

RESEARCH

Open Access



# Analysis of imprecise measurement data utilizing z-test for correlation

Muhammad Aslam<sup>1\*</sup>

\*Correspondence:  
aslam\_ravian@hotmail.com

<sup>1</sup> Department of Statistics,  
Faculty of Science, King  
Abdulaziz University,  
21551 Jeddah, Saudi Arabia

## Abstract

The conventional Z-test for correlation, grounded in classical statistics, is typically employed in situations devoid of vague information. However, real-world data often comes with inherent uncertainty, necessitating an adaptation of the Z-test using neutrosophic statistics. This paper introduces a modified Z-test for correlation designed to explore correlations in the presence of imprecise data. We will present the simulation to check the effect of the measure of indeterminacy on the evolution of type-I error and the power of the test. The application of this modification is illustrated through an examination of heartbeat and temperature data. Upon analyzing the heartbeat and temperature data, it is determined that, in the face of indeterminacy, the correlation between heartbeat and temperature emerges as significant. This highlights the importance of accounting for imprecise data when investigating relationships between variables.

**Keywords:** Correlation, Medical data, Statistics, Imprecise data, Z-test

## Introduction

In medical science, correlation analysis has been used to investigate the degree of dependence between two medical variables. The correlation analysis tells about the strength of the relationship between two variables. The Z-test of correlation has been applied to investigating the significance of the correlation between two variables. For example, in medical science, the decision-makers are interested to see the significance of the relationship between blood pressure and diet. Based on correlation analysis, the medical decision-makers can suggest a suitable medication. Therefore, the statistical test investigates the correlation between the significance of the two variables under study. The null hypothesis that the correlation between two variables is insignificant is tested versus the alternative hypothesis that two variables are significantly correlated. Statistical tests have been widely used in medical science for decision-making. Gordon Lan et al. [13] introduced the weighted Z-tests and applied them in medical science. Bello et al. [7] provided a detailed discussion on the suitability of the statistical tests for medical studies. Mukaka [19] discussed the suitability of correlation analysis for medical data. Pandey et al. [21] discussed the importance of t-tests in medical-related problems. Schober et al. [23] discussed the correlation analysis for anesthesia data. Kc [17] wrote



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

a review on the applications of statistical methods for medical science. Janse et al. [15] discussed the limitations of correlation analysis using medical data. More applications of statistical methods in the medical field can be seen in [10, 18, 29, 32].

Statistical tests have been widely used for the analysis of measurement data. Grzesiek et al. [14] used the statistical test for the analysis of the temperature data. Avuçlu [6] presented the work on the detection covid-19 using statistical measurements. More applications of statistical analysis for the measurement data can be seen in [22, 27, 31].

The neutrosophic statistics were developed by [26] using the idea of neutrosophic logic developed by [25], and its efficiency of fuzzy logic and interval-analysis is shown by [9]. The applications of neutrosophic logic in medical science can be read in [8, 30]. Neutrosophic statistics are used for the collection of imprecise and interval data, analysis and interpretation of the imprecise data. The efficiency of neutrosophic statistics over classical statistics was discussed by [5, 11, 12]. Later on, the applications of neutrosophic statistics in the field of medical science were given by [1, 3, 5, 24].

The existing Z-test for correlation cannot be applied when the data is expressed in intervals or when uncertainty in parameters or level of significance is noted. To overcome this issue, in this paper, the Z-test for a single correlation coefficient using neutrosophic statistics will be presented. The test statistic for the proposed test will be developed and the application will be given using the heartbeat and body temperature data. It is expected that the proposed test will be efficient in investigating the significance of the correlation between variables expressed in intervals.

