต้สถานการณ์ที่เหมาะสม

คำสำคัญ: ตัวประมาณอัตราส่วน, การถดถอยที่มีความแกร่ง, การสุ่มตัวอย่างแบบง่าย

Adjusted Ratio Estimator for Estimating Population Mean Using Robust Regression

Sasinun Choeychit[1,*] and Nuanpan Lawson[2]
[1] *Department of Applied Statistics, Faculty of Applied Science*
King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
** c.sasinun94@gmail.com*

Abstract

This paper aims to propose the new ratio estimators for estimating population mean when an outlier occurs in the study using robust regression by adjusting the Nangsue (2009) estimator. We consider the Huber M method in the study and also the bias and mean square error of the estimator are investigated. The new estimators are compared with the existing estimator using mean square error of the estimator. A simulation study shows that the proposed estimators perform well when compared to the existing estimator under suitable conditions.

Keywords: Ratio estimator, Robust regression, Simple random sampling

1. บทนำ

ตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรใช้ประโยชน์ของความสัมพันธ์ระหว่างตัวแปรช่วย (Auxiliary variable) และตัวแปรที่สนใจศึกษา (Variable of interest) ซึ่งจะมีประสิทธิภาพเมื่อตัวแปรดังกล่าวมีความสัมพันธ์เชิงบวกต่อกัน คณะผู้วิจัยหลายคณะได้เสนอตัวประมาณอัตราส่วนแบบต่าง ๆ มากมาย เช่น ตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากร (Cochran, 1977) ซึ่งเป็นตัวประมาณที่เอนเอียง (Bias) แต่สามารถละความเอนเอียงได้ถ้าสุ่มตัวอย่างขนาดใหญ่ แสดงดังสมการต่อไปนี้

$$\hat{\bar{Y}}_R = \left(\frac{\bar{y}}{\bar{x}}\right)\bar{X},$$

โดยที่ $\bar{y}$ คือ ค่าเฉลี่ยตัวอย่างของตัวแปรที่สนใจศึกษา, $\bar{x}$ คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย, $\bar{X}$ คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย

ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Errors: MSE) ของตัวประมาณอัตราส่วน แสดงดังสมการต่อไปนี้

$$MSE\left(\hat{\bar{Y}}_R\right) = \gamma\bar{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_x C_y\right),$$

เมื่อ $\gamma=\left(\frac{1}{n}-\frac{1}{N}\right)$, $C_x=\frac{\sigma_x}{\overline{X}}$, $\sigma_x^2=\frac{1}{(N-1)}\sum_{i=1}^{N}\left(X_i-\overline{X}\right)^2$, $\rho=\frac{S_{xy}}{S_xS_y}$

และ $S_{xy}=\frac{\sum_{i=1}^{n}\left(y_i-\overline{Y}\right)\left(x_i-\overline{X}\right)}{(N-1)}$ *โดยที่* $y=\sum_{i=1}^{n}y_i$ *คือ ยอดรวมตัวอย่างของตัวแปรที่สนใจศึกษา,* $x=\sum_{i=1}^{n}x_i$ *คือ ยอดรวมตัวอย่างของตัวแปรช่วย,* $\overline{x}=\frac{\sum_{i=1}^{n}x_i}{n}$ *คือ ค่าเฉลี่ยตัวอย่างของตัวแปรช่วย,* $\overline{X}=\frac{\sum_{i=1}^{N}X_i}{N}$ *คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย,* $\overline{Y}=\frac{\sum_{i=1}^{N}Y_i}{N}$ *คือ ค่าเฉลี่ยประชากรของตัวแปรตาม,* C_x *คือ สัมประสิทธิ์การแปรผัน (Coefficient of variation) ของตัวแปรช่วย,* C_y *คือ สัมประสิทธิ์การแปรผันของตัวแปรศึกษา,* ρ *คือ สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ของตัวแปรศึกษากับตัวแปรช่วย*

