Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(4): 8946–8964.
DOI: 10.3934/math.2025410
Received: 15 January 2025
Revised: 26 March 2025
Accepted: 08 April 2025
Published: 18 April 2025

Research article

Efficient classes of estimators for estimating indeterminate population mean using neutrosophic ranked set sampling

Anoop Kumar[1], **Priya**[1,*] **and Abdullah Mohammed Alomair**[2]

[1] Department of Statistics, Central University of Haryana, Mahendergarh, Haryana 123031, India
[2] Department of Quantitative Methods, School of Business, King Faisal University, Al-Ahsa 31982, Saudi Arabia

* **Correspondence:** Email: priyav2698@gmail.com.

Abstract: Estimating the population mean with accuracy is frequently challenged by uncertain and inaccurate data in survey sampling. This study presents some efficient classes of estimators for estimating the indeterminate population mean using neutrosophic ranked set sampling (NRSS). The study establishes the bias and mean squared error (MSE) of the suggested estimators and compares their performance with the existing neutrosophic estimators. Analytical comparisons show considerable efficiency benefits over the existing neutrosophic estimators. Simulation research and executions on real-life datasets confirm the accuracy of the proposed neutrosophic estimators when dealing with uncertainty. The findings highlight the potential of proposed NRSS estimators as a powerful tool for population mean estimation by providing new insights into statistical methodologies for uncertain and imprecise situations.

Keywords: mean square error; neutrosophic ranked set sampling; efficient classes of estimators; efficiency; indeterminacy
Mathematics Subject Classification: 62D05

1. Introduction

The use of auxiliary information in sampling surveys significantly improves the accuracy of an estimator because it provides additional information that is directly associated with the variable of interest. When auxiliary variables are appropriately utilized, they help minimize the variance or mean square error of the estimator, resulting in lesser standard errors and more precision in parameter estimations. For instance, in survey sampling, auxiliary variables such as population means, totals, or variances, which are often accessible, may be integrated into estimation methods utilizing ratio, regression, or logarithmic estimators. These methods utilize the association between the study and auxiliary variables to modify and elevate the estimation process. One may adjust the potential biases and produce more accurate outcomes by aligning the estimator with the auxiliary information. Several studies are available in the literature for the estimation of parameters based on auxiliary information. [1] utilized the auxiliary information and introduced the ratio estimator for population mean. [2] developed a product estimator for population mean using the auxiliary information. Several modified ratio estimators were introduced by [3], employing the coefficient of variation as auxiliary information. Some transformed auxiliary variables were used by [4] to estimate the population mean. [5] suggested the estimation of finite population distribution functions utilizing auxiliary information in both simple and stratified random sampling. [6] proposed a two-fold utilization of auxiliary information to estimate the finite population mean under stratified random sampling. [7] proposed the robust regression-type estimators to improve the mean estimation of sensitive variables utilising auxiliary information. [8] evaluated the performance of ratio-exponential-log-type estimators with two auxiliary variables. [9] designed the quantile regression-ratio-type estimators for mean estimation utilizing complete and partial auxiliary data. [10] suggested an enhanced estimation of population mean using simple random sampling. [11] proposed the novel logarithmic imputation procedures based on multi-auxiliary information under RSS. These studies emphasize advances in estimation methodologies, the use of auxiliary information, novel sampling strategies, and robust estimators to a meliorate the accuracy and efficiency in the population parameter estimation.

In survey sampling, the conventional sampling strategies frequently struggle with data that are unclear, vague, or imprecise. These complexities are increasingly encountered in real-life issues, where uncertain or indeterminate information arises due to ambiguous definitions, measurement errors, or inconsistent observations. New strategies that surpass the traditional statistical frameworks are required to address these issues. Neutrosophic logic was first introduced by [12], which is based on three components, such as truth, indeterminacy, and falsity. This adaptable structure is perfect for handling uncertain or indeterminate data, which makes it a perfect tool for contemporary statistical problems. In the context of survey sampling, [13] firstly introduced the neutrosophic ratio type estimators for uncertain population mean estimation under SRS. To improve the population mean estimation under uncertainty, a generalized neutrosophic estimator was developed by [14] using SRS. [15] introduced the neutrosophic estimators under uncertainty using two-phase sampling, while [16] proposed the neutrosophic factor-type exponential estimators for enhanced population mean estimation with the use of auxiliary data. For the estimation of the confined population mean, [17] suggested a neutrosophic robust ratio-type estimator. [18] suggested using robust parameters of the auxiliary variable to estimate the neutrosophic finite median.

The classical ranked set sampling (RSS) relies on accurate ranking, which may not be practicable in real-life situations, including uncertain or incomplete knowledge. Neutrosophic ranked set sampling (NRSS) is important for enhancing data collection and analysis, especially when handling imprecise or uncertain information. NRSS incorporates neutrosophic logic to handle uncertainty, incompleteness, and indeterminacy in ranking algorithms. This method improves sampling efficiency, generates more representative datasets, and increases the reliability and validity of statistical conclusions, making it important in environmental research, quality control, and social sciences. Very limited studies on NRSS are available in the literature for estimating the population parameters. For the very first time, [19] introduced the concept of NRSS and computed the uncertain population mean using the generalized estimators. Later on, [20] proposed neutrosophic estimators for population mean under NRSS. To efficiently estimate the indeterminate population mean, this paper introduces some efficient class of estimators under NRSS, which is an improved optimization strategy.

The RSS methodology has been thoroughly investigated and proven to be a more efficient option than SRS methodology in situations where accurate measurements are difficult, costly, or time-consuming. This is partly because RSS utilizes supplementary population information in the form of ranks, resulting in more accurate estimates with fewer measurements. However, while handling imprecise, inconsistent, or inconclusive data, which is common in a neutrosophic context, measuring issues become much more apparent. The statistical literature contains different kinds of neutrosophic data, including quantitative neutrosophic data, which is based on a number existing in an unknown interval $[p, q]$. This unknown interval $[p, q]$ based on neutrosophic numbers can be expressed in different forms. We have taken the neutrosophic interval values as $W_{rssN} = W_{rssL} + W_{rssU} I_{rssN}$ with $I_{rssN} \in [I_{rssL},\ I_{rssU}]$, N is here for the neutrosophic number. This shows that the notations utilized for neutrosophic data are in an interval form, $W_{rssN} \in [p, q]$, where 'p' and 'q' are lower and upper values of the neutrosophic data.

The NRSS methodology is based on selecting $m_N \in [m_L, m_U]$ bivariate random samples of size $m_N \in [m_L, m_U]$ units from a population of size N. The ranking is performed within each sample on the auxiliary variable $x_N \in [x_L,\ x_U]$ associated with the study variable $y_N \in [y_L,\ y_U]$. Following [12], the ranking of neutrosophic numbers can easily be done. In the NRSS, from the first set, choose the first smallest unit and consider it as the first measurement unit, and discard the rest of the units of the set. Similarly, from the second set, choose the second smallest unit and consider it as the second measurement unit, and discard the rest of the units of the set. Continuing this process, from the m_N^{th} set, choose the m_N^{th} smallest unit and consider it as the m_N^{th} measurement unit, and discard the rest of the units of the set. This process completes a cycle. If this whole cycle is iterated r times, then this yields a neutrosophic ranked set sample of size $n_N = m_N r \in [n_L,\ n_U]$. In the extraction of NRSS data, there is a total of $m_N^2 r$ units, but only $n_N = m_N r \in [n_L, n_U]$ units are required for actual computation. Let $X_{j(i)N} \in [X_{j(i)L}, X_{j(i)U}]$, $Y_{j[i]N} \in [Y_{j[i]L},\ Y_{j[i]U}]$; $j = 1, 2, ..., r$; $i = 1, 2, ..., m_N$, be the pair of neutrosophic bivariate quantified sets of the i^{th} units in the j^{th} cycle. The '[]' and '()' used in the subscript of variables y_N and x_N, respectively, denote the imperfect and perfect ranking of the units.

