Hindawi
Journal of Probability and Statistics
Volume 2020, Article ID 1693612, 13 pages
https://doi.org/10.1155/2020/1693612

Research Article

On Estimation of Distribution Function Using Dual Auxiliary Information under Nonresponse Using Simple Random Sampling

Saddam Hussain,[1] **Mi Zichuan**,[1] **Sardar Hussain**,[2] **Anum Iftikhar**,[1] **Muhammad Asif**,[3] **Sohail Akhtar**,[4] **and Sohaib Ahmad**[2]

[1]*Department of Statistics, Shanxi University of Finance and Economics, Taiyuan, China*
[2]*Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan*
[3]*Department of Statistics University of Malakand, Chakdara, Pakistan*
[4]*Department of Statistics Government College University Lahore, Lahore, Pakistan*

Correspondence should be addressed to Saddam Hussain; saddamhussain.stat885@gmail.com

Received 12 June 2020; Revised 18 July 2020; Accepted 1 September 2020; Published 26 September 2020

Academic Editor: Marek T. Malinowski

In this paper, we proposed two new families of estimators using the supplementary information on the auxiliary variable and exponential function for the population distribution functions in case of nonresponse under simple random sampling. The estimations are done in two nonresponse scenarios. These are nonresponse on study variable and nonresponse on both study and auxiliary variables. As we have highlighted above that two new families of estimators are proposed, in the first family, the mean was used, while in the second family, ranks were used as auxiliary variables. Expression of biases and mean squared error of the proposed and existing estimators are obtained up to the first order of approximation. The performances of the proposed and existing estimators are compared theoretically. On these theoretical comparisons, we demonstrate that the proposed families of estimators are better in performance than the existing estimators available in the literature, under the obtained conditions. Furthermore, these theoretical findings are braced numerically by an empirical study offering the proposed relative efficiencies of the proposed families of estimators.

1. Introduction

It is a well-known phenomenon that the known auxiliary information in the study of sample survey gives us an efficient estimate of population parameters, i.e., the population mean and population distribution function, under some essential conditions. This information (auxiliary) may be used for drawing a random sample using SRSWR or SRSWOR. Also, simple random sampling can be improved using the following sampling methods.

Stratification, systematic, nonresponse sampling, and probability proportional sampling schemes are used for estimating the population parameter. Auxiliary information gives us some sort of techniques by means of the ratio, product, regression, and other methods. In a practical situation, one of the important issues in surveys is that it suffers from nonresponse. Nonresponse is a common problem which may crawl with sampling survey. Nonresponse has many ways of occurrence. Examples are linguistic problems, illness, nonresponse, nonacceptance, process of return address misguided, and capture by another person. Research has labelled that various types of nonresponse may have different effects on estimators. A lot of work has been done on the estimation of population mean under nonresponse to control the nonresponse bias and to increase the efficiency of the estimators by different authors. The problem of nonresponse in sample surveys is more common and more prevalent in mail surveys than in special interview surveys. Hansen and Hurwitz [1] assumed that a part of sample of earlier nonrespondents to be recommunicated with a more expensive system; they attempted the first effort by mail questionnaire and performed the second attempt by a personal interview. However, Hansen and Hurwitz [1] have not used any kind of supplementary information to increase the efficiency of the estimator. For the first time, the author of [2] used the auxiliary information for estimating the population mean. Cochran [3] used the auxiliary information for estimating the population mean under nonresponse. Then, work on nonresponse extended by many authors (cf., [4–7]) recommends various types of estimators for estimation of population mean and distribution function using the secondary information under nonresponse. Okafor and Lee [8] presented ratio and regression estimation with partial sampling of the nonrespondents for estimating the population mean. Furthermore, the authors of [9, 10] proposed estimators for estimating population mean using multiauxiliary information in different directions and Zhao et al. [11] used the idea of robust estimation of the distribution function and quantiles with nonignorance missing data.

Also, for estimating population mean under the two-phase sampling strategy in the presence of nonresponse, the authors of [12–15] have made significant contributions. Diana and Perri [16] suggested a class of estimators in two-phase sampling with subsampling of nonrespondents in estimating the finite population mean. In this paper, we introduce the use of sample distribution functions of the study variable and auxiliary variable along with the mean of the auxiliary variable and also the ranks of the auxiliary variable for estimating the population distribution function.

Extensive literature has been published on estimation of population mean under nonresponse; however, no effort has been dedicated to the development of efficient methods for population cumulative distribution function. In survey sampling, the statisticians are often interested in proportion size of the study variable, i.e., proportion of units in population with values less than or equal to a specified value of y; for instance, we may be interested to know the proportion of the population in which 31% or more people are educated.

Motivated by $\hat{F}_{R,D}(y)$, $\hat{F}_S(y)$, and average of $\hat{F}_{\mathrm{BT},R}(y)$ and $\hat{F}_{\mathrm{BT},P}(y)$, two new families of estimators are proposed for estimating distribution function in the presence of nonresponse. By numerical results, we will show that the proposed family of estimators is more precise than the existing estimators.

We planned the paper as follows: In Section 2, some notations are introduced. In Section 3, the existing estimators are reviewed briefly. Two new families of estimators are introduced in Section 4, respectively. The existing and proposed estimators are compared (theoretically and numerically) in Sections 5 and 6. In Section 7, the concluding remarks of the paper are discussed.

2. Notations

Consider a finite population $\Omega = \{V_1, V_2, \ldots, V_N\}$ of N distinct units, which is partitioned into respondents $\Omega_1 = \{V_1, V_2, \ldots, V_{N1}\}$ and nonrespondents $\Omega_2 = \{V_{N1+1}, V_{N1+2}, \ldots, V_N\}$ groups with sizes N_1 and N_2, respectively, for estimating the CDF, where $N = N_1 + N_2$. A sample of size n has been drawn from this population by simple random sampling (SRSWOR), out of which n_1 units respond and $n_2 = n - n_1$ do not respond. It is assumed that the sample size n_1 is drawn from the response group of Ω_1 and n_2 is drawn from the nonresponse group of Ω_2. Moreover, a sample of size $r = n_2/k\,(k > 1)$ is drawn by simple random sampling (SRSWOR) from n_2, and this time response is obtained from all r units. Let Y and X be the study and auxiliary variables, respectively. Let Z be used for the ranks of the X and $I(Y \le y)$ and $I(X \le x)$ be the indicator variables based on Y and X. Furthermore, $F(y) = \sum_{i=1}^{N} I(Y_i \le y)/N$ and $\hat{F}(y) = \sum_{i=1}^{n} I(Y_i \le y)/n$ and $F(x) = \sum_{i=1}^{N} I(X_i \le x)/N$ and $\hat{F}(x) = \sum_{i=1}^{n} I(X_i \le x)/n$ are the population and sample distribution functions of Y and X, respectively. Similarly, let $\bar{X} = \sum_{i=1}^{N} X_i/N$ and $\hat{\bar{\bar{X}}} = \sum_{i=1}^{n} X_i/n$ and $\overline{Z} = \sum_{i=1}^{N} Z_i/N$ and $\hat{\bar{\bar{Z}}} = \sum_{i=1}^{n} Z_i/n$ be the population and sample means of X and Z,respectively.Furthermore,$F_2(y) = \sum_{i=N_1+1}^{N} I(Y_i \le y)/N_2$ and $F_2(x) = \sum_{i=N_1+1}^{N} I(X_i \le x)/N_2$ are the population distribution functions of $I(Y \le y)$ and $I(X \le x)$ for the nonresponse group and $\bar{X}_2 = \sum_{i=N_1+1}^{N} X_i/N_2$ and $\overline{Z}_2 = \sum_{i=N_1+1}^{N} Z_i/N_2$ are the population means of X and Z for the nonresponse group, respectively.

Here, ($x = \bar{X}$ and $\Theta_2(x)$) and ($y = \overline{Y}$ and $\Theta_2(y)$), where $\bar{X}$ and $\overline{Y}$ are the population means of $X(Y)$. Similarly, $\Theta_2(x)$ and $\Theta_2(y)$ are the population second quartiles of $X(Y)$, respectively.

To obtain the bias and MSE of the proposed estimator, we consider the following error terms. Let

$$
\begin{aligned}
e_1^* &= \frac{\hat{F}_H^*(y) - F(y)}{F(y)},\\
e_2^* &= \frac{\hat{F}_H^*(x) - F(x)}{F(x)},\\
e_3^* &= \frac{\hat{\bar{\bar{X}}}_H^* - \bar{X}}{\bar{X}},\\
e_4^* &= \frac{\hat{\bar{\bar{Z}}}_H^* - \overline{Z}}{\overline{Z}},\\
e_2 &= \frac{\hat{F}_H(x) - F(x)}{F(x)},\\
e_3 &= \frac{\hat{\bar{\bar{X}}}_H - \bar{X}}{\bar{X}},\\
e_4 &= \frac{\hat{\bar{\bar{Z}}}_H - \overline{Z}}{\overline{Z}}.
\end{aligned}
\tag{1}
$$

Here, $\hat{F}_H^*(y)$, $\hat{F}_H^*(x)$, and $\hat{\bar{\bar{X}}}^*$ and $\hat{\bar{\bar{Z}}}^*$ are the notations used for CDFs, mean, and mean of ranks when there are no responses on both study and auxiliary variables. And, $\hat{F}_H(x)$, $\hat{\bar{\bar{X}}}$, and $\hat{\bar{\bar{Z}}}$ are the notations used for CDF, mean, and mean of ranks when there are no responses on only auxiliary variable, shown in Table 1.