หลังจากนั้นได้มีคณะผู้วิจัยหลายคณะพัฒนาตัวประมาณอัตราส่วนขึ้นใหม่ เช่น ตัวประมาณอัตราส่วนขึ้นใหม่โดยใช้สัมประสิทธิ์การแปรผันของตัวแปรช่วย (Sisodia & Dwivedi, 1981) ต่อมา มีการเสนอตัวประมาณอัตราส่วนขึ้นใหม่โดยใช้สัมประสิทธิ์การแปรผันและสัมประสิทธิ์ความโด่งของประชากรของตัวแปรช่วย (Correlation of kurtosis: $\beta_2(x)$) (Upadhyaya & Singh, 1999) และ ตัวประมาณที่ใช้สัมประสิทธิ์สหสัมพันธ์ของตัวแปรช่วย x (Singh & Tailor, 2003) มาช่วยในการเพิ่มประสิทธิภาพของตัวประมาณ นอกจากนี้ยังมีนักวิจัยอีกมากมายได้เสนอตัวประมาณที่ใช้ประโยชน์ของสารสนเทศของตัวแปรช่วยตัวอื่นหลายตัวประมาณจนสามารถสรุปได้เป็นรูปแบบทั่วไป เช่น ตัวประมาณอัตราส่วนที่เสนอโดย Jaroengeratikun and Lawson (2018) และยังมีคณะผู้วิจัยที่ได้นำสัมประสิทธิ์การถดถอย (Regression coefficient) มาปรับปรุงตัวประมาณเพื่อให้มีประสิทธิภาพสูงขึ้น เช่น

ตัวประมาณอัตราส่วนและตัวประมาณการถดถอยมาใช้ในการประมาณค่าเฉลี่ยของประชากรเมื่อข้อมูลเกิดการสูญหาย (Nangsue, 2009) แสดงดังสมการ (1)

$$\hat{\overline{Y}}_N=\overline{y}\left(\frac{\overline{X}}{\overline{x}}\right)^{b_1}, \qquad (1)$$

เมื่อ b_1 *คือ ค่าสัมประสิทธิ์การถดถอยของตัวอย่างโดย* $b_1=\frac{rS_y}{S_x}$, r *คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร* x *และ* y *ของตัวอย่าง,* S_x *คือ ส่วนเบี่ยงเบนมาตรฐานของตัวอย่างสำหรับตัวแปรช่วย,* S_y *คือส่วนเบี่ยงเบนมาตรฐานของตัวอย่างสำหรับตัวแปรศึกษา*

และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณดังกล่าว แสดงดังสมการต่อไปนี้

$$MSE\left(\hat{\overline{Y}}_N\right)=\gamma\overline{Y}^2\left(C_y^2+b_1^2C_x^2-2b_1\rho C_xC_y\right),$$

ในกรณีที่ข้อมูลมีค่าผิดปกติปนอยู่การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุด (Ordinary least square method : OLS) อาจนำไปสู่การสรุปผลที่ผิดพลาดเนื่องจากไม่สอดคล้องกับข้อสมมติที่ว่าประชากรมีการแจกแจงแบบปกติและความแปรปรวนมีค่าคงที่ ดังนั้นเพื่อแก้ไขปัญหาดังกล่าว จึงมีผู้วิจัยพัฒนาวิธีการประมาณค่าพารามิเตอร์ของการถดถอยที่มีความแกร่ง (Robust regression) ขึ้นมาซึ่งสามารถลดอิทธิพลความผิดปกติของข้อมูลลงได้เมื่อเปรียบเทียบกับการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดโดย (Kadilar & Candan & Cingi, 2007)

ได้เสนอให้ใช้วิธีประมาณค่าพารามิเตอร์ของการถดถอยที่มีความแกร่งด้วยวิธี Huber-M โดยปรับจากตัวประมาณของ Sisodia and Dwivedi (1981); Singh and Kakran (1993); Upadhyaya and Singh (1999) และตัวประมาณของ Kadilar and Cingi (2004) ต่อมา Zaman and Bulut (2018) ได้เสนอตัวประมาณอัตราส่วนของประชากรของ (Kadilar & Candan & Cingi, 2007) โดยใช้การถดถอยที่มีความแกร่งอื่น ๆ เช่น วิธีการประมาณแบบ Least absolute deviations method (Lad) และนำมาเปรียบเทียบกับตัวประมาณของ (Kadilar & Candan & Cingi, 2007)

