Ratio Estimator for Population Variations Using Additional Information on Simple Random Sampling

Haposan Sirait, Stepi Karolin
Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Riau, Pekanbaru, Indonesia.
haposansirait96@gmail.com

Sukono
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Bandung, Indonesia.
sukono@unpad.ac.id

Kalfin
Doctor Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia
kalfin17001@mail.unpad.ac.id

Abdul Talib Bon
Department of Production and Operations, Universiti Tun Hussein Onn Malaysia (UTHM)
Parit Raja 86400, Johor, Malaysia
talibon@gmail.com

Abstract

Variance is a parameter that is often estimated in a research, because variance can give a picture of the degree of homogeneity of a population. To estimate the variance for rare and population has been the main problem in survey sampling Variance obtained based on simple random sampling, there further modified by utilizing additional characters that are assumed to be related to the character under study. This article discusses the population mean estimator by utilizing additional character variances, and then the modified estimator will be compared to the mean square error to obtain an efficient estimator

Keywords:
Simple random sampling, Mena square error, efficient estimator

1. Introduction

In manufacturing industries and pharmaceutical laboratories sometimes researchers are interested in the variation of their products. To measure the variations within the values of study variable y, the problem of estimating the population variance of S_y^2 of study variable Y received a considerable attention of the statistician in survey sampling including Isaki (1983), Singh and Singh (2001, 2003), Jhajj et al. (2005), Kadliar and Cingi (2006), Singh et al. (2008), Grover (2010), Singh et al. (2011) and Singh and Solanki (2012) have suggested improved estimator for estimation of S_y^2 .

Simple random sampling is a method for taking *n* units of population of size *N*, where each element has the same opportunity to be taken as a sample member. Sampling can be done with a refund (with replacement) or without refund (without replacement). Cochran (1977) states that in simple random sampling without returning, the number of samples to be formed is $\boldsymbol{C_n^N}$.

Let us consider a finite population $U = (U_1, U_2, U_3, \dots, U_N)$ having N units and let Yand X are the study and auxiliary variable with means X and Y respectively. Let us suppose that a sample of size n is drawn from the

population using simple random sampling without replacement (SRSWOR) method. Let
For a simple random sample

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2,$$

is an unbiased estimate of

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{n}(y_i - \bar{Y})^2,$$

In other words, The usual sample variance estimator of the population variance was defined as:

$$\hat{S}_y^2 = s_y^2$$

The variance of the usual unbiased estimator $\hat{\boldsymbol{S}}_{\boldsymbol{y}}^{\boldsymbol{2}}$ is given by :

$$V(\hat{S}_y^2) = \frac{N-n}{Nn} S_y^4 (\frac{\mu_{40}}{\mu_{20}^2} - 1)$$

where $\mu_{rs} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^r (x_i - \bar{X})^s$
Ratio-type estimators take advantage of the correlation between the auxiliary variable, X , Z and the study variable, Y. When information is available on the auxiliary variable that is positively correlated with the study variable, the ratio estimator is a suitable estimator to estimate the population variance. For ratio estimators in sampling theory, population information of the auxiliary variable, such as the variansi , is often used to increase the efficiency of the estimation for a population variance. Abid et al.(2016) suggested a number of new ratio estimators that were modified in simple random samples Yadav et al.(2013) proposed an estimator for variance with one additional variable, According to Perri (2007), additional variables are usually used in sample survey practices to get better designs and to achieve higher accuracy in estimating population parameters. Ismail et al.(2018) modified the estimator proposed by Isaki (1983) using two additional variables. Furthermore, it is shown that the new ratio is biased so that a square error (MSE) comparison is performed. Getting smaller. The more MSE obtained the more efficient the estimator. This method can be used to increase accuracy in searching for data security. The author is interested in reviewing a new method contained in the article Ismail et al.(2018). The existing ratio estimator for estimating the population variance Y , of the study variable Y is defined as :

$$\hat{S}_R^2 = \frac{s_y^2}{s_x^2} S_x^2$$

Bias and Mean square error of ratio type variance estimator was :

$$Bias(\hat{S}_y^2) \approx \frac{1}{n} S_y^2 \left((\beta_{2(x)} - 1) - (\lambda_{22} - 1) \right)$$

and

$$MSE(\hat{S}_y^2) \approx \frac{1}{n} S_y^4 \left((\beta_{2(x)} - 1) - 2(\lambda_{22} - 1) + (\beta_{2(x)} - 1) \right)$$