Let a finite population (U=U_1, U_2,...,U_N) be based on N identifiable units from which a neutrosophic sample of size $n_N \in [n_L,\ n_U]$ is randomly selected. Let $y_N(i)$ be the i^{th} observation of the sample for the neutrosophic study variable y_N expressed as $y_N(i) \in [y_L,\ y_U]$, whereas corresponding to the neutrosophic study variable, data on the neutrosophic auxiliary variable x_N are expressed as $x_N(i) \in [x_L,\ x_U]$. Let $\bar{y}_{[n]N} \in [\bar{y}_{[n]L},\ \bar{y}_{[n]U}]$ and $\bar{x}_{(n)N} \in [\bar{x}_{(n)L},\ \bar{x}_{(n)U}]$ be the neutrosophic ranked set sample means corresponding to the neutrosophic population means $\bar{Y}_N \in [\bar{Y}_L,\ \bar{Y}_U]$ and $\bar{X}_N \in [\bar{X}_L,\ \bar{Y}_U]$ for the neutrosophic study and auxiliary variables y_N and x_N, respectively. The neutrosophic variation coefficients of variables y_N and x_N are denoted as $C_{y_N} \in [C_{y_L},\ C_{y_U}]$ and $C_{x_N} \in [C_{x_L},\ C_{x_U}]$, respectively. The neutrosophic correlation coefficient between the neutrosophic variables x_N and y_N is denoted by $\rho_{xy_N} \in [\rho_{xy_L}, \rho_{xy_U}]$. The neutrosophic skewness and kurtosis coefficients for x_N are denoted by $\beta_1(x_N) \in [\beta_1(x_L),\ \beta_1(x_U)]$ and $\beta_2(x_N) \in [\beta_2(x_L),\ \beta_2(x_U)]$, respectively.

To obtain neutrosophic $Bias_N \in [Bias_L,\ Bias_U]$ and neutrosophic $MSE_N \in [MSE_L,\ MSE_U]$ of the neutrosophic estimators, we take neutrosophic error terms together with their expectations as

$$\left.\begin{aligned}
e_{0N} &= \frac{(\bar{y}_{[n]N} - \bar{Y}_N)}{\bar{Y}_N},\\
e_{1N} &= \frac{(\bar{x}_{(n)N} - \bar{X}_N)}{\bar{X}_N},\\
E(e_{0N}) &= E(e_{1N}) = 0,\\
E(e_{0N}^2) &= \frac{C_{y_N}^2}{n_N} - \frac{1}{m_N^2 r}\sum_{i=1}^{m_N}\frac{(\mu_{y_{[i]N}} - \bar{Y}_N)^2}{\bar{Y}_N^2} = V_{0N},\\
E(e_{1N}^2) &= \frac{C_{x_N}^2}{n_N} - \frac{1}{m_N^2 r}\sum_{i=1}^{m_N}\frac{(\mu_{x_{(i)N}} - \bar{X}_N)^2}{\bar{X}_N^2} = V_{1N},\\
E(e_{0N}e_{1N}) &= \frac{\rho_{xy_N} C_{x_N} C_{y_N}}{n_N} - \frac{1}{m_N^2 r}\sum_{i=1}^{m_N}\frac{(\mu_{x_{(i)N}} - \bar{X}_N)(\mu_{y_{[i]N}} - \bar{Y}_N)}{\bar{X}_N \bar{Y}_N} = V_{01N},
\end{aligned}\right\} \tag{1.1}$$

where $e_{0N} \in [e_{0L},\ e_{0U}]$, $e_{1N} \in [e_{1L},\ e_{1U}]$, $\bar{y}_{[n]N} = \sum_{i=1}^{m_N} y_{[i]N}/m_N r$, $\bar{x}_{(i)N} = \sum_{i=1}^{m_N} x_{(i)N}/m_N r$, $C_{x_N}^2 = \sigma_{x_N}^2/\bar{X}_N^2$, $C_{y_N}^2 = \sigma_{y_N}^2/\bar{Y}_N^2$, $\rho_{xy_N} = \sigma_{xy_N}/\sigma_{x_N}\sigma_{y_N}$, $\sigma_{x_N}^2 \in [\sigma_{x_L}^2, \sigma_{x_U}^2]$, $\sigma_{y_N}^2 \in [\sigma_{y_L}^2, \sigma_{y_U}^2]$, and $\sigma_{xy_N} \in [\sigma_{xy_L}, \sigma_{xy_U}]$.

The coming section offers an extensive review of the available and adapted neutrosophic estimators under NRSS together with their properties. The proposed class of neutrosophic estimators and their properties are established in Section 3. Section 4 contains analytical comparisons, a comprehensive simulation study based on a hypothetically drawn population, and the applications of the proposed estimators based on a neutrosophic real data sets. The results are discussed in Section 5. The article ends with the conclusions in Section 6.

2. Existing estimators

In survey sampling, many times, the neutrosophic auxiliary information is not available. In such cases, the sample mean $\bar{y}_{[n]N} = \sum_{i=1}^{m_N} y_{[i]N}/m_N r$ of the neutrosophic study variable serves as an unbiased estimate of the neutrosophic population mean $\bar{Y}_N$. The variance of the sample mean $\bar{y}_{[n]N}$ is given under NRSS as

$$V(\bar{y}_{[n]N}) = \bar{Y}_N^2 V_{0N}.$$

Employing neutrosophic auxiliary data and following [19, 21], the neutrosophic ratio estimator t_{r_N} for population mean $\bar{Y}_N$ under NRSS was prescribed as

$$t_{r_N} = \bar{y}_{[n]N}\left(\frac{\bar{X}_N}{\bar{x}_{(n)N}}\right).$$

The bias and MSE of the estimator t_{r_N} are given by

$$\begin{aligned}
Bias(t_{r_N}) &= \bar{Y}_N(V_{1N} - V_{01N}),\\
MSE(t_{r_N}) &= \bar{Y}_N^2(V_{0N} + V_{1N} - 2V_{01N}).
\end{aligned}$$

Following [22], we prescribe the neutrosophic regression estimator t_{lr_N} for the population mean $\bar{Y}_N$ under NRSS as

$$t_{lr_N} = \bar{y}_{[n]N} + \beta_N(\bar{X}_N - \bar{x}_{(n)N}),$$

where β_N is the neutrosophic regression coefficient of y_N on x_N. The bias and MSE of the estimator t_{lr_N} are given by

$$Bias(t_{lr_N}) = 0,$$
$$MSE(t_{lr_N}) = \bar{Y}_N^2 V_{0N} + \beta_N^2 \bar{X}_N^2 V_{1N} - 2\beta_N \bar{Y}_N \bar{X}_N V_{01N}.$$

The minimum MSE of the estimator t_{lr_N} at the optimum value of $\beta_{N(opt)} = \bar{Y}_N V_{01N}/\bar{X}_N V_{1N}$ is given by

$$min.MSE(t_{lr_N}) = \bar{Y}_N^2 \left(V_{0N} - \frac{V_{01N}^2}{V_{1N}}\right).$$

Referring to the work of [23], we prescribe the neutrosophic logarithmic estimator for $\bar{Y}_N$ under NRSS as

$$t_{bk_N} = \bar{y}_{[n]N} \left(1 + \log \frac{\bar{x}_{(n)N}}{\bar{X}_N}\right)^{\eta_N},$$

where η_N is a suitably chosen scalar. The bias and MSE of the estimator t_{bk_N} are given as

$$Bias(t_{bk_N}) = \bar{Y}_N \left(1 - \frac{V_{1N}}{2} + V_{01N}\right),$$
$$MSE(t_{bk_N}) = \bar{Y}_N^2 (V_{0N} + V_{1N} + 2V_{01N}).$$

The minimum MSE of the estimator t_{bk_N} at the optimum value of $\eta_{N(opt)} = -V_{01N}/V_{1N}$ is given by

$$min.MSE(t_{bk_N}) = \bar{Y}_N^2 \left(V_{0N} - \frac{V_{01N}^2}{V_{1N}}\right).$$

Inspired by [19], [24] utilized different neutrosophic auxiliary information and introduced some ratio-type estimators under NRSS as

$$t_{vs_{1N}} = \bar{y}_{[n]N} \left(\frac{\bar{X}_N + C_{x_N}}{\bar{x}_{(n)N} + C_{x_N}}\right),$$
$$t_{vs_{2N}} = \bar{y}_{[n]N} \left(\frac{\bar{X}_N + \beta_2(x_N)}{\bar{x}_{(n)N} + \beta_2(x_N)}\right),$$
$$t_{vs_{3N}} = \bar{y}_{[n]N} \left(\frac{\beta_2(x_N)\bar{X}_N + C_{x_N}}{\beta_2(x_N)\bar{x}_{(n)N} + C_{x_N}}\right),$$
$$t_{vs_{4N}} = \bar{y}_{[n]N} \left(\frac{C_{x_N}\bar{X}_N + \beta_2(x_N)}{C_{x_N}\bar{x}_{(n)N} + \beta_2(x_N)}\right).$$