Table 1: Estimators, variances, covariances, and correlation under nonresponse situations.

	Estimator		
Situation	=	I	II
$\widehat{F}(y)$	=	$\widehat{F}_H^*(y)$	$\widehat{F}_H^*(y)$
$\widehat{F}(x)$	=	$\widehat{F}_H^*(x)$	$\widehat{F}_H(x)$
$\widehat{\overline{X}}$	=	$\widehat{\overline{X}}_H^*$	$\widehat{\overline{X}}_H$
$\widehat{\overline{Z}}$	=	$\widehat{\overline{Z}}_H^*$	$\widehat{\overline{Z}}_H$
	Variance/covariance		
Θ_{rstu}	=	ψ_{rstu}	Ψ_{rstu}
Θ_{2000}	=	ψ_{2000}	Ψ_{2000}
Θ_{0200}	=	ψ_{0200}	Ψ_{0200}
Θ_{0020}	=	ψ_{0020}	Ψ_{0020}
Θ_{0002}	=	ψ_{0002}	Ψ_{0002}
Θ_{1100}	=	ψ_{1100}	Ψ_{1100}
Θ_{1010}	=	ψ_{1010}	Ψ_{1010}
Θ_{1001}	=	ψ_{1001}	Ψ_{1001}
Θ_{0110}	=	ψ_{0110}	Ψ_{0110}
Θ_{0101}	=	ψ_{0101}	Ψ_{0101}
	Coefficient of correlation		
ϱ_{12}	=	$\rho_{12(2)}$	ρ_{12}
ϱ_{13}	=	$\rho_{13(2)}$	ρ_{13}
ϱ_{23}	=	$\rho_{23(2)}$	ρ_{23}
ϱ_{14}	=	$\rho_{14(2)}$	ρ_{14}
ϱ_{24}	=	$\rho_{24(2)}$	ρ_{24}
	Coefficient of multiple determination		
$\mathfrak{R}^2_{1.23}$	=	$\Phi^2_{1.23}$	$\Phi^2_{1.23}$
$\mathfrak{R}^2_{1.24}$	=	$\Phi^2_{1.24}$	$\Phi^2_{1.24}$

Let $E(e_i^*) = 0$ for $i^* = 1,2,3,4$ and $E(e_i) = 0$ for $i = 2,3,4$, where $E(\cdot)$ is the mathematical expectation of $(\cdot)$. Let

$$\Psi_{\text{rstu}} = E\left[e_1^r e_2^s e_3^t e_4^u\right], \quad \psi_{\text{rstu}} = E\left[e_1^{*r} e_2^{*s} e_3^{*t} e_4^{*u}\right], \tag{2}$$

where $r,s,t,u = 1,2,3,4$. Here,

$$\begin{aligned}
E\left(e_1^{*2}\right) &= \lambda C_1^2 + \lambda_2 C_{1(2)}^2 = \psi_{2000}^2,\\
E\left(e_2^{*2}\right) &= \lambda C_2^2 + \lambda_2 C_{2(2)}^2 = \psi_{0200}^2,\\
E\left(e_3^{*2}\right) &= \lambda C_3^2 + \lambda_2 C_{3(2)}^2 = \psi_{0020}^2,\\
E\left(e_4^{*2}\right) &= \lambda C_4^2 + \lambda_2 C_{4(2)}^2 = \psi_{0002}^2,\\
E\left(e_2^2\right) &= \lambda C_2^2 = \Psi_{0200},\\
E\left(e_3^2\right) &= \lambda C_3^2 = \Psi_{0020},\\
E\left(e_4^2\right) &= \lambda C_4^2 = \Psi_{0002},\\
E\left(e_1^* e_2^*\right) &= \lambda \rho_{12} C_1 C_2 + \lambda_2 \rho_{12(2)} C_{1(2)} C_{2(2)} = \psi_{1100},\\
E\left(e_1^* e_3^*\right) &= \lambda \rho_{13} C_1 C_3 + \lambda_2 \rho_{13(2)} C_{1(2)} C_{3(2)} = \psi_{1010},\\
E\left(e_1^* e_4^*\right) &= \lambda \rho_{14} C_1 C_4 + \lambda_2 \rho_{14(2)} C_{1(2)} C_{4(2)} = \psi_{1001},\\
E\left(e_2^* e_3^*\right) &= \lambda \rho_{23} C_2 C_3 + \lambda_2 \rho_{23(2)} C_{2(2)} C_{3(2)} = \psi_{0110},\\
E\left(e_2^* e_4^*\right) &= \lambda \rho_{24} C_2 C_4 + \lambda_2 \rho_{24(2)} C_{2(2)} C_{4(2)} = \psi_{0101},\\
E\left(e_1^* e_2\right) &= \lambda \rho_{12} C_1 C_2 = \Psi_{1100},\\
E\left(e_1^* e_3\right) &= \lambda \rho_{13} C_1 C_3 = \Psi_{1010},\\
E\left(e_1^* e_4\right) &= \lambda \rho_{14} C_1 C_4 = \Psi_{1001},\\
E\left(e_2 e_3\right) &= \lambda \rho_{23} C_2 C_3 = \Psi_{0110},\\
E\left(e_2 e_4\right) &= \lambda \rho_{24} C_2 C_4 = \Psi_{0101}.
\end{aligned} \tag{3}$$

Here,

$$\varrho_{1.23}^2 = \left(\frac{\psi_{1100}^2 \psi_{0020} + \psi_{1010}^2 \psi_{0200} - 2\psi_{1010}\psi_{1100}\psi_{0110}}{\psi_{2000}\left(\psi_{0200}\psi_{0020} - \psi_{0110}^2\right)}\right). \tag{4}$$

where it is the coefficient of multiple determination of $I(Y \le y)$ on $I(X \le x)$ and X with situation-I. Also,

$$\varrho_{1.23}^2 = \left(\frac{\Psi_{1100}^2 \Psi_{0020} + \Psi_{1010}^2 \Psi_{0200} - 2\Psi_{1010}\Psi_{1100}\Psi_{0110}}{\psi_{2000}\left(\Psi_{0200}\Psi_{0020} - \Psi_{0110}^2\right)}\right) \tag{5}$$

is the coefficient of multiple determination of $I(Y \le y)$ on $I(X \le x)$ and X with situation-II. And,

$$\varrho_{1.24}^2 = \left(\frac{\psi_{1100}^2 \psi_{0002} + \psi_{1001}^2 \psi_{0200} - 2\psi_{1001}\psi_{1100}\psi_{0101}}{\psi_{2000}\left(\psi_{0200}\psi_{0002} - \psi_{0101}^2\right)}\right) \tag{6}$$

is the coefficient of multiple determination of $I(Y \le y)$ on $I(X \le x)$ and Z with situation-I. Finally,

$$\varrho_{1.24}^2 = \left(\frac{\Psi_{1100}^2 \Psi_{0002} + \Psi_{1001}^2 \Psi_{0200} - 2\Psi_{1001}\Psi_{1100}\Psi_{0101}}{\psi_{2000}\left(\Psi_{0200}\Psi_{0002} - \Psi_{0101}^2\right)}\right) \tag{7}$$

is the coefficient of multiple determination of $I(Y \le y)$ on $I(X \le x)$ and Z with situation-II. Here, $\lambda = (1/n1/N)$, $\lambda_2 = W_2(k-1)/n$, $S_1^2 = \sum_{i=1}^N (I(Y_i \le y) - F(y))^2/(N-1)$, $S_2^2 = \sum_{i=1}^N (I(X_i \le x) - F(x))^2/(N-1)$, $S_3^2 = \sum_{i=1}^N (X_i - \bar{X})^2/(N-1)$, and $S_4^2 = \sum_{i=1}^N (Z_i - \overline{Z})^2/(N-1)$ are the population variances of $I(Y \le y)$, $I(X \le x)$, X, and Z for the response group, respectively.

Similarly, $S_{1(2)}^2 = \sum_{i=N_1+1}^N (I(Y_i \le y) - F(y_2))^2/(N_2-1)$, $S_{2(2)}^2 = \sum_{i=N_1+1}^N (I(X_i \le x_2) - F(x_2))^2/(N_2-1)$, $S_{3(2)}^2 = \sum_{i=N_1+1}^N (X_i - \overline{X_2})^2/(N_2-1)$, and $S_{4(2)}^2 = \sum_{i=N_1+1}^N (Z_i - \overline{Z_2})^2/(N_2-1)$ are the population variances of $I(Y \le y)$, $I(X \le x)$, X, and Z for the nonresponse group, respectively.

$C_1 = S_1/F(y)$, $C_2 = S_2/F(x)$, $C_3 = S_3/\bar{X}$, and $C_4 = S_4/\overline{Z}$ are the population coefficient of variations for the response group, and $C_{1(2)} = S_{1(2)}/F(y_2)$, $C_{2(2)} = S_{2(2)}/F(x_2)$, $C_{3(2)} = S_{3(2)}/\bar{X}_2$, and $C_{4(2)} = S_{4(2)}/\overline{Z_2}$ are the population coefficient of variations for the nonresponse group.

