SoftwareX 32 (2025) 102448

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Neutrostat: An R package for neutrosophic statistical models

Zahid Khan [a], Zsolt T. Kosztyán [a, b, *]

[a] Department of Quantitative Methods, University of Pannonia, Veszprém, H-8200, Hungary
[b] Institute of Advances Studies, Kőszeg (iASK), Charnel Street 14, Kőszeg, H-9730, Hungary

HIGHLIGHTS

- First open-source R implementation of statistical methods for neutrosophic data.
- Provides computation of neutrosophic measures and simulation of imprecise data.
- Bridges statistical computing and neutrosophic logic for modeling uncertain data.

ARTICLE INFO

Keywords:
Neutrosophic logic
Neutrosophic statistics
Neutrosophic distribution
Simulation

ABSTRACT

In many practical applications, we consider indeterminate and ambiguous information in data. Traditional statistical methods address data formed by precise knowledge, whereas neutrosophic statistical (NS) approaches analyze data that result from incomplete, vague, imprecise and partial information. Although several extensions of classical methods to neutrosophic data exist, they have not been implemented in an open-source computational R tool. To address this gap, we introduce the R package `neutrostat`, the first comprehensive package for neutrosophic statistics. The R package `neutrostat` provides functions for descriptive summaries of neutrosophic measures as well as tools for data visualization and reproducibility of essential statistical characteristics. By making NS methods computationally accessible, `neutrostat` facilitates the analysis of imprecise data and supports the reproducibility of many essential statistical characteristics of neutrosophic probabilistic models.

1. Metadata

Nr.	Code metadata description	Metadata
C1	Current code version	v0.0.2
C2	Permanent link to the code/repository used for this code version	https://github.com/kzst/neutrostat
C3	Permanent link to the reproducible capsule	N/A
C4	Legal code license	GNU General Public License (GPL ≥ 2)
C5	Code versioning system used	git
C6	Software code languages, tools, and services used	R (≥ 4.0.0), Roxygen2
C7	Compilation requirements, operating environments & dependencies	R packages: `moments`, `ntsDists`, `ggplot2`, `stats`
C8	If available, link to developer documentation/manual	https://cran.r-project.org/web/packages/neutrostat/neutrostat.pdf; https://github.com/kzst/neutrostat/blob/0.0.2/README.md
C9	Support email for questions	`kosztyan.zsolt@gtk.uni-pannon.hu`

* Corresponding author at: Department of Quantitative Methods, University of Pannonia, Veszprém, H-8200, Hungary.
Email address: kosztyan.zsolt@gtk.uni-pannon.hu (Z.T. Kosztyán).

https://doi.org/10.1016/j.softx.2025.102448
Received 26 August 2025; Received in revised form 7 November 2025; Accepted 12 November 2025
Available online 19 November 2025
2352-7110/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

2. Motivation and significance

We encounter vagueness and various degrees of certainty in most aspects of our day-to-day lives, where not all possible decisions can be accurately defined. The persistence of such uncertainties highlights the limitations of binary thinking and suggests the importance of adopting flexible methods, such as neutrosophic logic, to address vague or imprecise data [1]. The concept of neutrosophy in statistical methods was first introduced by Smarandache for handling imprecise information in data analysis cases [2]. Neutrosophic statistical methods assist in analyzing real-world situations where the data or model parameters are based on incomplete or imprecise information [3]. The classic statistical methods have been structured into neutrosophic frameworks in various data analysis approaches. For example, numerous authors have investigated the conversion of classic descriptive measures under the neutrosophic framework [4,5]. Several studies have advanced neutrosophic inferential statistics. For example, Aslam [6] used a z-test in an uncertain environment, whereas Miari et al. [7] converted the classical two-way ANOVA approach to analyze neutrosophic data. Similarly, Yadav et al. [8] and Masood et al. [9] extended the classical methods to neutrosophic frameworks for estimating the mean and median of the population, respectively. Many authors have explored the transformation of classical probability distributions under the neutrosophic framework. For example, Duan et al. [10] proposed a neutrosophic extension of the exponential distribution, while Ibrahim et al. [11] and Aslam [12] studied neutrosophic adaptations of geometric and negative binomial models, respectively. Additionally, some recent works by Alduais and Khan [13], Alamri and Aslam [14], Latif et al. [15] and Bashir et al. [16] further extended these classical models to handle imprecise and uncertain data.