The bias and MSE of the estimators $t_{vs_{iN}},\ i = 1, 2, 3, 4$, are given by

$$Bias(t_{vs_{iN}}) = \bar{Y}_N \lambda_{iN}(\lambda_{iN} V_{1N} - V_{01N}),$$

$$MSE(t_{vs_{iN}}) = \bar{Y}_N^2 \left(V_{0N} + \lambda_{iN}^2 V_{1N} - 2\lambda_{iN} V_{01N}\right),$$

where $\lambda_{1N} = \bar{X}_N/(\bar{X}_N + C_{x_N})$, $\lambda_{2N} = \bar{X}_N/(\bar{X}_N + \beta_2(x_N))$, $\lambda_{3N} = \beta_2(x_N)\bar{X}_N/(\beta_2(x_N)\bar{X}_N + C_{x_N})$, and $\lambda_{4N} = C_{x_N}\bar{X}_N/(C_{x_N}\bar{X}_N + \beta_2(x_N))$.
To estimate the population mean $\bar{Y}_N$ under indeterminacy, [20] suggested the following neutrosophic estimator under NRSS as

$$t_{sk_{1N}} = \bar{y}_{[n]N}(g_{1N} + 1) + \bar{y}_{[n]N} g_{2N} \log\left(\frac{\bar{x}_{(n)N}}{\bar{X}_N}\right),$$

where g_{1N} and g_{2N} are suitably chosen scalars. The bias and MSE of the estimator $t_{sk_{1N}}$ are given by

$$Bias(t_{sk_{1N}}) = \bar{Y}_N \left\{g_{1N} + g_{2N}\left(V_{01N} - \frac{1}{2}V_{1N}\right)\right\},$$
$$MSE(t_{sk_{1N}}) = \bar{Y}_N^2 \left(L_1 + g_{1N}^2 M_1 + g_{2N}^2 N_1 + 2g_{1N} L_1 + 2g_{2N} O_1 + 2g_{1N} g_{2N} E_1\right),$$

where $L_1 = V_{0N}$, $M_1 = (1 + V_{0N})$, $N_1 = V_{1N}$, $O_1 = V_{01N}$, and $P_1 = 2V_{01N} - \frac{1}{2}V_{1N}$.
The minimum MSE of the estimator $t_{sk_{1N}}$ at the optimum values of $g_{1N} = (O_1P_1 - L_1N_1)/(M_1N_1 - P_1^2)$, $g_{2N} = (L_1P_1 - O_1M_1)/(M_1N_1 - P_1^2)$ is given as

$$min.MSE(t_{sk_{1N}}) = \bar{Y}_N^2 \left\{L_1 - \frac{M_1O_1^2 + L_1^2N_1 - 2L_1O_1P_1}{(L_1N_1 - P_1^2)}\right\}.$$

Further, to efficiently estimate the population mean $\bar{Y}_N$ under indeterminacy, [20] also suggested two more neutrosophic estimators for $\bar{Y}_N$ under NRSS as

$$t_{sk_{2N}} = g_{3N}\bar{y}_{[n]N} + g_{4N}\bar{y}_{[n]N} \exp\left(\frac{\bar{X}_N - \bar{x}_{(n)N}}{\bar{X}_N + \bar{x}_{(n)N}}\right)\left(1 + \log\frac{\bar{x}_{(n)N}}{\bar{X}_N}\right),$$
$$t_{sk_{3N}} = g_{5N}\bar{y}_{[n]N} + g_{6N}\bar{y}_{[n]N}\left(\frac{\bar{X}}{\bar{x}_{(n)N}}\right)\exp\left(\frac{\bar{X}_N - \bar{x}_{(n)N}}{\bar{X}_N + \bar{x}_{(n)N}}\right).$$

where g_{3N}, g_{4N}, g_{5N}, and g_{6N} are suitably chosen neutrosophic scalars. The bias and MSE of the estimators t_{sk_i}, $i = 2, 3$, are given by

$$Bias(t_{sk_{2N}}) = \bar{Y}_N \left\{g_{3N} + g_{4N}\left(1 + \frac{1}{2}V_{01N} - \frac{5}{8}V_{1N}\right) - 1\right\},$$
$$Bias(t_{sk_{3N}}) = \bar{Y}_N \left\{g_{5N} + g_{6N}\left(1 + \frac{15}{8}V_{1N}^2 - \frac{3}{2}V_{01N}\right) - 1\right\},$$
$$MSE(t_{sk_{2N}}) = \bar{Y}_N^2 \left(1 + g_{3N}^2 L_2 + g_{4N}^2 M_2 + 2g_{3N}g_{4N}N_2 - 2g_{3N}O_2 - 2g_{4N}P_2\right),$$
$$MSE(t_{sk_{3N}}) = \bar{Y}_N^2 \left(1 + g_{5N}^2 L_3 + g_{6N}^2 M_3 + 2g_{5N}g_{6N}N_3 - 2g_{5N}O_3 - 2g_{6N}P_3\right),$$

where

$$L_2 = 1 + V_{0N},\ M_2 = 1 + V_{0N} - V_{1N} + 2V_{01N},\ N_2 = \left(1 + V_{0N} - \frac{5}{8}V_{1N} + V_{01N}\right),$$

$$O_2 = 1,\ P_2 = 1 - \frac{5}{8}V_{1N} + \frac{1}{2}V_{01N},\ L_3 = 1 + V_{0N},\ M_3 = (1 + V_{0N} + 6V_{1N} - 6V_{01}),$$
$$O_3 = 1,\ N_3 = \left(1 + V_{0N} + \frac{15}{8}V_{1N} - 3V_{01}\right),\ \text{and}\ P_3 = \left(1 + \frac{15}{8}V_{1N} - \frac{3}{2}V_{01}\right).$$

The minimum MSE of the estimators $t_{sk_{iN}}$, $i = 2, 3$, at the optimal values of $g_{3N} = (M_2O_2 - N_2P_2)/(L_2M_2 - N_2^2)$, $g_{4N} = (P_2L_2 - N_2O_2)/(L_2M_2 - N_2^2)$, $g_{5N} = (O_3M_3 - N_3P_3)/(L_3M_3 - N_3^2)$ and $g_{6N} = (L_3P_3 - N_3O_3)/(L_3M_3 - N_3^2)$ is given by

$$min.MSE(t_{sk_{iN}}) = \bar{Y}_N^2 \left\{1 - \frac{(L_iP_i^2 + M_iO_i^2 - 2N_iO_iP_i)}{(L_iM_i - N_i^2)}\right\}.$$

3. Proposed classes of estimators

The purpose of proposing the efficient classes of estimators under NRSS is due to the limits of the existing traditional estimators in handling the uncertainty and imprecision inherent in real-life data. The traditional estimation approaches fail to produce efficient estimates in the presence of uncertain or inconsistent data. Moreover, most of the existing estimators utilizing the NRSS framework were found to be inefficient or perform poorly due to their inability to properly utilize the additional information furnished by NRSS. We achieve improved efficiency under uncertain environments by proposing the following efficient classes of estimators utilizing NRSS.

$$t_{a_{1N}} = \left[\varphi_{1N}\bar{y}_{[n]N} + \delta_{1N}\bar{y}_{[n]N}\left\{\frac{\bar{X}_N}{\theta\bar{x}_{(n)N} + (1-\theta)\bar{X}_N}\right\}^g\right]\left(1 + \log\frac{\bar{x}_{(n)N}}{\bar{X}_N}\right)^{\eta_N},$$
$$t_{a_{2N}} = \left\{\varphi_{2N}\bar{y}_{[n]N} + \delta_{2N}\left(\bar{x}_{(n)N} - \bar{X}_N\right)\right\}\left(1 + \log\frac{\bar{x}_{(n)N}}{\bar{X}_N}\right)^{\eta_N},$$

where $\varphi_{iN},\ i = 1, 2, \delta_{iN}$, and η_N are the suitably chosen neutrosophic scalars. Also, g and θ are constants taking real values to generate different estimators.

