Journal of Statistics: Advances in Theory and Applications
Volume 25, Number 1, 2021, Pages 1-12
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jsata_7100122173

EXPONENTIAL TYPE ESTIMATOR FOR THE POPULATION MEAN UNDER RANKED SET SAMPLING

KHALID UL ISLAM RATHER and CEM KADILAR

Division of Statistics and Computer Science
Main Campus SKUAST-J
Chatha Jammu-180009
India
e-mail: khalidstat34@gmail.com

Department of Statistics
Hacettepe University
Beytepe, Ankara 06800
Turkey
e-mail: kadilar@hacettepe.edu.tr

2020 Mathematics Subject Classification: 62D05, 62F07.
Keywords and phrases: exponential type estimators, ranked set sampling, ratio estimator, population mean, efficiency.
Received February 11, 2021; Revised March 3, 2021

Abstract

We propose a new exponential type estimator for the population mean by adapting the estimator suggested by Kadilar [12] to the Ranked Set Sampling (RSS). Theoretically and numerically, we show that the proposed exponential type estimator is more efficient than the classical ratio estimator in the RSS and the estimator of Kadilar et al. [11].

1. Introduction

It is well known that the information of the auxiliary variable is commonly used in order to increase efficiency and precision in sample surveys. It has also a role in the related methods of estimation, such as ratio, product, and regression. If the correlation between the study variable (Y) and the auxiliary variable (X) is highly positive, the ratio method of estimation is used. If not, the product method of estimation is employed effectively provided that this correlation is highly negative. In recent years, there have been many articles on estimators for the population mean in the Sampling Theory Literature, such as Qureshi et al. [16], Zaman [22], Zaman and Kadilar [23, 24, 25], Irfan et al. [6, 7], Cekim and Kadilar [4], Qureshi et al. [17], Singh et al. [20].

In addition to the Simple Random Sampling (SRS) method, Ranked Set Sampling (RSS), which may be considered as a controlled random sampling design, was first introduced by McIntyre [14] to estimate the pasture yield. The RSS procedure involves randomly drawing n sets of $n$units each from the population for which the mean is to be estimated. It is assumed that the units in each set can be ranked visually. From the first set of n units, the lowest unit ranked is measured. From the second set of n units, the second lowest unit ranked is measured. This process continues until the n-th ranked unit is measured. McIntyre [14] illustrated the gain in efficiency by a computation involving five distributions. As a simple introduction to the concept of RSS, when X is a random variable with a density function $F(x)$ and $(x_1, x_2, \ldots, x_n)$ are the unobserved values from n units, we may then rank them by visual inspection or based on a concomitant variable. RSS involves selecting one unit among every ranked set consisting of m units for quantification.

The RSS method can be briefly described step by step as follows:

Step 1: Randomly select m^2 units from the target population.

Step 2: Allocate the m^2 selected units as randomly as possible into m sets, each of size m.

Step 3: Without knowing any values of the variable of interest, rank the units within each set with respect to variable of interest. This may be based on personal professional judgment or done with concomitant variable correlated with the variable of interest.

Step 4: Choose a sample for actual quantification by including the smallest ranked unit in the first set, the second smallest ranked unit in the second set and this process continues in this way until the largest ranked unit is selected from the last set.

Step 5: Repeat Steps 1 through 4 for n cycles to obtain a sample of size mn for actual quantification. (Al-Omari and Bouza [1]).

When it is ranked on the auxiliary variable, let $(y_{[i]}, x_{(i)})$ denote an i-th judgement ordering in the i-th set for the study variable and the i-th order statistic in the i-th set for the auxiliary variable, respectively.

In the remaining part of this article, the estimators for the population mean under RSS are mentioned in Section 2, the adapted estimator from the SRS to RSS is given in Section 3, theoretical and numerical comparisons of the adapted estimator are performed with the existing adapted estimators in literature in Sections 4 and 5, respectively.