## Method

Let  $X_N = X_L + X_U I_{X_N}$ ;  $I_{X_N} \in [I_{X_L}, I_{X_U}]$  and  $Y_N = Y_L + Y_U I_{Y_N}$ ;  $I_{Y_N} \in [I_{Y_L}, I_{Y_U}]$  be two neutrosophic random variables of size  $n_N = n_L + n_U I_{n_N}$ ;  $I_{n_N} \in [I_{n_L}, I_{n_U}]$  follow the neutrosophic normal distribution with the neutrosophic means  $\mu_{X_N} = \mu_{X_L} + \mu_{X_U} I_{\mu_{X_N}}$ ;  $I_{\mu_{X_N}} \in [I_{\mu_{X_L}}, I_{\mu_{X_U}}]$  and  $\mu_{Y_N} = \mu_{Y_L} + \mu_{Y_U} I_{\mu_{Y_N}}$ ;  $I_{\mu_{Y_N}} \in [I_{\mu_{Y_L}}, I_{\mu_{Y_U}}]$  and neutrosophic standard deviation  $\sigma_{X_N} = \sigma_{X_L} + \sigma_{X_U} I_{\sigma_{X_N}}$ ;  $I_{\sigma_{X_N}} \in [I_{\sigma_{X_L}}, I_{\sigma_{X_U}}]$  and  $\sigma_{Y_N} = \sigma_{Y_L} + \sigma_{Y_U} I_{\sigma_{Y_N}}$ ;  $I_{\sigma_{Y_N}} \in [I_{\sigma_{Y_L}}, I_{\sigma_{Y_U}}]$  respectively. Note that  $X_L, Y_L, n_L, \mu_{X_L}, \mu_{Y_L}$  are the determinate parts denote the classical statistics,  $X_U I_{X_N}, Y_U I_{Y_N}, I_{n_N} \in [I_{n_L}, I_{n_U}], n_U I_{n_N}, \mu_{X_U} I_{\mu_{X_N}}, \mu_{Y_U} I_{\mu_{Y_N}}$  are indeterminate parts and  $I_{X_N} \in [I_{X_L}, I_{X_U}], I_{Y_N} \in [I_{Y_L}, I_{Y_U}], I_{\mu_{X_N}} \in [I_{\mu_{X_L}}, I_{\mu_{X_U}}], I_{\mu_{Y_N}} \in [I_{\mu_{Y_L}}, I_{\mu_{Y_U}}]$  are measures of indeterminacy. For designing of the proposed Z-test of a correlation coefficient, it is assumed that variance in  $X_N = X_L + X_U I_{X_N}$ ;  $I_{X_N} \in [I_{X_L}, I_{X_U}]$  should be independent from the variance in  $Y_N = Y_L + Y_U I_{Y_N}$ ;  $I_{Y_N} \in [I_{Y_L}, I_{Y_U}]$ . Suppose that  $r_N = r_L + r_U I_{r_N}$ ;  $I_{r_N} \in [I_{r_L}, I_{r_U}]$  is neutrosophic correlation between  $X_N = X_L + X_U I_{X_N}$ ;  $I_{X_N} \in [I_{X_L}, I_{X_U}]$  and  $Y_N = Y_L + Y_U I_{Y_N}$ ;  $I_{Y_N} \in [I_{Y_L}, I_{Y_U}]$  that is given by the following [2] as

$$r_N \in \left\{ \frac{\sum_{i=1}^{n_L} (X_{iL} - \bar{X}_L)(Y_{iL} - \bar{Y}_L)}{\sqrt{\sum_{i=1}^{n_L} (X_{iL} - \bar{X}_L)^2 \sum_{i=1}^{n_L} (Y_{iL} - \bar{Y}_L)^2}}, \frac{\sum_{i=1}^{n_U} (X_{iU} - \bar{X}_U)(Y_{iU} - \bar{Y}_U)}{\sqrt{\sum_{i=1}^{n_U} (X_{iU} - \bar{X}_U)^2 \sum_{i=1}^{n_U} (Y_{iU} - \bar{Y}_U)^2}} \right\}; \quad (1)$$

$$I_{r_N} \in [I_{r_L}, I_{r_U}]$$

where  $\bar{X}_L$  and  $\bar{X}_U$  are the lower and upper values of the neutrosophic sample average. To investigate either correlation coefficient  $r_N \in [r_L, r_U]$  differs significantly from the

specified correlation  $r_{0N} \in [r_{0L}, r_{0U}]$ , the null hypothesis that the correlation coefficient  $r_N \in [r_L, r_U]$  is at least  $r_{0N} \in [r_{0L}, r_{0U}]$  vs. the alternative hypothesis correlation coefficient  $r_N \in [r_L, r_U]$  is at most  $r_{0N} \in [r_{0L}, r_{0U}]$ . Using the Fisher's transformation, the value of quantity  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  is calculated as

$$Z_{1N} = Z_{1L} + Z_{1U}I_{Z_{1N}}; I_{Z_{1N}} \in [I_{Z_{1L}}, I_{Z_{1U}}] \quad (2)$$