Theorem 3.1. *The bias, MSE, and minimum MSE of the proposed estimators $t_{a_{iN}},\ i = 1, 2$ at the optimum values $\varphi_{iN(opt)} = (B_{iN}D_{iN} - C_{iN}E_{iN})/(A_{iN}B_{iN} - C_{iN}^2)$ and $\delta_{iN(opt)} = (A_{iN}E_{iN} - C_{iN}D_{iN})/(A_{iN}B_{iN} - C_{iN}^2)$ are given up to first-order approximation as follows:*

$$Bias(t_{a_{iN}}) = \bar{Y}_N(\varphi_{iN}D_{iN} + \delta_{iN}E_{iN} - 1),$$
$$MSE(t_{a_{iN}}) = \bar{Y}_N^2\left(1 + \varphi_{iN}^2A_{iN} + \delta_{iN}^2B_{iN} + 2\varphi_{iN}\delta_{iN}C_{iN} - 2\varphi_{iN}D_{iN} - 2\delta_{iN}E_{iN}\right),$$
$$min.MSE(t_{a_{iN}}) = \bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\},$$

where

$$A_{1N} = 1 + V_{0N} + 2\eta_N(\eta_N - 1)V_{1N} + 4\eta_NV_{01N},$$
$$B_{1N} = 1 + V_{0N} + \left(2\eta_N^2 - 2\eta_N + g^2\theta^2 + g(g+1)\theta^2 - 4\eta_Ng\theta\right)V_{1N} + 4(\eta_N - g\theta)V_{01N},$$
$$C_{1N} = 1 + V_{0N} + \left(2\eta_N^2 - 2\eta_Ng\theta - 2\eta_N + \tfrac{g(g+1)}{2}\theta^2\right)V_{1N} + 2(2\eta_N - g\theta)V_{01N},$$

$$D_{1N} = 1 + \left(\frac{\eta_N^2}{2} - \eta_N\right) V_{1N} + \eta_N V_{01N},$$

$$E_{1N} = 1 + \left(\frac{\eta_N^2}{2} - \eta_N - \eta_N g\theta + \frac{g(g+1)}{2}\theta^2\right) V_{1N} + (\eta_N - g\theta) V_{01N},$$

$$A_{2N} = 1 + V_{0N} + 2\eta_N(\eta_N - 1)V_{1N} + 4\eta_N V_{01N},$$

$$B_{2N} = \frac{V_{1N}}{R_N^2},$$

$$C_{2N} = \frac{1}{R_N}(2\eta_N V_{1N} + V_{01N}),$$

$$D_{2N} = 1 + \left(\frac{\eta_N^2}{2} - \eta_N\right) V_{1N} + \eta_N V_{01N},$$

$$E_{2N} = \frac{\eta_N V_{1N}}{R_N},$$

$$R_N = \frac{\bar{Y}_N}{\bar{X}_N}.$$

Proof. Refer to Appendix A. □

4. Analytical comparisons

Analytical comparisons of the MSEs of proposed and available estimators under NRSS are crucial for determining their relative efficiency. Such comparisons produce a clear quantitative estimate of the accuracy and precision of the estimators under different settings. The proposed estimators may outperform by exhibiting lower MSEs, which validates their theoretical benefits and practical usability. The comparison also illustrates the conditions in which the proposed estimators excel, providing useful insights for opting the best estimator for real-life applications, especially while handling uncertainty in the context of NRSS.

i. The proposed estimators $t_{a_{iN}}$, $i = 1, 2$, outperform the neutrosophic sample mean $\bar{y}_{[n]N}$, when

$$min.MSE(t_{a_{iN}}) < V(\bar{y}_{[n]N})$$

$$\bar{Y}_N^2 \left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2 V_{0N}$$

$$1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} < V_{0N}$$

$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > 1 - V_{0N}.$$

ii. The proposed estimators $t_{a_{iN}}$, $i = 1, 2$, outperform the neutrosophic ratio estimator t_{r_N}, when

$$min.MSE(t_{a_{iN}}) < MSE(t_{r_N})$$

$$\bar{Y}_N^2 \left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2 (V_{0N} + V_{1N} - 2V_{01N})$$

$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > 1 - (V_{0N} + V_{1N} - 2V_{01N}).$$

iii. The proposed estimators $t_{a_{iN}},\ i = 1, 2$, outperform the neutrosophic regression estimator t_{lr_N}, when

$$min.MSE(t_{a_{iN}}) < min.MSE(t_{lr_N})$$
$$\bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2\left(V_{0N} - \frac{V_{01N}^2}{V_{1N}}\right)$$
$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > 1 - \left(V_{0N} - \frac{V_{01N}^2}{V_{1N}}\right).$$

iv. The proposed estimators $t_{a_{iN}},\ i = 1, 2$, outperform the neutrosophic estimators $t_{vs_{iN}},\ i = 1, 2, 3, 4$, envisaged by [19], when

$$min.MSE(t_{a_{iN}}) < MSE(t_{vs_{iN}})$$
$$\bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2\left(V_{0N} + \lambda_{iN}^2 V_{1N} - 2\lambda_{iN}V_{01N}\right)$$
$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > 1 - \left(V_{0N} + \lambda_{iN}^2 V_{1N} - 2\lambda_{iN}V_{01N}\right).$$

v. The proposed estimators $t_{a_{iN}},\ i = i = 1, 2$, outperform the neutrosophic estimator $t_{sk_{1N}}$ envisaged by [20], when

$$min.MSE(t_{a_{iN}}) < min.MSE(t_{sk_{1N}})$$
$$\bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2\left\{L_1 - \frac{M_1O_1^2 + L_1^2N_1 - 2L_1O_1P_1}{(L_1N_1 - P_1^2)}\right\}$$
$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > 1 - \left\{L_1 - \frac{M_1O_1^2 + L_1^2N_1 - 2L_1O_1P_1}{(L_1N_1 - P_1^2)}\right\}.$$

vi. The proposed estimators $t_{a_{iN}},\ i = 1, 2$, outperform the neutrosophic estimator $t_{sk_{2N}}$ envisaged by [20] when

$$min.MSE(t_{a_{iN}}) < min.MSE(t_{sk_{2N}})$$
$$\bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2\left\{1 - \frac{L_2P_2^2 + M_2O_2^2 - 2N_2O_2P_2}{(L_2M_2 - N_2^2)}\right\}$$
$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > \frac{L_2P_2^2 + M_2O_2^2 - 2N_2O_2P_2}{(L_2M_2 - N_2^2)}.$$

vii. The proposed estimators $t_{a_{iN}},\ i = 1, 2$, outperform the neutrosophic estimator $t_{sk_{3N}}$ envisaged by [20], when

$$min.MSE(t_{a_{iN}}) < min.MSE(t_{sk_{3N}})$$
$$\bar{Y}_N^2\left\{1 - \frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)}\right\} < \bar{Y}_N^2\left\{1 - \frac{L_3P_3^2 + M_3O_3^2 - 2N_3O_3P_3}{(L_3M_3 - N_3^2)}\right\}$$
$$\frac{(A_{iN}E_{iN}^2 + B_{iN}D_{iN}^2 - 2C_{iN}D_{iN}E_{iN})}{(A_{iN}B_{iN} - C_{iN}^2)} > \frac{L_3P_3^2 + M_3O_3^2 - 2N_3O_3P_3}{(L_3M_3 - N_3^2)}.$$

The above efficiency conditions will further be evaluated through a simulation study and a real data illustration.

4.1. Simulation study

In this section, a simulation study is conducted to evaluate the performance of the neutrosophic estimators by artificially generating a population from neutrosophic normal (NN) distribution using R-Software. For this population, an artificial dataset of size N=1000 was constructed by assuming that $x_N \sim NN([14,18],[49,81])$ and $y_N \sim NN([10,12],[36,64])$ with different correlations among auxiliary and study variables $\rho_{xy_N} \in [(0.2,0.3),(0.4,0.5),(0.6,0.7),(0.8,0.9)]$. From this population, a neutrosophic ranked set sample of size $n_N = (12,12)$ is drawn with set size $m_N = [3,3]$ and number of cycles $r = 4$. The necessary descriptive statistics are computed for this dataset. With 5000 iterations, the simulated dataset enabled the estimation of $MSE_N \sim [MSE_L, MSE_U]$ along with relative efficiency (RE) $RE_N \sim [RE_L, RE_U]$, indicating the comparative performance of the proposed and existing neutrosophic estimators. The $MSE_N \sim [MSE_L, MSE_U]$ and $RE_N \sim [RE_L, RE_U]$ are computed by utilizing the following formulas:

$$MSE_N(T^*) = \frac{1}{5,000}\sum_{i=1}^{5,000}(T_i^* - \bar{Y}_N)^2, \tag{4.1}$$

$$RE_N = \frac{MSE_N(\bar{y}_{[n]N})}{MSE_N(T^*)}. \tag{4.2}$$

where $T^* = \bar{y}_{[n]N},\ t_{r_N},\ t_{lr_N},\ t_{bk_N},\ t_{vs_{iN}},\ i = 1,2,3,4,\ t_{sk_{iN}},\ i = 1,2,3$, and $t_{a_{iN}},\ i = 1,2$.