While neutrosophic logic has attracted significant interest, computational tools and statistical methods for handling imprecise data are still under development. There remains a need for efficient, user-friendly, and reliable statistical software capable of analyzing neutrosophic data. Recently, several computational tools have been developed to support neutrosophic set operations. For example, PyIVNS [17] provides a Python-based environment for performing operations on interval-valued neutrosophic sets, including matrix operations and normalization techniques. Similarly, a Python-based tool has been proposed for bipolar neutrosophic matrices and their algebraic operations [18]. The Python neutrosophic package developed in [19] focuses on implementing basic neutrosophic set concepts, such as single-valued and interval-valued neutrosophic numbers and sets. Additionally, MATLAB toolkits have been developed to perform basic operations on single-valued and interval-valued neutrosophic sets [20,21]. However, these tools are limited primarily to algebraic operations on neutrosophic fuzzy matrices and do not support essential statistical features such as descriptive summaries of neutrosophic data or random data generation from neutrosophic probability distributions. To bridge this gap, the present work introduces a new package `neutrostat` that specifically addresses these missing features.

2.1. Our contributions

In this work, we develop the `neutrostat` R package, which is a pioneering computational R package that is designed primarily for conducting descriptive analyses and generating neutrosophic statistical distributions with imprecise distributional parameters. The `neutrostat` package addresses the critical limitations of the classic statistical methods, which often fail to handle imprecision, vagueness, and inconsistency issues that are related to data processing. By integrating neutrosophic principles into statistical computations, the proposed package offers a flexible and powerful tool for analyzing neutrosophic data.

This package provides a suite of core functions for basic neutrosophic descriptive measures and computations that are related to commonly used statistical distributions. These functions include functions for performing neutrosophic data algebraic operations; functions for computing neutrosophic basic descriptive measures such as the mean, median, mode, skew and kurtosis coefficients; density functions; cumulative distribution functions; quantile functions; and functions for randomly generating data via commonly used statistical models.

We apply the `neutrostat` package to real-world datasets to demonstrate its effectiveness and implementation methodology. By bridging the gap between applied statistics and neutrosophic logic, the `neutrostat` package represents a significant advancement in statistical computing and uncertainty modeling.

3. Statistical background

In this section, we illustrate neutrosophic numbers and their algebraic operations. Neutrosophic data are generalizations of classical data that are available in continuous or discrete forms. Neutrosophic data include observations that are not exactly defined because of indeterminacy or imprecision. When zero indeterminacy is exhibited by the data, the neutrosophic data converge to the classical data. We represent a neutrosophic number as follows [2]:

$$z = d + I \tag{1}$$

where $d \in \mathbb{R}$ represents the determinate part and I represents the indeterminate part, which is modeled as $I_k \in [\alpha_k, \beta_k] \subseteq \mathbb{R}$.

In a discrete context, a datum is an isolated value (not a continuous interval) that has embedded uncertainty.

Examples include the following:

$$z_1 = 4 + I_1, \quad \text{where } I_1 \in [0, 1]$$

$$z_2 = 30 + I_2, \quad \text{where } I_2 \in [3, 5]$$

Formally, a discrete neutrosophic set can be defined as:

$$D = \left\{ d_k + I_k \mid d_k \in \mathbb{R},\ I_k \in [\alpha_k, \beta_k] \right\}$$

In a continuous context, each datum is a full interval of imprecision, which expresses continuous uncertainty.

Examples include the following:

$$z \in [0.5, 1.5], \quad \text{or } z \in [0.3, 0.8]$$

Notably, the indeterminacy factor in the data context creates interval values on which operations can be performed by using interval mathematics [22]. We consider two interval numbers:

$$M = [m_1, m_2], \quad N = [n_1, n_2]$$

Here, $m_1, m_2, n_1, n_2 \in \mathbb{R}$, and the main operations executed on intervals are defined as follows:

1. Addition

$$[m_1, m_2] + [n_1, n_2] = [m_1 + n_1,\ m_2 + n_2] \tag{2}$$

2. Subtraction

$$[m_1, m_2] - [n_1, n_2] = [m_1 - n_2,\ m_2 - n_1] \tag{3}$$

3. Multiplication

We compute all possible products between endpoints:

$$P = \{m_1 n_1,\ m_1 n_2,\ m_2 n_1,\ m_2 n_2\}$$

$$[m_1, m_2] \times [n_1, n_2] = [\min(P),\ \max(P)] \tag{4}$$

4. Division

Provided that $0 \notin [n_1, n_2]$ (i.e., neither $n_1 = 0$ nor $n_2 = 0$),

$$Q = \left\{ \frac{m_1}{n_1},\ \frac{m_1}{n_2},\ \frac{m_2}{n_1},\ \frac{m_2}{n_2} \right\}$$

$$[m_1, m_2] \div [n_1, n_2] = [\min(Q),\ \max(Q)] \tag{5}$$

Notably, two intervals are equal if their endpoints are equal, i.e., $m_1 = n_1$ and $m_2 = n_2$. Moreover, the intersection of two intervals M and

