Neutrosophic Sets and Systems, Vol. 76, 2025
University of New Mexico

Construction Of Almost Unbiased Estimator For Population Mean Using Neutrosophic Information

Rajesh Singh[1], Anamika Kumari[2], Florentin Smarandache[3] and Shobh Nath Tiwari[1*]
[1]Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi-221005, Uttar Pradesh, India

[2]Department of Data Science and Computing, Manipal Academy of Higher Education

[3]Department of Mathematics, University of New Mexico, 705 Gurley Ave, Gallup, USA

rsinghstat@gmail.com, anamika.kumari@manipal.edu, fsmarandache@gmail.com, shobh7285@gmail.com;

*Corresponding: shobh7285@gmail.com

Abstract: In classical statistics, the population mean is estimated using determinate, crisp data value when auxiliary information is known. These estimates can often be biased. The main objective of this study is to introduce the neutrosophic estimator with the minimum mean squared error (MSE) for the unknown value of the population mean as well as overcome the limitations of classical statistics when dealing with ambiguous or indeterminate data. Neutrosophic statistics was introduced by Florentin Smarandache. It is a generalisation of classical statistics that addresses ambiguous, unclear, vague, and indeterminate data. In this study, we have proposed neutrosophic almost unbiased estimator that use known neutrosophic auxiliary parameters to estimate the neutrosophic population mean of the primary variable. Equations for bias and mean squared error are calculated for the suggested estimators up to the first order of approximation. The proposed estimator performs better than the other existing estimators with respect to the MSE and percent relative efficiency (PRE) criteria. The estimator with the highest PRE or lowest MSE is advised for practical utility in various kinds of application areas. The theoretical conclusions are validated by the empirical analysis, which made use of the real data sets.

Keywords: Classical statistics, neutrosophic statistics, auxiliary information, bias, mean squared error, percent relative efficiency.

---

1. Introduction

One of the purpose of sampling theory is to estimate study variable on the basis of auxiliary information with more precision. In sampling theory, we estimate population parameter on the basis of auxiliary variable that has a correlation with study variable. Some of the popular methods used to obtain these objectives to estimate population parameter using ratio, product, regression method of estimation.

Ratio and product estimators are frequently used in survey sampling for getting an improved estimator of population mean given that the auxiliary information is known. The concept of ratio estimator was introduced by [1] as a method for utilising auxiliary information. Ratio estimator compute ratios between the sample means of the study variable and the auxiliary variable and using known auxiliary data on a variable that is linearly related to the variable under study, is used to estimate the population mean.

Whereas, the product estimator, an additional technique for adding auxiliary information into estimation, was put up by [2]. In the product method of estimation, the sample means of auxiliary variable and study variable are multiplied, and the result is used as the foundation for estimation. Over time, several improvements have been made to the classical ratio and product estimators.

One improvement was proposed by [3], they proposed exponential ratio and product type estimators for estimating unknown population mean of the study variable using auxiliary information.

Several researchers such as [1,5,6,7] accomplished some outstanding work in this area. However, our primary focus here is on proposing almost unbiased estimator using neutrosophic information.

Classical statistics deals with data form by crisp numbers and there is no uncertainty in measurement of the observation. But there are so many cases when indeterminacy, vagueness, ambiguousness exists in the data. In day-to-day life where we have indeterminate data, for illustration, the measurement of temperature anywhere, decision support systems, information systems, the weather forecast, economy growth and more. In such cases we need to introduce new techniques to handle with such type of indeterminate data. Fuzzy logic is one method for resolving these kinds of issues regarding ambiguous, imprecise or uncertain observation but still, it ignores indeterminacy. In this situation, neutrosophic techniques are more suitable which address both randomness and indeterminacy.

While classical statistics deals with determinate and crisp values, neutrosophic statistics deals with sets values. As an extension of classical statistics, neutrosophic statistics is identical as classical statistics in the absence of indeterminacy, [9]. The evaluation of both determinate and indeterminate components of the observations is possible with neutrosophic logic, which is an extension of fuzzy logic. It is particularly useful for analysing data that is imprecise or unclear, as noted by [9].

Neutrosophic statistics can be used to solve a number of real-world issues as neutrosophic statistics to the analysis of road traffic accidents, [10]. [11] introduced a goodness-of-fit test within the framework of neutrosophic statistics and [12] concentrated on using neutrosophic analysis of variance on data from university students. In neutrosophic statistics, the sample size may not be precisely determined and not known as the exact number, [13]. Methods based on fuzzy logic are rapidly evolving and frequently used in decision-making settings, [14]. Further advancement of fuzzy sets is complex fuzzy sets, and a complex neutrosophic set is their generalized form. Estimation using neutrosophic techniques are a relatively new and mostly unexplored area compared to estimation challenges in classical probability sampling designs with definite and well-defined data, [15].

Classical statistics struggled to handle indeterminacies that arise from ambiguous, imprecise or incomplete frequency-related data, because it involved values that were uncertain and could include a range of possible measurements. Therefore, neutrosophic statistics is used as a more adaptable and extension of classical statistics in uncertain or indeterminacy situations. However, there has been limited attention given to neutrosophic sampling as compare to classical statistics for estimating the population parameters.

Frame of our paper is constructed in the following manner: Section 1 mentioned Introduction. Section 2 discussed about observation and terminology of neutrosophic statistics. Section 3 outlines some existing neutrosophic estimators. Section 4 presents the proposed neutrosophic almost unbiased estimator. Section 5 mentioned some member of the proposed estimator. Section 6 presents empirical study. Section 7 is devoted to discussion whereas section 8 discusses the conclusion.