2. Estimators in Literature under RSS

Samawi and Muttlak [18] defined the estimator of the population ratio using the RSS as

$$\hat{R}_{RSS} = \frac{\bar{y}_{[n]}}{\bar{x}_{(n)}}, \tag{2.1}$$

where $\bar{y}_{[n]} = \frac{1}{n}\sum_{i=1}^{n} y_{[i]}$ and $\bar{x}_{(n)} = \frac{1}{n}\sum_{i=1}^{n} x_{(i)}$. Note that the estimator in (2.1) can also be used for the population total and mean. Then, the estimator for the population mean can be written as follows:

$$\bar{y}_{rRSS} = \frac{\bar{y}_{[n]}}{\bar{x}_{(n)}}\bar{X}, \tag{2.2}$$

where it is assumed that the population mean $\bar{X}$ of the auxiliary variable x is known and the MSE equation of the estimator in (2.2) can be given by

$$\mathrm{MSE}(\bar{y}_{rRSS}) \cong \frac{1}{mr}\left(S_y^2 - 2RS_{yx} + R^2 S_x^2\right) - \frac{1}{m^2 r}\left(\sum_{i=1}^{m}\tau_{y[i]}^2 - 2R\sum_{i=1}^{m}\tau_{yx(i)} + R^2\sum_{i=1}^{m}\tau_{x(i)}^2\right), \tag{2.3}$$

where $R = \frac{\bar{Y}}{\bar{X}}$, S_x^2 is the population variance of the auxiliary variable, S_y^2 is the population variance of the study variable, S_{yx} is the population covariance between the auxiliary and study variables, $\tau_{x(i)} = \mu_{x(i)} - \bar{X}$, $\tau_{y[i]} = \mu_{y[i]} - \bar{Y}$, and $\tau_{yx(i)} = \left(\mu_{y[i]} - \bar{Y}\right)\left(\mu_{x(i)} - \bar{X}\right)$. Here, $\bar{Y}$ is the population mean of the study variable. Note that the values of $\mu_{x(i)}$ and $\mu_{y[i]}$ depend on the order statistics. The details about (2.3) can be found in Kadilar et al. [11]. We would like to remind that the values of $\mu_{x(i)}$ and $\mu_{y[i]}$ can be taken to be same in the absence of judgment error if the variables have the same distribution (see the appendix of Dell and Clutter [5]).

We would also like to note that Bouza [3] mentioned some different developments of ratio type estimators under the RSS, Jeelani and Bouza [8] introduced using the linear combination of Median and Quartile Deviation of the auxiliary variable under the RSS and Kocyigit and Kadilar [13] studied on ratio estimators under the RSS when there is tie information.

Kadilar et al. [11] proposed the following estimator by adapting Prasad [15] to the RSS as follows:

$$\bar{y}_{\kappa RSS} = \frac{\kappa \bar{y}_{[n]}}{\bar{x}_{(n)}} \bar{X} = \hat{R}_{\kappa RSS} \bar{X}, \tag{2.4}$$

where κ is a constant and $\hat{R}_{\kappa RSS} = \kappa \frac{\bar{y}_{[n]}}{\bar{x}_{(n)}} = \kappa \frac{\bar{y}_{RSS}}{\bar{x}_{RSS}}$.

The MSE of the estimator in (2.4) is given by

$$\mathrm{MSE}_{\min}(\bar{y}_{\kappa RSS}) \cong \frac{1}{mr}\left(\kappa^{*2} S_y^2 - 2R\kappa^* S_{yx} + R^2 S_x^2\right) + \bar{Y}^2\left(\kappa^* - 1\right)^2$$
$$- \frac{1}{m^2 r}\left(\kappa^{*2} \sum_{i=1}^{m} \tau_{y[i]}^2 - 2R\kappa^* \sum_{i=1}^{m} \tau_{yx(i)} + R^2 \sum_{i=1}^{m} \tau_{x(i)}^2\right), \tag{2.5}$$

where $\kappa^* = \frac{1 + \gamma C_{yx} - W_{yx(i)}}{1 + \gamma C_y^2 - W_{y[i]}^2}$. Here, $\gamma = \frac{1}{mr}$, C_x, and C_y are the population coefficients of variation of the auxiliary and study variables, respectively, ρ is the population correlation between the auxiliary and study variables, $C_{yx} = \rho C_y C_x$, $W_{yx(i)} = \frac{1}{m^2 r \bar{X}\bar{Y}} \sum_{i=1}^{m} \tau_{yx(i)}$, and $W_{y[i]}^2 = \frac{1}{m^2 r \bar{Y}^2} \sum_{i=1}^{m} \tau_{y[i]}^2$.

3. The Suggested Estimator

Motivated by Bahl and Tuteja [2], Singh et al. [21], and Kadilar [12], we propose a modified exponential type estimator using the RSS method for the population mean as follows:

$$\bar{y}_{pro} = \bar{y}_{(n)} \left[\frac{\bar{x}_{[n]}}{\bar{X}}\right]^{\alpha} \exp\left[\frac{\bar{X} - \bar{x}_{[n]}}{\bar{X} + \bar{x}_{[n]}}\right], \tag{3.1}$$

where α is a constant.

