

## Research Article

# Estimating a Finite Population Mean Using Transformed Data in Presence of Random Nonresponse

Nelson Kiprono Bii <sup>1</sup>, Christopher Ouma Onyango,<sup>2</sup> and John Odhiambo <sup>1</sup>

<sup>1</sup>*Institute of Mathematical Sciences, Strathmore University, P.O. Box 59857-00200, Nairobi, Kenya*

<sup>2</sup>*Department of Statistics, Kenyatta University, P.O. Box 43844-00200, Nairobi, Kenya*

Correspondence should be addressed to Nelson Kiprono Bii; [nkiprono@strathmore.edu](mailto:nkiprono@strathmore.edu)

Received 10 March 2020; Accepted 22 June 2020; Published 8 July 2020

Academic Editor: Andrei I. Volodin

Copyright © 2020 Nelson Kiprono Bii et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Developing finite population estimators of parameters such as mean, variance, and asymptotic mean squared error has been one of the core objectives of sample survey theory and practice. Sample survey practitioners need to assess the properties of these estimators so that better ones can be adopted. In survey sampling, the occurrence of nonresponse affects inference and optimality of the estimators of finite population parameters. It introduces bias and may cause samples to deviate from the distributions obtained by the original sampling technique. To compensate for random nonresponse, imputation methods have been proposed by various researchers. However, the asymptotic bias and variance of the finite population mean estimators are still high under this technique. In this paper, transformation of data weighting technique is suggested. The proposed estimator is observed to be asymptotically consistent under mild assumptions. Simulated data show that the estimator proposed is much better than its rival estimators for all the different mean functions simulated.

## 1. Introduction

A lot of significance is attached to efficient and cost-effective survey sampling designs in sample surveys while estimating a finite population mean, see for instance [1, 2]. Careful design of samples based on random selection with known probabilities of population elements should be considered. This gives a target sample of intended respondents where each may provide responses to a set of survey questions that result in an array of responses. Van Buuren et al. [3] observed that nonresponse occurs if some of the expected responses are missing, for instance, where a whole vector of responses is missing for some sampled units or where responses are obtained for some questions and not to others in the sample selected. As noted in [4], nonresponse often occurs in human population surveys as people hesitate to respond in surveys and increases notably while studying sensitive issues. Moreover, the presence of nonresponse increases the bias in estimates, ultimately reducing their efficiency as observed in [4].

The basis for statistical inference is therefore formed by a sampling design that provides a link between a sample and the population. As observed in [1], a good sample survey practice and efficient methods of compensating for nonresponse should be adopted.

In sample surveys, nonresponse leads to biased results in the estimation of a finite population mean. It may force samples to deviate from the distributions originally established by the sampling design. The incorporation of regression models is acknowledged as one of the methods of minimizing bias resulting from nonresponse using auxiliary data, for details see [5]. In practice, knowledge on the study variables is unavailable for nonrespondents, whereas auxiliary data may be given. To minimize the bias and variance resulting from nonresponse, Liang and Zeger [6] noted that it is desirable to incorporate auxiliary data in the process of estimation where the probabilities of response are mostly assumed to be correlated with certain characteristics, for instance, age, race, and income in human population surveys. Sanaullah et al. [7] studied nonresponse under

stratified two-phase sampling using generalized exponential chain-ratio and chain-product estimators and concluded that their estimator is more efficient than that in [8], one of the pioneer studies of estimation of finite population parameters under nonresponse.

In the sequence of addressing the problem of nonresponse in sample surveys, Javaid et al. [9] derived a modified ratio estimator in systematic random sampling. They proposed the use of single auxiliary variable to estimate a finite population mean. However, as a departure from the work in [7, 9], this paper proposed the use of weights due to transformed data to compensate for nonresponse. The weighting method is highlighted in the following section.

This paper is organized as follows: introduction of the paper has been presented in Section 1; in Section 2, a review of weighting method has been discussed; the proposed estimator is derived in Section 3; in Sections 3.1–3.3, the bias, the variance, and the mean squared error of the estimator are derived, respectively; Section 4 presents results from the simulation experiment conducted; and the conclusion of the paper has been presented in Section 5.