Where $\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2} \mu_{02}^{s/2}}$, $\beta_{2(x)} = \frac{\mu_{40}}{\mu_{02}^2}$

2. **Proposed Modified Ratio Estimators**

Estimators for the population variance in the simple random sampling using some known auxiliary information on variance . They showed that their suggested estimators are more efficient than traditional ratio estimator in the estimation of the population variance. Ismail et al.(2018) estimators are given as :

$$\hat{S}_{rpG_1}^2 = (s_y^2 + d)\left(\omega \left(\frac{S_u^2}{s_u^2}\right)^{\tau} + (1 - \omega - \gamma)\left(\frac{s_u^2}{S_u^2}\right)^{\chi} + \psi \left(\frac{S_w^2}{s_w^2}\right)^{\phi} \right) - d.$$

and

$$\hat{S}_{rpG_2}^2 = (s_y^2 + d)\left(\omega \left(\frac{S_u^2}{s_u^2}\right)^{\tau} + (1 - \omega)\left(\frac{s_u^2}{S_u^2}\right)^{\chi} \right) - d$$

Where $\boldsymbol{s_w^2 = ms_z^2 + pS_y^2}$, $\boldsymbol{S_w^2 = (m+p)S_z^2}$, $\boldsymbol{s_u^2 = bs_x^2 + cS_x^2}$, **dan** $\boldsymbol{S_u^2 = (b+c)S_x^2}$.

Bias and Mean Square Error of Ratio Type Variance Estimator $\hat{S}^2_{rpG_1}$

In order to drive bias of ratio type variance estimator $\hat{S}^2_{rpG_1}$ by using , $e_0 = S_y^{-2}(s_y^2 - S_y^2)$, $e_1 = S_x^{-2}(s_x^2 - S_x^2)$ $e_2 = S_z^{-2}(s_z^2 - S_z^2)$ was written as:

$$\hat{S}_{rpG_1} = (S_y^2(1+e_0)+d)\left(\omega\left(\frac{b(1+e_1)S_x^2 + cS_x^2}{(b+c)S_x^2}\right)^{-\tau}\right.$$

$$+(1-\omega-\psi)\left(\frac{b(1+e_1)S_x^2 + cS_x^2}{(b+c)S_x^2}\right)^{\chi}$$

$$\left.+\psi\left(\frac{m(1+e_2)S_z^2 + pS_z^2}{(m+p)S_z^2}\right)^{-\phi}\right) - d.$$

By suppose, $b/(b+c) = \theta_1$ and $m/(m+p) = \theta_2$ obtained

$$\hat{S}_{rpG_1} = (S_y^2 + d + S_y^2 e_0)(\omega(1+\theta_1 e_1)^{-\tau} + (1-\omega-\psi)$$

$$(1+\theta_1 e_1)^{\chi} + \psi(1+\theta_2 e_2)^{-\phi}) - d.$$

By suppose $(\theta_1 e_1) = x$, let.

$$f(x) = \frac{1}{(1+\theta_1 e_1)^{\tau}},$$

$$f(x) = \frac{1}{(1+x)^{\tau}},$$

Furthermore, using the MacLaurin series approximation is obtained

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \cdots + \frac{x^n}{n!}f^n(0),$$

$$\frac{1}{(1+\theta_1 e_1)^{\tau}} = 1 + \tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2 - \cdots - \frac{\theta_1 e_1^n}{n!}f^n(0).$$