In order to find the MSE equation of the estimator in (3.1), we use the following notations:

$$\bar{y}_{(n)} = \bar{Y}(1+\in_0),\ \bar{x}_{(n)} = \bar{X}(1+\in_1),\ E(\in_0) = E(\in_1) = 0,$$

$$E(\in_0)^2 = V\left(\frac{\bar{y}_{(n)}}{\bar{Y}^2}\right) = \frac{1}{mr}\frac{1}{\bar{Y}^2}\left[S_y^2 - \frac{1}{m}\sum t_{y(i)}^2\right] = \left[\theta C_y^2 - W_{y(i)}^2\right],$$

$$E(\in_1)^2 = V\left(\frac{\bar{x}_{(n)}}{\bar{X}^2}\right) = \frac{1}{mr}\frac{1}{\bar{Y}^2}\left[S_x^2 - \frac{1}{m}\sum t_{y(i)}^2\right] = \left[\theta C_x^2 - W_{x(i)}^2\right],$$

$$E(\in_0\in_1) = \frac{1}{mr}\frac{1}{\bar{Y}\,\bar{X}}\left[S_{yx} - \frac{1}{m}\sum t_{y(i)}^2\right] = \left[\theta C_{yx} - W_{yx(i)}\right],$$

where $W_{x[i]}^2 = \dfrac{1}{m^2 r\bar{X}^2}\sum_{i=1}^{m}\tau_{x[i]}^2$.

After some simplifications, we finally obtain the MSE equation as follows:

$$\mathrm{MSE}(\bar{y}_{pro}) = \bar{Y}^2\left(A\alpha - \frac{B}{2}\alpha^2 + C\right), \tag{3.2}$$

where

$$A = 2\gamma C_{yx} + \gamma C_x^2 + 2W_{yx(i)} - W_{x(i)}^2,\ B = 2(W_{x(i)}^2 - \gamma C_x^2),$$

and

$$C = \gamma(C_y^2 + \frac{C_x^2}{4} + C_{yx}) - (W_{y(i)}^2 + \frac{W_{x(i)}^2}{4} + W_{yx(i)}).$$

The optimum value of α, to minimize the MSE of $\bar{y}_{pro}$, is found as

$$\frac{\partial \mathrm{MSE}(\bar{y}_{pro})}{\partial\alpha} = 0,$$

$$\alpha_{opt} = \frac{A}{2B}. \tag{3.3}$$

Then, we get the following minimum MSE of the proposed estimator, by replacing α in (3.2) with α_{opt} in (3.3)

$$\mathrm{MSE}_{\min}(\bar{y}_{pro}) = \bar{Y}^2\left(\frac{A^2}{4B} + C\right) \tag{3.4}$$

that will be compared with the MSE equations of other estimators theoretically in the next section.

4. Efficiency Comparisons

In this section, the performances of the proposed estimator have been demonstrated over the traditional ratio estimator in the RSS and the estimator of Kadilar et al. [11], respectively, as follows:

$$\mathrm{MSE}(\bar{y}_{rRSS}) - \mathrm{MSE}_{\min}(\bar{y}_{pro}) > 0,$$

$$A - B - W_{y[i]}^2 - 4\gamma C_{yx} + \frac{3C_x^2}{4} - \frac{A^2}{4B} > 0. \tag{4.1}$$

$$\mathrm{MSE}(\bar{y}_{\kappa RSS}) - \mathrm{MSE}_{\min}(\bar{y}_{pro}) > 0,$$

$$\gamma R^2 C_x^2 - \frac{1}{D} - \frac{A^2}{2B} + C > 0, \tag{4.2}$$

where $D = \dfrac{1}{1 + \gamma C_y^2(1 - \rho^2)}$.

When the efficiency conditions of (4.1) and (4.2) are satisfied, it is inferred that the proposed estimator is more efficient than the traditional ratio estimator in RSS and the estimator of Kadilar et al. [11], respectively. In other words, we obtain the efficiency conditions of the proposed estimators (4.1) and (4.2) with respect to the traditional ratio estimator in RSS and the estimator of Kadilar et al. [11], respectively. As these conditions are obtained in theory, it is not so useful for the researchers while choosing the auxiliary variable for the estimation of the parameters of the study variable.

