Sinopfbd, 2025; 10(1): 60-72 **https://doi.org/10.33484/sinopfbd.1569245**

Sinop Üniversitesi Fen Bilimleri Dergisi
Sinop Uni J Nat Sci

E-ISSN: 2564-7873

https://dergipark.org.tr/tr/pub/sinopfbd

Exponential Estimators Under Non-Response Cases

Ceren ÜNAL[1] and Cem KADILAR[2]

How to cite: Ünal, C., & Kadılar, C. (2025). Exponential estimators under non-response cases. *Sinop Üniversitesi Fen Bilimleri Dergisi, 10*(1), 60-72. https://doi.org/10.33484/sinopfbd.1569245

Research Article

Corresponding Author
Ceren ÜNAL
cerenunal@hacettepe.edu.tr

ORCID of the Authors
C.Ü: 0000-0002-9357-1771
C.K: 0000-0003-4950-9660

Received: 18.10.2024
Accepted: 23.01.2025

Abstract
This study proposes new families of estimators for the estimation of the population mean using the Hansen-Hurwitz method. This method is examined in two cases, referred to as Case I and Case II. According to both cases, the expressions for the proposed family of estimators are derived. After theoretical comparisons, a new data set on the magnitude and a simulation study are conducted to support these theoretical results. As a consequence of this study, the proposed families of estimators perform well under the obtained conditions for both non-response schemes and can be used successfully in the field of seismology.

Keywords: Population mean, non-response, efficiency, exponential type estimators

Cevapsızlık Durumunda Üstel Tip Tahmin Ediciler

[1]Hacettepe University, Institute of Science, Beytepe, Ankara, Türkiye

[2]Hacettepe University, Department of Statistics, Beytepe, Ankara, Türkiye

Öz
Bu çalışma, Hansen-Hurwitz yöntemini kullanarak kitle ortalamasının tahmini için yeni tahmin edici aileleri önermektedir. Bu yöntem, Durum I ve Durum II olarak adlandırılan farklı iki durumda incelenmiştir. Her iki duruma göre de önerilen tahmin edici aileleri için teorik çıkarsamalar elde edilmiştir. Yapılan teorik karşılaştırmalardan sonra, bu sonuçları desteklemek amacıyla deprem ile alakalı gerçek veri seti uygulaması ve simülasyon çalışması gerçekleştirilmiştir. Bu çalışma ile önerilen tahmin edici aileleri elde edilen koşullar altında karşılaştırılan tahmin edicilere göre daha iyi performans göstermekte olduğu ve bu tahmin edici ailelerinin deprem alanında da kullanılabileceği sonucuna ulaşılmıştır.

Anahtar Kelimeler: Kitle ortalaması, cevapsızlık durumu, etkinlik, üstel tip tahmin ediciler

Introduction

It is not feasible to reach the entire population on a consistent basis, and the process of working with a population can present challenges in terms of time, financial resources, and the availability of labor. At this case, it is more prudent to utilize a sample as a proxy for the population. It is crucial to ensure that the sample is capable of representing the population in question. The parameters of the population can be obtained through the use of estimators in sample surveys. By utilizing the selected sample, the mathematical equation for the estimation of the population parameters can be defined as an estimator.

One of the key characteristics of an estimator is its efficiency. To enhance this, the auxiliary variable (x) can be employed in sample surveys when introducing an estimator. Consequently, the auxiliary variable plays a pivotal role. The aforementioned information allows us to identify a number of different types of estimators, including those based on the product, ratio, ln, regression and exponential functions. In the presence of the x, these estimators can be classified into two categories: those with equal efficiencies, such as the usual regression, ratio and product types, and those with unequal efficiencies. The line of best fit does not pass through the origin in the majority of cases. Consequently, the ratio, regression and product types of estimators are not equally efficient. Consequently, exponential estimators become a prominent feature among other types [1]. The different types of estimators are proposed generally in case of information on some variables obtained completely. As examples of this situation, Oncel Cekim and Kadilar [2] proposed unbiased estimators in a stratified sampling method. Zaman and Kadilar [3] also proposed a population mean estimator under the same sampling method. In practice, this situation may not occur with every sample survey, but it is considered to be the most significant problem in such surveys [4]. Hansen and Hurwitz [5] defined a new technique using the sub-sampling method for the non-response units. In this technique, both the response and non-response units in the estimator are used for reducing the non-response effect. Recently, Kumar et al. [6] proposed a new class of estimator using exponential function based the concept of sub-sampling method. The two generalized class of estimators are proposed by Jaiswal et al. [7] for this situation. Singh et al. [8] considered a new estimator under stratified random sampling method. Singh and Singh [9] defined a class of estimators for population utilizing an auxiliary variable. Singh and Usman [10] proposed the ratio-product type difference cum-exponential estimators under non-response scheme. Under the situation of non-response, Pandey et al. [11] contributed significantly and suggested difference and ratio type estimators in two distinct situations of non-response using different sampling schemes. Singh and Nigam [12] proposed a class of estimators in case of non-response on study variable only under stratified random sampling. Khalid and Singh [13] made a significant contribution by proposing different classes of estimators in the case of non-response using different sampling schemes. Hussain et al. [14] proposed a new efficient class of estimators in the presence of non-response for estimating the population mean using dual auxiliary information in simple random sampling. This study addresses new estimators in case of non-response on both the study and the concomitant variables using simple random sampling.

