Exponential Ratio Type Estimator In The Presence of Non-Response

Ceren Unal (cerenunal@hacettepe.edu.tr)
Department of Statistics, Faculty of Science, Hacettepe University, Ankara, Turkey

Cem Kadilar
Department of Statistics, Faculty of Science, Hacettepe University,, Ankara, Turkey

ABSTRACT

We propose exponential ratio type estimator to estimate the population mean under non-response scheme in this article. The expressions for the Bias and Mean Square Error (MSE) are derived and theoretical comparisons are made to examine the efficiencies of the estimator with the estimators considered by many researchers in the literature. In the end of the article, the efficiencies of the proposed estimator is supported numerically by three different data sets. According to theoretical and numerical results, it is concluded that the proposed estimator is the most efficient estimator among all others and can be applied successfully.

Keywords: *Non-Response, Exponential Type Estimator, Ratio Estimator, Efficiency, Mean Square Error.*
JEL Classification: *C83*

1. INTRODUCTION

Many researches have used the information of auxiliary variable (x) to estimate the population mean. The use of information can considerably enhance the efficiency for product, ratio, regression and exponential type of estimation. Within this scope, to estimate the population mean, classical regression and ratio estimators was proposed by Cochran (1940, 1977). Reddy (1973) introduced an estimator using the suitable constant (k). Bahl and Tuteja (1991) proposed exp. type estimator for the first time in literature. After the pioneer study carried out by Bahl and Tuteja (1991), many authors have proposed the estimator taking the advantage of exponential function. These and then some estimators are proposed in a case that there is not any non-response units. However, in practice, this may be not always the case. For this situation, Hansen and Hurwitz (1946) introduced a new sub-sampling technique in order to deal with this non-response problem.

In this technique, a sample size of n units is drawn from N units as $S=(S_1,S_2,...,S_N)$ and (x_i, y_i) refers the auxiliary and study variables for the i^{th} unit of the population, respectively. The size of population N is composed of respondent and non-respondent units, N_1 and N_2, respectively. For this reason, $W_1 = N_1/N$ indicate the proportions of responding for the population. Similarly, n is divided in two parts as n_1 $(n_1 = n - n_2)$ units respond and n_2 units do not respond. A sub-sample size of $r = n_2/z$ $(z>1)$ is drawn from n_2 units by paying extra effort. Here, z refers the inverse sampling rate at the second phase sample of size n. Thus, the population mean can be estimated by using $(n_1 + r)$ units substituted for n in this technique.

In this method, they introduced the unbiased estimator, under non-response condition in order to estimate the population mean by using $(n_1 + r)$ units, given by

$$t_H = w_1\bar{y}_1 + w_2\bar{y}_{2(r)}. \tag{1.1}$$

Note that, in (1.8), $w_1 = \dfrac{n_1}{n}$ and $w_2 = \dfrac{n_2}{n}$ refer proportions of responding and non-responding for the sample, respectively, and also $\bar{y}_1$ and $\bar{y}_{2(r)}$ indicate the sample means of y based on n_1 and r units.

The variance of the unbiased estimator as

$$V(t_H) = \bar{Y}^2\left(\lambda C_y^2 + \frac{W_2(z-1)}{n}C_{y(2)}^2\right). \tag{1.2}$$

Here, $f = \dfrac{n}{N}$, $\lambda = \dfrac{1-f}{n}$, $C_y^2 = \dfrac{S_y^2}{\bar{Y}^2}$ and $C_{y(2)}^2 = S_{y(2)}^2/\bar{Y}^2$ and $W_2 = N_2/N$ indicate the coefficient of variation of y and the proportions, respectively, for N_2 units of the population.

Many authors suggest estimators for the population mean after the pioneer study of Hansen and Hurwitz (1946).

When incomplete information occurs on both x and y as well and the population mean of x is known, Cochran (1977) proposed the ratio and reg. estimators and Singh et al. (2009) adapted the exp. type estimator as follows, respectively,

$$t_R^{**} = \bar{y}^*\frac{\bar{X}}{\bar{x}^*}, \tag{1.3}$$

$$t_{reg}^{**} = \bar{y}^* + b^*\left(\bar{X} - \bar{x}^*\right), \tag{1.4}$$

$$t_{BT}^{**} = \bar{y}^{*} exp\left(\frac{\bar{X} - \bar{x}^{*}}{\bar{X} + \bar{x}^{*}}\right), \tag{1.5}$$

where $b^{*} = S_{xy}^{*} / S_{x}^{*2}$. Here, $\bar{x}^{*}$ and $\bar{y}^{*}$ are the sample mean of *x* and y, respectively, in case of non-response. The pop. mean of *x* is symbolized as $\bar{X}$.