2. Neutrosophic observation and terminology

Numerous kinds of neutrosophic observations were provided, such as quantitative neutrosophic data that suggested a number might fall within an unknown interval (a,b), [9]. There are several ways to represent the interval value of a neutrosphic number. [16] and [17] have used neutrosophic interval values as $Z_N = Z_L + Z_U I_N$, where, $I_N \in [I_L, I_U]$. In addition, we apply the same notations from [17] to the considered neutrosophic data, which are represented in the intervals form as $Z_N \in [Z_L, Z_U]$, where, 'Z_L'is the lower value and 'Z_U' is the upper value of the neutrosophic variable Z_N. We choose a neutrosophic sample of size $n_N \in [n_L, n_U]$ using the simple random sampling without replacement (SRSWOR) technique from a finite population of 'N' units $(P_1, P_2, ..., P_N)$. Let $y_N(i) \in [y_L, y_U]$ be the ith sample observation of our neutrosophic data. The auxiliary variable $x_N(i)$ is similarly represented as $x_N(i) \in [x_L, x_U]$.

Let $\bar{y}_N(i) \in [\bar{y}_L, \bar{y}_U]$ and $\bar{x}_N(i) \in [\bar{x}_L, \bar{x}_U]$ be sample mean for the neutrosophic y_N and x_N respectively and neutrosophic auxiliary variable x_N is correlated to our neutrosophic study variable y_N. The population means for the neutrosophic variables y_N and x_N, which represent the overall averages of the neutrosophic data set, are represented by $\bar{Y}_N \in [\bar{Y}_L, \bar{Y}_U]$ and $\bar{X}_N \in [\bar{X}_L, \bar{X}_U]$ respectively. For y_N and x_N, the neutrosophic coefficients of variation are $C_{yN} \in [C_{yNL}, C_{yNU}]$ and $C_{xN} \in [C_{xNL}, C_{xNU}]$ respectively. Furthermore, the neutrosophic correlation coefficient between neutrosophic variables x_N and y_N is denoted as $\rho_{yxN} \in \left[\rho_{yxL}, \rho_{yxU}\right]$. The neutrosophic coefficient of kurtosis for auxiliary variable x_N is denoted by $\beta_{2(x)N} \in [\beta_{2(x)\mathrm{L}}, \beta_{2(x)\mathrm{U}}]$. The neutrosophic mean errors are $\bar{\mathrm{e}}_{yN} \in [\bar{\mathrm{e}}_{yL}, \bar{\mathrm{e}}_{yU}]$ and $\bar{\mathrm{e}}_{xN} \in [\bar{\mathrm{e}}_{xL}, \bar{\mathrm{e}}_{xU}]$. In order to find the estimator's bias and MSE, we write

$$\bar{\mathrm{y}}_N(i) = \bar{Y}_N\left(1+\bar{e}_{yN}(i)\right),\quad \bar{x}_N(i) = \bar{X}_N\left(1+\bar{e}_{xN}(i)\right),$$

$$E(\bar{e}_{xN}) = E(\bar{e}_{yN}) = 0 \text{ and}$$

$$E(\bar{e}_{yN}^2) = \theta_N C_{yN}^2,\quad E(\bar{e}_{xN}^2) = \theta_N C_{xN}^2,\quad E(\bar{e}_{xN}\bar{e}_{yN}) = \theta_N \rho_{yxN} C_{yN} C_{xN} .$$

Where, $\bar{e}_{yN} \in [\bar{e}_{yL}, \bar{e}_{yU}]$, $\bar{e}_{xN} \in [\bar{e}_{xL}, \bar{e}_{xU}]$, $\bar{e}_{yN}\bar{e}_{xN} \in [\bar{e}_{yL}\bar{e}_{xL}, \bar{e}_{yU}\bar{e}_{xU}]$, $\bar{e}_{yN}^2 \in [\bar{e}_{yL}^2, \bar{e}_{yU}^2]$, $\bar{e}_{xN}^2 \in [\bar{e}_{xL}^2, \bar{e}_{xU}^2]$.

$$C_{xN}^2 = \frac{\sigma_{xN}^2}{\bar{X}_N^2} ,\quad C_{yN}^2 = \frac{\sigma_{yN}^2}{\bar{Y}_N^2},\quad C_{yxN} = \rho_{yxN} C_{yN} C_{xN},$$

where, $C_{xN}^2 \in [C_{xL}^2, C_{xU}^2]$, $C_{yN}^2 \in [C_{yL}^2, C_{yU}^2]$.