The structure of this article is organized in the following way. Firstly, this sub-sampling methods is examined in detail and the new families of estimators are introduced with their bias, Mean Square Error (MSE) and minimum MSEs under the non-response schemes. After that, the efficiency comparisons are conducted. In this part, theoretical comparisons, empirical study as well as simulation study are analyzed, respectively. Eventually, the obtained results are concluded.

Methodology, Notations and Literature Review

The sub-sampling technique involves the partitioning of a population (N), with N_1 and N_2 denoting the two resulting subsets. The response units, available in N_1 only, are not present in N_2 $(N_2 = N - N_1)$. Hence, this same situation applies to a sample of size n. Here, the response units, present in n_1 only, are not representative of the population, as evidenced by n_2 units. In the Hansen–Hurwitz technique, the sub-sample comprises $k = \frac{n_2}{r}$ $(r > 1)$ units obtained from n_2 units through additional efforts. Using $n_1 + k$ units enable the estimation of the population mean, in lieu of n units, as per this technique. In this context, r denotes the inverse of the sampling rate. Using the sub-sampling method, Hansen and Hurwitz [5] proposed the unbiased estimator for the estimation of the population mean. The estimator and its variance are, respectively,

$$t = w_1\bar{y}_1 + w_2\bar{y}_{2(k)} \tag{1}$$

and

$$V(t) = \left(\gamma S_y^2 + \lambda S_{y(2)}^2\right) \tag{2}$$

The weight of response units is $w_1 = \frac{n_1}{n}$ for sample, while the weights of non-response units are $w_2 = \frac{n_2}{n}$ and $W_2 = \frac{N_2}{N}$ for sample and population, respectively. Also, $\bar{y}_1$ and $\bar{y}_{2(k)}$ are the sample means of y $(\bar{y})$ according to n_1 and k $\left(k = \frac{n_2}{r}\right)$ units, respectively. In Eq. (2), $\gamma = \frac{1-\frac{n}{N}}{n}$, $S_y^2 = C_y^2\bar{Y}^2, S_{y(2)}^2 = C_{y(2)}^2\bar{Y}^2$, and $\lambda = \frac{W_2(r-1)}{n}$ while $\bar{Y}$ is the population mean of y. Two main categories, called Case I and Case II, are used to analyse the non-response situation. In Case I, while the population mean of x $(\bar{X})$ is known, the non-response units are only observed on y. Firstly, to estimate under this case, Rao [15] proposed a classical ratio and classical regression estimators as follows:

$$t_{Ratio1} = \bar{y}^*\frac{\bar{X}}{\bar{x}} \tag{3}$$

In the context of a non-response situation, $\bar{y}^*$ represents the sample mean of y. Additionally, $\bar{x}$ denotes the sample mean of x. The Mean Square Error (MSE) equation of this estimator is given as

$$MSE(t_{Ratio1}) = \bar{Y}^2\left(\gamma\left(C_x^2 - 2C_{yx} + C_y^2\right) + \lambda C_{y(2)}^2\right) \tag{4}$$

where $S_x^2 = \bar{X}^2C_x^2, C_{yx} = \rho_{xy}C_yC_x$ and ρ_{xy} is the population correlation coefficient between y and x. Taking advantage of exponential function, Singh et al. [16] introduced a new exponential type estimator using the method. The estimator and its MSE equation are, respectively,

$$t_{Exp1} = \bar{y}^*\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) \tag{5}$$

and

$$MSE(t_{Exp1}) = \bar{Y}^2\gamma\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right) + \bar{Y}^2\lambda C_{y(2)}^2 \tag{6}$$

In Case II, the non-response units are observed on both y and x, while the population mean of x $(\bar{X})$ is known. In this case, Cochran [17] proposed a classical ratio type estimator in literature as follows:

$$t_{Ratio2} = \bar{y}^* \frac{\bar{X}}{\bar{x}^*} \tag{7}$$

In the context of a non-response situation, $\bar{x}^*$ represents the sample mean of x and whose MSE of this estimator

$$MSE(t_{Ratio2}) = \bar{Y}^2\left(\gamma(C_x^2 - 2C_{yx} + C_y^2) + \lambda(C_{y(2)}^2 + C_{x(2)}^2 - 2C_{yx(2)})\right) \tag{8}$$

where $S_{x(2)}^2 = \bar{X}^2 C_{x(2)}^2$, $C_{yx(2)} = C_{y(2)}C_{x(2)}\rho_{xy(2)}$, and the $\rho_{xy(2)}$ is symbolized the coefficient of the correlation for the non-response group.