5. Numerical Example

We use the data of the Marmara Region of Turkey in Kadilar and Cingi [9]. We apply the estimators in (2.2), (2.4), and (3.1) to the data set concerning the level of apple production (as study variable) and number of apple trees (as auxiliary variable) in 106 villages in the Marmara Region in 1999 (Source: Institute of Statistics, Republic of Turkey). We have standardized these data in order to obtain the standard normal distribution provided that we can use means of normal order statistics for the values of $\mu_{(1)}$, $\mu_{(2)}$, and $\mu_{(3)}$. These statistics are given in Table 1. Using (2.3), (2.5), and (3.4), we compute the MSE values of the traditional ratio estimator in RSS, the ratio estimator in Kadilar et al. [11], and the proposed estimator, respectively. These computed MSE values are shown in Table 2. It is obvious that there is considerable gain in efficiency by using the proposed estimator over other estimators. From Table 2, we infer that the proposed ratio estimator is the most efficient estimator.

Table 1. Statistics of standardized values of the population

$N = 106$	$\overline{X} = -6.97 \times 10^{-17}$	$\alpha_{opt} = -0.837$
$n = 12$	$\overline{Y} = 1.67 \times 10^{-16}$	$\kappa^* = -1.55$
$m = 3$	$R = -2.398$	$\mu_{(1)} = -0.846$
$r = 4$	$S_y = 1$	$\mu_{(2)} = 0$
$\rho = 0.82$	$S_x = 1$	$\mu_{(3)} = 0.846$
$A = 2.55\text{E}31$	$B = -1.65E31$	$C = 1.17E31$

Table 2. MSE values of estimators

Estimators	MSE Values
Proposed $(\bar{y}_{pro})$	0.05211
Kadilar et al. $(\bar{y}_{\kappa RSS})$	0.14558
Samawi and Muttlak $(\bar{y}_{rRSS})$	0.42923

For detail investigation of the α_{opt} values, we obtain the ranges of α_{opt} where the proposed estimator is more efficient than the other estimators. These ranges for each estimator are shown in Table 3. We see that there is enough scope of choosing α, instead of α_{opt}, to obtain better estimates by the proposed estimator. From Table 3, we observe that when $\alpha \in (-1.2; -0.4)$, the proposed ratio estimator is more efficient than the traditional ratio estimator in the RSS and the ratio estimator in Kadilar et al. [11] for this data set.

Table 3. Efficiency ranges of α_{opt} for the proposed ratio estimator

Estimators	Ranges of α_{opt}
Kadilar et al. $(\bar{y}_{\kappa RSS})$	(– 1.2; – 0.4)
Samawi and Muttlak $(\bar{y}_{rRSS})$	(– 1.6; 0.1)

6. Conclusion

This article shows that the exponential type estimator can be used in the RSS for the population mean and that using the exponential type estimator improves the efficiency of the RSS estimators in literature. Therefore, in the forthcoming studies, we hope to adapt the estimator presented in this article to the stratified random sampling, utilizing the methods in Samawi and Siam [19] and Kadilar and Cingi [10].

References

[1] A. I. Al-Omari and C. N. Bouza, Review of ranked set sampling: Modifications and applications, Revista Investigacion Operacional 35(3) (2014), 215-240.

[2] S. Bahl and R. K. Tuteja, Ratio and product type exponential estimators, Journal of Information and Optimization Sciences 12(1) (1991), 159-164.

DOI: https://doi.org/10.1080/02522667.1991.10699058

[3] C. N. Bouza, Model-assisted ranked survey sampling, Biometrical Journal 43(2) (2001), 249-259.

DOI: https://doi.org/10.1002/1521-4036(200105)43:2<249::AID-BIMJ249>3.0.CO;2-U

[4] H. O. Cekim and C. Kadilar, New families of unbiased estimators in stratified random sampling, Journal of Statistics and Management Systems 21(8) (2018), 1481-1499.

DOI: https://doi.org/10.1080/09720510.2018.1530176

[5] T. R. Dell and J. L. Clutter, Ranked set sampling theory with order statistics background, Biometrics 28(2) (1972), 545-555.

DOI: https://doi.org/10.2307/2556166

[6] M. Irfan, M. Javed and Z. Lin, Enhanced estimation of population mean in the presence of auxiliary information, Journal of King Saud University Science 31(4) (2019), 1373-1378.