The MSE for the t_{R}^{**} , t_{reg}^{**} and t_{BT}^{**} , are given by

$$MSE\left(t_{R}^{**}\right) = \bar{Y}^{2}\left(\lambda\left(C_{x}^{2} - 2C_{yx} + C_{y}^{2}\right) + \frac{W_{2}(z-1)}{n}\left(C_{y(2)}^{2} + C_{x(2)}^{2} - 2C_{yx(2)}\right)\right), \tag{1.6}$$

$$MSE\left(t_{reg}^{**}\right) = \bar{Y}^{2}\left(\lambda C_{y}^{2}\left(1-\rho_{xy}^{2}\right) + \frac{W_{2}(z-1)}{n}\left(C_{y(2)}^{2} + \rho_{xy}^{2}\frac{C_{y}^{2}}{C_{x}^{2}}C_{x(2)}^{2} - 2\rho_{xy}\frac{C_{y}}{C_{x}}C_{yx(2)}\right)\right), \tag{1.7}$$

$$MSE\left(t_{BT}^{**}\right) = \bar{Y}^{2}\left(\lambda C_{y}^{2} + \lambda\frac{C_{x}^{2}}{4} - \lambda C_{yx} + \frac{W_{2}(z-1)}{n}\left(C_{y(2)}^{2} + \frac{C_{x(2)}^{2}}{4} - C_{yx(2)}\right)\right), \tag{1.8}$$

where $C_{x}^{2} = \frac{S_{x}^{2}}{\bar{X}^{2}}$, $C_{x(2)}^{2} = \frac{S_{x(2)}^{2}}{\bar{X}^{2}}$, $C_{xy} = \rho_{xy}C_{x}C_{y}$ and $C_{xy(2)} = \rho_{xy(2)}C_{x(2)}C_{y(2)}$. The ρ_{xy} is the population correlation coefficient between *y* and *x* while $\rho_{yx(2)} = \frac{C_{yx(2)}}{C_{y(2)}C_{x(2)}}$ is the pop. correlation coefficient for the non-response group between *y* and *x*. Also, $\bar{Y}$ refers the population mean of *y.*

After Singh et al. (2009), Kumar and Bhougal (2011), Kumar (2013), Yadav et. al (2016), Pal and Singh (2017), Singh and Usman (2019), Ünal and Kadilar (2020, 2021), Muneer et. al (2018), Sanaullah et. al (2019), Riaz et. al (2020), Sinha and Kumar (2017), Pal and Singh (2018) as well as Kumar and Kumar (2017) also proposed a new estimators taking advantage of exponential function under the non-response scheme.

In this study, we propose an estimator under non-response scheme and then theoretical comparisons of the proposed estimator are performed with the existing traditional estimators and then the results are supported by three empirical studies.

2. **THE ADAPTED ESTIMATOR**

We adapt the estimator which is proposed by Reddy (1973) to t_{R2} using the exponential function for the population mean under non response situation as follows:

The proposed t_{R2} estimator is as follows:

$$t_{R2} = \bar{y}^* \exp\left(\frac{\bar{X}}{\bar{X} + k_2\left(\bar{x}^* - \bar{X}\right)} - 1 \right), \tag{2.1}$$

where k_2 is a chosen constant to be determined that the MSE of t_{R2} is minimum. To obtain $B\left(t_{R2}\right)$ and $MSE\left(t_{R2}\right)$, we can write

$$\bar{x}^* = \left(\bar{X} + \bar{X}e_1^*\right), \bar{y}^* = \left(\bar{Y} + \bar{Y}e_0^*\right), \quad E\left(e_0^*\right) = E\left(e_1^*\right) = 0 \quad , \quad E\left(e_1^{*2}\right) = \lambda C_x^2 + \frac{W_2\left(z-1\right)}{n} C_{x(2)}^2 \quad ,$$

$$E\left(e_0^{*2}\right) = \lambda C_y^2 + \frac{W_2\left(z-1\right)}{n} C_{y(2)}^2 \text{ and } E\left(e_0^* e_1^*\right) = \lambda C_{yx} + \frac{W_2\left(z-1\right)}{n} C_{xy(2)}.$$

Expressing (2.1) in terms of e_0^* and e_1^*, we obtain

$$t_{R2} = \bar{Y}\left(e_0^* - k_2 e_1^* + \frac{3k_2^2 e_1^{*2}}{2} - k_2 e_0^* e_1^* + 1 \right). \tag{2.2}$$