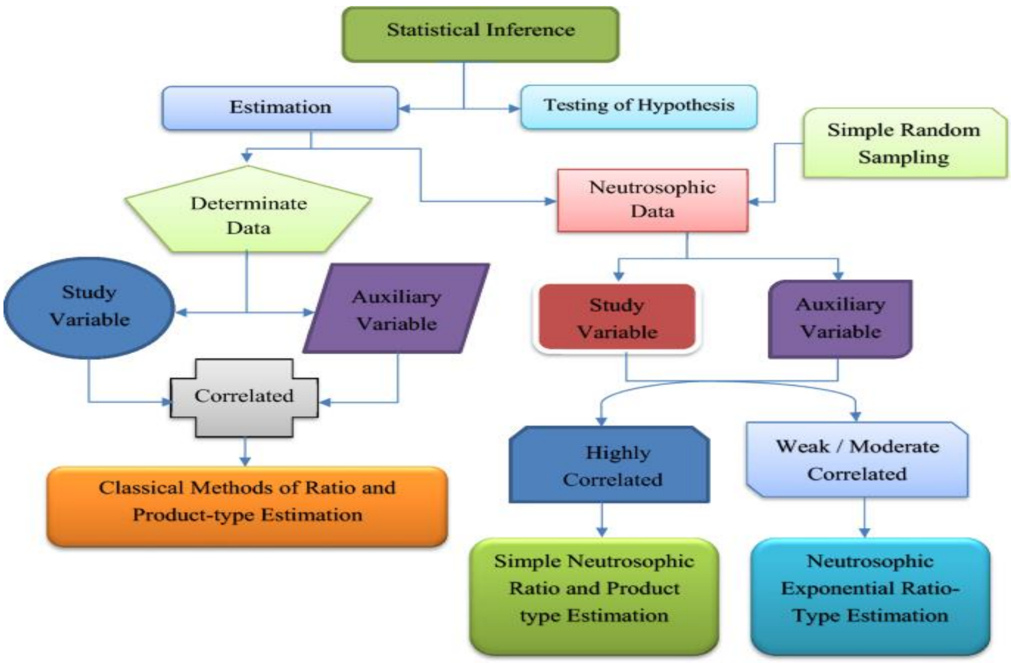
$$\rho_{yxN} = \frac{\sigma_{yxN}}{\sigma_{yN}\sigma_{xN}} \quad , \; \rho_{yxN} \in [\rho_{yxL}, \rho_{yxU}],$$

$$\theta_N = \frac{(1-f_N)}{n_N}, \text{ where, } \theta_N \in [\theta_L, \theta_U], \; n_N \in [n_L, n_U], \; \sigma_{xN}^2 \in [\sigma_{xL}^2, \sigma_{xU}^2], \; \sigma_{yN}^2 \in [\sigma_{yL}^2, \sigma_{yU}^2],$$

$$\sigma_{yxN} \in [\sigma_{yxL}, \sigma_{yxU}].$$

Likewise, in a neutrosophic environment, bias and MSE can be defined as $Bias(\bar{y}_N) \in [Bias_L, Bias_U]$, $MSE(\bar{y}_N) \in [MSE_L, MSE_U]$ respectively.

The graph below illustrates the flowchart of the proposed study employing neutrosophic numbers. The flowchart below has been recommended by [16].



3. Some existing neutrosophic estimators

The estimation of neutrosophic population mean $\bar{Y}_N$ using neutrosophic sample mean which is provided by,

$$t_{1N} = \bar{y}_N \tag{3.1}$$

The variance of the estimator t_{1N} to the first order of approximation is provided by,

$$V(t_{1N}) = \theta_N \bar{Y}_N^2 C_{yN}^2 \quad , \text{ where, } t_{1N} \in [t_{1L}, t_{1U}]. \tag{3.2}$$

[16] recommended the usual neutrosophic ratio estimator of neutrosophic population mean $\bar{Y}_N$ utilising the known neutrosophic population mean of X as,

$$t_{2N} = \bar{y}_N \left(\frac{\bar{X}_N}{\bar{x}_N} \right) \tag{3.3}$$

The bias term and MSE of the estimator t_{2N} are given by, respectively

$$Bias(t_{2N}) = \theta_N \bar{Y}_N (C_{xN}^2 - C_{yxN}) \tag{3.4}$$

$$MSE(t_{2N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + C_{xN}^2 - 2C_{yxN} \right) \tag{3.5}$$

Where, $C_{yxN} = \rho_{yxN} C_{xN} C_{yN}$ and $t_{2N} \in [t_{2L}, t_{2U}]$.

Motivated by [2], the neutrosophic product estimators is given by,

$$t_{3N} = \bar{y}_N \left(\frac{\bar{x}_N}{\bar{X}_N} \right) \tag{3.6}$$

The bias and MSE of the estimator's (t_{3N}) are provided by,

$$Bias(t_{3N}) = \theta_N \bar{Y}_N C_{yxN} \tag{3.7}$$

$$MSE(t_{3N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + C_{xN}^2 + 2C_{yxN} \right) \text{, where, } t_{3N} \in [t_{3L}, t_{3U}]. \tag{3.8}$$

Inspired by [4], [16] proposed the following neutrosophic ratio estimator that utilises the coefficient of variation (CV) of the neutrosophic variable X as,

$$t_{4N} = \bar{y}_N \left(\frac{\bar{X}_N + C_{xN}}{\bar{x}_N + C_{xN}} \right) \tag{3.9}$$

The bias and MSE of the estimator t_{4N}, up to the first order of approximation are respectively,

$$Bias(t_{4N}) = \theta_N \bar{Y}_N (\lambda_{1N}^2 C_{xN}^2 - \lambda_{1N} C_{yxN}) \tag{3.10}$$

$$MSE(t_{4N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \lambda_{1N}^2 C_{xN}^2 - 2\lambda_{1N} C_{yxN} \right) \tag{3.11}$$

where, $\lambda_{1N} = \dfrac{\bar{X}_N}{\bar{X}_N + C_{xN}}$ and $t_{4N} \in [t_{4L}, t_{4U}]$.

