Hacettepe Journal of Mathematics and Statistics
Volume 32 (2003), 105 – 108

A STUDY ON THE CHAIN RATIO-TYPE ESTIMATOR

Cem Kadılar* and Hülya Çıngı*

Received 17. 03. 2003 : Accepted 16. 12. 2003

Abstract

We examine the chain ratio-type estimator and obtain its MSE equation. We prove that the chain ratio-type estimator is more efficient than the traditional ratio estimator under certain conditions. In addition, this proof is supported by an application with original data.

Keywords: Chain ratio-type estimator, Sampling, Efficiency.

1. Introduction

The classical ratio estimator for the population mean $\overline{Y}$ of the variate of interest y is defined by

$$(1) \qquad \bar{y}_r = \frac{\bar{y}}{\bar{x}}\overline{X},$$

where it is assumed that the population mean $\overline{X}$ of the auxiliary variate x is known. Here

$$(2) \qquad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n},$$

where n is the number of units in the sample [1].

The MSE of the classical ratio estimator is

$$(3) \qquad \text{MSE}(\bar{y}_r) \cong \frac{1-f}{n}(R^2S_x^2 - 2R\rho S_x S_y + S_y^2),$$

where $f = \frac{n}{N}$; N is the number of units in the population; $R = \frac{\overline{Y}}{\overline{X}}$ is the population ratio; S_x^2 is the population variance of the auxiliary variate and S_y^2 is the population variance of the variate of interest [2].

*Hacettepe University, Faculty of Science, Department of Statistics, 06532 Beytepe, Ankara, Turkey. E-mail: kadilar@hacettepe.edu.tr and hcingi@hacettepe.edu.tr

2. The Chain Ratio-Type Estimator

When $\bar{y}$ in (1) is replaced with $\bar{y}_r$, the chain ratio-type estimator is obtained as

(4) $$\bar{y}_{cr} = \frac{\bar{y}_r}{\bar{x}}\overline{X}.$$

We can re-write (4) using (1) as

(5) $$\bar{y}_{cr} = \frac{\bar{y}}{\bar{x}^2}\overline{X}^2.$$

In general, Srivastava [8] presents this type of estimator in the form

$$\bar{y}_{cr}(\alpha) = \bar{y}\left(\frac{\overline{X}}{\bar{x}}\right)^{\alpha},$$

where the optimal value of α is $\frac{\rho C_y}{C_x}$ [7]. Rao [7] calls this general form of the estimator the *repeated substitution estimator*. In addition, Kiregyera [5] examined the chain ratio-type estimator when there are two auxiliary variates.

The MSE of the chain ratio-type estimator can be obtained using the Taylor series method. In general, this method for k variables can be given as

$$h(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k) = h(\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_k) + \sum_{j=1}^{k} d_j(\bar{x}_j - \overline{X}_j) + R_k(\overline{X}_k, a),$$

where

$$d_j = \frac{\partial h(a_1, a_2, \ldots, a_k)}{\partial a_j}$$

and

$$R_k(\overline{X}_k, a) = \sum_{j=1}^{k}\sum_{i=1}^{k}\frac{1}{2!}\frac{\partial^2 h(\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_k)}{\partial \overline{X}_i \partial \overline{X}_j}(\bar{x}_j - \overline{X}_j)(\bar{x}_i - \overline{X}_i) + O_k,$$

where O_k represents the remainder of the Taylor series expansion having terms of degree higher than two [9]. When we omit the term $R_k(\overline{X}_k, a)$, we obtain the Taylor series method for two variables as

(6) $$h(\bar{x}, \bar{y}) - h(\bar{X}, \overline{Y}) \cong \left.\frac{\partial h(c,d)}{\partial c}\right|_{\overline{X},\overline{Y}}(\bar{x} - \overline{X}) + \left.\frac{\partial h(c,d)}{\partial d}\right|_{\overline{X},\overline{Y}}(\bar{y} - \overline{Y}),$$

see [6]. Here

(7) $$h(\bar{x}, \bar{y}) = \frac{\bar{y}}{\bar{x}^2} = \widehat{R}_c \quad \text{and} \quad h(\overline{X}, \overline{Y}) = \frac{\overline{Y}}{\overline{X}^2} = R_c.$$

From (6) and (7), we can write

$$\begin{aligned}
\widehat{R}_c - R_c &\cong \frac{-2\overline{X}\,\overline{Y}}{\overline{X}^4}(\bar{x} - \overline{X}) + \frac{1}{\overline{X}^2}(\bar{y} - \overline{Y}) \\
\mathrm{E}(\widehat{R}_c - R_c)^2 &\cong \frac{4\overline{X}^2\overline{Y}^2}{\overline{X}^8}V(\bar{x}) - \frac{4\overline{X}\,\overline{Y}}{\overline{X}^6}\mathrm{Cov}(\bar{x}, \bar{y}) + \frac{1}{\overline{X}^4}V(\bar{y}) \\
&\cong \frac{1}{\overline{X}^4}\left\{\frac{4\overline{Y}^2}{\overline{X}^2}V(\bar{x}) - \frac{4\overline{Y}}{\overline{X}}\mathrm{Cov}(\bar{x}, \bar{y}) + V(\bar{y})\right\} \\
&\cong \frac{1-f}{n\overline{X}^4}\{4R^2S_x^2 - 4R\rho S_x S_y + S_y^2\}.
\end{aligned}$$

From this equivalence, we obtain the MSE of the chain ratio-type estimator given by Equation 5 in the form

$$\text{MSE}(\bar{y}_{cr}) \cong \frac{1-f}{n}\{4R^2S_x^2 - 4R\rho S_x S_y + S_y^2\} \tag{8}$$

and the bias of this estimator as

$$\text{B}(\bar{y}_{cr}) \cong \frac{1-f}{n}(S_y^2 - 2RS_x^2).$$

This bias term is omitted for the first degree approximation in the Taylor series method while obtaining the MSE of the estimator.