The quantity  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  can be written as

$$Z_{1N} \in \left\{ \frac{1}{2} \log_e \left( \frac{1+r_L}{1-r_L} \right), \frac{1}{2} \log_e \left( \frac{1+r_U}{1-r_U} \right) \right\} \quad (3)$$

Note that  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  follows the neutrosophic normal distribution. The mean, say  $\mu_{Z_{1N}}$  of  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  is given by

$$\mu_{Z_{1N}} \in \left\{ \frac{1}{2} \log_e \left( \frac{1+\rho_{0L}}{1-\rho_{0L}} \right), \frac{1}{2} \log_e \left( \frac{1+\rho_{0U}}{1-\rho_{0U}} \right) \right\} \quad (4)$$

where  $\rho_{0N}$  represents the specified value of the correlation coefficient.

The variance, say  $\sigma_{Z_{1N}}$  of  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  is given by

$$\sigma_{Z_{1N}} \in \left\{ \frac{1}{\sqrt{n_L-3}}, \frac{1}{\sqrt{n_U-3}} \right\} \quad (5)$$

The neutrosophic test statistic  $Z_N \in [Z_L, Z_U]$  is defined as

$$Z_N = Z_L + Z_U I_{Z_N}; I_{Z_N} \in [I_{Z_L}, I_{Z_U}] \quad (6)$$

The statistic  $Z_N \in [Z_L, Z_U]$  can be written as

$$Z_N = \frac{Z_{1L} - \mu_{Z_{1L}}}{\sigma_{Z_{1L}}} + \frac{Z_{1U} - \mu_{Z_{1U}}}{\sigma_{Z_{1U}}} I_{Z_N}; I_{Z_N} \in [I_{Z_L}, I_{Z_U}] \quad (7)$$

Note that the proposed statistic  $Z_N \in [Z_L, Z_U]$  is an extension of the existing Z-test. The proposed  $Z_N \in [Z_L, Z_U]$  reduces to the existing Z-test under classical statistics when  $I_{Z_L}=0$ .

## Simulation studies

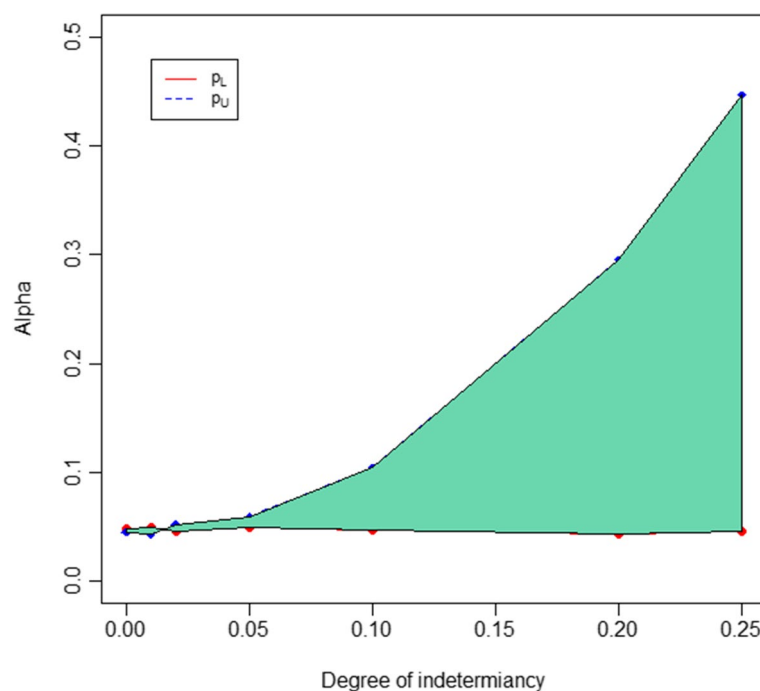
In this section, we will simulate the impact of the measure of indeterminacy on the type-I error, denoted by  $\alpha$ , which represents the probability of rejecting the null hypothesis when it is true. Additionally, we will investigate the way in which indeterminacy affects the power of the test  $(1 - \beta)$ , with  $\beta$  representing the probability of failing to reject the null hypothesis when it is false. Following the approach outlined by [28], we define the neutrosophic variable  $\rho_{0N}$  as  $\rho_{0L} + \rho_{0U}I_N$ , where  $\rho_{0L}$  represents the determinate value of the correlation coefficient,  $\rho_{0U}I_N$  is the indeterminate part, and  $I_N$  belongs to the interval  $[I_L, I_U]$ , representing the degree of indeterminacy. Our analysis will initially focus on the impact of  $I_N$  on the type-I error and subsequently examine its effect on the power of the test.