N is empty, which is denoted as $M \cap N = \emptyset$, if either $n_1 > m_2$ or $m_1 > n_2$.. Otherwise, the intersection of M and N is again an interval, which is expressed as follows:

$$M \cap N = \left[\max(m_1, n_1),\ \min(m_2, n_2)\right] \tag{6}$$

In the case of a nonempty intersection, the union of the intervals is given by:

$$M \cup N = \left[\min(m_1, n_1),\ \max(m_2, n_2)\right] \tag{7}$$

Since some data feature metrics, such as the median, quartile points, and percentile points, are based on ordering values, the ordering of interval values is very important. We extend the transitive order relation $<$ from a real line to intervals in the following manner:

$$M < N \iff m_2 < n_1 \tag{8}$$

These are basic algebraic operations that are typically used in neutrosophic data analyses. However, via other operations and on the basis of the signs of the interval points of the interval values, Eq. (4) can be extended to nine additional cases, as discussed by Moore [23]. Additional supporting equations related to neutrosophic statistics are presented in the Appendix A.

4. Software description

The `neutrostat` R package emerges as a dedicated tool for the analysis and modeling of neutrosophic data, which are characterized by inherent imprecision, indeterminacy, or vagueness. Grounded in the theoretical foundation of neutrosophic statistics that was introduced by Smarandache [24], the package equips practitioners and researchers with essential functions for computing interval-based summary statistics, visualizing uncertain information, and modeling neutrosophic distributions. At its core, the `neutrostat` package enables users to input neutrosophic samples as interval data and obtain descriptive statistics by applying well-integrated functions to neutrosophic numbers to accommodate uncertainty analyses of such data. These functions return interval-based results, thus supporting neutrosophic data interpretations where classic statistics may fall short.

The package also includes functions for performing interval-based arithmetic operations such as addition, subtraction, multiplication, division, and sorting. A distinctive feature of the `neutrostat` package is its extensive support for neutrosophic probability distributions. The package enables users to generate random data from a wide variety of neutrosophic distributions, including neutrosophic exponential, gamma, Rayleigh, Laplace, normal, and Weibull distributions. Users can easily generate visual representations of probability density and cumulative distribution functions for a variety of neutrosophic distributions to better understand the shape and spread of imprecise data. The design of the `neutrostat` package is supported by dependencies on established packages such as ggplot2, ntsDists and moments, thus ensuring both elegant visualization capabilities and statistical soundness. A variety of real-world datasets, such as gold price, dioxin level, and temperature datasets, are bundled with the package to demonstrate practical applications of neutrosophic data analyses.

The open-source availability of the package on CRAN and the development version on GitHub makes it an adaptable and accessible tool for advancing research on uncertainty modeling and neutrosophic statistics.

4.1. Software architecture

In this section, we present the core components of the `neutrostat` package that are used to analyze neutrosophic data. The `neutrostat` package has a modular and functional architecture that is designed to provide ease of use and clarity when handling neutrosophic data. At its core, the package processes neutrosophic data, which are typically provided by users in the form of intervals for analysis under neutrosophic statistics. The main architecture of the package is shown in Fig. 1.

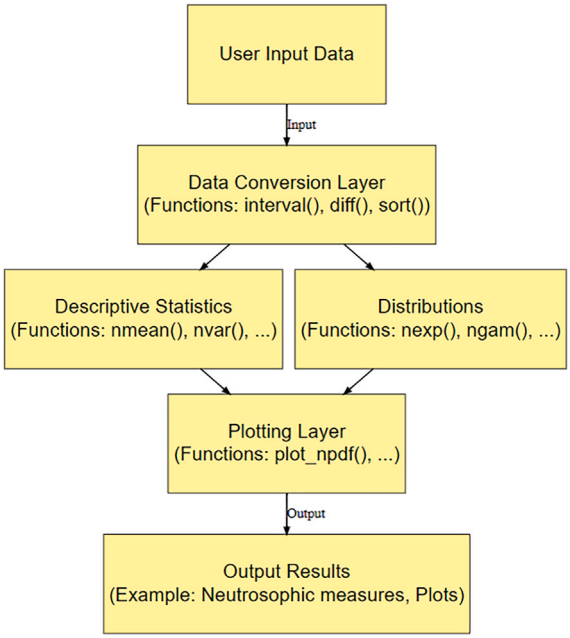


Fig. 1. Workflow of the `neutrostat` package for neutrosophic data processing.

As shown in Fig. 1, several functions for converting raw data into the standardized neutrosophic format are incorporated. The package then enables users to obtain neutrosophic statistics perform distributional modeling, visualize distributional data, and implement interval-based mathematical operations on the neutrosophic data.

4.2. Software functionalities

Various specific functions of the `neutrostat` package are described in this section. In accordance with the principles of neutrosophic statistics, the `neutrostat` package offers a comprehensive collection of functions for summarizing and analyzing neutrosophic data, which are typically characterized by imprecision, uncertainty and vagueness. The major functionalities include the following.