จากที่ได้กล่าวมาข้างต้น คณะผู้วิจัยจึงมีความประสงค์ที่จะนำเสนอตัวประมาณอัตราส่วนของประชากรขึ้นใหม่ในกรณีที่มีข้อมูลผิดปกติ โดยได้ทำการปรับจากตัวประมาณของ Nangsue (2009) พร้อมทั้งคำนวณค่าความเอนเอียง และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณที่นำเสนอขึ้นใหม่และทำการเปรียบเทียบประสิทธิภาพของตัวประมาณที่นำเสนอใหม่กับตัวประมาณเดิมโดยพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดด้วยข้อมูลที่ได้จากการจำลอง

2. ตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรที่นำเสนอขึ้นใหม่

ในการศึกษาครั้งนี้จะทำการเสนอตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรขึ้นมาใหม่โดยปรับจากตัวประมาณของ Nangsue (2009) โดยจะทำการประมาณค่าสัมประสิทธิ์การถดถอยด้วยตัวประมาณที่มีความแกร่งโดยใช้วิธี Huber M จากสมการที่ (1) แทนค่า b_1 ด้วย b_{rob} ที่ประมาณด้วยวิธี Huber M แสดงดังต่อไปนี้

$$\hat{\bar{Y}}_{Rob} = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{b_{rob}} \tag{2}$$

โดย b_{rob} คือ ค่าสัมประสิทธิ์การถดถอยที่มีความแกร่งด้วยวิธี Huber M

การประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอยที่มีความแกร่ง โดยใช้วิธี Huber M แสดงดังนี้

วิธีประมาณแบบ Huber-M เสนอโดย (Huber, 1973)

$$\min \sum_{i=1}^{n} \rho\left(e_i\right) \tag{3}$$

เมื่อ ρ เป็นฟังก์ชันความคลาดเคลื่อน

$$\rho(e) = \begin{cases} \dfrac{e^2}{2} & ,|e| \le k \\ k|e| - \dfrac{k^2}{2} & ,|e| > k \end{cases} \tag{4}$$

อนุพันธ์ของฟังก์ชันในสมการที่ (4) เป็นดังนี้

$$\varphi(y) = \begin{cases} e & ,|e| \le k \\ k\,\mathrm{sgn}(e) & ,|e| > k \end{cases} \tag{5}$$

เมื่อ $\mathrm{sgn}(.)$ *คือสัญลักษณ์แทนได้ดังนี้* $\mathrm{sgn}(e) = \begin{cases} -1 & ,e < k \\ 0 & ,e = k \\ 1 & ,e > k \end{cases}$ *และ k เป็นค่าคงที่มีค่าเท่ากับ 1.345 (Fox, 2002)*

จากตัวประมาณที่คณะผู้วิจัยได้นำเสนอมาข้างต้นนั้น คณะผู้วิจัยจะทำการคำนวณหาค่าความเอนเอียงและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยโดยแสดงดังต่อไปนี้

กำหนดให้ $\bar{y} = \bar{Y}\left(1+e_0\right)$ และ $\bar{x} = \bar{X}\left(1+e_1\right)$ เป็นตัวประมาณที่ไม่เอนเอียงของ $\bar{Y}$ และ $\bar{X}$ ตามลำดับ