The simulated MSE_N and RE_N are reported in Tables 1 and 2, respectively, for different values of correlation coefficient ρ_{xy_N}. These MSE_N and RE_N values are presented by bar diagrams in Figures 1 and 2, respectively.

Table 1. $MSE_N \in [MSE_L, MSE_U]$ of neutrosophic estimators using neutrosophic normal population.

	ρ_{xy_N}			
Estimators	(0.2, 0.3)	(0.4, 0.5)	(0.6, 0.7)	(0.8, 0.9)
$\bar{y}_{[n]N}$	(3.090, 5.499)	(3.086, 5.442)	(3.044, 5.344)	(2.973, 5.183)
t_{r_N}	(4.221, 6.874)	(3.945, 6.368)	(3.622, 5.817)	(3.278, 5.225)
t_{lr_N}	(2.796, 4.943)	(2.742, 4.772)	(2.620, 4.504)	(2.440, 4.127)
t_{bk_N}	(2.796, 4.943)	(2.742, 4.772)	(2.620, 4.504)	(2.440, 4.127)
$t_{vs_{1N}}$	(4.124, 6.773)	(3.859, 6.279)	(3.546, 5.741)	(3.210, 5.159)
$t_{vs_{2N}}$	(4.430, 7.085)	(4.132, 6.552)	(3.789, 5.976)	(3.424, 5.361)
$t_{vs_{3N}}$	(5.281, 7.582)	(4.801, 6.554)	(4.205, 6.191)	(4.298, 6.837)
$t_{vs_{4N}}$	(4.708, 7.347)	(4.388, 6.785)	(4.018, 6.177)	(3.626, 5.532)
$t_{sk_{1N}}$	(2.702, 4.735)	(2.651, 4.579)	(2.537, 4.334)	(2.369, 3.986)
$t_{sk_{2N}}$	(2.689, 4.708)	(2.634, 4.543)	(2.516, 4.292)	(2.345, 3.940)
$t_{sk_{3N}}$	(2.689, 4.708)	(2.634, 4.543)	(2.516, 4.292)	(2.345, 3.940)
$t_{a_{1N}}$	(2.685, 4.697)	(2.629, 4.531)	(2.510, 4.277)	(2.339, 3.924)
$t_{a_{2N}}$	(2.686, 4.699)	(2.630, 4.533)	(2.511, 4.279)	(2.340, 3.925)

Table 2. $RE_N \in [RE_L, RE_U]$ of neutrosophic estimators using neutrosophic normal population.

Estimators	ρ_{xy_N} (0.2, 0.3)	(0.4, 0.5)	(0.6, 0.7)	(0.8, 0.9)
$\bar{y}_{[n]N}$	(1.000, 1.000)	(1.000, 1.000)	(1.000, 1.000)	(1.000, 1.000)
t_{r_N}	(0.731, 0.799)	(0.782, 0.854)	(0.840, 0.918)	(0.907, 0.991)
t_{lr_N}	(1.105, 1.112)	(1.125, 1.140)	(1.161, 1.186)	(1.218, 1.255)
t_{bk_N}	(1.105, 1.112)	(1.125, 1.140)	(1.161, 1.186)	(1.218, 1.255)
$t_{vs_{1N}}$	(0.749, 1.168)	(0.799, 0.866)	(0.858, 0.930)	(0.926, 1.004)
$t_{vs_{2N}}$	(0.697, 0.776)	(0.746, 0.830)	(0.803, 0.894)	(0.868, 0.966)
$t_{vs_{3N}}$	(0.585, 0.725)	(0.642, 0.830)	(0.723, 0.863)	(0.691, 0.758)
$t_{vs_{4N}}$	(0.656, 0.748)	(0.703, 0.802)	(0.757, 0.865)	(0.820, 0.936)
$t_{sk_{1N}}$	(1.143, 1.161)	(1.164, 1.188)	(1.199, 1.233)	(1.254, 1.300)
$t_{sk_{2N}}$	(1.148, 1.168)	(1.171, 1.197)	(1.209, 1.245)	(1.267, 1.315)
$t_{sk_{3N}}$	(1.148, 1.168)	(1.171, 1.197)	(1.209, 1.245)	(1.267, 1.315)
$t_{a_{1N}}$	(1.151, 1.171)	(1.174, 1.201)	(1.213, 1.251)	(1.271, 1.321)
$t_{a_{2N}}$	(1.150, 1.170)	(1.173, 1.201)	(1.212, 1.250)	(1.271, 1.320)

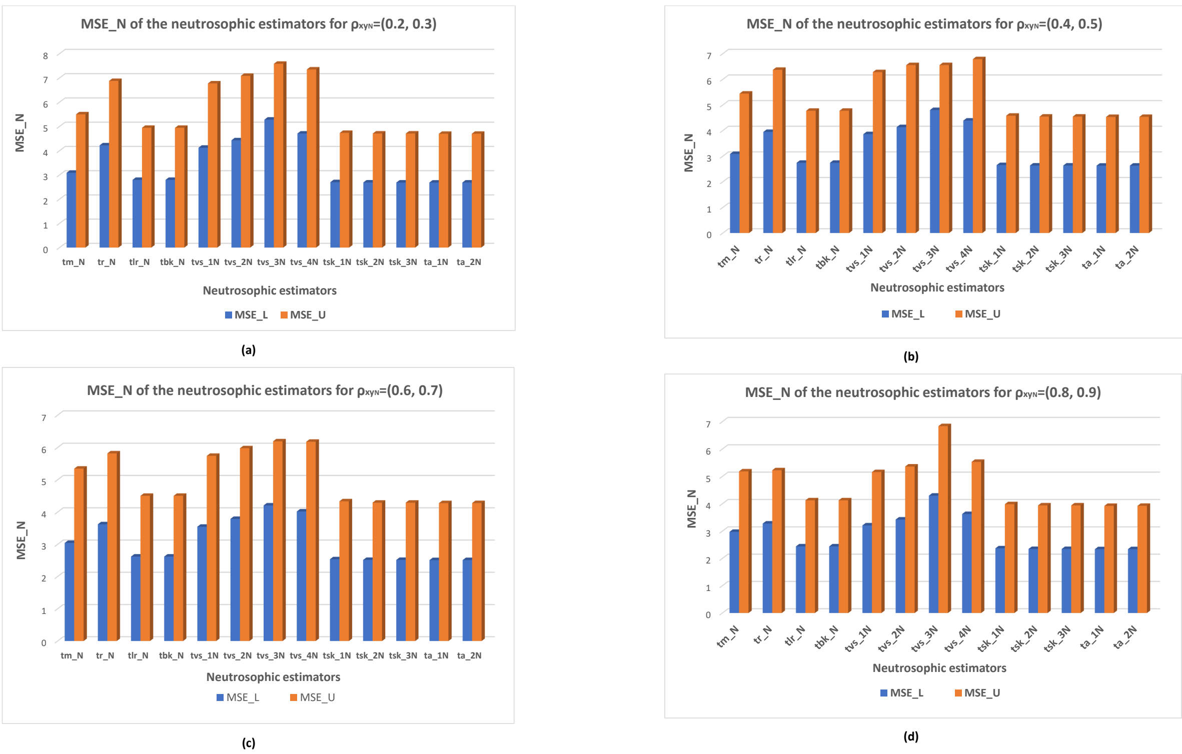

Figure 1. Diagrams of the neutrosophic estimators for the MSE_N of Table 1.