- **Interval Arithmetic**
  The package supports basic operations on interval or neutrosophic data. Functions such as `interval_add()`, `interval_sub()`, `interval_div()`, and `interval_mul()` carry out arithmetic operations on neutrosophic numbers. Additional functions such as `interval_sort()` and `interval_df()` help structure and organize the data.
- **Neutrosophic Descriptive Statistics**
  The `nmean()`, `nvar()`, `nstd()`, `nsk()`, and `nkur()` functions compute the neutrosophic mean, variance, standard deviation, skewness, and kurtosis, respectively, of neutrosophic data. Other supporting functions, such as `nmedian()`, `nquant()`, and `nsummary()`, are used to find distributional quantile points and provide a summary list of neutrosophic data.
- **Neutrosophic Probability Distributions**
  The package provides tools for modeling and analyzing neutrosophic versions of commonly used probability distributions, such as the gamma,exponential, normal, Rayleigh, Laplace, and Weibull distributions. These are implemented via the following functions:
  - Random generation function: `rn()`
  - Density function: `dn()`
  - Quantile function: `qn()`
  - Distribution function: `pn()`
  - Plotting functions: `plot_npdf()` and `plot_ncdf()`

For example, `dnwbl()`, `rnwbl()`, and `plot_npdfwbl()` are employed for the neutrosophic Weibull distribution.

- **Datasets for Practice**

  The package includes built-in real-world neutrosophic datasets for analysis and demonstration purposes.

  – `dioxin`: Dioxin ingestion levels from food in Japan.
  – `citytemp`: Daily temperature ranges in Pakistani cities.
  – `goldprice`: Monthly high and low gold prices in Indian cities.

- **Visualization Tools**

  Neutrosophic distributions are supported by shaded probability density function (PDF) and cumulative distribution function (CDF) plots to reflect interval-based uncertainty. We use `plot_npdf()` and `plot_ncdf()` for visualization purposes. For example, the functions `plot_npdfnorm()` and `plot_ncdfnorm()` are used to plot the PDF and CDF of the normal distribution, respectively.

The listed functionalities indicate that the `neutrostat` package provides researchers and analysts with a foundational statistical environment for working with neutrosophic data. It combines classic statistical techniques with neutrosophic logic to support analyses under uncertainty.

4.3. Sample code snippet analysis

In this section, sample code snippets that illustrate a practical application of the `neutrostat` functions for neutrosophic data analysis scenarios are presented.

```
# Neutrosophic data
> y <- list(c(2,6), c(5,8))
# Neutrosophic addition
> interval_add(y)
# Neutrosophic subtraction
> interval_sub(y)
# Neutrosophic multiplication
> interval_mul(y)
# Neutrosophic division
> interval_div(y)
```

Listing 1: Neutrosophic arithmetic operation.

```
#Neutrosophic data
>data <- list(8, c(3, 5), 20, c(18, 24),c(25,30), c(4,8),c(12,15),c(38,40))
# Basic neutrosophic measures
# Neutrosophic mean
>nmean
# Neutrosophic median
>nmedian
# Neutrosophic standard deviation
>nstd
# Neutrosophic coefficient of variation
>ncv
# Neutrosophic quantile points
>nquant
```

Listing 2: Basic Descriptive statistics of neutrosophic data.

```
# Simple data vector
>x <- c(2,4,6,8,9,10,14, 18,20)
#Data conversion
>y<-interval_df(x)
# neutrosophic summary
>nsummary(y)
```

Listing 3: Data conversion and neutrosophic summary.

```
# Generating 20 random values from the neutrosophic exponential distribution
>rnexp(10, rate_l = 2, rate_u = 4, stats = TRUE)
# Plotting the PDF and CDF
>plot_npdfexp(rate_l = 0.5, rate_u = 2, x = c(0, 8))
>plot_ncdfexp(rate_l = 0.5, rate_u = 2, x = c(0, 8))
```

Listing 4: Neutrosophic distribution and random data generation.

5. Illustrative examples

In this section, we introduce applications of neutrosophic data through both real and simulated datasets and explain the usage of the `neutrostat` package for analysis.

5.1. Real-world data examples

A gold price dataset, `goldprice`, is built into the `neutrostat` package and contains interval-valued monthly gold price data that were collected from six major Indian cities. Gold rates fluctuate significantly not only between different cities but also within the same day. Owing to this inherent uncertainty, the daily gold prices of 22-carat gold (in rupees per gram) were obtained. The data are originally from https://goo.gl/KoCNzt and have been used in neutrosophic analyses in previous studies [3,25].