ดังนั้น $E(e_0)=E(e_1)=0$

$$E(e_0^2)=\gamma C_y^2$$

$$E(e_1^2)=\gamma C_x^2$$

$$E(e_0e_1)=\gamma C_{xy}=\gamma\rho C_xC_y$$

แทนค่า $\bar{y}=\bar{Y}(1+e_0)$ และ $\bar{x}=\bar{X}(1+e_1)$ ในสมการ (2)

$$\hat{\bar{Y}}_{rob}=\bar{Y}(1+e_0)\left(\frac{\bar{X}}{\bar{X}(1+e_1)}\right)^{b_{rob}}$$

$$\hat{\bar{Y}}_{rob}=\bar{Y}(1+e_0)(1+e_1)^{-b_{rob}}$$

ในกรณีที่ตัวอย่างสุ่มมีขนาดใหญ่และสมมติว่า $|e_1|<1$ จะพบว่า e_0 และ e_1 ที่มีดีกรีมากกว่า 2 มีขนาดเล็กและสามารถละได้ ดังนั้นโดยทั่วไปในการพิจารณาค่าความเอนเอียงของตัวประมาณอัตราส่วนจะพิจารณาการประมาณค่าโดยใช้อนุกรมเทเลอร์อันดับที่ 1 (Taylor Series Approximation) ดังนั้นจะพบว่าสมการ $\hat{\bar{Y}}_{rob}=\bar{Y}(1+e_0)(1+e_1)^{-b_{rob}}$ สามารถจัดรูปแบบสมการได้ดังต่อไปนี้

$$\hat{\bar{Y}}_{rob}=\bar{Y}(1+e_0)\left(1-b_{rob}e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_1^2\right)$$

$$\hat{\bar{Y}}_{rob}-\bar{Y}=\bar{Y}\left(-b_{rob}e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_1^2+e_0-b_{rob}e_0e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_0e_1^2\right)$$

$$Bias\left(\hat{\bar{Y}}_{rob}\right)=E\left(\hat{\bar{Y}}_{rob}-\bar{Y}\right)$$

$$=E\left(\bar{Y}\left(-b_{rob}e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_1^2+e_0-b_{rob}e_0e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_0e_1^2\right)\right)$$

$$=\bar{Y}\gamma\left(\frac{b_{rob}(b_{rob}+1)}{2}C_x^2-b_{rob}\rho C_xC_y\right)$$

ดังนั้น $Bias\left(\hat{\bar{Y}}_{rob}\right)=\bar{Y}\gamma\left(\frac{b_{rob}(b_{rob}+1)}{2}C_x^2-b_{rob}\rho C_xC_y\right)$ (6)

และ $MSE\left(\hat{\bar{Y}}_{rob}\right)\cong E\left(\hat{\bar{Y}}_{rob}-\bar{Y}\right)^2$

$$=\bar{Y}^2E\left(-b_{rob}e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_1^2+e_0-b_{rob}e_0e_1+\frac{b_{rob}(b_{rob}+1)}{2}e_0e_1^2\right)^2$$

$$=\bar{Y}^2\gamma\left(b_{rob}^2C_x^2-2b_{rob}\rho C_xC_y+C_y^2\right)$$

ดังนั้น $MSE\left(\hat{\bar{Y}}_{rob}\right)=\bar{Y}^2\gamma\left(b_{rob}^2C_x^2-2b_{rob}\rho C_xC_y+C_y^2\right)$ (7)

ในการหา $MSE\left(\hat{\bar{Y}}_{rob}\right)$ ที่มีค่าต่ำสุด โดยการหาอนุพันธ์อันดับหนึ่งของ $MSE\left(\hat{\bar{Y}}_{rob}\right)$ เทียบกับ b_{rob} และเท่ากับ 0 จะได้

$$\frac{\partial MSE\left(\hat{\bar{Y}}_{rob}\right)}{\partial b_{rob}} = \frac{\partial}{\partial b_{rob}}\left[\bar{Y}^2\gamma\left(b_{rob}^2 C_x^2 - 2b_{rob}\rho C_x C_y + C_y^2\right)\right]$$

$$\bar{Y}^2\gamma\left[2b_{rob}C_x^2 - 2\rho C_x C_y\right] = 0$$

$$b_{rob} = \frac{2\rho C_x C_y}{2C_x^2}$$

$$b_{rob} = \frac{\rho C_y}{C_x} = b_{lopt}$$

แทนค่า ลงใน (7) จะได้

$$\min MSE\left(\hat{\bar{Y}}_{rob}\right) = \bar{Y}^2\gamma\left(b_{lopt}^2 C_x^2 - 2b_{lopt}\rho C_x C_y + C_y^2\right)$$