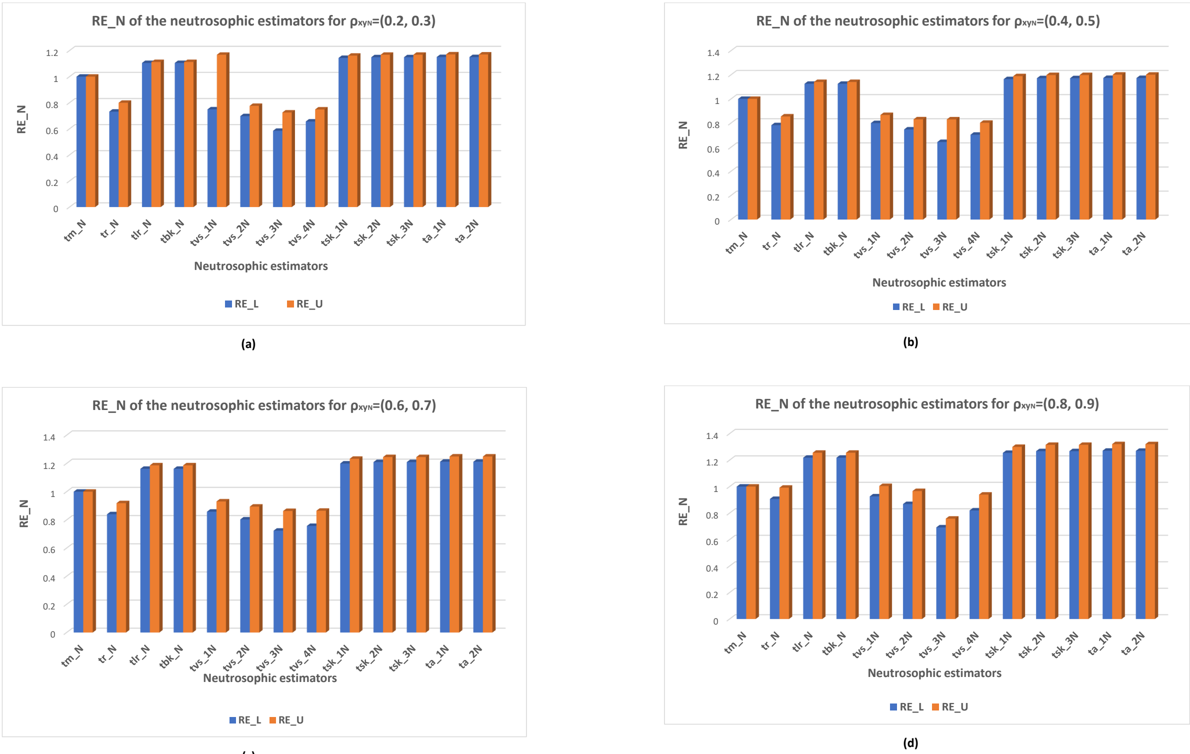

Figure 2. Diagrams of the neutrosophic estimators for the RE_N of Table 2.

4.2. *Real data applications*

Executing real-world data applications in any study greatly increases its relevance and effect. The use of real-world data helps to evaluate the theoretical results in practical contexts and demonstrates their usefulness and robustness under realistic conditions. In this study, we have used two different real-world indeterminate data sets.

Stock price data is related to indeterminate data due to the uncertainty and imprecision associated with the financial markets. The data set 1 is based on the daily price of the stock 'Moderna', which can be obtained from the publicly available website https://finance.yahoo.com/quote/MRNA/history/. The neutrosophic survey variable $y_N \in [y_L,\ y_U]$ is the varying price of the stock on each day from 1 September 2020 to 1 September 2021, while the neutrosophic auxiliary variable $x_N \in [x_L,\ x_U]$ is the varying price of the stock on each day from 1 September 2019 to 1 September 2020.

The sample registration system (SRS) offers accurate yearly estimates for natural growth rate, birth rate, death rate, and other fertility and mortality indicators at national and subnational levels. Every year, the office of the Registrar General of India (RGI) conducts a large-scale demographic census throughout all states and union territories. The data set 2 is based on the SRS bulletin 2020, which can be obtained from the publicly available website https://censusindia.gov.in/nada/index.php/catalog/42687. In this data, the neutrosophic survey variable $y_N \in [y_L,\ y_U]$ is the natural growth rate for the year 2020 for India, while the neutrosophic auxiliary variable $x_N \in [x_L,\ x_U]$ is the birth rate for the year 2020 for India.

The descriptive values of both data sets are given in Table 3. Utilizing these descriptive values, the MSE_N and RE_N are calculated for both data sets and reported in Table 4. The RE_N is calculated using the following formula:

$$RE_N = \frac{MSE(\bar{y}_{[n]N})}{MSE(T^*)}. \tag{4.3}$$

Table 3. Descriptive values of datasets 1 and 2.

Neutrosophic parameters	Dataset 1	Dataset 2
N	252	36
n_N	(12, 12)	(12, 12)
$\bar{Y}_N$	(162.973, 174.228)	(9.755, 11.766)
$\bar{X}_N$	(35.940, 38.762)	(14.708, 17.938)
C_{y_N}	(0.547, 0.559)	(0.353, 0.372)
C_{x_N}	(0.603, 0.607)	(0.225, 0.239)
ρ_{xy_N}	(0.829, 0.804)	(0.965, 0.955)
$\beta_1(x_N)$	(0.642, 0.622)	(0.675, 0.552)
$\beta_2(x_N)$	(1.778, 1.817)	(2.591, 2.239)

Table 4. MSE_N and RE_N of the neutrosophic estimators for real datasets.

	Dataset 1		Dataset 2	
Estimators	$(MSE_L,\ MSE_U)$	$(RE_L,\ RE_U)$	$(MSE_L,\ MSE_U)$	$(RE_L,\ RE_U)$
$\bar{y}_{[n]N}$	(662.606, 789.791)	(1.000, 1.000)	(0.975, 1.562)	(1.000, 1.000)
t_{r_N}	(122.853, 200.171)	(5.393, 3.945)	(0.173, 0.244)	(5.608, 6.382)
t_{lr_N}	(118.482, 190.140)	(5.592, 4.153)	(0.067, 0.084)	(14.421, 18.578)
t_{bk_N}	(118.482, 190.140)	(5.592, 4.153)	(0.067, 0.084)	(14.421, 18.578)
$t_{vs_{1N}}$	(121.274, 197.652)	(5.463, 3.995)	(0.180, 0.253)	(5.413, 6.162)
$t_{vs_{2N}}$	(119.278, 193.859)	(5.555, 4.074)	(0.243, 0.325)	(3.998, 4.800)
$t_{vs_{3N}}$	(121.916, 198.731)	(5.434, 3.974)	(0.176, 0.248)	(5.531, 6.282)
$t_{vs_{4N}}$	(118.509, 191.542)	(5.591, 4.123)	(0.428, 0.546)	(2.274, 2.861)
$t_{sk_{1N}}$	(117.842, 189.971)	(5.622, 4.157)	(0.064, 0.078)	(15.121, 19.990)
$t_{sk_{2N}}$	(114.796, 183.967)	(5.772, 4.293)	(0.067, 0.083)	(14.441, 18.616)
$t_{sk_{3N}}$	(114.796, 183.967)	(5.772, 4.293)	(0.067, 0.083)	(14.441, 18.616)
$t_{a_{1N}}$	(111.247, 179.592)	(5.956, 4.397)	(0.064, 0.078)	(15.234, 20.025)
$t_{a_{2N}}$	(88.456, 156.815)	(7.490, 5.036)	(0.063, 0.077)	(15.725, 20.037)

From the results of Table 4, it can be seen that the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, perform better than the neutrosophic mean estimator $\bar{y}_{[n]N}$, the neutrosophic ratio estimator t_{r_N}, the neutrosophic regression estimator t_{lr_N}, the neutrosophic logarithmic estimator t_{bk_N}, the neutrosophic ratio type estimators $t_{vs_{iN}},\ i = 1, 2, 3, 4$, suggested by [19], and the [20] estimators by minimum MSE_N and maximum RE_N.

5. Interpretation of results

The efficiency conditions obtained under analytical comparison show that the proposed neutrosophic estimators dominate the existing neutrosophic estimators. These conditions have been evaluated through a simulation study carried out on artificially generated uncertain data. Furthermore, these conditions are evaluated by two real-life applications: One involving the analysis of stock market data to evaluate the financial trends and decision-making processes and the other utilizing the demographic data to examine patterns and insights related to population dynamics and social behaviors. These results are interpreted in the following points:

(i) Table 1 contains the simulated MSE_N of the neutrosophic estimators, from which it can be seen that the MSE_N of the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, is minimum compared to the neutrosophic ratio estimator t_{r_N}, the neutrosophic regression estimator t_{lr_N}, the neutrosophic logarithmic estimator t_{bk_N}, [19] estimators $t_{vs_{iN}},\ i = 1, 2, 3, 4$, and [20] estimators $t_{sk_{iN}},\ i = 1, 2, 3$, for passably chosen values of ρ_{xy_N}. This shows the outperformance of the proposed class of neutrosophic estimators over the existing neutrosophic estimators. Moreover, the MSE_N values of the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, decrease as the correlation coefficient ρ_{xy_N} increases. This dominance can easily be seen from Figure 1.