Then, the same way is done for $1/(1+\theta_1 e_1)^{-\chi}$, obtained

$$\frac{1}{(1+\theta_1 e_1)^{-\chi}} = 1 + \chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2 - \cdots - \frac{\theta_1 e_1^n}{n!}f^n(0).$$

and for $1/(1+\theta_1 e_1)^{\phi}$ obtained

$$\frac{1}{(1+\theta_2 e_2)^{\phi}} = 1 + \phi\theta_2 e_2 - \frac{\phi(\phi+1)}{2}(\theta_2 e_2)^2 - \cdots - \frac{\theta_2 e_2^n}{n!}f^n(0).$$

assumption $|\theta_1 e_1| < 1$ and $|\theta_2 e_2| < 1$, so that

$$\left(\hat{S}_{rpG_1}\right) \approx \left((S_y^2 + d)(1+\theta_0 e_0)\left(1 - \omega\left(\tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2\right)\right.\right.$$

$$+(1-\omega-\psi)\left(\chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2\right)$$

$$\left.\left.-\psi\left(\phi\theta_2 e_2 - \frac{\phi(\phi+1)}{2}(\theta_2 e_2)^2\right) - d\right)\right).$$

Apply expectation on both side of $\hat{S}_{rpG_1}$,so that ,

$$E\left(\hat{S}_{rpG_1}\right) \approx E\left((S_y^2 + d)(1+\theta_0 e_0)\left(1 - \omega\left(\tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2\right)\right.\right.$$

$$+(1-\omega-\psi)\left(\chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2\right)$$

$$-\psi\left(\phi\theta_2 e_2 - \frac{\phi(\phi+1)}{2}(\theta_2 e_2)^2\right) - d\Big)$$

or it can be written in form,

$$E\left(\hat{S}_{rpG_1}\right) = E\big((S_y^2 + d)(1+\theta_0 e_0)\omega(1+\theta_1 e_1)^{-\tau} + (1-\omega)$$
$$(1+\theta_1 e_1)^{\chi} - d\big)$$

and using $E(e_0) = 0,\ E(e_1) = 0$,E($e_2) = 0$ we obtained,

$$E\left(\hat{S}_{rpG_1}\right) \approx S_y^2 - (\beta_2'(x) w_1^2 + (1-\omega-\psi)\chi^2$$
$$+\beta_2'(x)(\omega\tau^2 + (1-\omega-\psi)\chi) w_1(w_1 - 2v'))$$

Finally,

$$B\left(\hat{S}_{rpG_1}\right) \approx (\beta_2'(x) w_1^2 + (1-\omega-\psi)\chi^2$$
$$+\beta_2'(x)(\omega\tau^2 + (1-\omega-\psi)\chi) w_1(w_1 - 2v'))$$

In order to drive mean square error of $\hat{S}_{rpG_1}$,

$$MSE\left(\hat{S}_{rpG_1}\right) = E\left(\hat{S}_{rpG_1} - S_y^2\right)^2.$$

The ratio estimator is $\hat{S}_{rpG_1}$, expressed as a function $f(s_y^2, s_x^2, s_z^2)$ and the estimated parameter S_y^2 is $f(S_y^2, S_x^2, S_z^2)$, so that：

$$MSE\left(f(s_y^2, s_x^2, s_z^2)\right) = E\left(f(s_y^2, s_x^2, s_z^2) - f(S_y^2, S_x^2, S_z^2)\right)^2.$$

by approximating to $f(s_y^2, s_x^2, s_z^2)$ with the Taylor series and ignoring degrees greater than one, obtained :

$$f(s_y^2, s_x^2, s_z^2) = f(s_y^2, s_x^2, s_z^2)\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}} + \left((s_y^2 - S_y^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right.$$
$$+(s_x^2 - S_x^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}$$
$$\left.+(s_z^2 - S_z^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_z^2}\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)^2.$$

Equivalent to

$$f(s_y^2, s_x^2, s_z^2) = f(S_y^2, S_x^2, S_z^2) + \left((s_y^2 - S_y^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right.$$
$$+(s_x^2 - S_x^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Big|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}$$

$$+(s_z^2 - S_z^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_z^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\Bigg)^2$$

Thus.

