Available online at www.sciencedirect.com

SCIENCE DIRECT®

Applied Mathematics Letters 19 (2006) 75–79

Applied Mathematics Letters

www.elsevier.com/locate/aml

Improvement in estimating the population mean in simple random sampling

Cem Kadilar*, Hulya Cingi

Department of Statistics, Hacettepe University, Beytepe, Ankara 06800, Turkey

Received 2 February 2005; accepted 24 February 2005

Abstract

This paper proposes some estimators for the population mean using the ratio estimators presented in [C. Kadilar, H. Cingi, Ratio estimators in simple random sampling, Applied Mathematics and Computation 151 (2004) 893–902] and shows that all proposed estimators are always more efficient than the ratio estimators. This result is also supported by a numerical example.

Keywords: Ratio-type estimators; Simple random sampling; Mean square error; Auxiliary information; Efficiency

1. Introduction

Kadilar and Cingi [1] suggested the following ratio estimators for the population mean $\bar{Y}$ of the variate of interest y in simple random sampling:

$$\bar{y}_{\mathrm{KC1}} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X}, \tag{1}$$

$$\bar{y}_{\mathrm{KC2}} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x} (\bar{X} + C_x), \tag{2}$$

$$\bar{y}_{\mathrm{KC3}} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + \beta_2(x)} [\bar{X} + \beta_2(x)], \tag{3}$$

* Corresponding author. Tel.: +90 312 299 2016; fax: +90 312 297 7913.
E-mail addresses: kadilar@hacettepe.edu.tr (C. Kadilar), hcingi@hacettepe.edu.tr (H. Cingi).

0893-9659/$ - see front matter
doi:10.1016/j.aml.2005.02.039

$$\bar{y}_{KC4} = \frac{\bar{y} + b\,(\bar{X} - \bar{x})}{\bar{x}\beta_2(x) + C_x}[\bar{X}\beta_2(x) + C_x], \tag{4}$$

$$\bar{y}_{KC5} = \frac{\bar{y} + b\,(\bar{X} - \bar{x})}{\bar{x}C_x + \beta_2(x)}[\bar{X}\,C_x + \beta_2(x)], \tag{5}$$

where C_x and $\beta_2(x)$ are the population coefficient of variation and the population coefficient of the kurtosis, respectively, of the auxiliary variate, $\bar{y}$ is the sample mean of the variate of interest, $\bar{x}$ is the sample mean of the auxiliary variate and it is assumed that the population mean $\bar{X}$ of the auxiliary variate x is known, and $b = \frac{s_{xy}}{s_x^2}$ is the regression coefficient. Here s_x^2 is the sample variance of the auxiliary variate and s_{xy} is the sample covariance between the auxiliary variate and the variate of interest.

In [1], mean square error (MSE) equations of these ratio estimators were given by

$$\mathrm{MSE}(\bar{y}_{KC1}) \cong \frac{1-f}{n}[R_{KC1}^2 S_x^2 + S_y^2(1-\rho^2)], \tag{6}$$

$$\mathrm{MSE}(\bar{y}_{KC2}) \cong \frac{1-f}{n}[R_{KC2}^2 S_x^2 + S_y^2(1-\rho^2)], \tag{7}$$

$$\mathrm{MSE}(\bar{y}_{KC3}) \cong \frac{1-f}{n}[R_{KC3}^2 S_x^2 + S_y^2(1-\rho^2)], \tag{8}$$

$$\mathrm{MSE}(\bar{y}_{KC4}) \cong \frac{1-f}{n}[R_{KC4}^2 S_x^2 + S_y^2(1-\rho^2)], \tag{9}$$

$$\mathrm{MSE}(\bar{y}_{KC5}) \cong \frac{1-f}{n}[R_{KC5}^2 S_x^2 + S_y^2(1-\rho^2)], \tag{10}$$

respectively, where $f = \frac{n}{N}$, n is the sample size, N is the population size, $R_{KC1} = R = \frac{\bar{Y}}{\bar{X}}$ is the population ratio, $R_{KC2} = \frac{\bar{Y}}{\bar{X}+C_x}$, $R_{KC3} = \frac{\bar{Y}}{\bar{X}+\beta_2(x)}$, $R_{KC4} = \frac{\bar{Y}\,\beta_2(x)}{\bar{X}\beta_2(x)+C_x}$ and $R_{KC5} = \frac{\bar{Y}\,C_x}{\bar{X}\,C_x+\beta_2(x)}$, S_x^2 and S_y^2 are the population variances of the auxiliary variate and of the variate of interest, respectively, and ρ is the population coefficient of correlation between the auxiliary variate and the variate of interest.