Motivated by [3], [16] suggested the following neutrosophic exponential ratio-type estimator as follows:

$$t_{5N} = \bar{y}_N \exp\left(\frac{\bar{X}_N - \bar{x}_N}{\bar{x}_N + \bar{X}_N} \right) \tag{3.12}$$

The estimator's (t_{5N}) bias and MSE are provided by,

$$Bias(t_{5N}) = \theta_N \bar{Y}_N \left(\frac{3}{8} C_{xN}^2 - \frac{1}{2} C_{yxN} \right) \tag{3.13}$$

$$MSE(t_{5N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \frac{C_{xN}^2}{4} - C_{yxN} \right) \text{ , where, } t_{5N} \in [t_{5L}, t_{5U}]. \tag{3.14}$$

Motivated by [3], [18] suggested the neutrosophic product-type exponential estimator as,

$$t_{6N} = \bar{y}_N \exp\left(\frac{\bar{x}_N - \bar{X}_N}{\bar{x}_N + \bar{X}_N} \right) \tag{3.15}$$

The bias and MSE of the estimator t_{6N} are given by,

$$Bias(t_{6N}) = \theta_N \bar{Y}_N \left(\frac{3}{8} C_{xN}^2 + \frac{1}{2} C_{yxN} \right) \tag{3.16}$$

$$MSE(t_{6N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \frac{C_{xN}^2}{4} + C_{yxN} \right) \text{ , where, } t_{6N} \in [t_{6L}, t_{6U}]. \tag{3.17}$$

Motivated by [19], [17] suggested the neutrosophic ratio estimator t_{7N}, utilising the known population coefficient of correlation ρ_{yxN} is provided as,

$$t_{7N} = \bar{y}_N \left(\frac{\bar{X}_N + \rho_{yxN}}{\bar{x}_N + \rho_{yxN}} \right) \tag{3.18}$$

The bias and MSE of the estimator t_{7N} are respectively given as,

$$Bias(t_{7N}) = \theta_N \bar{Y}_N (\lambda_{2N}^2 C_{xN}^2 - \lambda_{2N} C_{yxN}) \tag{3.19}$$

$$MSE(t_{7N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \lambda_{2N}^2 C_{xN}^2 - 2\lambda_{2N} C_{yxN} \right) \tag{3.20}$$

where, $\lambda_{2N} = \dfrac{\bar{X}_N}{\bar{X}_N + \rho_{yxN}}$ and $t_{7N} \in [t_{7L}, t_{7U}]$.

Inspired by [20], [17] proposed the neutrosophic ratio estimators utilising CV and the coefficient of kurtosis of auxiliary variable X, expressed as follows:

$$t_{8N} = \bar{y}_N \left(\frac{\beta_{2(x)N} \bar{X}_N + C_{xN}}{\beta_{2(x)N} \bar{x}_N + C_{xN}} \right) \tag{3.21}$$

The estimator's (t_{8N}) bias and MSE are defined as,

$$Bias(t_{8N}) = \theta_N \bar{Y}_N (\lambda_{3N}^2 C_{xN}^2 - \lambda_{3N} C_{yxN}) \tag{3.22}$$

$$MSE(t_{8N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \lambda_{3N}^2 C_{xN}^2 - 2\lambda_{3N} C_{yxN} \right) \tag{3.23}$$

where, $\lambda_{3N} = \dfrac{\beta_{2(x)N}\bar{X}_N}{\beta_{2(x)N}\bar{X}_N + C_{xN}}$ and $t_{8N} \in [t_{8L}, t_{8U}]$.

Motivated by [8], [16] recommended the neutrosophic modified exponential ratio estimator defined as,

$$t_{9N} = \bar{y}_N \exp\left(\frac{(a\bar{X}_N + b) - (a\bar{x}_N + b)}{(a\bar{X}_N + b) + (a\bar{x}_N + b)}\right) \quad (3.24)$$

where a and b denotes the auxiliary neutrosophic parameters.

The bias and MSE of the neutrosophic exponential ratio estimator t_{9N} are respectively given as,

$$Bias(t_{9N}) = \theta_N \bar{Y}_N \left(\frac{3}{8}\lambda_{4N}^2 C_{xN}^2 - \frac{1}{2}\lambda_{4N} C_{yxN}\right) \quad (3.25)$$

$$MSE(t_{9N}) = \theta_N \bar{Y}_N^2 \left(C_{yN}^2 + \lambda_{4N}^2 C_{xN}^2 - 2\lambda_{4N} C_{yxN}\right) \quad (3.26)$$

where, $\lambda_{4N} = \dfrac{a\bar{X}_N}{2(a\bar{X}_N + b)}$ and $t_{9N} \in [t_{9L}, t_{9U}]$.

From (3.22) and (3.25), we see that the estimators t_{8N} and t_{9N} recommended by [17], [16] respectively, are biased estimator. Bias is disadvantageous in certain applications.

Motivated by [8] and [17], we have suggested an almost unbiased estimator for the population mean $\bar{Y}_N$ using neutrosophic auxiliary information in SRSWOR.

4. The proposed neutrosophic almost unbiased estimator

Suppose $t_0 = \bar{y}_N$, $t_1 = \left[\bar{y}_N \left(\dfrac{a\bar{X}_N + b}{a\bar{x}_N + b}\right)^K\right]$, $t_2 = \left[\bar{y}_N \exp\left(K\left(\dfrac{(a\bar{X}_N + b) - (a\bar{x}_N + b)}{(a\bar{X}_N + b) + (a\bar{x}_N + b)}\right)\right)\right]$

such that $t_0, t_1, t_2 \in A$, where A is the collection of all possible estimators for estimating the neutrosophic population mean $\bar{Y}_N$. Justification of using a and b in proposing estimators is given by [22]. By definition, the set A is a linear variety if

$$t^* = \sum_{i=0}^{2} \alpha_i t_i \in A. \quad (4.1)$$

for $\sum_{i=0}^{2} \alpha_i = 1$ and $\alpha_i \in R$. (4.2)

where $\alpha_i (i = 0,1,2)$ refers to the statistical constants, R stands for the set of real numbers and K takes distinct values, that is, $K = 1$ for the ratio estimators and $K = -1$ for the product estimators.