Singh et al. [16] also introduced a new exponential type estimator as well as for Case II and this estimator is given as

$$t_{Exp2} = \bar{y}^* \exp\left(\frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*}\right) \tag{9}$$

and whose MSE of this estimator

$$MSE(t_{Exp2}) = \bar{Y}^2\left(\gamma\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right) + \lambda\left(C_{y(2)}^2 + \frac{C_{x(2)}^2}{4} - C_{yx(2)}\right)\right) \tag{10}$$

The Proposed Family of Estimators in The Case of Non-Response Schemes

The generalized class of exponential type estimators is proposed by Grover and Kaur [18] as follows:

$$t_{GK} = [\alpha\bar{y} + (\bar{X} - \bar{x})\beta]\exp\left(\frac{\bar{Z} - \bar{z}}{\bar{Z} + \bar{z}}\right) \tag{11}$$

where $\bar{Z} = \delta\bar{X} + \varphi$ and $\bar{z} = \delta\bar{x} + \varphi$. Here, δ and φ are either functions of the known population parameters of x or constants. We re-write the class of exponential type estimators as

$$t_{GK} = [\alpha\bar{y} + (\bar{X} - \bar{x})\beta]\exp\left(\frac{(\bar{X} - \bar{x})\delta}{2\varphi + (\bar{X} + \bar{x})\delta}\right) \tag{12}$$

This class of estimators becomes prominent among exponential type estimators in literature. On getting the motivation of this class of estimators, for both cases in this study, we propose a new family of estimators in the presence of non-response to the estimation of the population mean.

Case I

For the Case I, we can write

$$t_{(\delta,\varphi),j} = [\upsilon_{11}\bar{y}^* + (\bar{X} - \bar{x})\upsilon_{12}]\exp\left[\frac{(\bar{X} - \bar{x})\delta}{2\varphi + (\bar{X} + \bar{x})\delta}\right], j \in \{1,2,\ldots,10\} \tag{13}$$

Using different δ and φ, some members of the proposed estimators can be derived. Besides, υ_{11} and υ_{12} are both constants that make the $MSE\left(t_{(\delta,\varphi),j}\right), j \in \{1,2,\dots,10\}$ minimum. For obtaining the bias, MSE as well minimum MSE of the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$, we use notations in the presence of Case I as follows:

$\bar{y}^* = \left(\bar{Y}e_y^* + \bar{Y}\right), \bar{x} = (\bar{X}e_x + \bar{X}),$
$E(e_x^*) = 0, E(e_x) = 0,$
$E\left({e_y^*}^2\right) = \left(\gamma C_y^2 + \lambda C_{y(2)}^2\right), E(e_x^2) = \gamma C_x^2$, and $E\left(e_y^* e_x\right) = \gamma C_{yx}$.

Using these notations, we can obtain bias and MSE of the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimator, respectively, as follows:

$$Bias\left(t_{(\delta,\varphi),j}\right) = E\left(t_{(\delta,\varphi),j} - \bar{Y}\right) = \left\{\bar{Y}\left((\upsilon_{11} - 1) + \upsilon_{11}\omega\gamma C_x\left(\frac{3\omega}{2}C_x - \rho_{xy}C_y\right)\right) + \upsilon_{12}\bar{X}\omega\gamma C_x^2\right\} \quad (14)$$

$$MSE\left(t_{(\delta,\varphi),j}\right) = \left\{\bar{Y}^2\left[(\upsilon_{11} - 1)^2 + \upsilon_{11}^2 A - \omega\upsilon_{11}\left(B + \gamma\left(\omega C_x^2 - C_{yx}\right)\right)\right] + \upsilon_{12}^2\bar{X}^2\gamma C_x^2 + 2\bar{X}\bar{Y}\upsilon_{12}(B\upsilon_{11} - \omega\gamma C_x^2)\right\} \quad (15)$$

where $A = \left(4\omega\gamma\left(\omega C_x^2 - C_{yx}\right) + \gamma C_y^2 + \lambda C_{y(2)}^2\right)$ and $B = \left(2\omega\gamma C_x^2 - \gamma C_{yx}\right)$ in Eq. (15). Here, ω is considered as $\omega = \frac{\delta\bar{X}}{2(\delta\bar{X}+\varphi)}$ in order to simplify the mathematical notation. According to these δ and φ values, some member of the ω are obtained, respectively, as follows:

$\omega_1 = \frac{\bar{X}}{2\bar{X}+2}, \omega_2 = \frac{\bar{X}}{2\bar{X}+2\beta_2(x)}, \omega_3 = \frac{\bar{X}}{2\bar{X}+2C_x}, \omega_4 = \frac{\bar{X}}{2\bar{X}+2\rho}, \omega_5 = \frac{\beta_2(x)\bar{X}}{2\beta_2(x)\bar{X}+2C_x}, \omega_6 = \frac{C_x\bar{X}}{2C_x\bar{X}+2\beta_2(x)},$
$\omega_7 = \frac{C_x\bar{X}}{2C_x\bar{X}+2\rho}, \omega_8 = \frac{\rho\bar{X}}{2\rho\bar{X}+2C_x}, \omega_9 = \frac{\beta_2(x)\bar{X}}{2\beta_2(x)\bar{X}+2\rho}$, and $\omega_{10} = \frac{\rho\bar{X}}{2\rho\bar{X}+2\beta_2(x)}$.