$S_{12} = \sum_{i=1}^N \{(I(Y_i \le y) - F(y))(I(X_i \le x) - F(x))\}/(N-1)$, $S_{13} = \sum_{i=1}^N \{(I(Y_i \le y) - F(y))(X_i - \bar{X})\}/(N-1)$, $S_{23} = \sum_{i=1}^N \{(I(X_i \le x) - F(x))(X_i - \bar{X})\}/(N-1)$, $S_{14} = \sum_{i=1}^N \{(I(Y_i \le y) - F(y))(Z_i - \overline{Z})\}/(N-1)$, and $S_{24} = \sum_{i=1}^N \{(I(X_i \le x) - F(x))(Z_i - \overline{Z})\}/(N-1)$ are the population covariances for the response group.

$S_{12(2)} = \sum_{i=N_1+1}^N \{(I(Y_i \le y) - \mathrm{F}(y_2))(I(X_i \le x) - F(x_2))\}/(N_2-1)$, $S_{13(2)} = \sum_{i=N_1+1}^N \{(I(Y_i \le y) - F(y_2))(X_i - \overline{X_2})\}/(N_2-1)$, $S_{23(2)} = \sum_{i=N_1+1}^N \{(I(X_i \le x) - F(x_2))(X_i - \overline{X_2})\}/(N_2-1)$, $S_{14(2)} = \sum_{i=N_1+1}^N \{(I(Y_i \le y) - F(y_2))(Z_i - \overline{Z})\}/(N_2-1)$, and $S_{24(2)} = \sum_{i=N_1+1}^N \{(I(X_i \le x) - F(x_2))(Z_i - \overline{Z_2})\}/(N_2-1)$ are the population covariances for the nonresponse group.

Similarly, $\rho_{12} = S_{12}/S_1S_2$, $\rho_{13} = S_{13}/S_1S_3$, $\rho_{23} = S_{23}/S_2S_3$, $\rho_{14} = S_{14}/S_1S_4$, and $\rho_{24} = S_{24}/S_2S_4$ are the population correlation coefficients for the response group, respectively.

$\rho_{12(2)} = S_{12(2)}/S_{1(2)}S_{2(2)}$, $\rho_{13(2)} = S_{13(2)}/S_{1(2)}S_{3(2)}$, $\rho_{23(2)} = S_{23(2)}/S_{2(2)}S_{3(2)}$, $\rho_{14(2)} = S_{14(2)}/S_{1(2)}S_{4(2)}$, and $\rho_{24(2)} = S_{24(2)}/S_{2(2)}S_{4(2)}$ are the population correlation coefficients for the nonresponse group.

Let $F(y) = W_1F_1(y) + W_2F_2(y)$, where $W_j = N_j/N$ and $F_j(y) = \sum_{i=1}^{N_j} I(Y_i \le y)/N_j$ for $j = 1, 2$. Also, $\hat{F}_1(y) = \sum_{i=1}^{n_1} I(Y_i \le y)/n_1$ denote the sample distribution function of n_1 responding units out of n units and $\hat{F}_{2r}(y) = \sum_{i=1}^{r} I(Y_i \le y)/r$ denote the sample distribution function of r responding units out of nonresponse units.

The existing Hansen and Hurwitz [1] unbiased estimator of $F(y)$ with its variance is

$$\begin{aligned}\hat{F}_H(y) &= w_1\hat{F}_1(y) + w_2\hat{F}_{2r}(y),\\ \mathrm{Var}\left(\hat{F}_H(y)\right) &= \lambda S_1^2 + \lambda_2 S_{1(2)}^2.\end{aligned} \tag{8}$$

Similarly, the unbiased estimators for $\hat{F}_H(x)$, $\hat{\overline{X}}_H$, and $\hat{\overline{Z}}_H$ and their corresponding variances are

$$\begin{aligned}\hat{F}_H(x) &= w_1\hat{F}_1(x) + w_2\hat{F}_{2r}(x),\\ \hat{\overline{X}}_H &= w_1\hat{\overline{X}}_1 + w_2\hat{\overline{X}}_{2r},\\ \hat{\overline{Z}}_H &= w_1\hat{\overline{Z}}_1 + w_2\hat{\overline{Z}}_{2r},\\ \mathrm{Var}\left(\hat{F}_H(x)\right)&\lambda S_1^2 + \lambda_2 S_{1(2)}^2,\\ \mathrm{Var}\left(\hat{\overline{X}}_H\right) &= \lambda S_2^2 + \lambda_2 S_{2(2)}^2,\\ \mathrm{Var}\left(\hat{\overline{Z}}_H\right) &= \lambda S_3^2 + \lambda_2 S_{3(2)}^2, \text{respectively.}\end{aligned} \tag{9}$$

In practice, we use three situations, occurring under nonresponse, but here, we use two situations which mostly occur, namely, nonresponse on both the study variable and the auxiliary variable (say situation-I) and nonresponse just on study variable only (say situation-II). For notational convenience, we follow the notations given in Table 1.

3. Existing Estimators

In this section, some estimators of finite population mean exist for estimating the finite CDF under nonresponse; the biases and MSEs of these existing estimators are derived under the first order of approximation.

(1) Cochran's [17] existing ratio estimator of $F(y)$ is

$$\hat{F}_R(y) = \hat{F}(y)\left(\frac{F(x)}{\hat{F}(x)}\right). \tag{10}$$

The bias and MSE of $\hat{F}_R(y)$, to the first order of approximation, are

$$\begin{aligned}\mathrm{bias}\left(\hat{F}_R(y)\right) &\cong F(y)\left(\Theta_{0200} - \Theta_{1100}\right),\\ \mathrm{MSE}\left(\hat{F}_R(y)\right) &\cong F^2(y)\left(\Theta_{2000} + \Theta_{0200} - 2\Theta_{1100}\right).\end{aligned} \tag{11}$$

(2) Murthy's [18] existing product estimator of $F(y)$ is

$$\hat{F}_P(y) = \hat{F}(y)\left(\frac{\hat{F}(x)}{F(x)}\right). \tag{12}$$

The bias and MSE of $\hat{F}_P(y)$, to the first order of approximation, are

$$\begin{aligned}\mathrm{bias}\left(\hat{F}_P(y)\right) &= F(y)\Theta_{1100},\\ \mathrm{MSE}\left(\hat{F}_P(y)\right) &\cong F^2(y)\left(\Theta_{2000} + \Theta_{0200} + 2\Theta_{1100}\right).\end{aligned} \tag{13}$$

(3) The existing regression estimator of $F(y)$ is

$$\hat{F}_{\mathrm{Reg}}(y) = \hat{F}(y) + k\left(F(x) - \hat{F}(x)\right), \tag{14}$$

where k is an unknown constant. Here, $\hat{F}_{\mathrm{Reg}}(y)$ is an unbiased estimator of $\hat{F}(y)$. The minimum variance of $\hat{F}_{\mathrm{Reg}}(y)$ at the optimum value $k_{(\mathrm{opt})} = (F(y)\Theta_{1100})/(F(x)\Theta_{0200})$ is

$$\mathrm{Var}_{\min}\left(\hat{F}_{\mathrm{Reg}}(y)\right) = \frac{F^2(y)\left(\Theta_{2000}\Theta_{0200} - \Theta_{1100}^2\right)}{\Theta_{0200}}. \tag{15}$$

Here, (15) may be written as

$$\mathrm{Var}_{\min}\left(\hat{F}_{\mathrm{Reg}}(y)\right) = F^2(y)\Theta_{2000}\left(1 - \varrho_{12}^2\right). \tag{16}$$

(4) Rao's [19] existing difference-type estimator of $F(y)$ is

$$\hat{F}_{R,D}(y) = k_1\hat{F}(y) + k_2\left(F(x) - \hat{F}(x)\right), \tag{17}$$

where k_1 and k_2 are unknown constants. The bias and MSE of $\hat{F}_{R,D}(y)$, to the first order of approximation, are

$$\begin{aligned}\mathrm{bias}\left(\hat{F}_{R,D}(y)\right) &= F(y)\left(k_1 - 1\right),\\ \mathrm{MSE}\left(\hat{F}_{R,D}(y)\right) &\cong F^2(y) - 2k_1F^2(y) + k_1^2F^2(y)\\ &\quad + k_1^2F^2(y)\Theta_{2000} - 2k_1k_2F(y)F(x)\Theta_{1100}\\ &\quad + k_2^2F^2(x)\Theta_{0200}.\end{aligned} \tag{18}$$

The optimum values of k_1 and k_2, determined by minimizing (18), are

$$\begin{aligned}k_{1(\mathrm{opt})} &= \frac{\Theta_{0200}}{\left(\Theta_{0200}\Theta_{2000} - \Theta_{1100}^2 + \Theta_{0200}\right)},\\ k_{2(\mathrm{opt})} &= \frac{F(y)\Theta_{1100}}{F(x)\left(\Theta_{2000}\Theta_{0200} - \Theta_{1100}^2 + \Theta_{0200}\right)}.\end{aligned} \tag{19}$$

The minimum MSE of $\hat{F}_{R,D}(y)$ at the optimum values of k_1 and k_2 is

$$\text{MSE}_{\min}\left(\hat{F}_{R,D}(y)\right) = \frac{F^2(y)\left(\Theta_{2000}\Theta_{0200} - \Theta_{1100}^2\right)}{\left(\Theta_{2000}\Theta_{0200} - \Theta_{1100}^2 + \Theta_{0200}\right)}. \quad (20)$$

Here, (20) may be written as

$$\text{MSE}_{\min}\left(\hat{F}_{R,D}(y)\right) = \frac{F^2(y)\Theta_{2000}\left(1 - \varrho_{12}^2\right)}{1 + \Theta_{2000}\left(1 - \varrho_{12}^2\right)}. \quad (21)$$