## 2. Review of Weighting Method

It has been observed by authors in [10] that nonresponse leads to reduced number of observations. Weighting thus implies that the weights are increased almost for all the elements that do not respond in a survey. For instance, the authors in [11, 12] explored a modified Horvitz–Thompson estimator to correct the problem of nonresponse using the weighting technique. The estimator used was defined as

$$\hat{y}_{HT} = N^{-1} \sum_{k=1}^N (\phi_k \hat{p}_{HT})^{-1} y_k \tau_k, \quad (1)$$

where  $\hat{p}_{HT}$  is given as

$$\hat{p}_{HT} = N^{-1} \sum_{k=1}^N \phi_k^{-1} \gamma_k, \quad (2)$$

such that  $y_k, k \in U$ , is the  $k^{\text{th}}$  survey value of the  $k^{\text{th}}$  study variable taken from a sample  $s$  selected from a finite population,  $U = (1, 2, \dots, N)$ ,  $\phi_k$  is the inclusion probability given by  $\phi_k = p_r (k \in s)$ , and  $\gamma_k$  is the value of the  $k^{\text{th}}$  respondent in the sample selected  $s$ . The estimator  $\hat{y}_{HT}$  adjusts the weights by an unbiased estimator  $\hat{p}_{HT}$ , of the response probabilities of the population mean, as shown in the following equation:

$$\bar{P}_N = N^{-1} \sum_{i=1}^N p_i. \quad (3)$$

An approximate bias of the estimator  $\hat{y}_{HT}$  is thus given as

$$\text{Bias}(\hat{y}_{HT}) = E(\hat{y}_{HT}) - \bar{Y} = \bar{Y}^* - \bar{Y} = \hat{P}_N^{-1} C_p Y, \quad (4)$$

where

$$\bar{Y}^* = N^{-1} P_N^{-1} \sum_{i=1}^N p_i y_i, \quad (5)$$

$$C_p Y = N^{-1} \sum_{i=1}^N (p_i - \bar{P}_N) (y_i - \bar{Y}).$$

If  $C_p Y$  tends to zero, the bias would be minimal, for more details see [13]. The adjusted Horvitz–Thompson estimator,  $\hat{y}_{HT}$ , is an illustration of reweighting measurements of respondents without using auxiliary information. However, in this paper, auxiliary information is used in the estimation procedure. To compensate for nonresponse, weights obtained from transformed data are used.

## 3. Finite Population Mean Estimator Proposed

Suppose a finite population of size  $N$  consists of  $M$  clusters having  $N_j$  elements in the  $i^{\text{th}}$  cluster.  $y_{ij}$  is taken to represent the survey value of the study variable,  $Y_{ij}$ , for the  $j^{\text{th}}$  unit in the  $i^{\text{th}}$  cluster, for  $i = 1, \dots, M; j = 1, \dots, N$ , for details see [14]. Let the mean of the finite population to be estimated be given by  $\bar{Y} = (1/N) \sum_{i=1}^M \sum_{j=1}^{N_j} Y_{ij}$ . Data are generated by a regression model used in [15, 16] and more recently in [17]; the model is given as

$$\hat{Y}_{ij} = m(\hat{x}_{ij}) + \hat{e}_{ij}, \quad (6)$$

where  $m(\cdot)$  is a function of the auxiliary data having continuous derivatives and  $\hat{e}_{ij}$  is a residual variable having a mean of zero and a nonnegative variance. Auxiliary information is assumed to be known in this study. To predict nonresponse values in the study variable, the following estimator due to transformed data is proposed:

$$m(\hat{x}_{ij}) = \frac{1}{mnb} \sum_{i=1}^n \sum_{j=1}^m \left\{ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right\}. \quad (7)$$

The original data  $X_{ij}$  are first transformed to  $g(X_{ij}), i = 1, 2, \dots, n, j = 1, 2, \dots, m$ , where  $g(\cdot)$  is a nonnegative, continuous, and monotonically increasing function from  $[0, \infty)$  to  $[0, \infty)$ . The transformed data are then reflected  $g(X_{11}), \dots, g(X_{mn})$  around the origin to obtain  $-g(X_{11}), \dots, -g(X_{mn})$ , for details see [18]. More recently, Bii et al. [14] used the same procedure while developing boundary bias correction under nonresponse. However, the weights resulting from the current study are easier to implement and give better estimates than those in [14]. Hence, utilizing the transformed sample data  $(g(X_{ij}), -g(X_{ij}), i = 1, \dots, n; j = 1, \dots, m)$ , the population mean estimator is defined as