### Effect of $I_N$ on type-I error

We will examine the impact of the measure of indeterminacy on the type-I error through a simulation conducted  $10^6$  times. Following the approach outlined in [20], the type-I error is computed as the ratio of rejecting the null hypothesis to the total number of replicates. The type-I error values for both the classical statistics test and the proposed test, across various values of  $I_N$ , are depicted in Fig. 1. The lower curve in Fig. 1 represents the type-I error for the test using classical statistics, while the upper curve displays the values for the proposed test. The observation from Fig. 1 is that, for the classical statistics test, the type-I error remains consistent across all levels of indeterminacy. Conversely, the higher curve indicates an increase in the type-I error as the values of  $I_N$  rise. This suggests a significant effect of the measure of indeterminacy on the evaluation of the type-I error, cautioning decision-makers to exercise care when making decisions regarding hypothesis testing in the presence of uncertainty.

### Effect of $I_N$ on type-II error

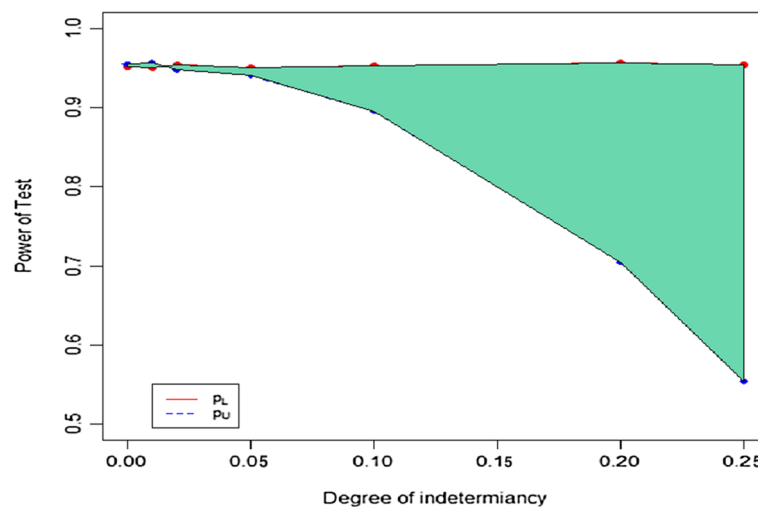
We will assess how the degree of indeterminacy influences the test's efficacy through a simulation conducted a million times. Following the methodology outlined by [20], the type-II error is calculated as the ratio of incorrect decisions to the total number of replicates. Table 1 presents the type-II error values for both the conventional statistical test and the proposed test across various levels of  $I_N$ . Figure 2 illustrates the trends in test power. In Fig. 2, the lower curve represents the power of the test using classical statistics, while the upper curve depicts the power of the test for the proposed method. Figure 2 reveals that, for the classical statistics test, the power remains consistent regardless of the level of indeterminacy. In contrast, the higher curve indicates a decline in test



**Fig. 1** Graphs illustrating the type-I error for both tests

**Table 1** The power of the test

$I_N$	$[p_L, p_U]$
0	[0.9523, 0.9551]
0.01	[0.9505, 0.9571]
0.02	[0.9545, 0.948]
0.05	[0.9503, 0.9408]
0.10	[0.9529, 0.8958]
0.20	[0.9567, 0.7047]
0.25	[0.9544, 0.5536]