(5) Grover and Kaur's [20] existing generalized class of ratio-type exponential estimator of $F(y)$ is

$$\hat{F}_{G,K}(y) = \left\{k_3\hat{F}(y) + k_4(F(x) - \hat{F}(x))\right\} \cdot \exp\left(\frac{a(F(x) - \hat{F}(x))}{a(F(x) + \hat{F}(x)) + 2b}\right), \quad (22)$$

where k_3 and k_4 are unknown constants. The bias and MSE of $\hat{F}_{G,K}(y)$, to the first order of approximation, are

$$\begin{aligned}
\text{bias}\left(\hat{F}_{G,K}(y)\right) &\cong F(y)(k_3 - 1) + \frac{3}{8}\theta^2 k_3 F(y) + \frac{1}{2}\theta k_4 F(x)\Theta_{0200} - \frac{1}{2}\theta F(y)\Theta_{1100}, \\
\text{MSE}\left(\hat{F}_{G,K}(y)\right) &\cong k_4^2 F^2(x)\Theta_{0200} + k_3^2 F^2(y)\Theta_{2000} + 2\theta k_3 k_4 F(y)F(x)\Theta_{0200} \\
&\quad - 2k_3 k_4 F(y)F(x)\Theta_{1100} + F^2(y) - 2k_3 F^2(y) + \theta k_3^2 F^2(y) \\
&\quad + k_3 F^2(y)\Theta_{1100} - \theta k_4 F(y)F(x)\Theta_{0200} - 2\theta k_3^2 F^2(y)\Theta_{1100} \\
&\quad - \frac{3}{4}\theta^2 k_3 F^2(y)\Theta_{0200} + \theta^2 k_3^2 F^2(y)\Theta_{0200}.
\end{aligned} \quad (23)$$

The optimum values of k_3 and k_4, determined by minimizing (15), are

$$\begin{aligned}
k_{3(\text{opt})} &= \frac{\Theta_{0200}\left(\theta^2\Theta_{0200} - 8\right)}{8\left(-\Theta_{2000}\Theta_{0200} + \Theta_{1100}^2 - \Theta_{0200}\right)}, \\
k_{4(\text{opt})} &= \frac{F(y)\left(\theta^3\Theta_{0200}^2 - \theta^2\Theta_{0200}\Theta_{1100} + 4\theta\Theta_{2000}\Theta_{0200} - 4\theta\Theta_{1100}^2 - 4\theta\Theta_{0200} + 8\Theta_{1100}\right)}{8F(x)\left(\Theta_{2000}\Theta_{0200} - \Theta_{1100}^2 + \Theta_{0200}\right)}.
\end{aligned} \quad (24)$$

The simplified minimum MSE of $\hat{F}_{G,K}(y)$ at the optimum values of k_3 and k_4 is

$$\text{MSE}_{\min}\left(\hat{F}_{G,K}(y)\right) \cong \frac{F^2(y)}{64}\left(64 - 16\theta^2\Theta_{0200} - \frac{\Theta_{0200}\left(-8 + \theta^2\Theta_{0200}\right)^2}{\Theta_{0200}\left(1 + \Theta_{2000}\right) - Q_{1100}^2}\right). \quad (25)$$

Here, (25) may be written as

$$\text{MSE}_{\min}\left(\hat{F}_{G,K}(y)\right) \cong \text{Var}_{\min}\left(\hat{F}_{\text{Reg}}(y)\right) - \frac{F^2(y)\left(\theta^2\Theta_{0200}^2 - 8\Theta_{1100}^2 + 8\Theta_{0200}\Theta_{2000}\right)^2}{64\Theta_{0200}^2\left\{1 + \Theta_{2000}\left(1 - \varrho_{12}^2\right)\right\}}, \quad (26)$$

which shows that $\hat{F}_{G,K}(y)$ is more precise than $\hat{F}_4(y)$.

4. Proposed Estimators

On the lines of $\hat{F}_{R,D}(y)$, $\hat{F}_S(y)$, and average of $\hat{F}_{\text{BT},R}(y)$ and $\hat{F}_{\text{BT},P}(y)$, the first proposed family of estimators for estimating $F(y)$ is given by

$$\hat{F}_{Pr_1}(y) = \left[\frac{\hat{F}(y)}{2}\left\{\exp\left(\frac{F(x)-\hat{F}(x)}{\hat{F}(x)+F(x)}\right)+\exp\left(\frac{\hat{F}(x)-F(x)}{\hat{F}(x)+F(x)}\right)\right\}+m_5\left(\frac{F(x)-\hat{F}(x)}{F(x)}\right)+m_6\hat{F}(y) + m_7\left(\frac{\bar{X}-\hat{\bar{X}}}{\bar{X}}\right)\right]\exp\left(\frac{a(F(x)-\hat{F}(x))}{a(F(x)+\hat{F}(x))+2b}\right), \quad (27)$$

where m_5, m_6, and m_7 are unknown constants and $a(\neq 0)$ and b are either two real numbers or functions of known population parameters of $I(X\leq x)$, such as ρ_{12}, β_2 (coefficient of kurtosis), and C_2.

The estimator $\hat{F}_{Pr_1}(y)$ can also be written as

$$\hat{F}_{Pr_1}(y) = \left\{F(y)(1+\xi_0)(1+m_6)-m_5\xi_1-m_7\xi_2+\frac{1}{8}\theta^2F(y)\xi_1^2\right\}\left(1-\frac{1}{2}\theta\xi_1+\frac{3}{8}\theta^2\xi_1^2+\cdots\right). \quad (28)$$

Simplifying (28) and keeping terms only up to the second power of ξs, we can write

$$\begin{aligned}\left(\hat{F}_{Pr_1}(y)-F(y)\right) \cong{}& m_6F(y)+F(y)\xi_0+m_6F(y)\xi_0-\frac{1}{2}\theta F(y)\xi_1+\frac{1}{2}\theta^2F(y)\xi_1^2\\ &-\frac{1}{2}\theta F(y)\xi_0\xi_1-m_5\xi_1+\frac{1}{2}\theta m_5\xi_1^2-\frac{1}{2}\theta m_6F(y)\xi_1\\ &+\frac{3}{8}\theta^2m_6F(y)\xi_1^2-\frac{1}{2}\theta m_6F(y)\xi_0\xi_1-m_7\xi_2+\frac{1}{2}\theta m_7\xi_1\xi_2.\end{aligned} \quad (29)$$

The bias and MSE of $\hat{F}_{Pr_1}(y)$, to the first order of approximation, respectively, are

$$\begin{aligned}\text{bias}\left(\hat{F}_{Pr_1}(y)\right) \cong{}& \frac{1}{2}\theta^2F(y)\Theta_{0200}-\frac{1}{2}\theta F(y)\Theta_{1100}+\frac{1}{2}m_5\theta\Theta_{0200}+m_6F(y)\\ &+\frac{3}{8}m_6\theta^2F(y)\Theta_{0200}-\frac{1}{2}m_6\theta F(y)\Theta_{1100}+\frac{1}{2}m_7\theta\Theta_{0110},\\ \text{MSE}\left(\hat{F}_{Pr_1}(y)\right) \cong{}& -\theta F^2(y)\Theta_{1100}+\frac{3}{2}m_6\theta^2F^2(y)\Theta_{0200}+m_6^2\theta^2F^2(y)\Theta_{0200}+m_5\theta F(y)\Theta_{0200}\\ &-2m_6m_7F(y)\Theta_{1010}+\frac{1}{4}\theta^2F^2(y)\Theta_{0200}+2m_6F^2(y)\Theta_{2000}+m_5^2\Theta_{0200}\\ &-2m_6^2\theta F^2(y)\Theta_{1100}+F^2(y)\Theta_{2000}+m_6^2F^2(y)+m_7\theta F(y)\Theta_{0110}\\ &+2m_5m_6\theta F(y)\Theta_{0200}-2m_5m_6F(y)\Theta_{1100}+2m_5m_7\Theta_{0110}\\ &-3m_6\theta F^2(y)\Theta_{1100}-2m_5F(y)\Theta_{1100}-2m_7F(y)\Theta_{1010}\\ &+m_6^2F^2(y)\Theta_{2000}+m_7^2\Theta_{0020}+2m_6m_7\theta F(y)\Theta_{0110}.\end{aligned} \quad (30)$$

The optimum values of m_5, m_6, and m_7, determined by minimizing (29), are

$$m_{5(\text{opt})} = \frac{F(y)\left[\theta^3\Theta_{0200}^{3/2}\left(\varrho_{23}^2-1\right)+\Theta_{2000}^{1/2}\left(-4+\theta^2\Theta_{0200}\right)\left(\varrho_{12}-\varrho_{23}\varrho_{13}\right)\right]+2\theta\Theta_{0200}^{1/2}\left(\varrho_{23}^2-1\right)\left\{-1+\Theta_{2000}\left(1-\varrho_{1.23}^2\right)\right\}}{8\Theta_{0200}^{1/2}\left(\varrho_{23}^2-1\right)\left\{-1+\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)\right\}},$$
$$m_{6(\text{opt})} = \frac{\left\{4\Theta_{2000}\left(-2\varrho_{12}\varrho_{13}\varrho_{23}+\varrho_{12}^2+\varrho_{13}^2+\varrho_{23}^2-1\right)+\theta^2\Theta_{0200}\left(\varrho_{23}^2-1\right)\right\}}{-4\left[\left(\varrho_{23}^2-1\right)\left\{1+\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)\right\}\right]}, \quad (31)$$
$$m_{7(\text{opt})} = \frac{F(y)\Theta_{2000}^{1/2}\left\{\theta^2\Theta_{0200}\left(-\varrho_{13}+\varrho_{12}\varrho_{23}\right)+4\left(\varrho_{13}-\varrho_{12}\varrho_{23}\right)\right\}}{-4\Theta_{0020}^{1/2}\left[\left(\varrho_{23}^2-1\right)\left\{1+\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)\right\}\right]}.$$

The simplified minimum MSE of $\hat{F}_{Pr_1}(y)$ at the optimum values of m_5, m_6, and m_7 is

$$\text{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) \cong \frac{F^2(y)\left\{16\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)-\theta^4\Theta_{0200}^2-8\theta^2\Theta_{0200}\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)\right\}}{16\left\{1+\Theta_{2000}\left(1-\mathfrak{R}_{1.23}^2\right)\right\}}, \quad (32)$$

where $\mathfrak{R}_{1.23}^2 = ((\Theta_{1100}^2\Theta_{0020}+\Theta_{1010}^2 \quad \Theta_{0200}-2\Theta_{1010}\Theta_{1100}\Theta_{0110})/(\Theta_{2000}(\Theta_{0200}\Theta_{0020}-\Theta_{0110}^2)))$.