The dataset spans from February 1, 2022, to January 31, 2023. For each month and city, the minimum and maximum gold prices were extracted, forming interval-valued observations. These intervals effectively represent the neutrosophic uncertainty that is inherent in daily price variations. The `goldprice` dataset can be loaded and viewed in R by using the following commands:

```
# Load the package
library(neutrostat)
# Load data
data(goldprice)
# View the first few rows
head(goldprice)
```

The existing format of the data is not compatible with the `neutrostat` package. We can now run the following command to make it compatible for further neutrosophic data analysis.

```
>data <- mapply(function(low, high) c(low, high), goldprice$Delhi_Low,goldprice$Delhi_High,
                SIMPLIFY = FALSE)
```

This dataset can be summarized directly by using a function (Listing 3) such as `nsummary()` to compute neutrosophic descriptive statistics.

```
>nsummary(data)
```

The output results are as follows:

```
    Statistic                Value
1    Neutro_min   [4897.65, 5434.28]
2   Neutro_mean [5239.0433, 5479.59]
3     Neutro_Q1   [5146.25, 5269.85]
4 Neutro_median  [5188.385, 5437.62]
5     Neutro_Q3  [5284.295, 5416.68]
6    Neutro_std [210.7693, 402.5725]
7    Neutro_max    [5668.74, 5959.2]
```

In addition to `goldprice`, `dioxin` represents another built-in dataset in the `neutrostat` package. The dataset contains interval-valued observations of daily dioxin ingestion levels (in picograms per kilogram of body weight per day) for a sample of individuals in Japan. Dioxin is a toxic environmental pollutant that is found in food, and its exposure levels often vary because of differences in contamination levels and dietary intake, which makes it a suitable candidate for neutrosophic analysis. We can use the `nsummary()` function to analyze the neutrosophic characteristics of this dataset, which provides an interval-valued summary of the neutrosophic data.

```
> data("dioxin")
> nsummary(dioxin)
      Statistic             Value
1    Neutro_min      [0.02, 0.91]
2   Neutro_mean [0.5488, 1.3747]
3     Neutro_Q1     [0.13, 1.135]
4 Neutro_median      [0.36, 1.47]
5     Neutro_Q3    [0.9145, 1.84]
6    Neutro_std [0.7889, 1.2473]
7    Neutro_max      [2.36, 2.44]
```

This output provides users with an informative statistical profile of neutrosophic health-related environmental data for better understanding.

5.2. Neutrosophic probability distributions

Classic probability models assume that all distributional parameters and observations are exactly known. Neutrosophic probability distributions, on the other hand, generalize classic models by incorporating uncertainty and indeterminacy not only in data but also in distributional parameters such as the mean, standard deviation, rate, scale, and shape. The `neutrostat` package offers functions as computational tools for working with such distributions. By setting values for their imprecision, users can simulate random data from neutrosophic versions of commonly used distributions, including the gamma (`ngam`) exponential (`nexp`), Laplace (`nlap`), normal (`nnorm`), Weibull (`nwbl`), and Rayleigh (`nray`) distributions. Each distribution function includes support for computing the neutrosophic PDF, the CDF, quantile values, and random variates. Additionally, graphical functions such as `plot_npdf()` and `plot_ncdf()` enable users to visualize the indeterminacy that is present in a distribution.

For example, the following commands simulate 10 observations of random data that follow a neutrosophic gamma distribution with an imprecise shape parameter ($\alpha = [0.5, 1.5]$) and an imprecise scale parameter ($\beta = [2, 2.5]$):

```
> set.seed(123)
> rngam(10, scale_l = 2, scale_u = 3, shape_l = 0.5, shape_u = 1.5, stats=FALSE)
```

The output result is as follows:

```
$samples
             [,1]       [,2]
 [1,] 0.13586934  2.0560148
 [2,] 1.55979756  6.7600260
 [3,] 0.28874937  2.8669227
 [4,] 2.45726123  8.8380570
 [5,] 3.55035490 11.1366212
 [6,] 0.00326357  0.4938222
 [7,] 0.51753632  3.7782715
 [8,] 2.58940186  9.1264229
 [9,] 0.57427536  3.9769665
[10,] 0.36930149  3.2135807
```

In the `rngam()` function, setting the argument `stats` to "true" yields additional outputs that contain summary statistics of the generated data.

```
$stats
  Statistic Simulated_Value
1      Mean    1.205, 5.225
2        SD    1.252, 3.516
3        Q1    0.309, 2.954
4    Median    0.546, 3.878
5        Q3    2.233, 8.319
6  Skewness    0.412, 0.758
7  Kurtosis     1.846, 2.08
```

The output indicates that the descriptive statistics of the simulated data are also ambiguous because of the assumed indeterminacy of the parameters of the neutrosophic gamma distribution. Similarly, users can determine the values of the PDF and CDF at specified values of the input parameters. For example, the function

```
dngam(x=2, scale_l=2, scale_u=3, shape_l=0.5, shape_u=1.5)
```

returns the following interval-valued PDF:

$$PDF = [PDF_l, PDF_u] = [0.1037, 0.1576]$$

Similarly, the function

```
pngam(q = 1.5, scale_l = 2, scale_u = 3, shape_l = 0.5, shape_u = 1.5)
```

returns the following interval-valued CDF:

$$CDF = [CDF_l, CDF_u] = [0.1987, 0.7793].$$

In addition to computing point values, users can also visualize the complete shape of the neutrosophic gamma distribution using these functions