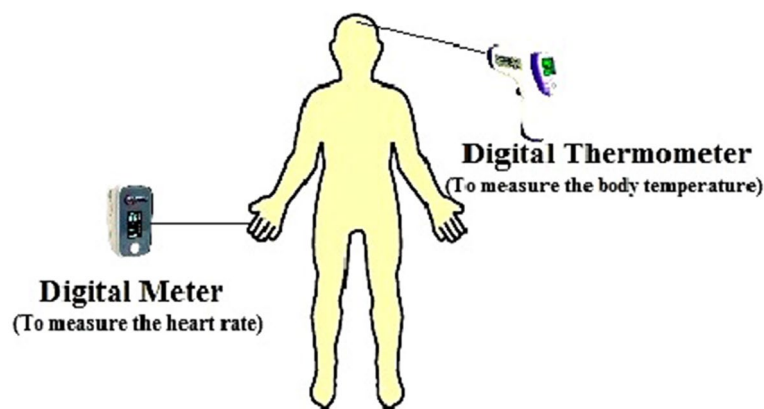


**Fig. 2** Power of the test curves

power as  $I_N$  values increase. This implies a significant impact of the measure of indeterminacy on test performance. This study suggests that, unlike the classical statistics test, the proposed test's power is affected by the degree of indeterminacy. Consequently, it is concluded that relying on the existing test under classical statistics may lead decision-makers astray when making decisions in the presence of uncertainty.

## Application

This section presents the application of the proposed Z-test for correlation using the heartbeat (HBT) and temperature (TMP) data. The medical decision-makers are interested in investigating the relationship between the HBT and TMP. The primary and secondary healthcare department is responsible for the delivery of essential and effective health services in the province of Punjab, Pakistan. Punjab Health Facilities Management Company (PHFMC) on behalf of the health department engages in providing the required services. Basic Health Unit (BHU) is the first level health care unit under the supervision of qualified doctors which usually covers around 10,000 to 25,000 population. Three months (June 2021 to August 2021) patients' daily data who visited BHU with reporting gastritis (authenticated and reported by a qualified medical doctor). The minimum and maximum values of the patients visited in a day



**Fig. 3** The schematic diagram

are recorded and the data of two variables is arranged in intervals. The schematic diagram is shown in Fig. 3. The interval data of HBT and TMP is recorded and reported in Table 2. From the data given in Table 2, it can be seen that the medical decision-makers cannot apply the existing Z-test to investigate the significance of the correlation between HBT and TMP. The use of the proposed Z-test for correlation seems suitable for analyzing HBT and TMP data.

The proposed Z-test for correlation using the HBT and TMP data is carried out as: the neutrosophic correlation  $r_N \in [r_L, r_U]$  is calculated as follows  $r_N \in [0.2024, 0.4333]$  and expressed in neutrosophic form as  $r_N = 0.2024 + 0.4333I_{r_N}$ ;  $I_{r_N} \in [0, 0.5329]$ . The quantity  $Z_{1N} \in [Z_{1L}, Z_{1U}]$  is calculated as  $Z_{1N} \in [0.2052, 0.4640]$ . The mean and standard deviation are calculated as  $\mu_{Z_{1N}} \in [0.5493, 0.5493]$  and  $\sigma_{Z_{1N}} \in [0.1222, 0.1222]$ , respectively. The proposed test statistic  $|Z_N \in [Z_L, Z_U]|$  is calculated as  $Z_N = 2.8162 - 0.6981I_{Z_N}$ ;  $I_{Z_N} \in [0, 3.0341]$ . Suppose that the value of the level of significance  $\alpha = 0.05$ . The proposed test for investigating the relationship between HBT and TMP is implemented as.

Step 1: State null hypothesis  $H_{0N}$  : correlation between HBT and TMP is  $r_{0N} = 0.50$  vs. the alternative hypothesis  $H_{1N}$  : correlation between HBT and TMP is less than  $r_{0N} = 0.50$ .

Step 2: For  $\alpha = 0.10$ , the tabulated value from [16] is 1.64.

Step 3: Compare  $Z_N \in [2.8162, 0.6981]$  with 1.64 and reject  $H_{0N}$  if  $Z_N \in [2.8162, 0.6981] > 1.64$ .

By comparing the values of  $Z_N \in [2.8162, 0.6981]$  with 1.64, it is clear that the lower value of  $Z_L$  is larger than 1.64, so, the null hypothesis  $H_{0N}$  will be rejected in favor of  $H_{1N}$ . On the other hand, the upper value of statistic  $Z_U$  is smaller than 1.64 which leads to rejection of  $H_{1N}$ . From the analysis, it is clear that the determinate part which presents the statistic using classical statistics indicates that the correlation between HBT and TMP is less than 0.50. On the other hand, the indeterminate part shows that the correlation between HBT and TMP is 0.50. Under uncertainty, it is expected that there will be significant correlation between HBT and TMP.