$$\hat{\bar{Y}} = \frac{1}{M} \left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \sum_{i=n+1}^N \sum_{j=m+1}^M \hat{y}_{ij} \right\}, \quad (8)$$

where  $\hat{y}_{ij}$  estimates the nonresponse units. Similarly,  $\hat{y}_{ij}$  can be represented as

$$\hat{y}_{ij} = m(\hat{x}_{ij}) = \sum_i \sum_j w_{ij}^*(x_{ij}) Y_{ij}, \quad (9)$$

where  $w_{ij}^*(x_{ij})$  are the weights obtained by transformation of the data. These weights are given as

$$w_{ij}^*(x_{ij}) = \frac{\sum_i \sum_j K((x_{ij} - g(X_{ij}))/b)}{\sum_{i=n+1}^N \sum_{j=m+1}^M \{K((x_{ij} - g(X_{ij}))/b)\}}, \quad (10)$$

so that  $m(\hat{x}_{ij})$  in equation (9) can be rewritten as

$$m(\hat{x}_{ij}) = \frac{\sum_{i=n+1}^N \sum_{j=m+1}^M \{K((x_{ij} - g(X_{ij}))/b)\} Y_{ij}}{\sum_{i=n+1}^N \sum_{j=m+1}^M \{K((x_{ij} - g(X_{ij}))/b)\}}. \quad (11)$$

Hence, using equation (11), equation (8) becomes

$$\hat{\bar{Y}} = \frac{1}{M} \left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + m(\hat{x}_{ij}) \right\}, \quad (12)$$

which is equivalent to

$$\hat{\bar{Y}} = \frac{1}{M} \left\{ \sum_{i=1}^n \sum_{j=1}^m Y_{ij} + \frac{\sum_{i=n+1}^N \sum_{j=m+1}^M \{K((x_{ij} - g(X_{ij}))/b)\} Y_{ij}}{\sum_{i=n+1}^N \sum_{j=m+1}^M \{K((x_{ij} - g(X_{ij}))/b)\}} \right\}. \quad (13)$$

The following section gives some properties of the estimator proposed.

3.1. *The Bias of the Estimator.* Assumptions in [18] are used in deriving the bias and the variance of the estimator proposed. More recently, these assumptions have also been used in [14] in boundary bias correction in presence of nonresponse.

$m''(\hat{x}_{ij})$  and  $g''(\hat{x}_{ij})$  are assumed to exist and are continuous; given  $g(x_{ij}) = x_{ij} + dx_{ij}^2 + Ad^2x_{ij}^3$  such that  $A > 0$  and  $d = (m'(0)/m(0))$ ;  $g^{-1}(0) = 0$ ,  $g'(0) = 1$ ,  $g''(0) = (2m'(0)/m(0))$ ,  $m^{(0)} = m$ , and  $g^{(0)} = g$ , where  $g^{-1}$  is the inverse function of  $g$ , whereas  $m^{(i)}$  and  $g^{(i)}$  are the  $i^{\text{th}}$  derivatives of  $m$  and  $g$ , respectively, for  $i \geq 0$ , where  $b$  is the bandwidth such that  $-\infty \leq w \leq a$ , where  $w = ((x_{ij} - g(x_{ij}))/b)$ . Furthermore, let  $x = ab$ , where  $0 \leq a \leq 1$  such that  $-1 \leq w \leq a$  for  $a \in [0, 1]$ . Besides, a kernel function  $K$  is assumed to be nonnegative and symmetric function with support  $[-1, 1]$  such that  $\int K(w)dw = 1$ ,  $\int wK(w)dw = 0$ , and  $0 < \int w^2k(w)dw < \infty$ . From these assumptions, the following equation is thus obtained:

$$\begin{aligned} E(m(\hat{x}_{ij})) &= \frac{1}{b} E \left[ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right] \\ &= m(x_{ij}) \int_{-1}^a K(w)dw - bm'(x_{ij}) \int_{-1}^a wK(w)dw \\ &\quad + \frac{m''(x_{ij})b^2}{2} \int_{-1}^a w^2K(w)dw + o(b^2). \end{aligned} \quad (14)$$