Kadilar and Cingi [1] concluded that all ratio estimators, given above, were more efficient than traditional estimators, presented in [2,3] and [4], under certain conditions. In addition, this result was satisfied with the aid of a numerical example, whose data will also be used in this paper. Note that Kadilar and Cingi [5] adapted these traditional estimators in the simple random sampling to the stratified random sampling and then Kadilar and Cingi [6] proposed a new ratio estimator that was always more efficient than these adapted estimators in stratified random sampling.

In the next section, we develop new estimators combining ratio estimators in [1] and obtain the MSE equation of these new estimators. We compare the efficiencies, based on MSE equations, between the proposed estimators and ratio estimators, theoretically, in Section 3 and also we show this theoretical comparison numerically in Section 4. In the last section, we give a hint to obtain different estimators by a similar method presented in this study.

2. Suggested estimators

We propose the estimator combining ratio estimators (1) and (2) as follows:

$$\bar{y}_{pr1} = \omega_1 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}}\bar{X} + \omega_2 \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x}(\bar{X} + C_x), \tag{11}$$

where ω_1 and ω_2 are weights that satisfy the condition: $\omega_1 + \omega_2 = 1$.

The MSE of this estimator can be found using the first degree approximation in the Taylor series method defined by

$$\mathrm{MSE}(\bar{y}_{pr1}) \cong \boldsymbol{d\Sigma d}', \tag{12}$$

where

$$\boldsymbol{d} = \left[\left.\frac{\partial h(a,b)}{\partial a}\right|_{\bar{Y},\bar{X}} \quad \left.\frac{\partial h(a,b)}{\partial b}\right|_{\bar{Y},\bar{X}} \right]$$

$$\boldsymbol{\Sigma} = \frac{1-f}{n}\begin{bmatrix} S_y^2 & S_{yx} \\ S_{xy} & S_x^2 \end{bmatrix}$$

(see [7]). Here $h(a,b) = h(\bar{y},\bar{x}) = \bar{y}_{pr1}$. According to this definition, we obtain $\boldsymbol{d}$ for the proposed estimator as follows:

$$\boldsymbol{d} = \left[1 \quad -\omega_1(B+R) - \omega_2(B+R_{\mathrm{KC2}})\right],$$

where $B = \frac{S_{xy}}{S_x^2} = \frac{\rho S_x S_y}{S_x^2} = \frac{\rho S_y}{S_x}$. Note that we omit the difference: $b - B$ [8].

We obtain the MSE of the proposed estimator using (12) as

$$\mathrm{MSE}(\bar{y}_{pr1}) = \frac{1-f}{n}(S_y^2 - 2\eta S_{yx} + \eta^2 S_x^2), \tag{13}$$

where

$$\eta = \omega_1(B+R) + \omega_2(B+R_{\mathrm{KC2}}). \tag{14}$$

We also propose the estimator combining ratio estimators (1) and (3) as

$$\bar{y}_{pr2} = \omega_1\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}}\bar{X} + \omega_2\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}+\beta_2(x)}[\bar{X}+\beta_2(x)]. \tag{15}$$

The MSE of this estimator is the same as (13) but R_{KC2} in (14) is replaced with R_{KC3}.

In addition, we propose the following estimator combining ratio estimators (1) and (4):

$$\bar{y}_{pr3} = \omega_1\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}}\bar{X} + \omega_2\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}\beta_2(x)+C_x}[\bar{X}\beta_2(x)+C_x]. \tag{16}$$

The MSE of this estimator is again the same as (13) but R_{KC2} in (14) is replaced with R_{KC4}.

Lastly, we propose the estimator combining ratio estimators (1) and (5) as

$$\bar{y}_{pr4} = \omega_1\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}}\bar{X} + \omega_2\frac{\bar{y}+b(\bar{X}-\bar{x})}{\bar{x}C_x+\beta_2(x)}[\bar{X}C_x+\beta_2(x)]. \tag{17}$$

The MSE of this estimator is also the same as (13) but R_{KC2} in (14) is replaced with R_{KC5}.