**Table 2** The HBT and TEMP data

Observation#	HBT	TEM	Observation#	HBT	TEM
1	[70, 85]	[98, 99]	36	[72, 84]	[98, 102]
2	[70, 90]	[98, 100]	37	[72, 74]	[99, 98]
3	[78, 90]	[98, 100]	38	[72, 72]	[99, 98]
4	[70, 85]	[98, 98]	39	[72, 84]	[98, 100]
5	[78, 82]	[100, 98]	40	[72, 70]	[98, 99]
6	[78, 82]	[98, 98]	41	[70, 72]	[98, 98]
7	[74, 70]	[98, 98]	42	[72, 84]	[98, 101]
8	[73, 75]	[99, 99]	43	[72, 74]	[98, 99]
9	[72, 82]	[98, 99]	44	[72, 72]	[98, 98]
10	[70, 85]	[98, 100]	45	[72, 70]	[99, 98]
11	[70, 75]	[99, 101]	46	[72, 84]	[98, 100]
12	[72, 85]	[98, 101]	47	[74, 85]	[98, 100]
13	[78, 90]	[99, 100]	48	[72, 84]	[98, 100]
14	[72, 80]	[99, 101]	49	[70, 72]	[98, 98]
15	[69, 80]	[99, 99]	50	[70, 74]	[98, 98]
16	[68, 81]	[98, 101]	51	[72, 84]	[98, 101]
17	[70, 74]	[98, 99]	52	[80, 74]	[99, 98]
18	[70, 80]	[99, 98]	53	[72, 70]	[98, 98]
19	[72, 74]	[98, 99]	54	[72, 84]	[99, 100]
20	[70, 82]	[99, 100]	55	[72, 74]	[99, 99]
21	[70, 79]	[98, 101]	56	[72, 74]	[99, 98]
22	[70, 82]	[99, 101]	57	[72, 82]	[98, 101]
23	[70, 85]	[98, 100]	58	[72, 85]	[98, 99]
24	[78, 74]	[98, 100]	59	[74, 78]	[99, 98]
25	[74, 85]	[100, 101]	60	[76, 85]	[98, 98]
26	[72, 46]	[98, 99]	61	[76, 86]	[98, 100]
27	[74, 90]	[99, 100]	62	[72, 84]	[98, 100]
28	[72, 84]	[98, 100]	63	[70, 74]	[98, 98]
29	[72, 76]	[98, 99]	64	[72, 82]	[98, 100]
30	[70, 74]	[98, 100]	65	[74, 84]	[98, 100]
31	[70, 80]	[98, 100]	66	[72, 84]	[98, 101]
32	[72, 74]	[99, 98]	67	[72, 80]	[98, 99]
33	[74, 88]	[98, 101]	68	[74, 74]	[100, 98]
34	[74, 72]	[98, 100]	69	[74, 85]	[98, 100]
35	[70, 70]	[98, 100]	70	[74, 88]	[98, 98]

### Comparative studies based on HBT and TMP data

Based on the analysis of HBT and TMP data, the comparisons of the proposed Z-test for correlation are carried out with the existing Z-test for correlation using the classical statistics, fuzzy-based test and interval-statistics in terms of information and flexibility of the results. The neutrosophic forms of the correlation  $r_N$  and the test statistic  $|Z_N \in [Z_L, Z_U]|$  are presented as  $r_N = 0.2024 + 0.4333I_{r_N}; I_{r_N} \in [0, 0.5329]$  and  $Z_N = 2.8162 - 0.6981I_{Z_N}; I_{Z_N} \in [0, 3.0341]$ , respectively. From the results, it can be analyzed that under indeterminacy, the correlation between HBT and TMP may vary from 0.2024 to 0.4333. The values of statistic  $Z_N \in [Z_L, Z_U]$  may vary from 2.8162 to 0.6981. Note that as mentioned before the first values 0.2024, and 2.8162 present the