It can be seen that $\hat{F}_{Pr_1}(y)$ is more precise than $\hat{F}_{\text{Reg}}(y)$.

On similar lines, the second proposed family of estimators for estimating $F(y)$ is given by

$$\hat{F}_{Pr_2}(y) = \left[\frac{\hat{F}(y)}{2}\left\{\exp\left(\frac{F(x)-\hat{F}(x)}{\hat{F}(x)+F(x)}\right)+\exp\left(\frac{\hat{F}(x)-F(x)}{\hat{F}(x)+F(x)}\right)\right\}+m_8\left(\frac{F(x)-\hat{F}(x)}{F(x)}\right)+m_9\hat{F}(y)\right.$$
$$\left.+m_{10}\left(\frac{\overline{Z}-\hat{\overline{Z}}}{\overline{Z}}\right)\right]\exp\left(\frac{a(F(x)-\hat{F}(x))}{a(F(x)+\hat{F}(x))+2b}\right), \quad (33)$$

where m_8, m_9, and m_{10} are unknown constants and $a(\neq 0)$ and b are either two real numbers or functions of known population parameters of $I(X\leq x)$, such as ρ_{12}, β_2 (coefficient of kurtosis), and C_2.

The estimator $\hat{F}_{Pr_2}(y)$ can also be written as

$$\hat{F}_{Pr_2}(y) = \{m_8F(y)(1+e_1)-m_9e_2-m_{10}e_4\}\left(1-\frac{1}{2}\theta e_2+\frac{3}{8}\theta^2e_2^2+\cdots\right). \quad (34)$$

Simplifying (34) and keeping terms only up to the second power of e_is, we can write

$$\left(\hat{F}_{Pr_2}(y)-F(y)\right) = -F(y)+m_8F(y)+m_8F(y)e_1-\frac{1}{2}\theta m_8F(y)e_2-m_9e_2-m_{10}e_4$$
$$+\frac{3}{8}\theta^2m_8F(y)e_1^2+\frac{1}{2}\theta m_9e_1^2-\frac{1}{2}\theta m_8F(y)e_1e_2+\frac{1}{2}\theta m_{10}e_2e_4. \quad (35)$$

The bias and MSE of $\hat{F}_9(y)$, to the first order of approximation, are

$$
\begin{aligned}
\operatorname{bias}\left(\hat{F}_{Pr_2}(y)\right) \cong\ & \frac{1}{2}\theta^2 F(y)\Theta_{0200} - \frac{1}{2}\theta F(y)\Theta_{1100} + \frac{1}{2}m_8\theta\Theta_{0200} + m_9 F(y) \\
& + \frac{3}{8}m_9\theta^2 F(y)\Theta_{0200} - \frac{1}{2}m_9\theta F(y)\Theta_{1100} + \frac{1}{2}m_{10}\theta\Theta_{0101}, \\
\operatorname{MSE}\left(\hat{F}_{Pr_2}(y)\right) \cong\ & -\theta F^2(y)\Theta_{1100} + \frac{3}{2}m_9\theta^2 F^2(y)\Theta_{0200} + m_9^2\theta^2 F^2(y)\Theta_{0200} + m_8\theta F(y)\Theta_{0200} \\
& - 2m_9 m_{10} F(y)\Theta_{1001} + \frac{1}{4}\theta^2 F^2(y)\Theta_{0200} + 2m_9 F^2(y)\Theta_{2000} + m_8^2\Theta_{0200} \\
& - 2m_9^2\theta F^2(y)\Theta_{1100} + F^2(y)\Theta_{2000} + m_9^2 F^2(y) + m_{10}\theta F(y)\Theta_{0101} \\
& + 2m_8 m_9\theta F(y)\Theta_{0200} - 2m_8 m_9 F(y)\Theta_{1100} + 2m_8 m_{10}\Theta_{0101} \\
& - 3m_9\theta F^2(y)\Theta_{1100} - 2m_8 F(y)\Theta_{1100} - 2m_{10}F(y)\Theta_{1001} \\
& + m_9^2 F^2(y)\Theta_{2000} + m_{10}^2\Theta_{0002} + 2m_9 m_{10}\theta F(y)\Theta_{0101}.
\end{aligned}
\tag{36}
$$

The optimum values of k_8, k_9, and k_{10}, determined by minimizing (36), are

$$
\begin{aligned}
m_{8(\text{opt})} &= \frac{F(y)\left[\theta^3\Theta_{0200}^{3/2}\left(\varrho_{24}^2 - 1\right) + \Theta_{2000}^{1/2}\left(-4 + \theta^2\Theta_{0200}\right)\left(\varrho_{12} - \varrho_{24}\varrho_{14}\right) + 2\theta\Theta_{0200}^{1/2}\left(\varrho_{24}^2 - 1\right)\left\{-1 + \Theta_{2000}\left(1 - \varrho_{1.24}^2\right)\right\}\right]}{8\Theta_{0200}^{1/2}\left(\varrho_{24}^2 - 1\right)\left\{-1 + \Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right)\right\}}, \\
m_{9(\text{opt})} &= \frac{\left\{4\Theta_{2000}\left(-2\varrho_{12}\varrho_{14}\varrho_{24} + \varrho_{12}^2 + \varrho_{14}^2 + \varrho_{24}^2 - 1\right) + \theta^2\Theta_{0200}\left(\varrho_{24}^2 - 1\right)\right\}}{-4\left[\left(\varrho_{24}^2 - 1\right)\left\{1 + \Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right)\right\}\right]}, \\
m_{10(\text{opt})} &= \frac{F(y)\Theta_{2000}^{1/2}\left\{\theta^2\Theta_{0200}\left(-\varrho_{14} + \varrho_{12}\varrho_{24}\right) + 4\left(\varrho_{14} - \varrho_{12}\varrho_{24}\right)\right\}}{-4\Theta_{0020}^{1/2}\left[\left(\varrho_{24}^2 - 1\right)\left\{1 + \Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right)\right\}\right]}.
\end{aligned}
\tag{37}
$$

The simplified minimum MSE of $\hat{F}_{Pr_2}(y)$ at the optimum values of k_8, k_9, and k_{10} is

$$
\operatorname{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) \cong \frac{F^2(y)\left\{16\Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right) - \theta^4\Theta_{0200}^2 - 8\theta\Theta_{0200}\Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right)\right\}}{16\left\{1 + \Theta_{2000}\left(1 - \mathfrak{R}_{1.24}^2\right)\right\}},
\tag{38}
$$

where $\mathfrak{R}_{1.24}^2 = (\Theta_{1100}^2\Theta_{0002} + \Theta_{1001}^2\Theta_{0200} - 2\ \Theta_{1001}\Theta_{1100}\Theta_{0101}/\Theta_{2000}(\Theta_{0200}\Theta_{0002} - \Theta_{0101}^2))$.

It can be seen that $\hat{F}_{Pr_2}(y)$ is more precise than $\hat{F}_{\text{Reg}}(y)$.

In Table 2, we put some members of the Grover and Kaur [20] and proposed families of estimators with selected choices of a and b.

5. Efficiency Comparisons

In this section, the adapted and proposed estimators of $F(y)$ are compared in terms of the minimum MSEs.