Expanding the right hand side of (2.2) and neglecting the terms involving powers of e_0^* and e_1 greater than two, we have

$$\left(t_{R2} - \bar{Y}\right) = \bar{Y}\left(e_0^* - k_2 e_1^* + \frac{3k_2^2 e_1^{*2}}{2} - k_2 e_0^* e_1^* \right). \tag{2.3}$$

We take the expectation on both sides of (2.3), we obtain the $B\left(t_{R2}\right)$ as

$$B\left(t_{R2}\right) = \bar{Y}\left(\lambda C_x^2\left(\frac{3k_2^2}{2} - k_2 \rho_{xy} \frac{C_y}{C_x} \right) + \frac{W_2\left(z-1\right)}{n} C_{x(2)}^2\left(\frac{3k_2^2}{2} - k_2 \rho_{xy(2)} \frac{C_{y(2)}}{C_{x(2)}} \right) \right). \tag{2.4}$$

We take square of (2.3) firstly and then expectation, we derive the $MSE\left(t_{R2}\right)$ as

$$MSE\left(t_{R2}\right) = \bar{Y}^2\left(\lambda\left(C_y^2 + k_2^2 C_x^2 - 2k_2 C_{yx}\right) + \frac{W_2\left(z-1\right)}{n}\left(C_{y(2)}^2 + k_2^2 C_{x(2)}^2 - 2k_2 C_{yx(2)}\right) \right). \tag{2.5}$$

By the minimization of (2.5) with respect to k_2, we obtain the k_2^* as

$$k_2^* = \frac{\left(\lambda C_{yx} + \frac{W_2\left(z-1\right)}{n} C_{yx(2)} \right)}{\left(\lambda C_x^2 + \frac{W_2\left(z-1\right)}{n} C_{x(2)}^2 \right)} = \frac{E\left(e_0^* e_1^*\right)}{E\left(e_1^{*2}\right)}. \tag{2.6}$$

Replacing k_2 in the MSE of the estimator in (2.5) with k_2^*, we get the $MSE_{\min}\left(t_{R2}\right)$ as follows:

$$MSE_{\min}\left(t_{R2}\right)=\bar{Y}^2\left[\left(\lambda C_y^2+\frac{W_2\left(z-1\right)}{n}C_{y(2)}^2\right)-\frac{\left(\lambda C_{xy}+\frac{W_2\left(z-1\right)}{n}C_{yx(2)}\right)^2}{\left(\lambda C_x^2+\frac{W_2\left(z-1\right)}{n}C_{x(2)}^2\right)}\right]. \tag{2.7}$$

3. EFFICIENCY COMPARISONS

The theoretical comparisons for the efficiencies of the adapted estimator t_{R2} is examined under non-response scheme. We use the Eq. (1.2), (1.6), (1.8) and (1.7) to compare the t_{R2} estimator with the t_H, t_R^{**}, t_{BT}^{**}, t_{reg}^{**} and we obtain as follows:

$$i)\quad \left[MSE\left(t_H\right)-MSE_{\min}\left(t_{R2}\right)\right]=\left(\lambda C_{xy}+\frac{W_2\left(z-1\right)}{n}C_{yx(2)}\right)^2>0 \tag{3.1}$$

$$ii)\quad \left[MSE\left(t_R^{**}\right)-MSE_{\min}\left(t_{R2}\right)\right]=\left(\left(\lambda C_x^2+\frac{W_2\left(z-1\right)}{n}C_{x(2)}^2\right)-\left(\lambda C_{yx}+\frac{W_2\left(z-1\right)}{n}C_{yx(2)}\right)\right)^2>0 \tag{3.20}$$

$$iii)\quad \left[MSE\left(t_{BT}^{**}\right)-MSE_{\min}\left(t_{R2}\right)\right]=\left(\left(\lambda C_{xy}+\frac{W_2\left(z-1\right)}{n}C_{yx(2)}\right)-\frac{1}{2}\left(\lambda C_x^2+\frac{W_2\left(z-1\right)}{n}C_{x(2)}^2\right)\right)^2>0 \tag{3.3}$$

$$iv)\quad \left[MSE\left(t_{reg}^{**}\right)-MSE_{\min}\left(t_{R2}\right)\right]=\left(\left(\frac{W_2\left(z-1\right)}{n}C_{x(2)}^2\rho_{xy}\frac{C_y}{C_x}\right)-\left(\frac{W_2\left(z-1\right)}{n}C_{yx(2)}\right)\right)^2>0 \tag{3.4}$$

The t_{R2} estimator is the most efficient estimator according to the compared estimators because all the conditions (3.1) – (3.4) are always satisfied.