results of the Z-test for correlation using classical statistics. The statistical test using classical statistics states that the probability of rejecting  $H_{0N}$  : the correlation between HBT and TMP is  $r_{0N} = 0.50$  when it is true is 0.05 and the probability of accepting  $H_{0N}$  : a correlation between HBT and TMP is  $r_{0N} = 0.50$  is 0.95. On the other hand, the proposed test for correlation states that the probability of rejecting  $H_{0N}$  : the correlation between HBT and TMP is  $r_{0N} = 0.50$  when it is true is 0.05, the probability of accepting  $H_{0N}$  : correlation between HBT and TMP is  $r_{0N} = 0.50$  is 0.95 and the measure of indeterminacy/uncertainty associated with the decision is 3.0341. Similarly, the Z-test using fuzzy-logic gives the information about the statistic  $Z_N \in [Z_L, Z_U]$  in intervals only. According to the fuzzy-based statistical test, it can be expected that the values of  $Z_N \in [Z_L, Z_U]$  may vary from 2.8162 to 0.6981. The fuzzy-based analysis and interval-analysis only give information in intervals and are unable to give any information about the measure of indeterminacy. From the comparative studies, it is concluded that the proposed Z-test for correlation is more informative than the test using classical statistics, fuzzy-based analysis and interval-based analysis.

### Discussions based on HBT and TMP data

The main aim of the paper is to investigate the significance of the relationship between HBT and TMP. The neutrosophic form correlation analysis of HBT and TMP is  $r_N = 0.2024 + 0.4333I_{r_N}$ ;  $I_{r_N} \in [0, 0.5329]$ . The correlation analysis of HBT and TMP shows that the correlation between HBT and TMP may vary from 0.2024 to 0.4333. As mentioned earlier, the correlation value 0.2024 denotes the correlation using classical statistics. The value  $0.4333I_{r_N}$  denotes the correlation related to the indeterminate part. From the correlation analysis of the determined part that is 0.2024, it can be seen that there is weak correlation between HBT and TMP. It means that the increase in TMP does not increase the HBT significantly. Under indeterminacy, the correlation of the indeterminate part is 0.4333. This means that there is a moderate correlation between HBT and TMP. It means that the increase in TMP may increase HBT. From the correlation analysis, it can be seen that although the correlation of the determinate part is insignificant as the measure of indeterminacy increases, it can increase the correlation between HBT and TMP. Therefore, the decision makers should be careful in dealing with patents having the diseases of HBT and TMP.

### Concluding remarks

The paper discussed the adaptation of the Z-test of correlation through the application of neutrosophic statistics. It provided an explanation of the rationale behind employing the proposed Z-test and detailed the neutrosophic test statistic along with the corresponding implementation steps. The simulation study conducted in the paper led to the conclusion that there is a notable impact of indeterminacy on both the type-I error and the power of the test. The paper demonstrated the application of the proposed test using data from HBT and TMP intervals. The findings revealed that, in the presence of indeterminacy or when dealing with interval data, the correlation between HBT and TMP increases as the measure of indeterminacy rises. The analysis suggests that decision-makers can effectively use the proposed test to explore correlations between variables



in diverse fields such as medical science, business, and industry. Additionally, the paper suggested avenues for future research, including the exploration of the proposed test using a resampling scheme. It also recommended further investigation into additional statistical properties as potential areas for future research.

#### Acknowledgements

We express our gratitude to the editor and reviewers for their valuable suggestions, which have significantly contributed to enhancing the quality and presentation of the paper.

#### Author contributions

MA wrote the paper.

#### Funding

None.

#### Availability of data and materials

The data is given in the paper.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

No conflict of interest regarding the paper.