$$\left(f(S_y^2, S_x^2, S_z^2) - f(S_y^2, S_x^2, S_z^2)\right)^2 = \Bigg((s_y^2 - S_y^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}$$

$$+(s_x^2 - S_x^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}$$

$$+(s_z^2 - S_z^2)\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_z^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\Bigg)^2$$

Apply expectation on both side, so that,

$$MSE\left(f(s_y^2, s_x^2, s_z^2)\right) = E\left(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)^2 (s_y^2 - S_y^2)^2\right)$$

$$+E\left(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)^2 (s_x^2 - S_x^2)^2\right)$$

$$+E\left(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_z^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)^2 (s_z^2 - S_z^2)^2\right)$$

$$+2E\Bigg(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)$$

$$\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_y^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)\left((s_y^2 - S_y^2)\right.$$

$$\left.(s_x^2 - S_x^2)\right)\Bigg) + 2E\Bigg(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)$$

$$\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\Bigg|_{\substack{s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2}}\right)\left((s_x^2 - S_x^2)\right.$$

$$(s_x^2 - S_x^2))\Big) + 2E\left(\left(\frac{\partial f(s_y^2, s_x^2, s_z^2)}{\partial s_x^2}\middle|\begin{matrix} s_y^2 = S_y^2 \\ s_x^2 = S_x^2 \\ s_z^2 = S_z^2 \end{matrix}\right)\right.$$

$$\left((s_y^2 - S_y^2)(s_x^2 - S_x^2)\right)$$

In connection $f(s_y^2, s_x^2, s_z^2) = \hat{S}_{rpG_1}$, then :

$$MSE(\hat{S}_{rpG_1}) \approx E\left((s_y^2 - S_y^2)^2\right) + E\left(\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2\right.$$

$$(s_x^2 - S_x^2)^2\Big) + E\left(\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2\right.$$

$$(s_z^2 - S_z^2)^2\Big) + 2E\left(\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2\right.$$

$$\left.(s_y^2 - S_y^2)(s_x^2 - S_x^2)\right) + 2E$$

$$\left(\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)\right.$$

$$(s_x^2 - S_x^2)(s_z^2 - S_z^2)\Big) + 2E\left(\left(\frac{-S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)\right.$$

$$\left.(s_y^2 - S_y^2)(s_z^2 - S_z^2)\right)$$

Therefore,

$$MSE(\hat{S}_{rpG_1}) \approx \alpha S_y^4(\beta_{2(y)} - 1) + \left(\frac{S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2$$

$$\alpha S_y^4(\beta_{2(x)} - 1) + \left(\frac{S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2$$

$$\alpha S_y^4(\beta_{2(z)} - 1) - 2\left(\frac{S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)$$

$$\alpha s_y^2 s_x^2(\lambda_{22} - 1) + 2\left(\frac{S_y^2}{(S_x^2 + S_z^2 + kv)^2}\right)^2$$

or it can be written in form

$$MSE(\hat{S}_{rpG_1}^2) = \alpha S_y^4\left((\beta_{2(y)} - 1) + A_1\left(A_1(\beta_{2(x)} - 1)\right) + (\beta_{2(z)} - 1) + 2(\lambda_{22} - 1)\right) - 4(\lambda_{22} - 1),$$

$$MSE(\hat{S}_{rpG_2}^2) = \alpha S_y^4\left((\beta_{2(y)} - 1) + A_2\left(A_2(\beta_{2(x)} - 1)\right) + (\beta_{2(z)} - 1) + 2(\lambda_{22} - 1)\right),$$

Where $\boldsymbol{A_1 = S_y^2/(S_x^2 + S_z^2 + kv)}$ and $\boldsymbol{A_2 = S_y^2/(S_y^2 + kv)}$.