4. NUMERICAL APPLICATIONS

In Section 3, we show that the t_{R2} is always the most efficient estimator among compared estimators in case of using the optimal value of k_2. In this section, we obtain the ranges of k_2 value, based on the different values of *z*, with the purpose of determining the ranges of k_2 value that make the proposed estimator is more efficient than others under non-response scheme.

For this reason, we use three different data sets in Ünal and Kadilar (2020), considered by many researchers in literature, to examine the appropriateness of the t_{R2} estimator.

The descriptive statistics for Population 1-3 are given as follows:

Population 1. Khare and Sinha (2009)
y: The number of agriculture labors
x: The area of the village

$N = 96$	$\bar{X} = 144.87$	$\rho_{yx(2)} = 0.72$	C_{yx}=0.8232	$C_{yx(2)} = 1.4077$
n=40	$\bar{Y} = 137.92$	$\rho_{yx} = 0.77$	$C_x = 0.81$	$C_{x(2)} = 0.94$
W_2=0.25	$\lambda = 0.01458,\ f = 0.42$	$\beta_2(x) = 1.19$	$C_y = 1.32$	$C_{y(2)} = 2.08$

Population 2. Khare and Srivastava (1993)
y: The cultivated area (in acres)
x: Population of the village

$N = 70$	$\bar{X} = 1755.53$	$\rho_{yx(2)} = 0.445$	$C_{yx} = 0.3896$	$C_{yx(2)} = 0.104$
n=35	$\bar{Y} = 981.29$	$\rho_{yx} = 0.778$	$C_x = 0.801$	$C_{x(2)} = 0.574$
W_2=0.2	$\lambda = 0.0143,\ f = 0.50$	$\beta_2(x) = 0.34$	$C_y = 0.6254$	$C_{y(2)} = 0.4087$

Population 3. Satici and Kadilar (2011)
y: The number of successful students
x: The number of teachers

$N = 261$	$\bar{X} = 306.435$	$\rho_{yx(2)} = 0.9733$	C_{yx}=3.18535	$C_{yx(2)} = 1.458$
n=90	$\bar{Y} = 222.5766$	$\rho_{yx} = 0.9705$	$C_x = 1.7595$	$C_{x(2)} = 1.2285$
W_2=0.25	$\lambda = 0.00728,\ f = 0.35$	$\beta_2(x) = 21.36$	$C_y = 1.8654$	$C_{y(2)} = 1.2196$

When non-response exists both on x and y, the ranges of k_2 values for the efficiency of the t_{R2} estimator among other estimators, based on the different values of z, are given in Tables 1 – 3 for Populations 1-3, respectively.

Ranges of k_2 values for the proposed estimators t_{R2} for Population 1

Table 1

z=2	*z=3*	*z=4*	*z=5*	*z=6*
(1.255; 1.502)	(1.255; 1.617)	(1.255; 1.683)	(1.255; 1.727)	(1.255; 1.757)

Ranges of k_2 values for the proposed estimators t_{R2} for Population 2

Table 2

z=2	*z=3*	*z=4*	*z=5*	*z=6*
(0.509; 0.607)	(0.501; 0.545)	(0.494; 0.500)	(0.453; 0.500)	(0.421; 0.500)

Ranges of k_2 values for the proposed estimators t_{R2} for Population 3

Table 3

z=2	*z=3*	*z=4*	*z=5*	*z=6*
(1.010; 1.028)	(1.001; 1.023)	(1.001; 1.012)	(1.001; 1.004)	(0.998; 0.999)

These ranges show that the t_{R2} has the minimum MSE value among the compared estimators in case of using any k_2 values in the range for Populations 1-3, respectively. According to the Tables 1–3, it is also surprising that there is no important difference for the ranges of k_2 values according to various z values for Populations 1-3 under the non-response scheme.

5. CONCLUSION

We introduced a new exponential ratio type estimator for the population mean using the information of the x. The properties of the t_{R2} is studied under non-response scheme. Theoretically, we demonstrate that the t_{R2} is always the most efficient estimators among the basic mentioned estimators. In addition, we obtain the efficiency intervals of k_2 for the proposed estimator using three data sets and we show that the t_{R2} performs better than others using any k_2 values in the ranges, numerically. Consequently, the proposed estimator is recommended based on the obtained results in the presence of non-response.
Conflicts of Interest: The authors declare no conflicts of interest.