The bias and MSE of the estimators t_1 and t_2 are, respectively, given by (up to the first order of approximation)

$$B(t_1)=\theta_N\bar{Y}_N\left[\lambda^2C_{xN}^2\left(\frac{K(K+1)}{2}\right)-\lambda K\rho_{yxN}C_{yN}C_{xN}\right] \tag{4.3}$$

$$MSE(t_1)=\theta_N\bar{Y}_N^2\left[C_{yN}^2+\lambda^2K^2C_{xN}^2-2\lambda K\rho_{yxN}C_{yN}C_{xN}\right] \tag{4.4}$$

$$B(t_2)=\theta_N\bar{Y}_N\left[\lambda^2C_{xN}^2\left(\frac{2K+K^2}{8}\right)-\frac{\lambda K}{2}\rho_{yxN}C_{yN}C_{xN}\right] \tag{4.5}$$

$$MSE(t_2)=\theta_N\bar{Y}_N^2\left[C_{yN}^2+\frac{\lambda^2K^2}{4}C_{xN}^2-\lambda K\rho_{yxN}C_{yN}C_{xN}\right] \tag{4.6}$$

where, $\lambda=\dfrac{a\bar{X}_N}{a\bar{X}_N+b}$; a and b denotes the auxiliary neutrosophic parameters.

The bias and MSE of the estimator t^* are determined using,

$\bar{y}_N(i)=\bar{Y}_N\left(1+\bar{e}_{yN}(i)\right)$, $\bar{x}_N(i)=\bar{X}_N\left(1+\bar{e}_{xN}(i)\right)$.

Such that, $E(\bar{e}_{xN})=E(\bar{e}_{yN})=0$.

Defining the estimator t^* in terms of $\bar{e}_{xN}$ and $\bar{e}_{yN}$, we obtain,

$$\begin{aligned}
t^* &= \alpha_0\bar{Y}_N\left(1+\bar{e}_{yN}\right)+\alpha_1\left[\bar{Y}_N\left(1+\bar{e}_{yN}\right)\left(\frac{a\bar{X}_N+b}{a\bar{X}_N\left(1+\bar{e}_{xN}\right)+b}\right)^K\right] \\
&\quad +\alpha_2\left[\bar{Y}_N\left(1+\bar{e}_{yN}\right)\exp\left(K\left(\frac{(a\bar{X}_N+b)-\left(a\bar{X}_N\left(1+\bar{e}_{xN}\right)+b\right)}{(a\bar{X}_N+b)+\left(a\bar{X}_N\left(1+\bar{e}_{xN}\right)+b\right)}\right)\right)\right] \\
&= \alpha_0\bar{Y}_N\left(1+\bar{e}_{yN}\right)+\alpha_1\left[\bar{Y}_N\left(1+\bar{e}_{yN}\right)\left(1+\lambda\bar{e}_{xN}\right)^{-K}\right] \\
&\quad +\alpha_2\bar{Y}_N\left(1+\bar{e}_{yN}\right)\exp\left[K\left(-\frac{\lambda}{2}\bar{e}_{xN}\left(1+\frac{\lambda}{2}\bar{e}_{xN}\right)^{-1}\right)\right]
\end{aligned} \tag{4.7}$$

where, $\lambda=\dfrac{a\bar{X}_N}{a\bar{X}_N+b}$.

Expanding the terms on the right side of the equation (4.7) and having the terms up to the first order of approximation.

$$t^{*}=\bar{Y}_{N}\begin{bmatrix}(\alpha_0+\alpha_1+\alpha_2)+(\alpha_0+\alpha_1+\alpha_2)\bar{e}_{yN}-\lambda K\bar{e}_{xN}\left(\alpha_1+\frac{\alpha_2}{2}\right)+\alpha_1\lambda^2\bar{e}_{xN}{}^2\left(\frac{K(K+1)}{2}\right)\\ +\alpha_2\lambda^2\bar{e}_{xN}{}^2\left(\frac{2K+K^2}{8}\right)-K\alpha_1\lambda\bar{e}_{xN}\bar{e}_{yN}-\frac{K}{2}\alpha_2\lambda\bar{e}_{xN}\bar{e}_{yN}\end{bmatrix} \tag{4.8}$$

The bias of the estimator t^*, up to the second power of $e's$, can be computed by subtracting $\bar{Y}_N$ from both sides of equation (4.8) after taking the expectation of each side.

$$B(t^*)=\theta_N\bar{Y}_N C_{xN}^2\left[\alpha_1\lambda^2\frac{K(K+1)}{2}+\alpha_2\lambda^2\left(\frac{2K+K^2}{8}\right)-\left(K\alpha_1\lambda+\frac{K}{2}\alpha_2\lambda\right)\rho_{yxN}\frac{C_{yN}}{C_{xN}}\right] \tag{4.9}$$

From equation (4.8), we have

$$t^*-\bar{Y}_N\cong\bar{Y}_N\left[\bar{e}_{yN}-\lambda\alpha K\bar{e}_{xN}\right] \tag{4.10}$$

where, $\alpha=\left(\alpha_1+\frac{\alpha_2}{2}\right)$. (4.11)

By taking square of each side of (4.10) and then taking expectations, we get MSE of the estimator t^*, to the first order of approximation as,

$$MSE(t^*)=\theta_N\bar{Y}_N^2\left[C_{yN}^2+C_{xN}^2\lambda K\alpha\left(\lambda K\alpha-2\delta_N\right)\right] \tag{4.12}$$

where, $\delta_N=\rho_{yxN}\frac{C_{yN}}{C_{xN}}$.