DOI: https://doi.org/10.1016/j.jksus.2018.12.005

[7] M. Irfan, M. Javed and Z. Lin, Efficient ratio-type estimators of finite population mean based on correlation coefficient, Scientia Iranica 25(4) (2018), 2361-2372.

DOI: https://doi.org/10.24200/SCI.2017.4455

[8] M. I. Jeelani and C. N. Bouza, New ratio method of estimation under ranked set sampling, Revista Investigacion Operacional 36(2) (2015), 151-155.

[9] C. Kadilar and H. Cingi, Ratio estimators in stratified random sampling, Biometrical Journal 45(2) (2003), 218-225.

DOI: https://doi.org/10.1002/bimj.200390007

[10] C. Kadilar and H. Cingi, A new ratio estimator in stratified random sampling, Communications in Statistics: Theory and Methods 34(3) (2005), 597-602.

DOI: https://doi.org/10.1081/STA-200052156

[11] C. Kadilar, Y. Unyazici and H. Cingi, Ratio estimator for the population mean using ranked set sampling, Statistical Papers 50(2) (2009), 301-309.

DOI: https://doi.org/10.1007/s00362-007-0079-y

[12] G. O. Kadilar, A new exponential type estimator for the population mean in simple random sampling, Journal of Modern Applied Statistical Methods 15(2) (2016), 207-214.

DOI: https://doi.org/10.22237/jmasm/1478002380

[13] E. G. Kocyigit and C. Kadilar, Ratio estimators for ranked set sampling in the presence of tie information, Communications in Statistics: Simulation and Computation (2020). (Article in Press)

DOI: https://doi.org/10.1080/03610918.2020.1815777

[14] G. A. McIntyre, A method for unbiased selective sampling using ranked sets, Australian Journal of Agricultural Research 3(4) (1952), 385-390.

DOI: https://doi.org/10.1071/AR9520385

[15] B. Prasad, Some improved ratio type estimators of population mean and ratio in finite population sample surveys, Communications in Statistics: Theory and Methods 18(1) (1989), 379-392.

DOI: https://doi.org/10.1080/03610928908829905

[16] M. N. Qureshi, C. Kadilar and M. Hanif, Estimation of rare and clustered population mean using stratified adaptive cluster sampling, Environmental and Ecological Statistics 27(1) (2020), 151-170.

DOI: https://doi.org/10.1007/s10651-019-00438-z

[17] M. N. Qureshi, C. Kadilar, M. Noor Ul Amin and M. Hanif, Rare and clustered population estimation using the adaptive cluster sampling with some robust measures, Journal of Statistical Computation and Simulation 88(14) (2018), 2761-2774.

DOI: https://doi.org/10.1080/00949655.2018.1486842

[18] H. M. Samawi and H. A. Muttlak, Estimation of ratio using rank set sampling, Biometrical Journal 38(6) (1996), 753-764.

DOI: https://doi.org/10.1002/bimj.4710380616

[19] H. M. Samawi and M. I. Siam, Ratio estimation using stratified ranked set sample, METRON: International Journal of Statistics 61(1) (2003), 75-90.

[20] G. N. Singh, A. K. Singh and C. Kadilar, Almost unbiased estimation procedures of population mean in two-occasion successive sampling, Hacettepe Journal of Mathematics and Statistics 47(5) (2018), 1268-1280.

[21] R. Singh, P. Chauhan, N. Sawan and F. Smarandache, Improvement in estimating the population mean using exponential estimator in simple random sampling, International Journal of Statistics and Economics 3(A09) (2009), 13-18.

[22] T. Zaman, Generalized exponential estimators for the finite population mean, Statistics in Transition 21(1) (2020), 159-168.

DOI: https://doi.org/10.21307/stattrans-2020-009

[23] T. Zaman and C. Kadilar, Novel family of exponential estimators using information of auxiliary attribute, Journal of Statistics and Management Systems 22(8) (2019), 1499-1509.

DOI: https://doi.org/10.1080/09720510.2019.1621488

[24] T. Zaman and C. Kadilar, On estimating the population mean using auxiliary character in stratified random sampling, Journal of Statistics and Management Systems 23(8) (2020), 1415-1426.

DOI: https://doi.org/10.1080/09720510.2020.1723924

[25] T. Zaman and C. Kadilar, New class of exponential estimators for finite population mean in two-phase sampling, Communications in Statistics: Theory and Methods 50(4) (2021), 874-889.

DOI: https://doi.org/10.1080/03610926.2019.1643480

■