The min MSE of the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimators is obtained using the optimal values of υ_{11} and υ_{12}, υ_{11}^* and υ_{12}^*, respectively, as follows:

$$\upsilon_{11}^* = \frac{-E(e_x^2)\left(2-\omega B+\omega\left(\omega E(e_x^2)-E(e_y^* e_x)\right)\right)}{2\left(B^2-(1+A)E(e_x^2)\right)} \text{ and } \upsilon_{12}^* = \frac{\bar{Y}\left[B\left\{2+\omega B+\omega\left(\omega E(e_x^2)-E(e_y^* e_x)\right)\right\}-2\omega E(e_x^2)(1+A)\right]}{2\bar{X}\left(B^2-(1+A)E(e_x^2)\right)}$$

Substituting υ_{11}^* and υ_{12}^* values, instead of υ_{11} and υ_{12}, respectively, in Eq. (15), we obtain the min $MSE\left(t_{(\delta,\varphi),j}\right), j \in \{1,2,\dots,10\}$ in the presence of Case I as

$$MSE_{min}\left(t_{(\delta,\varphi),j}\right) = \frac{\omega^4\left(\gamma C_x^2\right)^3 + 4\left(\left(\gamma C_{yx}\right)^2 - \left(\gamma C_y^2+\lambda C_{y(2)}^2\right)\left(\gamma C_x^2\right)\right)\left(1-\omega^2\gamma C_x^2\right)}{4\left(\left(\left(\gamma C_{yx}\right)^2 - \left(\gamma C_y^2+\lambda C_{y(2)}^2\right)\left(\gamma C_x^2\right)\right) - \gamma C_x^2\right)}, j \in \{1,2,\dots,10\} \quad (16)$$

Case II

For the Case II, we can write

$$t^{**}_{(\delta,\varphi),j} = [\upsilon_{21}\bar{y}^* + \upsilon_{22}(\bar{X} - \bar{x}^*)] \exp\left[\frac{\delta(\bar{X}-\bar{x}^*)}{2\varphi+\delta(\bar{X}+\bar{x}^*)}\right], j \in \{1,2,\dots,10\} \tag{17}$$

where υ_{21} and υ_{22} are used instead of υ_{11} and υ_{12}, respectively, in Eq. (17) for the first proposed family of estimators, given in Eq. (13).

We use notations in order to obtain bias, MSE as well minimum MSE of the $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ for the Case II as follows:

$\bar{y}^* = \left(\bar{Y}e^*_y + \bar{Y}\right), \bar{x}^* = (\bar{X}e^*_x + \bar{X}),$

$E(e^*_x) = 0, E\left(e^*_y\right) = 0,$

$E\left({e^*_x}^2\right) = \left(\gamma C^2_x + \lambda C^2_{x(2)}\right), E\left({e^*_y}^2\right) = \left(\gamma C^2_y + \lambda C^2_{y(2)}\right)$, and $E\left(e^*_y e^*_x\right) = \left(\gamma C_{yx} + \lambda C_{yx(2)}\right)$.

Using these notations, we can obtain bias and MSE of the $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimator, respectively, as follows:

$$Bias\left(t^{**}_{(\delta,\varphi),j}\right) = \left\{\bar{Y}\left((\upsilon_{21}-1) + \upsilon_{21}\omega\gamma\left(\frac{3\omega}{2}C^2_x - C_{xy}\right) + \lambda\upsilon_{21}\omega\left(\frac{3\omega}{2}C^2_{x(2)} - C_{xy(2)}\right)\right) + \upsilon_{22}\bar{X}\omega\left(\gamma C^2_x + \lambda C^2_{x(2)}\right)\right\} \tag{18}$$

$$MSE\left(t^{**}_{(\delta,\varphi),j}\right) = \left\{\bar{Y}^2\left[(\upsilon_{21}-1)^2 - \omega\upsilon_{21}\left(D + \left(\omega\left(\gamma C^2_x + \lambda C^2_{x(2)}\right) - \left(\gamma C_{yx} + \lambda C_{yx(2)}\right)\right)\right)\right] + \upsilon^2_{21}C + \upsilon^2_{22}\bar{X}^2\left(\gamma C^2_x + \lambda C^2_{x(2)}\right) + 2\bar{X}\bar{Y}\upsilon_{22}\left(B\upsilon_{21} - \omega\left(\gamma C^2_x + \lambda C^2_{x(2)}\right)\right)\right\} \tag{19}$$

where $C = \gamma\left(4\omega^2C^2_x + C^2_y - 4\omega C_{yx}\right) + \lambda\left(4\omega^2C^2_{x(2)} + C^2_{y(2)} - 4\omega C_{yx(2)}\right)$ and $D = \left(2\omega\left(\gamma C^2_x + \lambda C^2_{x(2)}\right) - \left(\gamma C_{yx} + \lambda C_{yx(2)}\right)\right)$

The optimal values of υ_{21} and υ_{22} are obtained, respectively, as follows:

$$\upsilon^*_{21} = \frac{-E\left({e^*_x}^2\right)\left(2-\omega D+\omega\left(\omega E\left({e^*_x}^2\right)-E\left(e^*_y e^*_x\right)\right)\right)}{2\left(D^2-(1+C)E\left({e^*_x}^2\right)\right)} \text{ and } \upsilon^*_{22} = \frac{\bar{Y}\left[D\left\{2+\omega D+\omega\left(\omega E\left({e^*_x}^2\right)-E\left(e^*_y e^*_x\right)\right)\right\}-2\omega E\left({e^*_x}^2\right)(1+C)\right]}{2\bar{X}\left(D^2-(1+C)E\left({e^*_x}^2\right)\right)}$$

The υ^*_{21} and υ^*_{22} values are substituted in $MSE\left(t^{**}_{(\delta,\varphi),j}\right), j \in \{1,2,\dots,10\}$ and we obtain the minimum $MSE\left(t^{**}_{(\delta,\varphi),j}\right), j \in \{1,2,\dots,10\}$ in the presence of Case II as

$$MSE\left(t^{**}_{(\delta,\varphi),j}\right)_{min} = \frac{\omega^4\left(\gamma C^2_x+\lambda C^2_{x(2)}\right)^3+4G\left(1-\omega^2\left(\gamma C^2_x+\lambda C^2_{x(2)}\right)\right)}{4\left[G-\left(\gamma C^2_x+\lambda C^2_{x(2)}\right)\right]} \tag{20}$$

where $G = \left(\gamma C_{yx} + \lambda C_{yx(2)}\right)^2 - \left(\gamma C^2_y + \lambda C^2_{y(2)}\right)\left(\gamma C^2_x + \lambda C^2_{x(2)}\right)$

Using various δ and φ, we can present some members of the proposed families of estimators for the both cases.