(i) From (8) and (32),

Table 2: Some members of the adapted and proposed distribution function estimators.

a	b	$\widehat{F}_{G,K}(y)$	$\widehat{F}_8(y)$	$\widehat{F}_9(y)$
1	C_2	$\widehat{F}_{GK}^{(1)}(y)$	$\widehat{F}_{Pr_1}^{(1)}(y)$	$\widehat{F}_{Pr_2}^{(1)}(y)$
1	β_2	$\widehat{F}_{GK}^{(2)}(y)$	$\widehat{F}_{Pr_1}^{(2)}(y)$	$\widehat{F}_{Pr_2}^{(2)}(y)$
β_2	C_2	$\widehat{F}_{GK}^{(3)}(y)$	$\widehat{F}_{Pr_1}^{(3)}(y)$	$\widehat{F}_{Pr_2}^{(3)}(y)$
C_2	β_2	$\widehat{F}_{GK}^{(4)}(y)$	$\widehat{F}_{Pr_1}^{(4)}(y)$	$\widehat{F}_{Pr_2}^{(4)}(y)$
1	ρ_{12}	$\widehat{F}_{GK}^{(5)}(y)$	$\widehat{F}_{Pr_1}^{(5)}(y)$	$\widehat{F}_{Pr_2}^{(5)}(y)$
C_2	ρ_{12}	$\widehat{F}_{GK}^{(6)}(y)$	$\widehat{F}_{Pr_1}^{(6)}(y)$	$\widehat{F}_{Pr_2}^{(6)}(y)$
ρ_{12}	C_2	$\widehat{F}_{GK}^{(7)}(y)$	$\widehat{F}_{Pr_1}^{(7)}(y)$	$\widehat{F}_{Pr_2}^{(7)}(y)$
β_2	ρ_{12}	$\widehat{F}_{GK}^{(8)}(y)$	$\widehat{F}_{Pr_1}^{(8)}(y)$	$\widehat{F}_{Pr_2}^{(8)}(y)$
ρ_{12}	β_2	$\widehat{F}_{GK}^{(9)}(y)$	$\widehat{F}_{Pr_1}^{(9)}(y)$	$\widehat{F}_{Pr_2}^{(9)}(y)$
1	$NF(x)$	$\widehat{F}_{GK}^{(10)}(y)$	$\widehat{F}_{Pr_1}^{(10)}(y)$	$\widehat{F}_{Pr_2}^{(10)}(y)$

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{Var}(\hat{F}(y)), \quad \text{if } \mathrm{Var}(\hat{F}(y)) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{39}$$

(ii) From (11) and (32)

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{MSE}\left(\hat{F}_R(y)\right), \quad \text{if } \mathrm{MSE}\left(\hat{F}_R(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{40}$$

(iii) From (13) and (32),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{MSE}\left(\hat{F}_P(y)\right), \quad \text{if } \mathrm{MSE}\left(\hat{F}_P(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{41}$$

(iv) From (16) and (32),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{Var}_{\min}\left(\hat{F}_{\mathrm{Reg}}(y)\right), \quad \text{if } \mathrm{Var}_{\min} \cdot \left(\hat{F}_{\mathrm{Reg}}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{42}$$

(v) From (21) and (32),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{MSE}_{\min}\left(\hat{F}_{R,D}(y)\right), \quad \text{if } \mathrm{MSE}_{\min} \cdot \left(\hat{F}_{R,D}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{43}$$

(vi) From (26) and (32),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) < \mathrm{MSE}_{\min}\left(\hat{F}_{G,K}(y)\right), \quad \text{if } \mathrm{MSE}_{\min} \cdot \left(\hat{F}_{G,K}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_1}(y)\right) > 0. \tag{44}$$

(vii) From (8) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{Var}(\hat{F}(y)), \quad \text{if } \mathrm{Var}(\hat{F} - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{45}$$

(viii) From (11) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{MSE}\left(\hat{F}_R(y)\right), \quad \text{if } \mathrm{MSE}\left(\hat{F}_R(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{46}$$

(ix) From (13) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{MSE}\left(\hat{F}_P(y)\right), \quad \text{if } \mathrm{MSE}\left(\hat{F}_P(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{47}$$

(x) From (16) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{Var}_{\min}\left(\hat{F}_{\mathrm{Reg}}(y)\right), \quad \text{if } \mathrm{Var}_{\min} \cdot \left(\hat{F}_{\mathrm{Reg}}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{48}$$

(xi) From (21) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{MSE}_{\min}\left(\hat{F}_{R,D}(y)\right), \quad \text{if } \mathrm{MSE}_{\min} \cdot \left(\hat{F}_{R,D}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{49}$$

(xii) From (26) and (38),

$$\mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) < \mathrm{MSE}_{\min}\left(\hat{F}_{G,K}(y)\right), \quad \text{if } \mathrm{MSE}_{\min} \cdot \left(\hat{F}_{G,K}(y)\right) - \mathrm{MSE}_{\min}\left(\hat{F}_{Pr_2}(y)\right) > 0. \tag{50}$$

The proposed families of estimators are always more precise than the adapted estimators as conditions (i)–(xii) are always true.

6. Empirical Study

In this section, we conduct a numerical study to see the performance of the existing and proposed distribution function estimators. For this purpose, three populations are considered. The summary statistics of these populations are reported in Tables 3–5. The percentage relative efficiency PRE of an estimator $\hat{F}_i(y)$ with respect to $\hat{F}_H(y)$ is where $i = R, P, \mathrm{Reg}, R, D, \ldots, Pr_2$.

$$\mathrm{PRE}\left(\hat{F}_i(y), \hat{F}_H(y)\right) = \frac{\mathrm{Var}\left(\hat{F}_H(y)\right)}{\mathrm{MSE}_{\min}\left(\hat{F}_i(y)\right)} \times 100, \tag{51}$$

The PREs of distribution function estimators, computed from three populations, are given in Tables 6 and 7.

Table 3: Summary statistics for population I.

Parameter	Value	Parameter	Value
N	30	$\overline{Z}$	15.5000
n	8	S_3	9.23238
λ	0.09167	S_4	8.79557
$\overline{X}$	67.2667	ρ_{34}	0.98899
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$F(y)$	—	0.60630	0.16535
$F(x)$	—	0.76923	0.20513
S_1	—	0.50855	0.50855
S_2	—	0.50855	0.50742
ρ_{12}	—	−0.73333	−0.80178
ρ_{13}	—	0.71975	0.71975
ρ_{23}	—	−0.83726	−0.84109
ρ_{14}	—	0.73622	0.73622
ρ_{24}	—	−0.86728	−0.86535
Nonresponse			
Parameter	Value	Parameter	Value
N_2	8	$S_{3(2)}$	9.38749
W_2	0.26667	$S_{4(2)}$	2.44949
λ_2	0.03333	$\rho_{34(2)}$	0.91015
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$S_{1(2)}$	—	0.51755	0.51755
$S_{2(2)}$	—	0.51755	0.51755
$\rho_{12(2)}$	—	−0.60000	−0.60000
$\rho_{13(2)}$	—	0.36387	0.36387
$\rho_{23(2)}$	—	−0.81228	−0.81228
$\rho_{14(2)}$	—	0.28172	0.28172
$\rho_{24(2)}$	—	−0.84515	−0.84515

Table 4: Summary statistics for population II.

Parameter	Value	Parameter	Value
N	50	$\overline{Z}$	25.5000
n	15	S_3	21.3175
λ	0.04667	S_4	14.5756
$\overline{X}$	78.2900	ρ_{34}	0.94677
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$F(y)$	—	0.50000	0.16535
$F(x)$	—	0.50000	0.20513
S_1	—	0.50508	0.47121
S_2	—	0.50508	0.49856
ρ_{12}	—	−0.12000	−0.14941
ρ_{13}	—	0.22925	0.28411
ρ_{23}	—	−0.78936	−0.80938
ρ_{14}	—	0.18435	0.25257
ρ_{24}	—	−0.86630	−0.85514
Nonresponse			
Parameter	Value	Parameter	Value
N_2	12	$S_{3(2)}$	18.2593
W_2	0.24000	$S_{4(2)}$	3.60555
λ_2	0.01600	$\rho_{34(2)}$	0.97952
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$S_{1(2)}$	—	0.51493	0.38924
$S_{2(2)}$	—	0.52223	0.51493
$\rho_{12(2)}$	—	−0.16903	−0.37796
$\rho_{13(2)}$	—	0.25695	0.13750
$\rho_{23(2)}$	—	−0.81370	−0.84530
$\rho_{14(2)}$	—	0.22034	0.12955
$\rho_{24(2)}$	—	−0.86905	−0.85689

Table 5: Summary statistics for population III.

Parameter	Value	Parameter	Value
N	50	$\overline{Z}$	25.5000
n	15	S_3	22.18052
λ	0.04667	S_4	14.57598
$\overline{X}$	75.8720	ρ_{34}	0.95742
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$F(y)$		0.50000	0.66000
$F(x)$	—	0.50000	0.58000
S_1	—	0.50508	0.50508
S_2	—	0.50508	0.50508
ρ_{12}	—	−0.20000	−0.18306
ρ_{13}	—	0.30094	0.34288
ρ_{23}	—	−0.79517	−0.81844
ρ_{14}	—	0.25781	0.33356
ρ_{24}	—	−0.86628	−0.85512
Nonresponse			
Parameter	Value	Parameter	Value
N_2	12	$S_{3(2)}$	19.5392
W_2	0.24000	$S_{4(2)}$	3.60555
λ_2	0.01600	$\rho_{34(2)}$	0.98710
Parameter		$\overline{X}$, $\overline{Y}$	$\Theta_2(y)$, $\Theta_2(x)$
$S_{2(2)}$	—	0.52223	0.52223
$\rho_{12(2)}$	—	−0.50709	−0.44721
$\rho_{13(2)}$	—	0.35848	0.18467
$\rho_{23(2)}$	—	−0.82900	−0.82900
$\rho_{14(2)}$	—	0.36724	0.19433
$\rho_{24(2)}$	—	−0.86905	−0.86905

Table 6: PREs of distribution function estimators using populations I, II and III with situation-I and situation-II, when $\{x, y = \Theta_2(x), \Theta_2(y)\}$.