Received: 14 February 2023 Accepted: 20 December 2023

Published online: 02 January 2024

#### References

1. Aslam M. A new method to analyze rock joint roughness coefficient based on neutrosophic statistics. *Measurement*. 2019;146:65–71.
2. Aslam M, Albassam M. Application of neutrosophic logic to evaluate correlation between prostate cancer mortality and dietary fat assumption. *Symmetry*. 2019;11(3):330.
3. Aslam M, Arif OH, Sherwani RAK. New diagnosis test under the neutrosophic statistics: an application to diabetic patients. *BioMed Res Int*. 2020. <https://doi.org/10.1155/2020/2086185>.
4. Aslam M. Chi-square test under indeterminacy: an application using pulse count data. *BMC Med Res Methodol*. 2021;21:1–5.
5. Aslam M. Assessing the significance of relationship between metrology variables under indeterminacy. *Mapan*. 2021;37:119–24.
6. Avuçlu E. COVID-19 detection using X-ray images and statistical measurements. *Measurement*. 2022;201: 111702.
7. Bellolo MF, Serrano LA, Stead LG. Understanding statistical tests in the medical literature: which test should I use? *Int J Emerg Med*. 2008;1(3):197–9.
8. Broumi S, Deli I. Correlation measure for neutrosophic refined sets and its application in medical diagnosis. *Infinite Study*; 2015.
9. Broumi S, Smarandache F. Correlation coefficient of interval neutrosophic set. *Appl Mech Mater*. 2013;436:511–7.
10. Chen J, Talha M. Audit data analysis and application based on correlation analysis algorithm. *Comput Math Methods Med*. 2021;2021:1–11.
11. Chen J, Ye J, Du S. Scale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics. *Symmetry*. 2017;9(10):208.
12. Chen J, Ye J, Du S, Yong R. Expressions of rock joint roughness coefficient using neutrosophic interval statistical numbers. *Symmetry*. 2017;9(7):123.
13. Gordon Lan K, Soo Y, Siu C, Wang M. The use of weighted Z-tests in medical research. *J Biopharm Stat*. 2005;15:625–39.
14. Grzesiek A, Zimroz R, Śliwiński P, Gomolla N, Wylomańska A. Long term belt conveyor gearbox temperature data analysis—statistical tests for anomaly detection. *Measurement*. 2020;165: 108124.
15. Janse RJ, Hoekstra T, Jager KJ, Zoccali C, Tripepi G, Dekker FW, van Diepen M. Conducting correlation analysis: important limitations and pitfalls. *Clin Kidney J*. 2021;14:2332–7.
16. Kanji GK. 100 statistical tests. Sage; 2006.
17. Kc B. A note on the application of advanced statistical methods in medical research. *Biomed J Sci Tech Res*. 2018;11(2):8476–9.
18. Lin L, Wu F, Chen W, Zhu C, Huang T. Research on urban medical and health services efficiency and its spatial correlation in china: based on panel data of 13 cities in Jiangsu Province. Paper presented at the Healthcare; 2021.

19. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24(3):69–71.
20. Nosakhare UH, Bright AF. Statistical analysis of strength of W/S test of normality against non-normal distribution using Monte Carlo simulation. *Am J Theor Appl Stat.* 2017;6(5–1):62–5.
21. Pandey R. Commonly used t-tests in medical research. *J Pract Cardiovasc Sci.* 2015;1(2):185.
22. Rivas T, Pozo-Antonio J, Barral D, Martínez J, Cardell C. Statistical analysis of colour changes in tempera paints mock-ups exposed to urban and marine environment. *Measurement.* 2018;118:298–310.
23. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg.* 2018;126(5):1763–8.
24. Sherwani RAK, Shakeel H, Saleem M, Awan WB, Aslam M, Farooq M. A new neutrosophic sign test: an application to COVID-19 data. *PLoS ONE.* 2021;16(8): e0255671.
25. Smarandache F. Neutrosophy Neutrosophic probability, set, and logic, proQuest information & learning, vol. 105. *Infinity Study*; 1998. p. 118–23.
26. Smarandache F. Introduction to neutrosophic statistics. *Infinite Study*; 2014.
27. Sofińska K, Cieśla M, Barbasz J, Wilkosz N, Lipiec E, Szymoński M, Białas P. Double-strand breaks quantification by statistical length analysis of DNA fragments imaged with AFM. *Measurement.* 2022;198: 111362.
28. Wu B, Chang S-K. On testing hypothesis of fuzzy sample mean. *Jpn J Ind Appl Math.* 2007;24:197–209.
29. Yazici AC, Ögüş E, Ankarali H, Gürbüz F. An application of nonlinear canonical correlation analysis on medical data. *Turk J Med Sci.* 2010;40(3):503–10.
30. Zhang D, Zhao M, Wei G, Chen X. Single-valued neutrosophic TODIM method based on cumulative prospect theory for multi-attribute group decision making and its application to medical emergency management evaluation. *Econ Res Ekonomska Istraživanja.* 2021;35:4520–36.
31. Zhang Y, Chen Z, Zhu Z, Wang X. A sampling method for blade measurement based on statistical analysis of profile deviations. *Measurement.* 2020;163: 107949.
32. Zhuang X, Yang Z, Cordes D. A technical review of canonical correlation analysis for neuroscience applications. *Hum Brain Map.* 2020;41(13):3807–33.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)