Efficiency Comparisons

In this section, the proposed families of estimators are compared theoretically and numerically with other recent estimators in literature for the Cases I and II, respectively, to show the efficiency of the proposed estimators. The simulation study is also performed as well.

Theoretical Comparisons

Theoretical Comparisons for the Case I

We compare the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimators with the t_H, t_{Ratio1}, and t_{Exp1} estimators, and comparison of efficiency between MSE equations is acquired as follows:

$$MSE_{min}\left(t_{(\delta,\varphi),j}\right) < V(t_H), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4(\gamma C_x^2)^3+4\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)(1-\omega^2\gamma C_x^2)}{4\left(\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)-\gamma C_x^2\right)} - \bar{Y}^2\left(\lambda C_y^2 + \frac{W_2(r-1)}{n}C_{y(2)}^2\right) < 0 \quad (21)$$

$$MSE_{min}\left(t_{(\delta,\varphi),j}\right) < MSE_{min}(t_{Ratio1}), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4(\gamma C_x^2)^3+4\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)(1-\omega^2\gamma C_x^2)}{4\left(\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)-\gamma C_x^2\right)} - \bar{Y}^2\left(\lambda\left(C_x^2 - 2C_{yx} + C_y^2\right) + \frac{W_2(r-1)}{n}C_{y(2)}^2\right) < 0 \quad (22)$$

$$MSE_{min}\left(t_{(\delta,\varphi),j}\right) < MSE_{min}\left(t_{Exp1}\right), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4(\gamma C_x^2)^3+4\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)(1-\omega^2\gamma C_x^2)}{4\left(\left((\gamma C_{yx})^2-\left(\gamma C_y^2+\lambda C_{y(2)}^2\right)(\gamma C_x^2)\right)-\gamma C_x^2\right)} - \bar{Y}^2\left(\lambda\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right) + \frac{W_2(r-1)}{n}C_{y(2)}^2\right) < 0 \quad (23)$$

Under the conditions obtained for Case I, the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimator is more effective than the compared estimators.

Theoretical Comparisons for the Case II

We compare the $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimators with the t_H, t_{Ratio2}, and t_{Exp2} estimators, and comparison of efficiency between MSE equations is acquired as follows:

$$MSE_{min}\left(t^{**}_{(\delta,\varphi),j}\right) < V(t_H), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)^3+4G\left(1-\omega^2\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right)}{4\left[G-\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right]} - \bar{Y}^2\left(\lambda C_y^2 + \frac{W_2(r-1)}{n}C_{y(2)}^2\right) < 0 \quad (24)$$

$$MSE_{min}\left(t^{**}_{(\delta,\varphi),j}\right) < MSE(t_{Ratio2}), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)^3+4G\left(1-\omega^2\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right)}{4\left[G-\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right]} - \bar{Y}^2\left(\lambda\left(C_x^2 - 2C_{yx} + C_y^2\right) + \frac{W_2(r-1)}{n}\left(C_{y(2)}^2 + C_{x(2)}^2 - 2C_{yx(2)}\right)\right) < 0 \quad (25)$$

$$MSE_{min}\left(t^{**}_{(\delta,\varphi),j}\right) < MSE\left(t_{E\,xp\,2}\right), j \in \{1,2,\dots,10\}$$

$$\frac{\omega^4\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)^3+4G\left(1-\omega^2\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right)}{4\left[G-\left(\gamma C_x^2+\lambda C_{x(2)}^2\right)\right]}-\bar{Y}^2\left(\lambda C_y^2+\lambda\frac{C_x^2}{4}-\lambda C_{yx}+\frac{W_2(r-1)}{n}\left(C_{y(2)}^2-C_{yx(2)}+\frac{C_{x(2)}^2}{4}\right)\right)<0 \quad (26)$$

Under the conditions obtained for Case II, the $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimator is more effective than the compared estimators.

Numerical Comparisons

The Real Data Example

In this section, the percent relative efficiency (PRE) values, for which the reference estimator is t_{HH} estimator, are calculated for the comparison and the proposed families of estimators, $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ and $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ in the presence of both cases, respectively, using the new data set connected with magnitude. The PRE-values are obtained through proportioning the mean square errors as follows:

$$PRE_i = \frac{V(t_H)}{MSE(t_i)}x100 \quad (27)$$

This new data set is obtained within the scope of TUBITAK (The Scientific and Technological Research Council of Turkey) Scientific and Technological Research Projects Funding Program – 1001 Project [19]. Using the catalogs prepared by the KOERI (Bogazici University Kandilli Observatory and Earthquake Research Institute) and the General Directorate of Disaster Affairs-Earthquake Research Department, the data was obtained between the years 1900-2021 for this area. Because of the high seismicity in the Aegean Region, the examined data is included the earthquakes whose wave magnitude values (Mw) are 4.0 or greater between the years 2000 and 2021 for this region of this study. In this data set, the magnitude of the main shock and the magnitude of the largest aftershock of this main shock are considered as y and x, respectively. Here, the catalog includes 452 different shocks specified in this period. Similar to the studies in the literature, the last 25% of units $(W_2 = 0.25)$ are determined as the non-response group. The data set is presented in Table 1.