Estimators			Population I			Population II			Population III		
			$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$	$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$	$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$
$\widehat{F}^{(1)}_{\text{GK}}(y)$	$\widehat{F}^{(1)}_{Pr_1}(y)$	$\widehat{F}^{(1)}_{Pr_2}(y)$	208.21	211.36	216.86	108.45	113.58	109.94	115.27	118.13	115.63
$\widehat{F}^{(2)}_{\text{GK}}(y)$	$\widehat{F}^{(2)}_{Pr_1}(y)$	$\widehat{F}^{(2)}_{Pr_2}(y)$	208.22	211.39	216.90	108.46	113.58	109.95	115.28	118.13	115.64
$\widehat{F}^{(3)}_{\text{GK}}(y)$	$\widehat{F}^{(3)}_{Pr_1}(y)$	$\widehat{F}^{(3)}_{Pr_2}(y)$	208.21	211.36	216.86	108.45	113.58	109.94	115.27	118.13	115.63
$\widehat{F}^{(4)}_{\text{GK}}(y)$	$\widehat{F}^{(4)}_{Pr_1}(y)$	$\widehat{F}^{(4)}_{Pr_2}(y)$	208.24	211.43	216.93	108.46	113.59	109.95	115.28	118.14	115.65
$\widehat{F}^{(5)}_{\text{GK}}(y)$	$\widehat{F}^{(5)}_{Pr_1}(y)$	$\widehat{F}^{(5)}_{Pr_2}(y)$	270.64	579.94	610.90	111.76	121.61	117.66	121.61	134.60	131.64
$\widehat{F}^{(6)}_{\text{GK}}(y)$	$\widehat{F}^{(6)}_{Pr_1}(y)$	$\widehat{F}^{(6)}_{Pr_2}(y)$	284.55	855.14	921.21	111.73	121.54	117.60	121.50	134.27	131.32
$\widehat{F}^{(7)}_{\text{GK}}(y)$	$\widehat{F}^{(7)}_{Pr_1}(y)$	$\widehat{F}^{(7)}_{Pr_2}(y)$	209.72	214.62	220.21	108.27	113.18	109.56	115.09	117.74	115.26
$\widehat{F}^{(8)}_{\text{GK}}(y)$	$\widehat{F}^{(8)}_{Pr_1}(y)$	$\widehat{F}^{(8)}_{Pr_2}(y)$	270.64	579.94	610.90	111.76	121.61	117.66	121.61	134.60	131.64
$\widehat{F}^{(9)}_{\text{GK}}(y)$	$\widehat{F}^{(9)}_{Pr_1}(y)$	$\widehat{F}^{(9)}_{Pr_2}(y)$	209.85	579.94	220.51	108.27	113.18	109.56	115.09	117.75	115.26
$\widehat{F}^{(10)}_{\text{GK}}(y)$	$\widehat{F}^{(10)}_{Pr_1}(y)$	$\widehat{F}^{(10)}_{Pr_2}(y)$	207.47	209.83	215.29	108.26	113.17	109.55	115.07	117.70	115.22
	$\widehat{F}(y)$			100.00			100				100
	$\widehat{F}_R(y)$			29.47			43.97				38.87
	$\widehat{F}_P(y)$			164.94			57.46				69.33
	$\widehat{F}_{\text{Reg}}(y)$			194.40			101.8				108.61
	$\widehat{F}_{R,D}(y)$			207.46			108.26				115.07
$\widehat{F}^{(1)}_{\text{GK}}(y)$	$\widehat{F}^{(1)}_{Pr_1}(y)$	$\widehat{F}^{(1)}_{Pr_2}(y)$	177.62	185.80	186.47	107.67	111.58	109.81	109.64	114.27	112.09
$\widehat{F}^{(2)}_{\text{GK}}(y)$	$\widehat{F}^{(2)}_{Pr_1}(y)$	$\widehat{F}^{(2)}_{Pr_2}(y)$	177.63	185.82	186.49	107.68	111.58	109.81	109.64	114.28	112.09
$\widehat{F}^{(3)}_{\text{GK}}(y)$	$\widehat{F}^{(3)}_{Pr_1}(y)$	$\widehat{F}^{(3)}_{Pr_2}(y)$	177.62	185.80	186.47	107.67	111.58	109.81	109.64	114.27	112.09
$\widehat{F}^{(4)}_{\text{GK}}(y)$	$\widehat{F}^{(4)}_{Pr_1}(y)$	$\widehat{F}^{(4)}_{Pr_2}(y)$	177.65	185.84	186.52	107.68	111.59	109.82	109.65	114.28	112.09
$\widehat{F}^{(5)}_{\text{GK}}(y)$	$\widehat{F}^{(5)}_{Pr_1}(y)$	$\widehat{F}^{(5)}_{Pr_2}(y)$	207.48	295.08	296.44	109.98	116.91	115.04	113.76	124.47	122.04
$\widehat{F}^{(6)}_{\text{GK}}(y)$	$\widehat{F}^{(6)}_{Pr_1}(y)$	$\widehat{F}^{(6)}_{Pr_2}(y)$	212.84	327.94	329.57	109.96	116.87	115.00	113.69	124.28	121.86
$\widehat{F}^{(7)}_{\text{GK}}(y)$	$\widehat{F}^{(7)}_{Pr_1}(y)$	$\widehat{F}^{(7)}_{Pr_2}(y)$	178.54	187.79	188.48	107.54	111.29	109.53	109.51	114.00	111.82
$\widehat{F}^{(8)}_{\text{GK}}(y)$	$\widehat{F}^{(8)}_{Pr_1}(y)$	$\widehat{F}^{(8)}_{Pr_2}(y)$	207.48	295.08	296.44	109.98	116.91	115.04	113.76	124.47	122.04
$\widehat{F}^{(9)}_{\text{GK}}(y)$	$\widehat{F}^{(9)}_{Pr_1}(y)$	$\widehat{F}^{(9)}_{Pr_2}(y)$	178.62	187.97	188.65	107.54	111.29	109.53	109.51	114.00	111.82
$\widehat{F}^{(10)}_{\text{GK}}(y)$	$\widehat{F}^{(10)}_{Pr_1}(y)$	$\widehat{F}^{(10)}_{Pr_2}(y)$	177.17	184.83	185.50	107.53	111.28	109.52	109.50	113.97	111.79
	$\widehat{F}(y)$			100.00			100.00				100.00
	$\widehat{F}_R(y)$			35.82			52.24				49.21
	$\widehat{F}_P(y)$			151.28			64.09				69.33
	$\widehat{F}_{\text{Reg}}(y)$			164.11			101.07				103.04
	$\widehat{F}_{R,D}(y)$			177.16			107.53				109.50

Population I (source: [21]).

Y: duration of sleep of persons with age more than 50 years

X: the age of persons in years. The proportion of the non-response units in the given population is considered to be the last 25% units

Population II (source: [22]).

Y: the eggs produced in 1990 (millions)

X: the price per dozen (cents) in 1990. The proportion of the non-response units in the given population is considered to be the last 25% units

Population III (source: [22]).

Y: the eggs produced in 1990 (millions)

X: the price per dozen (cents) in 1991. The proportion of the non-response units in the given population is considered to be the last 25% units

From the numerical results, presented in Tables 6 and 7, it is observed that the PREs of all families of estimators change with the choices of a and b. It is further noted that the proposed families of estimators are more precise than the existing distribution function estimators of Hansen and Hurwitz [1]; Cochran [17]; Murthy [18]; Rao [19]; and Grover and Kaur [20], in terms of PRE under both situations.

Table 7: PREs of distribution function estimators using populations I, II, and III with situation-I and situation-II, when $\{x, y = \overline{X}, \overline{Y},\}$.