Minimum MSE of the estimator t^* would occur when

$$\alpha=\frac{\delta_N}{\lambda K}. \tag{4.13}$$

Applying this value of $\alpha=\frac{\delta_N}{\lambda K}$ in equation (4.1), we get the optimum value of estimator t^* (optimum).

Hence, the minimum MSE of the estimator t^* is provided by

$$\min.MSE\left(t^*\right)=\theta_N\bar{Y}_N^2C_{yN}^2\left(1-\rho_{yxN}^2\right) \tag{4.14}$$

This MSE is equivalent to the classical linear regression estimator.

By using equation (4.11) and (4.13) we have

$$\alpha_1+\frac{\alpha_2}{2}=\alpha=\frac{\delta_N}{\lambda K}. \tag{4.15}$$

There are only two equations with three unknowns in equations (4.2) and (4.15). The unique values for α_i's , $(i=0,1,2)$ can't be found. In order to get unique values for α_i's , we will apply the linear constraint

$$\sum_{i=0}^{2}\alpha_i B(t_i)=0. \tag{4.16}$$

where $B(t_i)$ represents the bias of i^{th} estimator.

The matrix form of equations (4.2), (4.15), and (4.16) is as follows:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & B(t_1) & B(t_2) \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}=\begin{pmatrix} 1 \\ \alpha \\ 0 \end{pmatrix} \tag{4.17}$$

where, $\alpha=\dfrac{\delta_N}{\lambda K}$.

By using equation (4.17), The unique values of α_i's $(i=0,1,2)$ are obtained as,

$$\left.\begin{aligned} \alpha_0 &= \frac{B(t_2)-\frac{B(t_1)}{2}-\alpha B(t_2)+\alpha B(t_1)}{B(t_2)-\frac{B(t_1)}{2}} \\ \alpha_1 &= \frac{\alpha B(t_2)}{B(t_2)-\frac{B(t_1)}{2}} \\ \alpha_2 &= \frac{-\alpha B(t_1)}{B(t_2)-\frac{B(t_1)}{2}} \end{aligned}\right] \tag{4.18}$$

such that,

$$\sum_{i=0}^{2}\alpha_i=1 \text{ and } \alpha_i\in R.$$

By using these α_i's , $(i=0,1,2)$ eliminate the bias up to terms of order $o(n^{-1})$ at (4.1).

5. Some member of the suggested estimator 't^*'

Several estimators of the population mean are provided by choosing the appropriate combination of constants a, b and α_i's ,$(i=0,1,2)$, as shown in the following scheme.

Estimator	Values of

	α_0	α_1	α_2	a	b	K
$t_{1N} = \bar{y}_N$ The mean per unit estimator	1	0	0	0	0	1
$t_{2N} = \bar{y}_N \left(\frac{\bar{X}_N}{\bar{x}_N} \right)$ [16] estimator	0	1	0	1	0	1
$t_{3N} = \bar{y}_N \left(\frac{\bar{x}_N}{\bar{X}_N} \right)$ Neutrosophic product estimator	0	1	0	1	0	-1
$t_{4N} = \bar{y}_N \left(\frac{\bar{X}_N + C_{xN}}{\bar{x}_N + C_{xN}} \right)$ [16] estimator	0	1	0	1	C_{xN}	1
$t_{5N} = \bar{y}_N \exp\left(\frac{\bar{X}_N - \bar{x}_N}{\bar{x}_N + \bar{X}_N} \right)$	0	0	1	1	0	1
$t_{6N} = \bar{y}_N \exp\left(\frac{\bar{x}_N - \bar{X}_N}{\bar{x}_N + \bar{X}_N} \right)$ [18] estimator	0	0	1	1	0	-1
$t_{7N} = \bar{y}_N \left(\frac{\bar{X}_N + \rho_{yxN}}{\bar{x}_N + \rho_{yxN}} \right)$	0	1	0	1	ρ_{yxN}	1
$t_{8N} = \bar{y}_N \left(\frac{\beta_{2(x)N}\bar{X}_N + C_{xN}}{\beta_{2(x)N}\bar{x}_N + C_{xN}} \right)$ [17] estimator	0	1	0	$\beta_{2(x)N}$	C_{xN}	1
$t_{9N} = \bar{y}_N \exp\left(\frac{(a\bar{X}_N + b) - (a\bar{x}_N + b)}{(a\bar{X}_N + b) + (a\bar{x}_N + b)} \right)$	0	0	1	a	b	1

In addition, the estimators mentioned above, by simply changing the values of a, b and α_i's ($i = 0,1,2$). From the proposed estimator t^* at (4.1), several estimators can be created.

6. Empirical Study

Population 1 : Here, we use the data provided by [21], which is birth rate and data on the natural growth rate obtained from the sample registration system (SRS)(2020) is used to demonstrate the performance of different estimators of neutrosophic study variable $Y_N \in [Y_L, Y_U]$. The neutrosophic auxiliary variable and neutrosophic study variable are given by,

X_N : Birth rate Y_N : Natural growth rate

The PRE formula is given by

$$PRE = \frac{MSE(\bar{y}_N)}{MSE(\text{estimator})} \times 100 \tag{6.1}$$

Table 1. Descriptive statistics of birth rate and natural growth rate data as per SRS 2020.