3. Efficiency Comparison

If we compare the MSE of the chain ratio-type estimator with the MSE of the traditional ratio estimator we have the condition

$$\begin{aligned}
&\frac{1-f}{n}\{4R^2S_x^2 - 4R\rho S_x S_y + S_y^2\} < \frac{1-f}{n}(R^2S_x^2 - 2R\rho S_x S_y + S_y^2)\\
\iff\ & 4R^2S_x^2 - 4R\rho S_x S_y < R^2S_x^2 - 2R\rho S_x S_y\\
\iff\ & 3R^2S_x^2 - 2R\rho S_x S_y < 0\\
\iff\ & 3 - \frac{2\rho S_y}{RS_x} < 0\\
\iff\ & \frac{3}{2} < \frac{\rho C_y}{C_x}\\
\iff\ & \frac{3}{2} < \alpha,
\end{aligned}$$

where $C_x = \frac{S_{\bar{x}}}{\overline{X}}$ is the population coefficient of variation of the auxiliary variate, and $C_y = \frac{S_{\bar{y}}}{\overline{Y}}$ is the population coefficient of variation of the variate of interest. When this restrictive condition is satisfied, the chain ratio estimator is more efficient than the traditional ratio estimator. It was because Rao [7] took 20 populations with similar values for C_x and C_y in each population, in other words, populations not satisfying the restrictive condition given above, that the chain ratio estimator was found to be worthless.

4. An Application

In this section, we apply the classical and chain ratio-type estimators to data concerning the level of apple production (as the variate of interest) and the number of apple trees (as the auxiliary variate) in 106 villages in the Marmara region in Turkey in 1999 (Source: Institute of Statistics, Republic of Turkey). We take the sample size as $n = 20$ using simple random sampling [2]. The MSE of the classical and chain ratio-type estimators are computed as given in (3) and (8), respectively, and these estimators are compared to each other with respect to their MSE values.

Table 1. Data Statistics

$N = 106$	$S_x = 49189.08$
$n = 20$	$S_y = 6425.09$
$\rho = 0.82$	$S_{xy} = 257778692.26$
$\overline{X} = 24375.59$	$C_y = 4.18$
$\overline{Y} = 1536.77$	$C_x = 2.02$

In Table 1, we observe the statistics about the population. Note that the correlation between the variates is 82%. In Table 2, the values of the MSE are given. From Table 2, it is seen that the chain ratio-type estimator has a smaller MSE than the traditional ratio estimator. Therefore, it is concluded that the chain ratio-type estimator is more efficient than the traditional ratio estimator for this data set. It is an expected result, since because

$$1.5 < \frac{\rho C_y}{C_x} = 1.69$$

the condition mentioned in section 3 is satisfied.

Table 2. The MSE Values of the Ratio Estimators

Chain Ratio	598071.11
Ratio	746223.95

5. Conclusion

We have analyzed the chain ratio-type estimator and obtained its MSE equation. Using this equation, the MSE of the chain ratio type estimator has been compared with that of the classical ratio estimator on a theoretical basis, and from this comparison a condition has been found under which the chain ratio-type estimator has a smaller MSE than the classical ratio estimator. This theoretical condition is shown to hold for the results of an application with original data. This study corrects the inference of Rao [7] that the chain ratio-type estimator is generally worthless. In forthcoming studies, it is hoped to adapt the chain ratio estimator to ratio estimators in stratified random sampling as in [4], or to other ratio estimators given in [3].

References

[1] Cochran, W. G. *Sampling Techniques*, (John Wiley and Sons, New-York, 1977).

[2] Çıngı, H. *Örnekleme Kuramı*, (Hacettepe University Press, 1994).

[3] Kadılar, C. and Çıngı, H. *Oransal ve çarpımsal tahmin ediciler*, DİE Araştırma Dergisi **1** (1), 101–112, 2002.

[4] Kadılar, C. and Çıngı, H. *Ratio estimators in stratified random sampling*, Biometrical Journal **45** (2), 218–225, 2003.

[5] Kiregyera, B. *A chain ratio-type estimator in finite population double sampling using two auxiliary variables*, Metrika **27** (4), 217–223, 1980.

[6] Lohr, S. L. *Sampling: Design and Analysis*, (Duxbury Press, 1999).

[7] Rao, T. J. *On Certain methods of improving ratio and regression estimators*, Communications in Statistics: Theory and Methods **20** (10), 3325–3340, 1991.

[8] Srivastava, S. K. *An estimator using auxiliary information in sample surveys*, Calcutta Statistical Association Bulletin **16**, 121–132, 1967.

[9] Wolter, K. M. *Introduction to Variance Estimation*, (Springer-Verlag, 1985).