Table 1. *Descriptive statistics of the population*

$n = 210, N = 452$	$W_2 = 0.25$	$\lambda = 0.002$	$\rho_{yx} = 0.63$	$C_x = 0.0841$
$\bar{X} = 4.306$	$C_{x(2)} = 0.0854$	f $= 0.5221$	$\rho_{yx(2)} = 0.83$	$C_y = 0.1111$
$\bar{Y} = 4.6058$	$C_{y(2)} = 0.1203$	$C_{yx(2)} = 0.0085$	$\beta_2(x) = 5.755$	$C_{yx} = 0.0058$

Due to the positive correlation between magnitude of a main shock and magnitude of the largest aftershock (ρ=0.63), it would be appropriate to use ratio type estimators instead of product type. The kurtosis value is calculated as 5.755 and this means that a leptokurtic distribution has positive kurtosis value, higher peaked and possesses thick tails. Using this data set, the PRE-values of the all mentioned estimators are obtained for the Case I as in Table 2. Table 2 shows that all the members of the proposed family of estimators have the highest PRE-values among the other estimators found in the literature for the Case I. Furthermore, it can be concluded that the $t_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$ estimator is the optimal choice for estimating the population mean in the presence of Case I. Due to very close PRE values, any of the family members can be preferred to estimate the $(\bar{Y})$of the magnitude of a main shock. Secondly, for Case II, as in Table 3, the PRE-values of all the estimators mentioned are obtained.

Table 2. *PRE-values of the estimators (Case I)*

Estimators	r=2	r=3	r=4	r=5	r=6	r=7
t_H	100.00	100.00	100.00	100.00	100.00	100.00
t_{Ratio1}	128.28	119.48	115.07	112.33	110.36	108.96
t_{Exp1}	124.22	116.84	113.10	110.75	109.05	107.84
$t_{(\delta,\varphi),1}$	2757.92	2557.95	2458.53	2396.83	2352.81	2321.45
$t_{(\delta,\varphi),2}$	2757.91	2557.94	2458.53	2396.82	2352.81	2321.44
$t_{(\delta,\varphi),3}$	2757.92	***2557.95***	***2458.53***	2396.83	***2352.82***	***2321.45***
$t_{(\delta,\varphi),4}$	2757.92	2557.95	2458.53	2396.83	2352.81	2321.45
$t_{(\delta,\varphi),5}$	***2757.92***	***2557.95***	***2458.53***	***2396.83***	***2352.82***	***2321.45***
$t_{(\delta,\varphi),6}$	2757.91	2557.94	2458.52	2396.82	2352.81	2321.44
$t_{(\delta,\varphi),7}$	2757.91	2557.94	2458.52	2396.82	2352.81	2321.44
$t_{(\delta,\varphi),8}$	2757.92	***2557.95***	2458.53	2396.83	***2352.82***	***2321.45***
$t_{(\delta,\varphi),9}$	2757.92	***2557.95***	2458.53	2396.83	***2352.82***	***2321.45***
$t_{(\delta,\varphi),10}$	2757.91	2557.94	2458.52	2396.82	2352.81	2321.44

Table 3. *PRE-values of the estimators (Case II)*

Estimators	r=2	r=3	r=4	r=5	r=6	r=7
t_H	100.00	100.00	100.00	100.00	100.00	100.00
t_{Ratio2}	198.42	218.65	231.83	241.48	249.21	255.22
t_{Exp2}	162.37	168.00	171.31	173.58	175.31	176.61
$t^{**}_{(\delta,\varphi),1}$	162.37	4638.74	4927.11	5145.57	5325.37	5468.24
$t^{**}_{(\delta,\varphi),2}$	4218.26	4638.73	4927.08	5145.53	5325.33	5468.19
$t^{**}_{(\delta,\varphi),3}$	4218.25	4638.76	4927.13	5145.59	5325.40	5468.27
$t^{**}_{(\delta,\varphi),4}$	4218.27	4638.75	4927.12	5145.58	5325.38	5468.25
$t^{**}_{(\delta,\varphi),5}$	***4218.27***	***4638.76***	***4927.13***	***5145.59***	***5325.40***	***5468.27***
$t^{**}_{(\delta,\varphi),6}$	4218.25	4638.72	4927.07	5145.52	5325.31	5468.17
$t^{**}_{(\delta,\varphi),7}$	4218.25	4638.72	4927.08	5145.53	5325.32	5468.18
$t^{**}_{(\delta,\varphi),8}$	4218.27	4638.76	4927.13	5145.59	5325.39	5468.27
$t^{**}_{(\delta,\varphi),9}$	4218.27	4638.76	4927.13	5145.59	5325.40	5468.27
$t^{**}_{(\delta,\varphi),10}$	4218.25	4638.72	4927.08	5145.53	5325.32	5468.18

According to the obtained result in Table 4, all members of the proposed family of estimators have maximum PRE. Therefore, we conclude that the $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\ldots,10\}$ can be preferred on the estimation of the $(\bar{Y})$ in the presence of Case II as same as Case I. These results show that the proposed $t_{(\delta,\varphi),j}, j \in \{1,2,\ldots,10\}$ and $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\ldots,10\}$ estimators can be applied on the estimation of the $(\bar{Y})$ in the presences of Cases I and II as well. According to the obtained PRE-values, it is concluded that these values is higher especially for Case 2 compared to Case 1 as well.