Bias and Mean Square Error of Ratio Type Variance Estimator $\hat{S}_{rpG_2}^2$

In order to drive bias of ratio type variance estimator $\hat{S}_{rpG_2}^2$

$$\hat{S}_{rpG_2} = (s_y^2 + d)\left(\omega\left(\frac{(b + c)S_x^2}{bs_x^2 + cS_x^2}\right)^\tau + (1 - \omega)\left(\frac{bs_x^2 + cS_x^2}{(b + c)S_x^2}\right)^\chi\right) - d,$$

by using , $e_0 = S_y^{-2}(s_y^2 - S_y^2)$, $e_1 = S_x^{-2}(s_x^2 - S_x^2)$ $e_2 = S_z^{-2}(s_z^2 - S_z^2)$ was written as:

$$\hat{S}_{rpG_2} = (S_y^2(1 + e_0) + d)\left(\omega\left(\frac{b(1 + e_1)S_x^2 + cS_x^2}{(b + c)S_x^2}\right)^{-\tau}\right.$$

$$\left(\frac{b(1+e_1)S_x^2 + cS_x^2}{(b+c)S_x^2}\right)^{\chi}\Bigg) - d.$$

By suppose, $b/(b+c) = \theta_1$, obtained

$$\hat{S}_{rpG_2} = \left(S_y^2(1+e_0) + d\right)(\omega(1+\theta_1 e_1)^{-\tau} + (1-\omega)$$
$$(1+\theta_1 e_1)^{\chi}) - d.$$

Furthermore, using the MacLaurin series approximation is obtained

Assumption $|\theta_1 e_1| < 1$ and $|\theta_2 e_2| < 1$, so that

$$\hat{S}_{rpG_2} = \left(\left(S_y^2 + d\right)(1+\theta_0 e_0)1 - \omega\left(\tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2\right)\right.$$
$$\left.(1-\omega)\left(\chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2\right) - d - S_y^2\right).$$

Apply expectation on both side of $\hat{S}_{rpG_2}$,

$$E\left(\hat{S}_{rpG_2}\right) = E\left(\left(S_y^2 + d\right)(1+\theta_0 e_0)1 - \omega\left(\tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2\right)\right.$$
$$\left.(1-\omega)\left(\chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2\right) - d - S_y^2\right).$$
$$= E\left(\left(S_y^2 + d\right)\theta_0 e_0 - \omega\left(\tau\theta_1 e_1 - \frac{\tau(\tau+1)}{2}(\theta_1 e_1)^2\right)\right.$$
$$+(1-\omega)\left(\chi\theta_1 e_1 - \frac{\chi(\chi+1)}{2}(\theta_1 e_1)^2\right) - \omega\tau\theta_0\theta_1 e_0 e_1$$
$$\left.+(1-\omega)\chi\theta_0\theta_1 e_0 e_1 - d - S_y^2\right),$$

And using $E(e_{0)} = 0$, $E(e_1) = 0$, $\mathrm{E}(\ e_2) = 0$ we obtained,

$$E\left(\hat{S}_{rpG_2}\right) = S_y^2 - \left(\beta_2'(x)w_1(\omega\tau^2 + (1-\omega)\chi^2)w_1(w_1 2v')\right)$$

Finally,

$$Bias\left(\hat{S}_{rpG_2}\right) = \beta_2'(x)w_1(\omega\tau^2 + (1-\omega)\chi^2)w_1(w_1 2v').$$

In order to drive mean square error of $\hat{S}_{rpG_2}$, The ratio estimator is $\hat{S}_{rpG_2}$, expressed as a function $f(s_y^2, s_x^2, s_z^2)$ and the estimated parameter S_y^2 is $f(S_y^2, S_x^2, S_z^2)$, so that in the same way, it is earned,

$$MSE\left(f\left(s_y^2, s_x^2, s_z^2\right)\right) = E\left(\left(s_y^2 - S_y^2\right)^2\right) + E\left(\left(\frac{-S_y^2}{(S_x^2 + kv)^2}\right)^2\right.$$
$$\left.(s_x^2 - S_x^2)^2\right) + 2E\left(\left(\frac{-S_y^2}{(S_x^2 + kv)^2}\right)^2\right.$$
$$\left.\left(s_y^2 - S_y^2\right)(s_x^2 - S_x^2)\right)$$

$$MSE\left(\hat{S}_{rpG_2}\right) \approx \alpha S_y^4\left(\beta_{2(y)} - 1\right) + \left(\frac{-S_y^2}{(S_x^2 + kv)^2}\right)^2$$
$$\alpha S_y^4\left(\beta_{2(x)} - 1\right) - 2\left(\frac{-S_y^2}{(S_x^2 + kv)^2}\right)$$
$$\alpha s_y^2 s_x^2(\lambda_{22} - 1)$$

or it can be written in form

$$MSE(\hat{S}_{rpG_2}) \approx \alpha S_y^4\left((\beta_{2(y)}-1)+A_2^2(\beta_{2(x)}-1)-2A_2(\lambda_{22}-1)\right)$$