`plot_npdfgam()` and `plot_ncdfgam()`. These functions generate interval-valued plots for the PDF and CDF, respectively, based on the specified input parameters. For example,

```
> plot_ncdfgam(scale_l = 0.3, scale_u = 0.50, shape_l = 2.5, shape_u = 3)
```

displays the full range of possible CDF curves (see Fig. 2), whereas the function

```
> plot_npdfgam(scale_l = 0.3, scale_u = 0.50, shape_l = 2.5, shape_u = 3)
```

constructs the corresponding PDF curve, as shown in Fig. 3. The indeterminate region can be visualized as a shaded region. These visualizations are helpful for understanding the uncertainty inherent in the behavior of the distribution due to imprecisely defined parameters.

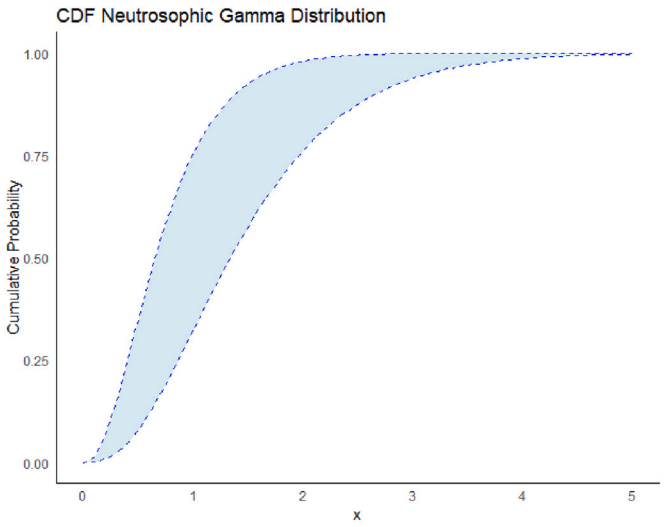


Fig. 2. CDF curve of the neutrosophic gamma distribution with the shaded band indicating the neutrosophic uncertainty range.

6. Impact of the software

The `neutrostat` package introduces a practical and comprehensive framework for analyzing neutrosophic data by providing tools that are grounded in neutrosophic logic. It fills a critical gap by providing neutrosophic statistical methods that can effectively handle the indeterminacies in data, whereas traditional models assume precise data information. A major research direction that is facilitated by `neutrostat`

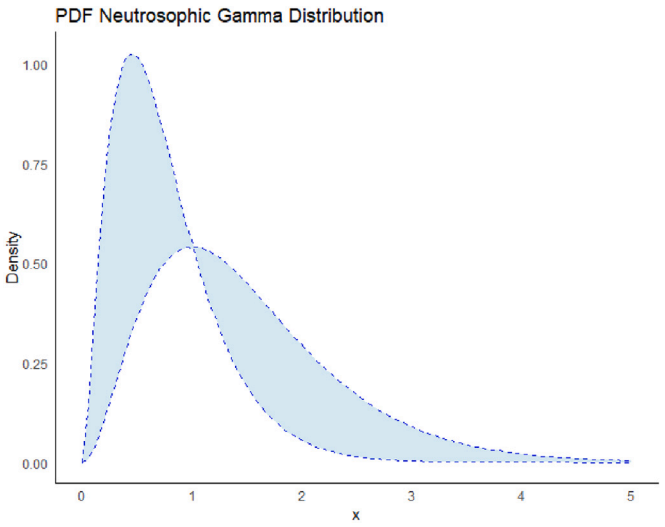


Fig. 3. PDF curve of the neutrosophic gamma distribution with the shaded band indicating the neutrosophic uncertainty range.

is the development and application of neutrosophic algebraic operations for imprecise data that commonly exist in real-world scenarios.

The package also supports the application of several advanced probability distributions to practical problems that involve indeterminate conditions. The statistical properties of neutrosophic distributions can now be explored both analytically and empirically through interval-valued functions, thus enabling users to study random phenomena under uncertainty. The software allows for the simulation and visualization of neutrosophic distributions through a user-friendly R interface, which makes it an invaluable resource for researchers who work with neutrosophic distribution theory. As a result, the package opens new research directions involving neutrosophic probability modeling, simulation, and inference, as many existing statistical distributions within the neutrosophic framework remain underdeveloped.

Moreover, built-in datasets such as goldprice, citytemp and dioxin enable users to apply methods directly to real-world scenarios, thereby facilitating their adoption in both academic research and applied settings. With citations in several recent neutrosophic statistics publications and the potential for integration into larger decision-support systems, `neutrostat` lays the groundwork for its adoption in applied research and possibly in commercial settings in the future.