Simulation Study

In this section, we perform a simulation study for supporting theoretical results and numerical illustrations. In the simulation code obtained by the R-Statistical software, we generated a population of size N=1000 and sample size is considered as n=300 from the bivariate normal distribution in which non-respondents and are respondents are considered for both cases. We suppose that the last 25% of units are symbolized as the non-response group. To show the performance of the proposed estimator, $t_{(\delta,\varphi),j}, j \in \{1,2,\ldots,10\}$, simulation study is conducted and we assume that they follow bivariate normal distribution with means (1, 1), and standard deviation (3, 0.5) for x and y, respectively, with the correlation coefficient that are designated as both 0.50 and 0.95. The calculated PRE-values are given in Tables 4 – 5 for both cases, respectively, as follows:

Table 4. *PRE-values of the estimators under simulation study (Case I)*

	PRE-Values					
	ρ_{yx} =0.50			ρ_{yx} =0.95		
	r=3	r=5	r=7	r=3	r=5	r=7
t_H	100.00	100.00	100.00	100.00	100.00	100.00
t_{Ratio1}	28.37	24.46	13.34	5.52	15.37	40.92
t_{Exp1}	139.89	208.21	***214.97***	168.60	194.13	147.20
$t_{(\delta,\varphi),1}$	144.05	156.05	151.07	264.14	237.74	161.50
$t_{(\delta,\varphi),2}$	139.43	152.67	152.92	260.95	236.38	158.67
$t_{(\delta,\varphi),3}$	140.51	152.88	152.77	260.97	236.51	159.46
$t_{(\delta,\varphi),4}$	146.17	157.87	150.00	264.32	237.83	161.60
$t_{(\delta,\varphi),5}$	144.96	156.24	150.83	264.17	238.04	161.99
$t_{(\delta,\varphi),6}$	143.05	155.85	151.32	264.11	237.48	160.90
$t_{(\delta,\varphi),7}$	147.79	159.64	149.41	268.40	240.39	162.65
$t_{(\delta,\varphi),8}$	137.84	150.73	153.30	260.83	236.47	159.31
$t_{(\delta,\varphi),9}$	***148.09***	***159.73***	149.41	***268.43***	***240.84***	***162.81***
$t_{(\delta,\varphi),10}$	136.97	150.55	153.35	260.80	236.33	158.53

Table 5. *PRE-values of the estimators under simulation study (Case II)*

	PRE-Values					
	ρ_{yx} =0.50			ρ_{yx} =0.95		
	r=3	r=5	r=7	r=3	r=5	r=7
t_H	100.00	100.00	100.00	100.00	100.00	100.00
t_{Ratio2}	1.98	2.45	0.97	2.75	1.66	1.68
t_{Exp2}	11.63	13.36	7.47	18.19	13.00	18.54
$t^{**}_{(\delta,\varphi),1}$	135.53	127.88	133.64	855.93	826.95	880.63
$t^{**}_{(\delta,\varphi),2}$	137.33	130.20	136.93	***866.15***	***839.12***	897.69
$t^{**}_{(\delta,\varphi),3}$	137.38	130.12	137.07	865.82	838.52	***898.51***
$t^{**}_{(\delta,\varphi),4}$	132.73	124.73	128.00	854.23	824.08	877.18
$t^{**}_{(\delta,\varphi),5}$	135.72	127.59	134.30	858.49	829.86	875.54
$t^{**}_{(\delta,\varphi),6}$	135.33	128.15	132.87	852.93	823.64	885.04
$t^{**}_{(\delta,\varphi),7}$	126.44	119.04	113.45	777.46	670.75	739.83
$t^{**}_{(\delta,\varphi),8}$	***137.71***	130.57	***137.45***	865.59	837.97	897.95
$t^{**}_{(\delta,\varphi),9}$	126.80	118.65	114.89	785.07	680.55	727.60
$t^{**}_{(\delta,\varphi),10}$	137.69	***130.60***	137.41	865.96	838.59	897.07

As seen in Tables 4 and 5, the proposed family of estimators, $t^{**}_{(\delta,\varphi),j}, j \in \{1,2,\dots,10\}$, has the highest PRE-values compared to the other estimators for both correlation coefficients and all r values as well except for $\rho_{yx} = 0.50$ and $r = 7$. It is found that, the t_{Exp1} estimator is the best estimator for this combination.

Conclusion

We consider a new family of estimators on the estimation of the unknown $(\bar{Y})$ by using the information of x. These estimators are defined under the two different non-response schemes. Firstly, we obtain the bias, MSE, and the minimum MSE for both schemes and then comparisons are given theoretically. After that, we use the application about the magnitude data for the numerical comparison. According to obtained results, the proposed estimators have higher PRE-values when comparing with the main estimators. We conduct a simulation study as well. In the simulation study as well, the proposed estimators are again more effective than the others, especially under the Case II. Therefore, we recommend the proposed families of estimators under the non-response cases.