$$= \bar{Y}^2\gamma\left(C_y^2\left(1-\rho^2\right)\right)$$

ดังนั้น $\min MSE\left(\hat{\bar{Y}}_{rob}\right) = \bar{Y}^2\gamma\left(C_y^2\left(1-\rho^2\right)\right)$ (เนื่องจากค่าที่ได้จะให้ค่าที่เป็นบวก ดังนั้นจึงให้ค่าต่ำสุดสัมพัทธ์)

2.1 การเปรียบเทียบประสิทธิภาพเชิงทฤษฎี

คณะผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรที่นำเสนอขึ้นใหม่กับตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรตัวดั้งเดิม (Nangsue, 2009), $\hat{\bar{Y}}_N$ โดยพิจารณาเงื่อนไข

$$MSE\left(\hat{\bar{Y}}_{rob}\right) < MSE\left(\hat{\bar{Y}}_N\right)$$

$$\bar{Y}^2\gamma\left(b_{rob}^2 C_x^2 - 2b_{rob}\rho C_x C_y + C_y^2\right) < \bar{Y}^2\gamma\left(b_1^2 C_x^2 - 2b_1\rho C_x C_y + C_y^2\right)$$

$$b_{rob}^2 C_x^2 - 2b_{rob}\rho C_x C_y + C_y^2 - b_1^2 C_x^2 + 2b_1\rho C_x C_y - C_y^2 < 0$$

$$\left(b_{rob} - b_1\right)\left(C_x^2\left(b_{rob} + b_1\right) - 2\rho C_x C_y\right) < 0$$

กรณี $\left(b_{rob} - b_1\right) > 0$ เมื่อ $b_{rob} > b_1$:

$$C_x^2\left(b_{rob} + b_1\right) - 2\rho C_x C_y < 0$$

$$b_{rob} < \frac{2\rho C_y}{C_x} - b_1$$

กรณี $\left(b_{rob} - b_1\right) < 0$ เมื่อ $b_{rob} < b_1$:

$$C_x^2\left(b_{rob} + b_1\right) - 2\rho C_x C_y > 0$$

$$b_{rob} > \frac{2\rho C_y}{C_x} - b_1$$

3. ผลจากการจำลองข้อมูล

ในการศึกษาครั้งนี้ผู้วิจัยได้ทำการเปรียบเทียบตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรที่นำเสนอขึ้นใหม่โดยทำการจำลองข้อมูลด้วยโปรแกรม R โดยกำหนดให้ตัวแปรอิสระ X มีการแจกแจงปกติ มีค่าเฉลี่ยเป็น 20 ความแปรปรวน 2 ประชากรขนาด 1000 ขนาดตัวอย่างเป็น 50, 100 และ 200 เป็นตัวแทนของข้อมูล ขนาดเล็ก ขนาดกลาง และขนาดใหญ่ ตามลำดับ โดยที่กำหนดให้ตัวอย่างนั้นมีข้อมูลผิดปกติปนอยู่ 5% ของขนาดตัวอย่างและมีค่าผิดปกติระดับปานกลาง โดยในการศึกษาครั้งนี้ผู้วิจัยใช้วิธีการจำลองข้อมูลด้วยโปรแกรม R ภายใต้กรณีที่กำหนด แต่ละกรณีจะทำซ้ำ 1000 รอบ ตัวประมาณที่ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดจะเป็นตัวประมาณที่มีประสิทธิภาพสูงสุด

ตารางที่ 1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

จำนวนตัวอย่าง	**วิธีประมาณ**	**MSE**
50	OLS	0.0301
	Huber M	0.0210
100	OLS	0.0097
	Huber M	0.0089
200	OLS	0.0045
	Huber M	0.0040