(ii) The simulated RE_N of the neutrosophic estimators is given in Table 2, which shows that the RE_N of the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, is higher compared to the neutrosophic ratio estimator t_{r_N}, the neutrosophic regression estimator t_{lr_N}, the neutrosophic logarithmic estimator t_{bk_N}, the modified neutrosophic ratio estimators $t_{vs_{iN}},\ i = 1, 2, 3, 4$, proposed by [19], and the [20] estimators $t_{sk_{iN}},\ i = 1, 2, 3$, for passably chosen values of ρ_{xy_N}. This shows the outperformance of the proposed class of neutrosophic estimators over the existing neutrosophic estimators. Additionally, the RE_N values of the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, increase as the correlation coefficient ρ_{xy_N} increases. This dominance can easily be seen from Figure 2.

(iii) Lastly, the real data findings of the neutrosophic estimators are given by MSE_N and RE_N in Table 4 for both datasets. From Table 4, for both datasets, the proposed neutrosophic estimators $t_{a_{iN}},\ i = 1, 2$, obtain the least MSE_N and the highest RE_N compared to the existing neutrosophic estimators. This shows that the proposed class of neutrosophic estimators is more efficient than the existing neutrosophic estimators.

6. Conclusions

Neutrosophic ranked set sampling is an extension of the traditional ranked set sampling. By introducing neutrosophic logic into the RSS framework, we effectively addressed the inherent uncertainties and imprecisions in ranking units that frequently occur in real-world scenarios. It is particularly useful when ranking errors occur due to vague or incomplete information, enhancing sampling efficiency in uncertain environments.

In this article, we proposed some efficient classes of estimators for estimating the indeterminate population mean using NRSS. The bias and MSE expressions of the proposed estimators were reported up to first-order approximation. The analytical comparisons of the proposed estimators with the existing competitors demonstrated the outperformance of the proposed estimators. The theoretical results were exemplified with the simulation study and real data illustrations. The simulation and real data findings showed that the proposed neutrosophic estimators outperform the existing neutrosophic estimators in terms of reduced MSE_N and maximum RE_N when dealing with uncertain data. The strength of this study is that the proposed estimators established a substantial improvement in the statistical techniques for indeterminate population mean estimation, providing a valuable tool for survey practitioners.

Future research might broaden the proposed work by investigating different scenarios of neutrosophy and applying it to real-world case studies in areas including environmental science, health, and social sciences.

Author contributions

Anoop Kumar: Methodology, simulation study, writing-review and editing, supervision; Priya: Writing-original manuscript, software; Abdullah Mohammed Alomair: Project administration, financial support. All authors have read and agreed to the published version of the manuscript.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgements

The authors are extremely grateful to the panel of reviewers and the Editor-in-Chief for their careful insights and constructive suggestions, which led to considerable improvement of the paper.

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. KFU251241].

Conflict of interest

The authors declare no competing interests.

References

1. W. G. Cochran, The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce, *J. Agric. Sci.*, **30** (1940), 262–275. https://doi.org/10.1017/S0021859600048012
2. M. N. Murthy, Product method of estimation, *Sankhya Indian J. St. Ser. A*, **26** (1964), 69–74. Available from: `https://www.jstor.org/stable/25049308`.
3. B. V. S. Sisodia, V. K. Dwivedi, A modified ratio estimator using coefficient of variation of auxiliary variable, *J. Indian Soc. Agri. Stat.*, **33** (1981), 13–18.
4. L. N. Upadhyaya, H. P. Singh, Use of transformed auxiliary variable in estimating the finite population mean, *Biometrical J.*, **41** (1999), 627–636. https://doi.org/10.1002/(SICI)1521-4036(199909)41:5<3C627::AID-BIMJ627>3E3.0.CO;2-W

5. S. Hussain, S. Ahmad, M. Saleem, S. Akhtar, Finite population distribution function estimation with dual use of auxiliary information under simple and stratified random sampling, *PLoS One*, **15** (2020), e0239098. https://doi.org/10.1371/journal.pone.0239098
6. S. Ahmad, S. Hussain, M. Aamir, U. Yasmeen, J. Shabbir, Z. Ahmad, Dual use of auxiliary information for estimating the finite population mean under the stratified random sampling scheme, *J. Math.*, **2021** (2021), 3860122. https://doi.org/10.1155/2021/3860122
7. N. Ali, I. Ahmad, M. Hanif, U. Shahzad, Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information, *Commun. Stat.-Theor. M.*, **50** (2021), 979–992. https://doi.org/10.1080/03610926.2019.1645857
8. J. Shabbir, S. Ahmed, A. Sanaullah, R. Onyango, Measuring performance of ratio-exponential-log type general class of estimators using two auxiliary variables, *Math. Probl. Eng.*, **2021** (2021), 5245621. https://doi.org/10.1155/2021/5245621
9. U. Shahzad, M. Hanif, I. Sajjad, M. M. Anas, Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information, *Sci. Iran.*, **29** (2022), 1705–1715. https://doi.org/10.24200/sci.2020.54423.3744
10. A. Kumar, A. S. Siddiqui, Enhanced estimation of population mean using simple random sampling, *Res. Stat.*, **2** (2024), 2335949. https://doi.org/10.1080/27684520.2024.2335949
11. A. Kumar, S. Bhushan, W. Emam, Y. Tashkandy, M. J. S. Khan, Novel logarithmic imputation procedures using multi auxiliary information under ranked set sampling, *Sci. Rep.*, **14** (2024), 18027. https://doi.org/10.1038/s41598-024-68940-4
12. F. Smarandache, Neutrosophic set a generalization of the intuitionistic fuzzy set, *Int. J. Pure Appl. Math.*, **24** (2005), 287.
13. Z. Tahir, H. Khan, M. Aslam, J. Shabbir, Mahmood, F. Smarandache, Neutrosophic ratio-type estimators for estimating the population mean, *Complex Intell. Syst.*, **7** (2021), 2991–3001. https://doi.org/10.1007/s40747-021-00439-1
14. S. K. Yadav, F. Smarandache, *Generalized neutrosophic sampling strategy for elevated estimation of population mean*, Infinite Study, 2023.
15. V. K. Yadav, S. Prasad, Neutrosophic estimators in two-phase survey sampling, *Neutrosophic Sets Sy.*, **61** (2023), 29. Available from: `https://digitalrepository.unm.edu/nss_journal/vol61/iss1/29`.
16. V. K. Yadav, S. Prasad, Neutrosophic estimators for estimating the population mean in survey sampling, *Meas.-Interdiscip. Res.*, **22** (2024), 373–397. https://doi.org/10.1080/15366367.2023.2267835
17. M. A. Alqudah, M. Zayed, M. Subzar, S. A. Wani, Neutrosophic robust ratio type estimator for estimating finite population mean, *Heliyon*, **10** (2024), 1–12. https://doi.org/10.1016/j.heliyon.2024.e28934
18. S. Masood, B. Ibrar, J. Shabbir, A. Shokri, Z. Movaheedi, Estimating neutrosophic finite median employing robust measures of the auxiliary variable, *Sci. Rep.*, **14** (2024), 10255. https://doi.org/10.1038/s41598-024-60714-2

19. G. K. Vishwakarma, A. Singh, Generalized estimator for computation of population mean under neutrosophic ranked set technique: An application to solar energy data, *J. Comput. Appl. Math.*, **41** (2022), 144. https://doi.org/10.1007/s40314-022-01820-7

20. R. Singh, A. Kumari, Neutrosophic ranked set sampling scheme for estimating population mean: An application to demographic data, *Neutrosophic Sets Sy.*, **68** (2024), 246–270.