7. Conclusions

The `neutrostat` R package represents a foundational and practical framework for performing statistical analyses under uncertainty by operationalizing the principles of neutrosophic statistics. It extends classic statistical functions to support interval-valued and indeterminate data, thus offering a versatile solution for users who address interval-valued data and indeterminate model parameters. Through its core functionalities, which range from descriptive summaries and visualization tools to modeling and randomly generating data from neutrosophic probability distributions, `neutrostat` fills a critical gap in neutrosophic data modeling within the R ecosystem.

This package supports new research directions in neutrosophic statistics, particularly regarding the development and application of neutrosophic probability distributions. By integrating neutrosophic probability distributions and supporting real-world datasets, `neutrostat` enables users to address problems that conventional statistical methods cannot fully capture. By making neutrosophic methods accessible through user-friendly R functions and well-documented examples, `neutrostat` opens doors for further theoretical developments and real-world applications of neutrosophic statistics.

Importantly, the current version of the package is limited to the descriptive aspects of neutrosophic statistics. Inferential methods such as hypothesis testing, confidence intervals, correlation, and regression fall beyond the present scope and may be considered in future extensions. Furthermore, based on the neutrosophic framework, the functionalities of the `neutrostat` package can be expanded by incorporating various classical data analysis approaches, such as neutrosophic clustering algorithms, which can increase its applicability to multivariate problems.

CRediT authorship contribution statement

Zahid Khan: Writing – original draft, Validation, Methodology, Investigation, Data curation, Conceptualization. **Zsolt T. Kosztyán:** Writing – original draft, Visualization, Validation, Software, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was implemented by the TKP2021-NVA-10 project with support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme.

Appendix A

The following additional mathematical expressions are included to support the theoretical development of the neutrosophic statistical methods.

Under the assumption that the given data consist of n neutrosophic observations $X_i = [x_{iL},\ x_{iU}]$, for i = 1, 2, …n, the mean, variance, skewness and kurtosis can be defined as follows in terms of the neutrosophic framework [26,27]:

$$\mu_N = \frac{\sum_{i=1}^{n} X_{iN}}{n} = \left[\frac{1}{n}\sum_{i=1}^{n} X_{iL},\ \frac{1}{n}\sum_{i=1}^{n} X_{iU}\right] \tag{A.1}$$

$$\sigma_N^2 = \frac{\sum_{i=1}^{n} (X_{iN} - \mu_N)^2}{n} = \left[\frac{1}{n}\sum_{i=1}^{n} (X_{iL} - \mu_U)^2,\ \frac{1}{n}\sum_{i=1}^{n} (X_{iU} - \mu_L)^2\right] \tag{A.2}$$

$$\mathrm{Skew}_N = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_{iL} - \mu_U}{\sigma_U}\right)^3,\ \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_{iU} - \mu_L}{\sigma_L}\right)^3\right] \tag{A.3}$$

$$\mathrm{Kurt}_N = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_{iL} - \mu_U}{\sigma_U}\right)^4,\ \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_{iU} - \mu_L}{\sigma_L}\right)^4\right] \tag{A.4}$$

We let $f_{Z_N}(z)$ be a probability distribution with imprecise parameters ρ_N. The neutrosophic distribution is defined as follows [28]:

$$f_{X_N}(Z) = \left[\min_{\rho\in[\rho_1,\rho_2]} f_X(Z),\ \max_{\rho\in[\rho_1,\rho_2]} f_X(Z)\right] \tag{A.5}$$

Eq. (A.5) shows that the neutrosophic probability mass or density function is evaluated under an indeterminate parameter $\rho \in [\rho_1, \rho_2]$. The function $f_{Z_N}(z)$ satisfies the fundamental properties of a probability density function and reflects model-level uncertainty.

References

[1] Rivieccio U. Neutrosophic logics: prospects and problems. Fuzzy Sets Syst 2008;159(14):1860–8. https://doi.org/10.1016/j.fss.2007.11.011

[2] Smarandache F. Neutrosophic statistics is an extension of interval statistics, while plithogenic statistics is the most general form of statistics (second version). Int J Neutrosophic Sci 2022;19(1):148–62. https://doi.org/10.54216/IJNS.190111

[3] Khan Z, Krebs KL. Enhancing neutrosophic data analysis: a review of neutrosophic measures and applications with neutrostat. Neutrosophic Sets and Systems 2025;78:181–90. https://doi.org/10.5281/zenodo.14232782

[4] Sherwani RAK, Aslam M, Shakeel H, Abbas K, Jamal F. Neutrosophic statistics for grouped data: theory and applications. In: Neutrosophic operational research: methods and applications. Springer; 2021. pp. 263–89. https://doi.org/10.1007/978-3-030-57197-9-14