Parameter	Neutrosophic Value	Parameter	Neutrosophic Value
N_N	[36, 36]	n_N	[10, 10]
μ_{xN}	[14.472, 18.053]	C_{yN}	[0.333, 0.411]
μ_{yN}	[10.894, 10.631]	ρ_{yxN}	[0.620, 0.849]
C_{xN}	[0.213, 0.236]	$\beta_{2(x)N}$	[3.034, 2.173]

Table 2. Neutrosophic MSEs and PRE of the different estimators.

Estimator	MSE	I_N	PRE
t_{1N}	[0.951886, 1.378714]	[0, 0.31]	[100, 100]
t_{2N}	[0.490003, 0.586746]	[0, 0.16]	[194.26, 234.98]
t_{3N}	[2.092575, 3.174567]	[0, 0.34]	[43.43, 45.49]
t_{4N}	[2.070519, 3.145647]	[0, 0.34]	[43.83, 45.97]
t_{5N}	[0.672372, 0.820965]	[0, 0.18]	[141.57, 167.94]
t_{6N}	[1.425287, 2.163247]	[0, 0.34]	[63.73, 66.79]
t_{7N}	[0.510456, 0.586471]	[0, 0.13]	[186.48, 235.09]
t_{8N}	[0.492618, 0.586646]	[0, 0.16]	[193.23, 235.02]
t_{9N} $(a=1, b=1)$	[0.684577, 0.844600]	[0, 0.19]	[139.05, 163.24]
t^*	**[0.385639, 0.586416]**	[0, 0.34]	**[235.11, 246.83]**

Population 2 : Here, we use the data sets considered by [18]. Numerous data points, such as the minimum, maximum and mean temperatures over a wide time duration, are included in this dataset regarding the All India Seasonal and Annual Temperature Series. The neutrosophic auxiliary variable and neutrosophic study variable are given by,

X_N : Minimum and maximum temperatures in January and February.

Y_N : Temperatures from March through May.

Table 3. Descriptive statistics of the population 2.

Parameter	Neutrosophic Value	Parameter	Neutrosophic Value
N_N	[119, 119]	n_N	[20, 20]
μ_{xN}	[13.908,24.674]	C_{yN}	[0.0007, 00.0006]
μ_{yN}	[20.694, 31.560]	ρ_{yxN}	[0.6274, 0.7202]

C_{xN}	[0.0017, 0.0015]	$\beta_{2(x)N}$	[4.973, 5.583]

Table 4. Neutrosophic MSEs and PRE of the different estimators.

Estimator	MSE	I_N	PRE
t_{1N}	[8.87822E-06, 1.72193E-05]	[0, 0.48]	[100.00, 100.00]
t_{2N}	[3.39108E-05, 5.58304E-05]	[0, 0.39]	[26.18, 30.84]
t_{3N}	[8.78263E-05, 0.000173977]	[0, 0.50]	[10.11, 9.90]
t_{4N}	[3.39013E-05, 5.58219E-05]	[0, 0.39]	[26.19, 30.85]
t_{5N}	[8.39692E-06, 1.21037E-05]	[0, 0.31]	[105.73, 142.26]
t_{6N}	[3.53547E-05, 7.1177E-05]	[0, 0.50]	[25.11, 24.19]
t_{7N}	[3.06832E-05, 5.20435E-05]	[0, 0.41]	[28.94, 33.09]
t_{8N}	[3.39089E-05, 5.58288E-05]	[0, 0.39]	[26.18, 30.84]
t_{9N} $(a=1, b=1)$	[7.61584E-06, 1.13888E-05]	[0, 0.33]	[116.58, 151.19]
t^*	**[5.38372E-06, 8.28831E-06]**	[0, 0.35]	**[164.91, 207.75]**

7. Discussion

In Table 2, as can be shown that the sampling variance of the neutrosophic estimator t_{1N} is [0.951886, 1.378714] and the neutrosophic exponential ratio estimator's (t_{5N}) MSE is [0.672372, 0.820965] whereas the range of the neutrosophic MSEs (t_{4N}) is [2.070519, 3.145647]. The reason behind the higher MSEs of the neutrosophic estimators t_{4N} over the neutrosophic estimator t_{1N} is that the low value of the neutrosophic CV of the neutrosophic auxiliary variable x. In comparison to the other population mean $(\bar{Y}_N)$ estimators in competition, the MSE of the proposed class of neutrosophic estimators (t^*) is the lowest at [0.385639, 0.586416] with higher PRE.

Similarly, the proposed and existing estimators' MSEs and PRE are given by Table 4. Additionally, it has shown that the suggested estimators' MSEs and PREs are lower and higher respectively, in comparison to those of the other estimators already in use. Its value is highlighted by the bolded text. Hence, Tables 2 and 4 confirmed these results, showing that the suggested estimators are preferable to the existing ones.

8. Conclusion

In this paper, we have recommended an almost unbiased neutrosophic estimator to estimate the neutrosophic population mean $\bar{Y}_N$ using some known value of auxiliary information. In order to evaluate their precision, we computed bias and MSE for the suggested estimators, concentrating on first-order approximations. We evaluated our suggested estimator with existing estimators using data from a natural environment. Via conducting empirical study, we have identified strong evidence that our suggested estimator performs better than the existing ones within the context of neutrosophic sampling. It should be emphasised that in cases where the study variable's observations are nondeterministic or uncertain, the neutrosophic estimators are best suited for enhanced population mean estimate. Therefore, it might be advised to use the proposed class of estimators to estimate the neutrosophic population mean $\bar{Y}_N$ in a variety of real-world settings, including those in the social sciences, economics, agriculture, biology and mathematics.