Expanding equation (14) leads to

$$\begin{aligned} E(m(\hat{x}_{ij})) &= m(x_{ij}) \int_{-1}^a K(w)dw + m(x_{ij}) \int_a^1 K(w)dw + \\ &\quad b \left\{ -m'(x_{ij}) \int_{-1}^a wK(w)dw + \int_a^1 (w-a)K(w)dw \right\} \\ &\quad + \frac{b^2}{2} m''(x_{ij}) \int_{-1}^a w^2K(w)dw + o(b^2). \end{aligned} \quad (15)$$

Simplifying equation (15) reduces to

$$\begin{aligned} E(m(\hat{x}_{ij})) &= m(x_{ij}) + b \int_a^1 (w-a)K(w)dw [2m'(x_{ij}) \\ &\quad - g''(x_{ij})m(x_{ij})] \\ &\quad + \frac{b^2}{2} m''(x_{ij}) \int_{-1}^1 w^2K(w)dw + o(b^2). \end{aligned} \quad (16)$$

Equation (16) is equivalent to

$$\begin{aligned} E(m(\hat{x}_{ij})) - m(x_{ij}) &= b \int_a^1 (w-a)K(w)dw [2m'(x_{ij}) \\ &\quad - g''(x_{ij})m(x_{ij})] \\ &\quad + \frac{b^2}{2} m''(x_{ij}) \int_{-1}^1 w^2K(w)dw + o(b^2). \end{aligned} \quad (17)$$

It is observed that  $E(m(\hat{x}_{ij}))$  is approximately equal to  $m(x_{ij})$  as  $b \rightarrow 0$  and  $mn \rightarrow \infty$  for all  $x_{ij}$ . Hence, the proposed estimator is asymptotically unbiased.

3.2. *Asymptotic Variance of the Estimator Proposed.* This estimator suggested in this study has variance given as

$$\begin{aligned} \text{Var}E(m(\hat{x}_{ij})) &= \frac{1}{(mn)^2 b^2} \text{Var} \left\{ \sum_{i=n+1}^N \sum_{j=m+1}^M K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right\}, \\ \text{Var}(m(\hat{x}_{ij})) &= \frac{1}{(mn)^2 b^2} \sum_{i=n+1}^N \sum_{j=m+1}^M E \left\{ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) - E \left[ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right] \right\}^2, \end{aligned} \tag{18}$$

which can similarly be expressed as

$$\begin{aligned} \text{Var}(m(\hat{x}_{ij})) &= \frac{(Mn - mn)^2}{(mn)^2 b^2} E \left[ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right]^2 \\ &\quad - \frac{(Mn - mn)^2}{(mn)^2 b^2} \left\{ E \left[ K \left( \frac{x_{ij} - g(X_{ij})}{b} \right) \right] \right\}^2. \end{aligned} \tag{19}$$

Using Taylor's series, the asymptotic expansion and simplification of the variance becomes

$$\text{Var}(m(\hat{x}_{ij})) = \frac{(Mn - mn)^2}{mnb} \left[ \int_{-1}^1 K(w)^2 dw + 2 \int_{-1}^a K(w)K(2a - w)dw \right] + o\left(\frac{1}{mnb}\right). \tag{20}$$

As  $mn \rightarrow \infty$ , the bandwidth  $b \rightarrow 0$ , and hence,  $mnb \rightarrow \infty$ . Hence, the variance,  $\text{Var}(m(\hat{x}_{ij}))$ , decreases in  $mnb$ . Thus, the larger the sample size, the smaller the variance.

3.3. *Mean Squared Error of the Estimator Proposed.* The mean squared error combines the variance and the squared bias terms of the estimator, that is,

$$\text{MSE}(m(\hat{x}_{ij})) = \text{Var}(m(\hat{x}_{ij})) + \text{Bias}^2(m(\hat{x}_{ij})). \tag{21}$$

Therefore, combining equations (17) and (19) lead to

$$\begin{aligned} \text{MSE}(m(\hat{x}_{ij})) &= \frac{(Mn - mn)^2}{mnb} \left[ \int_{-1}^1 K(w)^2 dw + 2 \int_{-1}^a K(w)K(2a - w)dw \right] + o\left(\frac{1}{mnb}\right) \\ &\quad + \left[ b \int_a^1 (w - a)K(w)dw \times [2m'(x_{ij}) - g''(x_{ij})m(x_{ij})] + \frac{b^2}{2} m''(x_{ij}) \int_{-1}^1 w^2 K(w)dw + o(b^2) \right]^2. \end{aligned} \tag{22}$$