[5] Riaz A, Sherwani RAK, Abbas T, Aslam M. Neutrosophic statistics and the medical data: a systematic review. Cogn Intell with Neutrosophic Stat in Bioinformatics 2023:357–72. https://doi.org/10.1016/B978-0-323-99456-9.00004-0

[6] Aslam M. Design of a new z-test for the uncertainty of Covid-19 events under neutrosophic statistics. BMC Med Res Methodol 2022;22(1):99. https://doi.org/10.1186/s12874-022-01593-x

[7] Miari M, Anan MT, Zeina MB. Neutrosophic two way Anova. Int J Neutrosophic Sci 2022;18(3):73–83. https://doi.org/10.54216/IJNS.180306

[8] Yadav VK, Prasad S. Neutrosophic estimators for estimating the population mean in survey sampling. Meas Interdiscip Res Perspect 2024;22(4):373–97. https://doi.org/10.1080/15366367.2023.2267835

[9] Masood S, Ibrar B, Shabbir J, Shokri A, Movaheedi Z. Estimating neutrosophic finite median employing robust measures of the auxiliary variable. Scientific Rep 2024;14(1):10255. https://doi.org/10.1038/s41598-024-60714-2

[10] Duan W-Q, Khan Z, Gulistan M, Khurshid A. Neutrosophic exponential distribution: modeling and applications for complex data analysis. Complexity 2021;2021(1):5970613. https://doi.org/10.1155/2021/5970613

[11] Ibrahim AMM, Alduais FS, Khan Z, Amin A, LaneKrebs K. Neutrosophic structure of the geometric model with applications. Neutrosophic Sets and Syst 2023;61:382–94.

[12] Aslam M. The neutrosophic negative binomial distribution: algorithms and practical application: accepted-August 2024. REVSTAT-Statistical Journal 2024. https://doi.org/10.57805/revstat.vi.781

[13] Alduais FS, Khan Z. Development of neutrosophic gamma distribution for modeling neonatal mortality data. Neutrosophic Sets and Syst 2025;79(1):48.

[14] Alamri FS, Aslam M. Data generation and application using the neutrosophic Erlang distribution. J Big Data 2025;12(1):48. https://doi.org/10.1186/s40537-025-01103-y

[15] Abd El-Latif AM, Almulhim FA, Noori NA, Khaleel MA, Alsaedi BSO. Properties with application to medical data for new inverse Rayleigh distribution utilizing neutrosophic logic. J Radiat Res Appl Sci 2025;18(2):101391. https://doi.org/10.1016/j.jrras.2025.101391

[16] Bashir S, Masood B, Shehzadi I, Al-Husseini Z, Aslam M. Neutrosophic lindley distribution: simulation, application, and comparative study. Contemporary Mathematics 2025:551–64. https://doi.org/10.37256/cm.6120256127

[17] Sleem A, Abdel-Baset M, El-Henawy I. Pyivns: a Python based tool for interval-valued neutrosophic operations and normalization. SoftwareX 2020;12:100632. https://doi.org/10.1016/j.softx.2020.100632

[18] Topal S, Broumi S, Bakali A, Talea M, Smarandache F. A Python tool for implementations on bipolar neutrosophic matrices. Neutrosophic Sets and Syst 2019;28:138–59.

[19] El-Ghareeb HA. Novel open source Python neutrosophic package. Neutrosophic Sets and Systems 2019;25:136–60.

[20] Broumi S, Le Hoang Son AB, Talea M, Smarandache F, Selvachandran G. Computing operational matrices in neutrosophic environments: a Matlab toolbox. Neutrosophic Sets and Syst 2017;18:58–66.

[21] Broumi S, Bakalı A, Talea M, Smarandache F. A Matlab toolbox for interval valued neutrosophic matrices for computer applications. International Journal of Management Information Systems and Computer Sciences 2017;1(1):1–21.

[22] Lee KH. First course on fuzzy theory and applications, vol. 27. Springer Science & Business Media; 2004.

[23] Moore RE. Methods and applications of interval analysis. SIAM; 1979.

[24] Smarandache F. Introduction to neutrosophic measure, neutrosophic integral, and neutrosophic probability. Sitech & Education Publisher, Craiova, Columbus; 2013.

[25] Mohan KR, Devi RN, Nagaram NB, Bharathi T, Rasappan S. Neutrosophic statistical analysis on gold rate. Neutrosophic Sets and Systems 2023;60:113–23.

[26] Smarandache F. Introduction to neutrosophic statistics. 2014. arXiv:1406.2000.

[27] Muhammad A. A study on skewness and kurtosis estimators of wind speed distribution under indeterminacy. Theor Appl Climatol 2021;143(3–4):1227–34. https://doi.org/10.1007/s00704-020-03509-5

[28] Alhabib R, Ranna MM, Farah H, Salama AA. Some neutrosophic probability distributions. Neutrosophic Sets and Systems 2018;22:30–8.