Acknowledgements This publication is a part of PhD thesis of the first author. The authors are very grateful to the blinded reviewers for their valuable comments and to Dr. Lovleen Kumar Grover for providing us all required information. This study was supported by Scientific and Technological Research Council of Turkey (TUBITAK) under the Grant Number 121F208. The authors thank to TUBITAK for their supports. The authors also thank to Assoc. Prof. Dr. Senem Tekin and Assoc. Prof. Dr. Tuba Eroğlu Azak for preparing the data set and thank to Prof. Dr. Tolga Çan for his valuable advice.

Funding/Financial Disclosure The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Ethics Committee Approval and Permissions Not applicable.

Conflict of Interests - The authors report no potential conflict of interest.

Authors Contribution Ceren Ünal designed and performed the experiments and wrote the paper. Cem Kadilar contributed to the paper writing, review and editing. All authors read and approved the final manuscript.

References

[1] Solanki, R. S., Singh, H. P., & Rathour, A. (2012). An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. *International Scholarly Research Notices*, 1-14. https://doi.org/10.5402/2012/657682

[2] Oncel Cekim, H., & Kadilar, C. (2018). New families of unbiased estimators in stratified random sampling. *Journal of Statistics and Management Systems*, *21*(8), 1481-1499. https://doi.org/10.1080/09720510.2018.1530176

[3] Zaman, T., & Kadilar, C. (2020). On estimating the population mean using auxiliary character in stratified random sampling. *Journal of Statistics and Management Systems*, *23*(8), 1415-1426. https://doi.org/10.1080/09720510.2020.1723924

[4] Dansawad, N. (2019). A class of exponential estimator to estimate the population mean in the presence of non-response. *Naresuan University Journal: Science and Technology*, *27*(4), 20-26.

[5] Hansen, M., & Hurwitz, N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, *41*(236), 517-529. https://doi.org/10.1080/01621459.1946.10501894

[6] Kumar, S., Kour, S. P., & Sharma, V. (2022). Modified exponential estimators using auxiliary information under response and non-response. *Revista Investigación Operacional*, *43*(4), 491-504

[7] Jaiswal, A. K., Singh, G. N., & Pandey, A. K. (2022). Improved procedures for mean estimation under non-response. *Alexandria Engineering Journal*, *61*(12), 12813-12828. https://doi.org/10.1016/j.aej.2022.06.031

[8] Singh, H. P., Yadav, A., & Pal, S. K. (2021). An exponential approach for estimating population mean using two auxiliary variables in stratified random sampling. *Revista Investigación Operacional*, *42*(4), 456-468.

[9] Singh, A. K., & Singh, V. K. (2022). A family of estimators for population mean under model approach in presence of non-response. *Journal of Reliability and Statistical Studies*, *15*(1), 1-20. https://doi.org/10.13052/jrss0974-8024.1511

[10] Singh, G. N., & Usman, M. (2021). Efficient combination of various estimators in the presence of non-response. *Communications in Statistics-Simulation and Computation*, *50*(8), 2432-2466. https://doi.org/10.1080/03610918.2019.1614618

[11] Pandey, A. K., Usman, M., & Singh, G. N. (2021). Optimality of ratio and regression type estimators using dual of auxiliary variable under non-response. *Alexandria Engineering Journal*, *60*(5), 4461–4471. https://doi.org/10.1016/j.aej.2021.03.031

[12] Singh, H. P., & Nigam, P. (2021). Efficient method of estimating the finite population mean based on two auxiliary variables in the presence of non-response under stratified sampling. *Journal of Reliability and Statistical Studies*, *14*(1), 223–242. https://doi.org/10.13052/jrss0974-8024.14111

[13] Khalid, M., & Singh, G. N. (2022). Some imputation methods to deal with the issue of missing data problems due to random non-response in two-occasion successive sampling. *Communications in Statistics-Simulation and Computation*, *51*(12), 7266–7286. https://doi.org/10.1080/03610918.2020.1828920

[14] Hussain, M., Zaman, Q., Khan, L., Metawa, A. E., Awwad, F. A., Ismail, E. A. A., Wasim, D., & Ahmad, H. (2024). Improved exponential type mean estimators for non-response case using two concomitant variables in simple random sampling. *Heliyon*, *10*(6), e27535. https://doi.org/10.1016/j.heliyon.2024.e27535

[15] Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents. *Survey Methodology*, *12*(2), 217–230.

[16] Singh, R.., Kumar, M., Chaudhary, M. K., & Smarandache, F. (2009). Estimation of mean in presence of non-response using exponential estimator. *Multispace & Multistructure. Neutrosophic Transdisciplinarity 100 Collected Papers of Sciences, North-European Scientific Publishers*, 2010, 758-768.

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

[18] Grover, L. K., & Kaur, P. (2014). A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable. *Communications in Statistics-Simulation and Computation*, *43*(7), 1552-1574. https://doi.org/10.1080/03610918.2012.736579

[19] TUBITAK 1001. (2021). The Scientific and Technological Research Projects Funding Program. *Artificial Intelligence and Probabilistic Model Based Earthquake Hazard Map*, Project Number: 121F208.