จากตารางที่ 1 พบว่าตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรที่เสนอขึ้นมาใหม่ด้วยวิธีการประมาณการถดถอยที่มีความแกร่งด้วยวิธี Huber M มีประสิทธิภาพดีกว่าวิธีการประมาณการถดถอยด้วยวิธีกำลังสองน้อยสุดในทุกระดับของขนาดตัวอย่างที่ทำการสุ่ม โดยเฉพาะในกรณีที่ตัวอย่างสุ่มมีขนาดเล็ก ซึ่งจะพบว่าถ้าสุ่มตัวอย่างขนาดใหญ่ขึ้นค่าที่ได้จากวิธี Huber M จะมีค่าใกล้เคียงกับวิธี OLS ดังนั้นจะพบว่าวิธี Huber M มีประสิทธิภาพดีกว่า OLS และยังช่วยประหยัดงบเวลาและงบประมาณได้ดีกว่าด้วย

4. สรุปผลการวิจัย

ในงานวิจัยครั้งนี้คณะผู้วิจัยเสนอตัวประมาณอัตราส่วนของค่าเฉลี่ยประชากรขึ้นใหม่โดยใช้วิธีการประมาณการถดถอยที่มีความแกร่งของ Huber M กับตัวประมาณของ Nangsue (2009) พบว่าตัวประมาณที่เสนอขึ้นใหม่มีประสิทธิภาพสูงกว่าตัวประมาณดั้งเดิม โดยในการศึกษาด้วยการใช้ข้อมูลที่ได้จากการจำลองได้พิจารณาระดับของข้อมูลที่มีความผิดปกติที่แตกต่างกัน ซึ่งสูตรที่นำเสนอขึ้นใหม่สามารถนำไปประยุกต์ใช้จริงได้กับข้อมูลที่มีค่าผิดปกติหลากหลาย ไม่ว่าจะเป็นทางภัยพิบัติ โรคระบาด สงคราม เป็นต้น จึงเป็นประโยชน์ให้กับนักศึกษา นักวิจัย และผู้ที่สนใจทั่วไปสามารถนำไปประยุกต์ใช้ได้ รวมทั้งตัวประมาณที่นำเสนอขึ้นใหม่มีประสิทธิภาพดีกว่าวิธี OLS โดยเฉพาะในกรณีที่ตัวอย่างสุ่มมีขนาดเล็กซึ่งจะเป็นประโยชน์กับนักวิจัยในการช่วยประหยัดเวลาและงบประมาณในการเก็บข้อมูลตัวอย่างที่มีจำนวนน้อยและมีประสิทธิภาพสูง

6. เอกสารอ้างอิง

Cochran, W. G. (1997). **Sampling Techniques.** New York: John Wiley and Sons.

Fox, J. (2002). **Robust Regression: Appendix to an R and S-PLUS companion to applied regression.** Accessed December 5, 2017.

Huber, P.J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo. **The Annals of**

Statistics. 799–821.

Jaroengeratikun, U., & Lawson, N. (2018). Improved ratio estimators of population mean using transformed variable in double sampling. **The Journal of Applied Science.** 17(2).

Kadılar, C., M. Candan, and H. Çıngı. (2007). Ratio estimators using robust regression. **Hacettepe Journal of Mathematics and Statistics.** 36(2): 181–88.

Nangsue, N. (2009). Adjusted Ratio and Regression Type Estimators for Estimation of Population Mean when some Observations are missing. **World Academy of Science.** 29.

Sisodia, B. V. S. & Dwivedi, V. K. (1981). A Modifiedratio Estimator Using Coefficient of Variation of Auxiliary Variable. **Journal of Indian Society Agricultural Statistics.** 33 : 13-18.

Singh, H.P. & Tailor, R. (2003**).** Use of Known Correlation Coefficient in Estimating the Finite Population Mean. **Statistics in Transition.** 6 : 555-560.

T. Zaman, H. Bulut. (2019). Improvement of modified ratio estimators using robust regression methods. **Applied Mathematics and Computation.** Vol.348 : 627-631.

Upadhyaya, L. N. & Singh, H. P. (1999). Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean. **Biometrical Journal.** Vol.41 No.5 : 627- 636.