As can be observed in equation (22), the mean squared error approaches zero as  $mn \rightarrow \infty$ ; this is true as the bandwidth becomes sufficiently smaller, that is, as  $b \rightarrow 0$ , in which  $m$  represents the sampled clusters, while  $n$  is the size of each sampled cluster. It is also noted that there is a trade-off between the variance and the bias terms of the estimator since as the bandwidth decreases, the variance increases, but the bias decreases. Optimal bandwidth ought to be developed to solve the bias-variance trade-off in the estimation process, see for details [19]. The next section presents a simulation study conducted to compare the performance of the finite population mean estimator suggested in this paper with those that exist in literature.

#### 4. Simulation Study

The simulation experiment was done using R code. The mean functions in [20] were used for simulation. Table 1 presents the different mean functions used for data simulation.

4.1. *Mean Functions of  $m(\hat{x}_{ij})$  Simulated.* The following steps were followed in data simulation:

- (1) The auxiliary data  $X_{ij}$  were generated as identical and independent uniform random variables on  $(0, 1)$ . The values of the residual term  $e_{ij}$  were

TABLE 1: Mean functions simulated.

Mean function	Equation
Bump	$m_1(\hat{x}_{ij}) = 1 + 2(x - 0.5) + \exp\{-200(x - 0.5)^2\}$
Exponential	$m_2(\hat{x}_{ij}) = \exp(-8x)$
Jump	$m_3(\hat{x}_{ij}) = 1 + 2(x - 0.5)I_{x \leq 0.65} + 0.65I_{x \geq 0.65}$
Linear	$m_4(\hat{x}_{ij}) = 1 + 2(x - 0.5)$
Quadratic	$m_5(\hat{x}_{ij}) = 1 + 2(x - 0.5)^2$
Sine	$m_6(\hat{x}_{ij}) = 2 + \sin(2\pi x)$

TABLE 2: Results of the bias.

Estimator	Estimator proposed	Nadaraya–Watson	Modified Nadaraya–Watson
Bump	-0.18757	-0.51888	-0.25989
Exponential	-0.01228	0.58383	0.29008
Jump	-0.2537	-0.30344	-0.15825
Linear	-0.16800	-0.40750	-0.20447
Quadratic	0.03000	0.07000	-0.17728
Sine	-0.29100	-1.17090	-0.5999

TABLE 3: Summary results of the mean squared error values.

Estimator	Estimator proposed	Nadaraya–Watson	Modified Nadaraya–Watson
Bump	0.06229	0.32087	0.08046
Exponential	0.00314	1.16734	0.29075
Jump	0.08790	0.94584	0.23835
Linear	0.04900	0.19020	0.04782
Quadratic	0.04400	0.56210	0.14017
Sine	0.13700	1.41030	0.36981

TABLE 4: Results of the confidence interval lengths.

Estimator	Estimator proposed	Nadaraya–Watson	Modified Nadaraya–Watson
Bump	0.97482	2.22051	1.91380
Exponential	0.21954	4.23531	2.11372
Jump	1.09441	3.81237	1.91380
Linear	0.92700	1.70900	0.85730
Quadratic	0.77000	3.0000	1.46763
Sine	1.52700	4.89000	2.54520

obtained from a normal distribution having mean 0 and variance 1. The population was made up of 80 clusters.

- (2) A sample of  $m_i = 20$  clusters was selected in stage one following the simple random sampling procedure with replacement.
- (3) From every cluster,  $i$ , picked in stage two, a sample  $m_{ij}, j = 1, 2, 3, \dots, 200$ , from  $j = 1, 2, 3, \dots, m_k, \dots, 2000$  was chosen from a total of  $M_k = 2000$  observations by a simple random sampling technique with replacement. Consider the study variable  $Y_{ij}, j = 1, 2, 3, \dots, m_k, \dots, M_k$ , that are given for the respondents in the sample only.
- (4) Using the auxiliary data,  $x_{ij}$ , for  $j = 1, 2, 3, \dots, m_k$ , nonresponse data were obtained from the regression equation  $\hat{Y}_{ij} = m(\hat{x}_{ij}) + \hat{e}_{ij}$  using simple random

sampling with replacement in the  $i^{\text{th}}$  cluster, where  $m(\hat{x}_{ij})$  is a function of auxiliary data obtained from the different mean functions given in Table 1;  $\hat{e}_{ij} \sim N(0, 1)$ , while  $K(u) \sim U[0, 1]$ .