References

1. **Bahl, S**. and **Tuteja, R. K**. (1991), *"Ratio and product type exponential estimators"*, Journal of Information and Optimization Sciences, 12(1), 159-164.
2. **Cochran, W. G** (1940), *"The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce"*, The Journal of Agricultural Science, 30(2), 262-275.
3. **Cochran, W. G**. (1977). *Sampling Techniques*, John Wiley and Sons, New-York.
4. **Hansen, M. H.** and **Hurwitz, W. N**. (1946), *"The problem of non-response in sample surveys"*, Journal of the American Statistical Association, 41(236), 517-529.

5. **Khare, B. B. and Sinha, R. R**. (2009), *"On class of estimators for population mean using multi-auxiliary characters in the presence of non-response"*, Statistics in Transition, 10(1), 3-14.
6. **Khare, B. B.** and **Srivastava, S.** (1993), *"Estimation of population mean using auxiliary character in presence of non-response"*, National Academy Science Letters, 16, 111-114.
7. **Kumar, S.** (2013), *"Improved exponential estimator for estimating the population mean in the presence of non-response"*, Communications for Statistical Applications and Methods, 20(5), 357-366.
8. **Kumar, S.** and **Bhougal, S**. (2011), *"Estimation of the population mean in presence of non-response"*, Communications for Statistical Applications and Methods, 18(4), 537-548.
9. **Kumar, K.** and **Kumar, M.** (2017), *"Improved exponential ratio and product type estimators for population mean in the presence of nonresponse"*, Bulletin of Mathematics and Statistics Research, 5(2), 68-76.
10. **Muneer, S., Shabbir, J**. and **Khalil, A.** (2018), *"A Generalized exponential type estimator of population mean in the presence of non-response"*, Statistics in Transition New Series, 19(2), 259-276.
11. **Pal, S. K.** and **Singh, H. P.** (2017), *"A class of ratio-cum-ratio-type exponential estimators for population mean with sub sampling the non-respondents"*, Jordan Journal of Mathematics and Statistics, 10(1), 73-94.
12. **Pal, S. K.** and **Singh H. P.** (2018), *"Estimation of finite population mean using auxiliary information in presence of non-response"*, Communications in Statistics - Simulation and Computation, 47(1), 143-165.
13. **Reddy, V. N.** (1973), *"On ratio and product methods of estimation"*, Sankhyā: The Indian Journal of Statistics, Series B, 35(3), 307-316.
14. **Riaz, S., Nazeer, A., Abbası, J.** and **Qamar, S.** (2020), *"On the generalized class of estimators for estimation of finite population mean in the presence of non-response problem"*, Journal of Prime Research in Mathematics, 16(1), 52-63.
15. **Sanaullah, A., Ehsan, I.** and **Noor-Ul-Amin, M**. (2019), *"Estimation of mean for a finite population using sub-sampling of non-respondents"*, Journal of Statistics and Management Systems, 22(6), 1015-1035.
16. **Satici, E.** and **Kadilar, C.** (2011), *"Ratio estimator for the population mean at the current occasion in the presence of non-response in successive sampling"*, Hacettepe Journal of Mathematics and Statistics, 40(1), 115–124.
17. **Sinha, R. R**. and **Kumar, V**. (2017), *"Regression cum exponential estimators for finite population mean under incomplete information"*, Journal of Statistics and Management Systems, 20(3), 355-368.
18. **Singh, R., Kumar, M., Chaudhary, M. K.** and **Smarandache, F.** (2009), *"Estimation of mean in presence of non-response using exponential estimator"*, Unpublished manuscript.
19. **Singh, G. N.** and **Usman, M**. (2019), *"Efficient combination of various estimators in the presence of non-response"*, Communications in Statistics-Simulation and Computation, 1-35, https://doi.org/10.1080/03610918.2019.1614618.
20. **Ünal, C.** and **Kadilar, C.** (2020), *"Exponential type estimator for the population mean in the presence of non-response"*, Journal of Statistics and Management Systems, 23(3), 603-615.
21. **Ünal, C.** and **Kadilar, C**. (2021), "Improved family of estimators using exponential function for the population mean in the presence of non-response", Communications in Statistics - Theory and Methods, 50(1), 237-248.
22. **Yadav, S. K., Subramani, J., Misra, S., Singh, L**. and **Mishra, S. S**. (2016), "Improved estimation of population mean in presence of non response using exponential estimator", International Journal of Agricultural and Statistical Sciences, 12(1), 271-276.