Where $S_y^2/(S_x^2+kv)=A_2,\ \alpha=1/n$

3. Comparison of the estimators .

We now compare the proposed estimator with other estimators considered here. These comparisons lead to the following obvious conditions. From the expressions of the MSE of the proposed estimators and the usually estimators, we have derived the conditions for which the proposed estimators are more efficient than the variance ratio estimators as follows :

(1). **Comparison between estimator $\hat{S}^2_{rpG_1}$ and estimator s^2_y**

$$MSE(\hat{S}^2_{rpG_1}) < V(\hat{S}^2_y)$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_1\left(A_1(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)-4(\lambda_{22}-1) < \frac{N-n}{Nn}S_y^4(\frac{\mu_{40}}{\mu_{20}^2}-1)$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_1\left(A_1(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)-4(\lambda_{22}-1)-\frac{N-n}{Nn}S_y^4\left(\frac{\mu_{40}}{\mu_{20}^2}-1\right)<0$$

By resolving this inequality it is obtained that estimator $\hat{S}^2_{rpG_1}$ is more efficient than s^2_y if

$$\beta_{2(x)} > \frac{(\beta_{2(z)}-1)+2A_1(\lambda_{22}-1)-4(\lambda_{22}-1)}{A_1}.$$

(2).**Comparison between estimator $\hat{S}^2_{rpG_2}$ and estimator s^2_y**

$$MSE(\hat{S}^2_{rpG_2}) < V(\hat{S}^2_y)$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_2\left(A_2(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right) < \frac{N-n}{Nn}S_y^4(\frac{\mu_{40}}{\mu_{20}^2}-1)$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_2\left(A_2(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)-\frac{N-n}{Nn}S_y^4\left(\frac{\mu_{40}}{\mu_{20}^2}-1\right)<0$$

By resolving this inequality it is obtained that estimator $\left(\hat{S}^2_{rpG_2}\right)$ is more efficient than s^2_y if

$$S_x^2 > \frac{(\beta_{2(x)}-1)-2(\lambda_{22}-1)}{A_2}.$$

(3).**Comparison between estimator $\hat{S}^2_{rpG_1}$ and estimator $\hat{S}^2_{rpG_2}$**

$$MSE(\hat{S}^2_{rpG_1}) < MSE(\hat{S}^2_{rpG_2})$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_1\left(A_1(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)-4(\lambda_{22}-1) < \alpha S_y^4\left((\beta_{2(y)}-1)+A_2\left(A_2(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)$$

$$\alpha S_y^4\left((\beta_{2(y)}-1)+A_1\left(A_1(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)-4(\lambda_{22}-1) - \left(\alpha S_y^4\left((\beta_{2(y)}-1)+A_2\left(A_2(\beta_{2(x)}-1)\right)+(\beta_{2(z)}-1)+2(\lambda_{22}-1)\right)\right)<0$$

By resolving this inequality it is obtained that estimator $\hat{S}^2_{rpG_1}$ is more efficient than $\hat{S}^2_{rpG_2}$ if

$$A_1 < A_2$$

and

$$\beta_{2(x)} > \frac{2(\lambda_{22}-1)-A_1\left(A_1(\beta_{2(z)}-1)-4(\lambda_{22}-1)\right)}{(A_1+A_2)}$$

4. Numerical illustration

We use the following data. Source: Basic Econometrics, Gujarati & Sangeetha (2007) Data collected from 64 countries regarding fertility and other factors: number of children died in a year under age 5 per one thousand live-births, literacy rate of females in percentage, and total fertility rate during 1980-1985.