- (5) This procedure was replicated to obtain the mean estimators,  $\hat{\hat{Y}}_{i1}, \dots, \hat{\hat{Y}}_{in}$ , in the  $i^{\text{th}}$  cluster.
- (6) At 95% level, confidence intervals were developed for the population mean estimators,  $\hat{\hat{Y}}_i, i = 1, 2, 3$ , which corresponds to the estimator proposed, Nadaraya–Watson estimator in [21, 22], and the modified Nadaraya–Watson estimator in [23], respectively.
- (7) A Gaussian kernel together with a locally adaptive bandwidth in [19] was used in the simulation of data.

The results are discussed in the following section.

**4.2. Simulation Results.** The simulated data for the estimator proposed, the estimator in [21, 22], and the estimator in [23] are given in Tables 2–4.

Table 2 presents a summary of the results of bias simulated from the mean functions in [20], as shown in Table 1. Negative values imply underestimation, while positive values of the bias indicate overestimation by the different estimators considered. The proposed estimator has got smaller values of the bias compared to the rest of the estimators considered as can be seen in Table 2 for all the mean functions simulated. For the exponential and quadratic mean functions, Nadaraya–Watson estimator overestimates the finite population mean, while the proposed estimator only overestimates the finite population mean for the quadratic mean function. Modified Nadaraya–Watson estimator underestimates the finite population mean in all the mean functions simulated except for exponential function. Generally, from Table 2, the proposed estimator has got smaller values of the bias in all the mean functions simulated.

The mean squared error values presented in Table 3 were generated using different mean functions, as indicated in Table 1. It can be noted that the estimator proposed has got smaller mean squared error values than those of the estimator in [21, 22] and the modified Nadaraya–Watson estimator in [23]. Nadaraya–Watson estimator has the largest mean squared error values than any other estimator considered. Comparing the mean squared error values for these three estimators shows that the estimator proposed is better than the rest of the estimators considered in this paper.

Confidence intervals are normally constructed around point estimators to provide properly calibrated measures of variability associated with estimators of population parameters of interest. In this paper, confidence intervals were obtained at 95% for the finite population mean estimators. Shorter confidence interval lengths mean the estimator is asymptotically equal to the true population parameter being predicted. From the results in Table 4, the estimator suggested in this paper is observed to have tighter confidence interval lengths than the other estimators considered. This means at 95% coverage rates, the estimator proposed is better than the Nadaraya–Watson and the improved Nadaraya–Watson estimators.

## 5. Conclusion

The estimator proposed in this paper is noted to be more desirable than the estimator in [21, 22] and the modified Nadaraya–Watson estimator in [23]. In Table 1, smaller bias values are observed for the proposed estimator for all the mean functions simulated compared to Nadaraya–Watson and the improved Nadaraya–Watson estimators. The mean squared error values in Table 2 indicate that the estimator proposed does much better than the rest considered in this study. Tighter confidence interval lengths can also be observed for the proposed estimator than the rest of the estimators, as given in Table 3. Hence, the proposed estimator provides a better estimation of the mean of a finite population compared to estimators in [21, 22] and that in [23].

The sampling procedure used in this study and the finite population mean estimator derived can be used to estimate the average health insurance coverage in a given population.