References

[1] W. G. Cochran, "The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce," J. Agric. Sci., vol. 30, no. 2, pp. 262–275, Apr. 1940, doi: 10.1017/S0021859600048012.

[2] M. N. Murthy, "Product method of estimation," Sankhyā Indian J. Stat. Ser. A. (1961-2002), Vol. 26, No. 1, pp. 69-74, Jul.1964, https://www.jstor.org/stable/25049308.

[3] S. Bahl and R. K. Tuteja, "Ratio and product type exponential estimators," J. Inf. Optim. Sci., vol. 12, no. 1, pp. 159–164, Jan. 1991, doi:10.1080/02522667.1991.10699058.

[4] B. V. S. Sisodia and V. K. Dwivedi, "A modified ratio estimator using coefficient of variation of auxiliary variable," J Indian Soc Agric Stat, 33(1), 13–18, 1981.

[5] C. Kadilar and H. Cingi, "Ratio estimators in simple random sampling," Appl. Math. Comput., vol.151, no. 3, pp. 893–902, Apr. 2004, doi:10.1016/S0096-3003(03)00803-8.

[6] R. Singh, R. Mishra, A. Kumari, and S. K. Yadav, "Almost unbiased estimators for population coefficient of variation using auxiliary information," Asian J. Probab. Stat., vol. 26, no. 5, pp. 1–18, May 2024, doi:10.9734/ajpas/2024/v26i5614.

[7] H. P. Singh and R. S. Solanki, "A new procedure for variance estimation in simple random sampling using auxiliary information," Stat. Pap., vol. 54, no. 2, pp. 479–497, May 2013, doi:10.1007/s00362-012-0445-2.

[8] R. Singh, P. Chauhan, F. Smarandache, and N. Sawan, "Auxiliary information and a priori values in construction of improved estimators," Infinite Study, 2007.

[9] F. Smarandache, "Introduction to neutrosophic statistics," Infinite Study, 2014.

[10] M. Aslam, "Monitoring the road traffic crashes using NEWMA chart and repetitive sampling," International Journal of Injury Control and Safety Promotion, 28(1), pp. 39–45. 2020, doi:10.1080/17457300.2020.1835990.

[11] M. Aslam, "A new goodness of fit test in the presence of uncertain parameters," Complex Intell. Syst., Oct. 2020, doi:10.1007/s40747-020-00214-8.

[12] M. Aslam, "Neutrosophic analysis of variance: application to university students," Complex Intell. Syst., vol. 5, no. 4, pp. 403–407, Dec. 2019, doi:10.1007/s40747-019-0107-2.

[13] F. Smarandache, "A unifying field in logics: Neutrosophic logic, neutrosophic set, neutrosophic probability, and statistics. arXiv preprint math/0101228 - 6th ed.". 2001.

[14] N. Jan, L. Zedam, T. Mahmood, K. Ullah, and Z. Ali, “Multiple attribute decision making method under linguistic cubic information,” J. Intell. Fuzzy Syst., vol. 36, no. 1, pp. 253–269, Jan. 2019, doi:10.3233/JIFS-181253.

[15] R. Varshney, A. Pal, Mradula, and I. Ali, “Optimum allocation in the multivariate cluster sampling design under gamma cost function,” J. Stat. Comput. Simul., vol. 93, no. 2, pp. 312–323, Jan. 2023, doi:10.1080/00949655.2022.2104845.

[16] Z. Tahir, H. Khan, M. Aslam, J. Shabbir, Y. Mahmood, and F. Smarandache, “Neutrosophic ratio-type estimators for estimating the population mean,” Complex Intell. Syst., vol. 7, no. 6, pp. 2991–3001, Dec. 2021, doi:10.1007/s40747-021-00439-1.

[17] S. K. Yadav and F. Smarandache, “Generalized neutrosophic sampling strategy for elevated estimation of population mean,” Neutrosophic Sets and Systems, Vol. 53, 2023 SSRN doi:10.2139/ssrn.4652035.

[18] V. K. Yadav and S. Prasad, “Neutrosophic estimators for estimating the population mean in survey sampling,” Meas. Interdiscip. Res. Perspect., pp. 1–25, Mar. 2024, doi: 10.1080/15366367.2023.2267835.

[19] H. P. Singh and R. Tailor, “Estimation of finite population mean using known correlation coefficient between auxiliary characters,” Statistica, vol. 65, no. 4, Art. no. 4, 2005, doi: 10.6092/issn.1973-2201/1300.

[20] L. N. Upadhyaya and H. P. Singh, “Use of transformed auxiliary variable in estimating the finite population mean,” Biom. J., vol. 41, no. 5, pp. 627–636, 1999, doi: 10.1002/(SICI)1521-4036(199909)41:5<627::AID-BIMJ627>3.0.CO;2-W.

[21] R. Singh and A. Kumari, “Neutrosophic ranked set sampling scheme for estimating population mean: An application to demographic data,” Neutrosophic Sets and Systems, Vol. 68, 2024, doi:10.5281/zenodo.11479519.

[22] R. Singh and M. Kumar, “A note on transformations on auxiliary variable in survey sampling,” Model Assist. Stat. Appl., vol. 6, no. 1, pp. 17–19, Mar. 2011, doi: 10.3233/MAS-2011-0154.

Received: July 26, 2024. Accepted: Oct 25, 2024