21. H. M. Samawi, H. A. Muttlak, Estimation of ratio using rank set sampling, *Biometrical J.*, **38** (1996), 753–764. https://doi.org/10.1002/bimj.4710380616

22. P. L. H. Yu, K. Lam, Regression estimator in ranked set sampling, *Biometrics*, **53** (1997), 1070–1080. https://doi.org/10.2307/2533564

23. S. Bhushan, A. Kumar, *Log type estimators of population mean under ranked set sampling, predictive analytics using statistics and big data: Concepts and modeling*, Bentham Science Publisher, **28** (2020), 47–74. https://doi.org/10.2174/97898114904911200100007

24. V. L. Mandowara, N. Mehta, Efficient generalized ratio-product type estimators for finite population mean with ranked set sampling, *Aust. J. Stat.*, **42** (2013), 137–148. https://doi.org/10.17713/ajs.v42i3.147

Appendix A

Proof of Theorem 3.1.

Proof. Take the proposed estimator $t_{a_{1N}}$ as

$$t_{a_{1N}} = \left[\varphi_{1N}\bar{y}_{[n]N} + \delta_{1N}\bar{y}_{[n]N}\left\{\frac{\bar{X}_N}{\theta\bar{x}_{(n)N} + (1-\theta)\bar{X}_N}\right\}^g\right]\left(1 + \log\frac{\bar{x}_{(n)N}}{\bar{X}_N}\right)^{\eta_N}.$$

Using the notations defined in (1.1), we rewrite the proposed estimator $t_{a_{1N}}$ as

$$\begin{aligned} t_{a_{1N}} &= \bar{Y}_N(1+e_{0N})\left\{\left(\varphi_{1N} + \delta_{1N}\left[\tfrac{\bar{X}_N}{\theta\bar{X}_N(1+e_{1N})+(1-\theta)\bar{X}_N}\right]^g\right)\times\left[1+\log\tfrac{\bar{X}_N(1+e_{1N})}{\bar{X}_N}\right]^{\eta_N}\right\} \\ &= \bar{Y}_N(1+e_{0N})\left[\left\{\varphi_{1N} + \delta_{1N}\left(\tfrac{1}{\theta+\theta e_{1N}+1-\theta}\right)^g\right\}\times\{1+\log(1+e_{1N})\}^{\eta_N}\right] \\ &= \bar{Y}_N(1+e_{0N})\{\varphi_{1N} + \delta_{1N}(1+\theta e_{1N})^{-g}\}\{1+\log(1+e_{1N})\}^{\eta_N}. \end{aligned}$$

Expanding the right side expression, multiplying out and neglecting the error terms with power greater than two, we get

$$t_{a_{1N}} = \bar{Y}_N(1+e_{0N})\left\{\begin{array}{l}\left[\varphi_{1N} + \delta_{1N}\left(1 - g\theta e_{1N} + \frac{g(g+1)}{2}\theta_N^2 e_{1N}^2\right)\right] \\ \left[1 + \left(\eta_N e_{1N} - \eta_N e_{1N}^2 + \frac{\eta_N^2}{2}e_{1N}^2\right)\right]\end{array}\right\}.$$

Simplifying and subtracting $\bar{Y}_N$ both side, we get

$$t_{a_{1N}} - \bar{Y}_N = \bar{Y}_N\left[\begin{array}{l}\varphi_{1N}\left\{1 + e_{0N} + \eta_N e_{1N} + \left(\frac{\eta_N^2}{2} - \eta_N\right)e_{1N}^2 + \eta_N e_{0N}e_{1N}\right\} + \\ \delta_{1N}\left\{\begin{array}{l}1 + e_{0N} + (\eta_N - g\theta)e_{1N} \\ +\left(\frac{\eta_N^2}{2} - \eta_N - \eta_N g\theta + \frac{g(g+1)}{2}\theta^2\right)e_{1N}^2 \\ +(\eta_N - g\theta)e_{0N}e_{1N}\end{array}\right\} - 1\end{array}\right]. \tag{6.1}$$

Taking expectation both side to (6.1), we get

$$Bias(t_{a_{1N}}) = \bar{Y}_N \left[\begin{array}{l} \varphi_{1N} \left\{ 1 + (\frac{\eta_N^2}{2} - \eta_N) V_{0N} + \eta_N V_{01N} \right\} \\ +\delta_{1N} \left\{ \begin{array}{l} 1 + \left(\frac{\eta_N^2}{2} - \eta_N - \eta_N g\theta + \frac{g(g+1)}{2}\theta^2 \right) V_{1N} \\ +(\eta_N - g\theta) V_{01N} \end{array} \right\} - 1 \end{array} \right].$$

Again, squaring both side to (6.1) and taking expectation, we get the MSE of the proposed class of estimators

$$\begin{aligned} MSE(t_{a_{1N}}) &= \bar{Y}_N^2 \left[\begin{array}{l} 1 + \varphi_{1N}^2 \{1 + V_{0N} + 2\eta_N(\eta_N - 1)V_{1N} + 4\eta_N V_{01N}\} \\ +\delta_{1N}^2 \left\{ \begin{array}{l} 1 + V_{0N} + \left(\begin{array}{l} 2\eta_N^2 - 2\eta_N + g^2\theta^2 \\ +g(g+1)\theta^2 - 4\eta_N g\theta \end{array} \right) V_{1N} \\ +4(\eta_N - g\theta)V_{01N} \end{array} \right\} \\ +2\varphi_{1N}\delta_{1N} \left\{ \begin{array}{l} 1 + V_{0N} + \left(\begin{array}{l} 2\eta_N^2 - 2\eta_N g\theta - 2\eta_N \\ +\frac{g(g+1)}{2}\theta^2 \end{array} \right) V_{1N} \\ +2(2\eta_N - g\theta)V_{01N} \end{array} \right\} \\ -2\varphi_{1N} \left\{ 1 + \left(\frac{\eta_N^2}{2} - \eta_N \right) V_{1N} + \eta_N V_{01N} \right\} \\ -2\delta_{1N} \left\{ \begin{array}{l} 1 + \left(\frac{\eta_N^2}{2} - \eta_N - \eta_N g\theta + \frac{g(g+1)}{2}\theta^2 \right) V_{1N} \\ +(\eta_N - g\theta)V_{01N} \end{array} \right\} \end{array} \right] \\ &= \left(1 + \varphi_{1N}^2 A_{1N} + \delta_{1N}^2 B_{1N} + 2\varphi_{1N}\delta_{1N} C_{1N} - 2\varphi_{1N} D_{1N} - 2\delta_{1N} E_{1N} \right), \end{aligned} \tag{6.2}$$

where,

$$\begin{aligned} A_{1N} &= 1 + V_{0N} + 2\eta_N(\eta_N - 1)V_{1N} + 4\eta_N V_{01N}, \\ B_{1N} &= 1 + V_{0N} + \left(2\eta_N^2 - 2\eta_N + g^2\theta^2 + g(g+1)\theta^2 - 4\eta_N g\theta \right) V_{1N} + 4(\eta_N - g\theta)V_{01N}, \\ C_{1N} &= 1 + V_{0N} + \left(2\eta_N^2 - 2\eta_N g\theta - 2\eta_N + \frac{g(g+1)}{2}\theta^2 \right) V_{1N} + 2(2\eta_N - g\theta)V_{01N}, \\ D_{1N} &= 1 + \left(\frac{\eta_N^2}{2} - \eta_N \right) V_{1N} + \eta_N V_{01N}, \\ E_{1N} &= 1 + \left(\frac{\eta_N^2}{2} - \eta_N - \eta_N g\theta + \frac{g(g+1)}{2}\theta^2 \right) V_{1N} + (\eta_N - g\theta)V_{01N}. \end{aligned}$$

The optimal values of φ_{1N} and δ_{1N} can be determined by minimizing (6.2) as

$$\varphi_{1N(opt)} = \frac{(B_{1N}D_{1N} - C_{1N}E_{1N})}{(A_{1N}B_{1N} - C_{1N}^2)} \text{ and } \delta_{1N(opt)} = \frac{(A_{1N}E_{1N} - C_{1N}D_{1N})}{(A_{1N}B_{1N} - C_{1N}^2)}.$$

Putting the above optimum values $\varphi_{1N(opt)}$ and $\delta_{1N(opt)}$ in (6.2), we get

$$min.MSE(t_{a_{1N}}) = \bar{Y}_N^2 \left\{ 1 - \frac{(A_{1N}E_{1N}^2 + B_{1N}D_{1N}^2 - 2C_{1N}D_{1N}E_{1N})}{(A_{1N}B_{1N} - C_{1N}^2)} \right\}. \tag{6.3}$$

□

The outline of the derivation of proposed estimator $t_{a_{2N}}$ can be done on the similar lines.