## Data Availability

To support the theoretical findings, data were generated using R statistical package.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] C. Fife-Schaw, “Surveys and sampling issues,” *Research Methods in Psychology*, vol. 2, pp. 88–104, 2000.
- [2] L. P. Lago and R. G. Clark, “Imputation of household survey data using linear mixed models,” *Australian & New Zealand Journal of Statistics*, vol. 57, no. 2, pp. 169–187, 2015.
- [3] S. Van Buuren, H. C. Boshuizen, and D. L. Knook, “Multiple imputation of missing blood pressure covariates in survival analysis,” *Statistics in Medicine*, vol. 18, no. 6, pp. 681–694, 1999.
- [4] M. Ismail, M. Q. Shahbaz, and M. Hanif, “A general class of estimator of population mean in presence of non-response,” *Pakistan Journal of Statistics*, vol. 27, no. 4, pp. 467–476, 2011.
- [5] R. R. Andridge and R. J. A. Little, “A review of hot deck imputation for survey non-response,” *International Statistical Review*, vol. 78, no. 1, pp. 40–64, 2010.
- [6] K. Y. Liang and S. L. Zeger, “Regression analysis for correlated data,” *Annual Review of Public Health*, vol. 14, no. 1, pp. 43–68, 1993.
- [7] A. Sanaullah, M. Noor-ul-Amin, M. Hanif, and C. Kadilar, “Generalized exponential chain ratio and chain product estimators under stratified two-phase random sampling for non-response,” *International Journal of Applied Mathematics & Statistics*, vol. 55, no. 2, pp. 57–79, 2016.
- [8] M. H. Hansen and W. N. Hurwitz, “The problem of non-response in sample surveys,” *Journal of the American Statistical Association*, vol. 41, no. 236, pp. 517–529, 1946.
- [9] A. Javaid, M. Noor-ul-Amin, and M. Hanif, “Modified ratio estimator in systematic random sampling under non-response,” *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, vol. 89, no. 4, pp. 817–825, 2019.
- [10] G. R. Pike, “Using weighting adjustments to compensate for survey nonresponse,” *Research in Higher Education*, vol. 49, no. 2, pp. 153–171, 2008.
- [11] J. G. Bethlehem, “Reduction of nonresponse bias through regression estimation,” *Journal of Official Statistics*, vol. 4, no. 3, pp. 251–260, 1988.
- [12] J. K. Kim and J. J. Kim, “Nonresponse weighting adjustment using estimated response probability,” *Canadian Journal of Statistics*, vol. 35, no. 4, pp. 501–514, 2007.
- [13] C. E. Särndal, “On  $\pi$ -inverse weighting versus best linear unbiased weighting in probability sampling,” *Biometrika*, vol. 67, no. 3, pp. 639–650, 1980.
- [14] N. K. Bii, C. O. Onyango, and J. Odhiambo, “Boundary bias correction using weighting method in presence of non-response in two-stage cluster sampling,” *Journal of Probability and Statistics*, vol. 2019, Article ID 6812795, 8 pages, 2019.
- [15] C. O. Onyango, R. O. Otieno, and G. O. Orwa, “Generalised model based confidence intervals in two stage cluster

- sampling,” *Pakistan Journal of Statistics and Operation Research*, vol. 6, no. 2, 2010.
- [16] C. Ouma and C. Wafula, “Bootstrap confidence intervals for model-based surveys,” *East African Journal of Statistics*, vol. 1, no. 1, pp. 84–90, 2005.
- [17] N. K. Bii, C. O. Onyango, and J. Odhiambo, “Estimating a finite population mean under random non-response in two stage cluster sampling with replacement,” *Open Journal of Statistics*, vol. 7, no. 5, pp. 834–848, 2017.
- [18] A. Cowling and P. Hall, “On pseudodata methods for removing boundary effects in kernel density estimation,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 3, pp. 551–563, 1996.
- [19] C. R. Loader, “Bandwidth selection: classical or plug-in?” *The Annals of Statistics*, vol. 27, no. 2, pp. 415–438, 1999.
- [20] J.-Y. Kim, F. J. Breidt, and J. D. Opsomer, “Nonparametric regression estimation of finite population totals under two-stage sampling,” Technical report, Colorado State University, Fort Collins, CO, USA, 2003.
- [21] E. A. Nadaraya, “On estimating regression,” *Theory of Probability & Its Applications*, vol. 9, no. 1, pp. 141–142, 1964.
- [22] G. S. Watson, “Smooth regression analysis,” *Sankhya: The Indian Journal of Statistics, Series A*, vol. 26, no. 4, pp. 359–372, 1964.
- [23] S. Demir and Ö. Toktamiş, “On the adaptive nadaraya-watson kernel regression estimators,” *Hacettepe Journal of Mathematics and Statistics*, vol. 39, no. 3, 2010.