The optimal values of ω_1 and ω_2 to minimize (13) can easily be found as follows:

$$\omega_1^* = \frac{R_{\mathrm{KC2}}}{R_{\mathrm{KC2}} - R} \qquad \text{and} \qquad \omega_2^* = \frac{R}{R - R_{\mathrm{KC2}}}. \tag{18}$$

When we use ω_1^* and ω_2^* instead of ω_1 and ω_2 in (14), respectively, we get $\eta = B$. As η is independent of R_{KC2}, all proposed estimators have the same minimum MSE as follows:

$$\mathrm{MSE}_{\min}(\bar{y}_{pr}) = \frac{1-f}{n}(S_y^2 - 2BS_{yx} + B^2 S_x^2).$$

Table 1
Data statistics

$N = 106$	$\bar{Y} = 2212.59$	$R = 0.0807$
$n = 20$	$\bar{X} = 27\,421.70$	$R_{\text{KC2}} = 0.0807$
$\rho = 0.86$	$S_y = 11\,551.53$	$R_{\text{KC3}} = 0.0806$
$C_y = 5.22$	$S_x = 57\,460.61$	$R_{\text{KC4}} = 0.0807$
$C_x = 2.10$	$\beta_2(x) = 34.57$	$R_{\text{KC5}} = 0.0806$
$S_{yx} = 568\,176\,176.10$		

Table 2
MSE values of estimators

$\bar{y}_{\text{KC1}}$	2318 722.45
$\bar{y}_{\text{KC2}}$	2318 589.19
$\bar{y}_{\text{KC3}}$	2316 527.82
$\bar{y}_{\text{KC4}}$	2318 718.59
$\bar{y}_{\text{KC5}}$	2317 674.08
Proposed	1446 719.34

We can also write this expression by

$$\text{MSE}_{\min}(\bar{y}_{pr}) \cong \frac{1-f}{n} S_y^2 (1-\rho^2). \tag{19}$$

3. Efficiency comparisons

In this section, we compare the MSE of proposed estimators in (19) with the MSE of ratio estimators in [1] given in (6)–(10). As we obtain the following condition by these comparisons

$$R_{\text{KC}\,i}^2 S_x^2 > 0, \qquad i = 1, 2, \ldots, 5 \tag{20}$$

we can infer that all proposed estimators are more efficient than all ratio estimators in [1] in all conditions, because the condition given in (20) is always satisfied.

4. Numerical illustration

We have used the same data, concerning the level of apple production (as the variate of interest) and number of apple trees (as the auxiliary variate), as in [1] to compare the efficiencies of the proposed estimators with the ratio estimators numerically.

In Table 1, we observe the statistics about the population. Note that we take the sample size as $n = 20$ [9]. We would like to recall that sample size has no effect on efficiency comparisons of estimators, as shown in Section 3.

In Table 2, values of MSE, which are computed using equations presented in Sections 1 and 2, are given. When we examine Table 2, we observe that the proposed estimators have the smallest MSE value among all ratio estimators given in Section 1. This is an expected result, as mentioned in Section 3.

From the result of this numerical illustration, we deduce that all proposed estimators are more efficient than ratio-type estimators that were more efficient than all traditional estimators for this data set in [1].

5. Conclusion

We have developed new estimators combining ratio estimators considered in [1] and obtained the minimum MSE equation for proposed estimators. Theoretically, we have demonstrated that all proposed estimators are always more efficient than ratio estimators. In addition, we support this theoretical result numerically using the same data set as in [1].

Some other estimators can also be derived combining ratio estimators given in (2)–(5) in the form (11), but all these estimators have again the same minimum MSE equation given in (19). We would like to recall that R and R_{KC2} in (14) and in (18) should be changed according to ratio estimators that are combined.

References

[1] C. Kadilar, H. Cingi, Ratio estimators in simple random sampling, Applied Mathematics and Computation 151 (2004) 893–902.
[2] B.V.S. Sisodia, V.K. Dwivedi, A modified ratio estimator using coefficient of variation of auxiliary variable, Journal of Indian Society Agricultural Statistics 33 (1981) 13–18.
[3] H.P. Singh, M.S. Kakran, A modified ratio estimator using known coefficient of kurtosis of an auxiliary character, 1993 (unpublished).
[4] L.N. Upadhyaya, H.P. Singh, Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal 41 (1999) 627–636.
[5] C. Kadilar, H. Cingi, Ratio estimators in stratified random sampling, Biometrical Journal 45 (2003) 218–225.
[6] C. Kadilar, H. Cingi, A new ratio estimator in stratified random sampling, Communications in Statistics: Theory and Methods 34 (2005) 597–602.
[7] K.M. Wolter, Introduction to Variance Estimation, Springer-Verlag, 1985.
[8] W.G. Cochran, Sampling Techniques, John Wiley and Sons, 1977.
[9] H. Cingi, Sampling Theory, Hacettepe University Press, 1994.