Estimators			Population I			Population II			Population III		
			$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$	$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$	$\widehat{F}_{\text{GK}}(y)$	$\widehat{F}_{Pr_1}(y)$	$\widehat{F}_{Pr_2}(y)$
$\widehat{F}_{\text{GK}}^{(1)}(y)$	$\widehat{F}_{Pr_1}^{(1)}(y)$	$\widehat{F}_{Pr_2}^{(1)}(y)$	239.81	240.75	243.79	107.15	110.21	109.02	109.41	114.28	114.10
$\widehat{F}_{\text{GK}}^{(2)}(y)$	$\widehat{F}_{Pr_1}^{(2)}(y)$	$\widehat{F}_{Pr_2}^{(2)}(y)$	239.73	240.59	243.63	107.09	110.08	108.90	109.35	114.16	113.97
$\widehat{F}_{\text{GK}}^{(3)}(y)$	$\widehat{F}_{Pr_1}^{(3)}(y)$	$\widehat{F}_{Pr_2}^{(3)}(y)$	239.83	240.79	243.84	107.18	110.26	109.08	109.44	114.34	114.16
$\widehat{F}_{\text{GK}}^{(4)}(y)$	$\widehat{F}_{Pr_1}^{(4)}(y)$	$\widehat{F}_{Pr_2}^{(4)}(y)$	239.68	240.48	243.52	107.07	110.02	108.85	109.33	114.09	113.91
$\widehat{F}_{\text{GK}}^{(5)}(y)$	$\widehat{F}_{Pr_1}^{(5)}(y)$	$\widehat{F}_{Pr_2}^{(5)}(y)$	287.13	431.28	439.07	109.76	116.82	115.55	112.69	122.96	122.76
$\widehat{F}_{\text{GK}}^{(6)}(y)$	$\widehat{F}_{Pr_1}^{(6)}(y)$	$\widehat{F}_{Pr_2}^{(6)}(y)$	270.32	338.77	343.87	110.19	118.07	116.77	113.44	125.30	125.09
$\widehat{F}_{\text{GK}}^{(7)}(y)$	$\widehat{F}_{Pr_1}^{(7)}(y)$	$\widehat{F}_{Pr_2}^{(7)}(y)$	243.94	250.01	253.20	106.96	109.80	108.63	109.22	113.88	113.70
$\widehat{F}_{\text{GK}}^{(8)}(y)$	$\widehat{F}_{Pr_1}^{(8)}(y)$	$\widehat{F}_{Pr_2}^{(8)}(y)$	296.41	502.56	512.87	109.55	116.22	114.96	112.35	121.92	121.72
$\widehat{F}_{\text{GK}}^{(9)}(y)$	$\widehat{F}_{Pr_1}^{(9)}(y)$	$\widehat{F}_{Pr_2}^{(9)}(y)$	242.79	247.30	250.44	106.95	109.79	108.61	109.21	113.86	113.67
$\widehat{F}_{\text{GK}}^{(10)}(y)$	$\widehat{F}_{Pr_1}^{(10)}(y)$	$\widehat{F}_{Pr_2}^{(10)}(y)$	238.92	238.93	241.95	106.94	109.77	108.59	109.20	113.83	113.65
	$\widehat{F}(y)$			100.00			100.00				100.00
	$\widehat{F}_R(y)$			30.55			31.00				31.42
	$\widehat{F}_P(y)$			208.91			45.86				50.75
	$\widehat{F}_{\text{Reg}}(y)$			225.86			104.18				106.19
	$\widehat{F}_{R,D}(y)$			238.91			106.94				109.20
$\widehat{F}_{\text{GK}}^{(1)}(y)$	$\widehat{F}_{Pr_1}^{(1)}(y)$	$\widehat{F}_{Pr_2}^{(1)}(y)$	201.21	203.59	203.62	104.76	111.86	110.04	105.97	116.90	117.29
$\widehat{F}_{\text{GK}}^{(2)}(y)$	$\widehat{F}_{Pr_1}^{(2)}(y)$	$\widehat{F}_{Pr_2}^{(2)}(y)$	201.16	203.50	203.53	104.72	111.77	109.95	105.93	116.80	117.19
$\widehat{F}_{\text{GK}}^{(3)}(y)$	$\widehat{F}_{Pr_1}^{(3)}(y)$	$\widehat{F}_{Pr_2}^{(3)}(y)$	201.22	203.62	203.65	104.77	111.90	110.08	105.99	116.94	117.33
$\widehat{F}_{\text{GK}}^{(4)}(y)$	$\widehat{F}_{Pr_1}^{(4)}(y)$	$\widehat{F}_{Pr_2}^{(4)}(y)$	201.13	203.43	203.46	104.70	111.73	109.91	105.91	116.76	117.15
$\widehat{F}_{\text{GK}}^{(5)}(y)$	$\widehat{F}_{Pr_1}^{(5)}(y)$	$\widehat{F}_{Pr_2}^{(5)}(y)$	224.45	274.00	274.05	106.51	116.30	114.38	108.12	122.64	123.06
$\widehat{F}_{\text{GK}}^{(6)}(y)$	$\widehat{F}_{Pr_1}^{(6)}(y)$	$\widehat{F}_{Pr_2}^{(6)}(y)$	217.00	246.45	246.49	106.79	117.08	115.15	108.58	124.07	124.49
$\widehat{F}_{\text{GK}}^{(7)}(y)$	$\widehat{F}_{Pr_1}^{(7)}(y)$	$\widehat{F}_{Pr_2}^{(7)}(y)$	203.62	208.84	208.87	104.62	111.56	109.75	105.84	116.60	116.99
$\widehat{F}_{\text{GK}}^{(8)}(y)$	$\widehat{F}_{Pr_1}^{(8)}(y)$	$\widehat{F}_{Pr_2}^{(8)}(y)$	228.30	290.50	290.56	106.38	115.92	114.01	107.90	121.99	122.41
$\widehat{F}_{\text{GK}}^{(9)}(y)$	$\widehat{F}_{Pr_1}^{(9)}(y)$	$\widehat{F}_{Pr_2}^{(9)}(y)$	202.96	207.35	207.39	104.61	111.55	109.74	105.83	116.58	116.97
$\widehat{F}_{\text{GK}}^{(10)}(y)$	$\widehat{F}_{Pr_1}^{(10)}(y)$	$\widehat{F}_{Pr_2}^{(10)}(y)$	200.67	202.49	202.52	104.61	111.54	109.72	105.82	116.56	116.95
	$\widehat{F}(y)$			100.00			100.00				100.00
	$\widehat{F}_R(y)$			36.69			39.25				40.01
	$\widehat{F}_P(y)$			183.15			51.37				55.81
	$\widehat{F}_{\text{Reg}}(y)$			187.61			101.84				102.81
	$\widehat{F}_{R,D}(y)$			200.66			104.61				105.82

7. Concluding Remarks

In this paper, we have proposed two new families of estimators for estimating the finite population distribution function. The proposed estimators needed supplementary data on the sample mean and ranks of the auxiliary variable. The biases and mean squared error of the proposed families of estimators were derived using the first order of approximation. Based on theoretical as well as numerical comparative studies, it is concluded that the proposed families of estimators are more precise than their existing counterparts under situation-I and situation-II. So, we recommend using the sample mean and ranks of the auxiliary variable with the proposed families of estimators for estimating the finite population distribution function.

Data Availability

The data used to support the numerical findings of this study are available from the corresponding author upon request. The data can also be obtained upon searching the given sources of data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The first three authors acknowledge the support of the National Social Science Fund of China (17BTJ010) and the National Fund for Shanxi "1331 Project" Key Innovative Research Team.

References

[1] 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.

[2] J. Neyman, "Lectures and conferences on mathematical statistics," *Nature*, vol. 142, p. 274, 1938.

[3] W. Cochran, *Sampling Techniques*, Wiley, New York, NY, USA, 3rd edition, 1977.

[4] R. L. Chambers, A. H. Dorfman, and P. Hall, "Properties of estimators of the finite population distribution function," *Biometrika*, vol. 79, no. 3, pp. 577–582, 1992.

[5] R. L. Chambers and R. Dunstan, "Estimating distribution functions from survey data," *Biometrika*, vol. 73, no. 3, pp. 597–604, 1986.

[6] P. Rao, "Ratio estimation with subsampling the nonrespondents," *Survey Methodology*, vol. 12, no. 2, pp. 217–230, 1986.

[7] B. B. Khare and S. Srivastava, "Transformed ratio type estimators for the population mean in the presence of nonresponse," *Communications in Statistics—Theory and Methods*, vol. 26, no. 7, pp. 1779–1791, 1997.

[8] F. C. Okafor and H. Lee, "Double sampling for ratio and regression estimation with sub-sampling the non-respondents," *Survey Methodology*, vol. 26, no. 2, pp. 183–188, 2000.

[9] B. Khare and R. Sinha, "Estimation of the ratio of the two population means using multi-auxiliary characters in the presence of non-response," *Statistical Techniques in Life Testing, Reliability, Sampling Theory and Quality Control*, vol. 1, pp. 63–171, 2007.

[10] B. Khare and R. Sinha, "Estimation of population mean using multi-auxiliary characters with subsampling the nonrespondents," *Statistics in Transition*, vol. 1, no. 12, pp. 45–56, 2011.

[11] P.-Y. Zhao, M.-L. Tang, and N.-S. Tang, "Robust estimation of distribution functions and quantiles with non-ignorable missing data," *Canadian Journal of Statistics*, vol. 41, no. 4, pp. 575–595, 2013.

[12] J. Shabbir and N. S. Khan, "On estimating the finite population mean using two auxiliary variables in two phase sampling in the presence of non response," *Communications in Statistics—Theory and Methods*, vol. 42, no. 22, pp. 4127–4145, 2013.

[13] H. P. Singh and S. Kumar, "A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information," *Statistics and Operations Research Transactions*, vol. 33, no. 1, pp. 71–84, 2009.

[14] H. P. Singh and S. Kumar, "Estimation of mean in presence of non-response using two phase sampling scheme," *Statistical Papers*, vol. 51, no. 3, pp. 559–582, 2010.

[15] R. Tabasum and I. Khan, "Double sampling ratio estimator for the population mean in presence of non-response," *Assam Statistical Review*, vol. 20, no. 1, pp. 73–83, 2006.

[16] G. Diana and P. F. Perri, "A class of estimators in two-phase sampling with subsampling the non-respondents," *Applied Mathematics and Computation*, vol. 219, no. 19, pp. 10033–10043, 2013.

[17] W. G. Cochran, "The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce," *The Journal of Agricultural Science*, vol. 30, no. 2, pp. 262–275, 1940.

[18] M. Murthy, "Product method of estimation," *Sankhya: The Indian Journal of Statistics, Series A*, vol. 26, no. 1, pp. 69–74, 1964.

[19] T. J. Rao, "On certail methods of improving ration and regression estimators," *Communications in Statistics—Theory and Methods*, vol. 20, no. 10, pp. 3325–3340, 1991.

[20] L. K. Grover and P. Kaur, "A generalized class of ratio type exponential estimators of population mean under linear transformation of auxiliary variable," *Communications in Statistics - Simulation and Computation*, vol. 43, no. 7, pp. 1552–1574, 2014.

[21] S. Singh, *Advanced Sampling Theory with Applications: How Michael "Selected" Amy*, Vol. 1, Springer Science & Business Media, vol. 1, Berlin, Germany, 2003.

[22] D. N. Gujarati, *Basic Econometrics*, Tata McGraw-Hill Education, New York, NY, USA, 2009.