Table 1. Data description

C_y	C_x	C_z	S_y^2	S_x^2	S_z^2
0.53695	0.50809	0.2719	5772.6667	676.4087	2.327706
ρ_{yx}	ρ_{yz}	ρ_{xz}	$\beta_{2(y)}$	$\beta_{2(x)}$	$\beta_{2(z)}$
0.81829	0.67114	0.62595	237833	1.6575111	2.81691

Table 2: MSE and PRE of the estimators

Estimators	MSE	PRE
$\hat{S}_y^2$	11627.2	100
$\hat{S}_{rpG_1}^2$	3927.17	296.0707
$\hat{S}_{rpG_2}$	3473.024	334.7860

where Percentage Relative Efficiency (PRE) is computed as $PRE(\hat{\theta}) = \frac{V(s_y^2)}{MSE(\hat{\theta})} x\ 100\%$

5. Conclusion

From the discussion the results showed that the proposed ratio estimator was more efficient compared to the usual sample variance estimator, assuming given conditions. has a higher precision than the normal estimator, and the more additional information content that is included in the modifiers of the estimator makes the estimator more efficient, of course if it meets the efficiency requirements

References

Isaki, C.T.(1983): Variance estimation using auxiliary information. Journal of American Statistical Association 78, 117–123.

Jhajj, H .S., Sharma, M. K. and Grover, L. K. (2005) : An efficient class of chain estimators of population variance under sub-sampling scheme. J. Japan Stat. Soc., 35(2), 273-286

Kadilar, C. and Cingi, H. (2006) : Improvement in variance estimation using auxiliary information. Hacettepe Journal of Mathematics and Statistics 35 (1), 111–115.

Kaur,(1985): An efficient regression type estimator in survey sampling, Biom. Journal 27 (1),107–110. Ray, S.K. and Sahai, A (1980) : Efficient families of ratio and product type estimators, Biometrika 67(1), 211–215

M. Abid, N. Abbas, R. A. K. Sherwani, dan H. Z. Nazir, *Improved ratio estimator for the population mean using non-conventional measures of dispension*, Pakistan Journal of Statistic and Operation Research, 12 (2016), 353-367.

W. G. Cochran, *Sampling Techniques, Third Edition*, John Wiley, New York, 1977.

Gujarati, D. N. & Sangeetha. (2007). Basic econometrics (4th Ed.). New Delhi: TATA McGRAW HILL Companies. Pp. 189.

G. Diana, dan P. F. Perri, *Estimation of finite population mean using multiauxilary information*, International Journal of Statistics, 65 (2007), 99- 112.

C. T. Isaki, *Variance estimation using auxiliary information*, Journal of the American Statistical Association, 78 (1983),117-123.

M. Ismail, N. Kanwal, dan M. Q. Shahbaz, *Generalized ratio-product-type estimator for variance using auxiliary information in simple random sampling*, Kuwait Journal of Science, 45 (1) (2018), 79-88.

R. Yadav, L. N. Upadhayaya, H. P. Singh, dan S. Chatterje, *A generalized family of transformed ratio-product estimator for variance in sample surveys*, Communications in Statistics-Theory and Methods, 42 (2013), 1839-1850.

Biographies

Haposan Sirait is a lecturer in the Study Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Riau, Pekanbaru, Indonesia. The field of applied mathematics, with a field of concentration of Statistics.

Stepi Karolin is a lecturer in the Study Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Riau, Pekanbaru, Indonesia. The field of applied mathematics, with a field of concentration of Statistics.

Sukono is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Bandung, Indonesia. Currently, as Chair of the Research Collaboration Community (RCC), the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

Kalfin is currently a Student at the Doctoral Program at Padjadjaran University, Indonesia since 2019. He received his M.Mat in Mathematics from Padjadjaran University (UNPAD), Indonesia in 2019. His current research focuses on financial mathematics and actuarial sciences.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a Ph.D. in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He's bachelor degree and diploma in Mechanical Engineering which he obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM, and MIM.