A New Way of Investigating the Relationship Between Fasting Blood Sugar Level and Drinking Glucose Solution

Muhammad Aslam* and Mohammed Albassam†

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

The existing t-test of a correlation coefficient works under a determinate environment. In uncertainty, the existing t-test of a correlation coefficient is unable to investigate the significance of correlation. The study presents a modification of the existing t-test of a correlation coefficient using neutrosophic statistics. The test statistic is designed to investigate the significance of correlation when imprecise observations or uncertainties in the level of significance are presented. The test is applied to data obtained from patients with diabetes. From the data analysis, the proposed t-test of a correlation coefficient is found to be more effective than existing tests.

Keywords: classical statistics, imprecise observations, medical data, neutrosophy, simulation

INTRODUCTION

Correlation analysis is conducted to see the degree of relationship between two variables. Correlation analysis helps in determining the positive or negative correlation between two variables. The value of a correlation coefficient lies between −1 and +1. The computed value of the correlation coefficient from data always lies in this interval. Statistical tests have been conducted in various fields for decision-making purposes. Among the statistical tests, the t-test for correlation is applied to investigate the significance of the correlation between two variables. In the t-test for correlation, the null hypothesis that there is no association between two variables is tested against the alternative hypothesis that two variables are associated. Values of the statistic of the t-test for correlation are calculated from given data and compared with tabulated values. The null hypothesis of no association between two variables is accepted if the calculated value is less than the tabulated value. Bartroff and Song (1) conducted a correlation analysis to investigate the relationship between impact factors and the ranking of electrical journals. Alexiandre-Benavent et al. (2) discussed the correlation between impact factors and published papers’ research data. McGillivray and Astell (3) provided a correlation between usage and citations of open access journals. For more information, the reader may refer to Tang and Landes (4) and e Silva et al. (5).

Correlation analysis has been widely applied in medical research. It is conducted to investigate the association between variables causing a specific disease. Schober et al. (6) applied correlation analysis on anesthesia data. Najmi and Balakrishnan Sadavisam (7) provided guidelines for medical students related to statistical tests. Statistical analysis has also been conducted to investigate the relationship among various causes of diabetes. Khan et al. (8) discussed a statistical analysis for patients with diabetes. Liu et al. (9) applied various statistical methods for data on diabetes. Wani et al. (10) investigated the effect of weight and smoking on type-2 diabetes. Nedyalkova et al. (11)
presented a statistical analysis on type-2 diabetes. Adaobi et al. (12) presented an analysis using blood glucose data. More information can be seen in Mukasheva et al. (13), Eynizadeh et al. (14), Balamurugan et al. (15), Alsaqr (16), Janse et al. (17), and Sun et al. (18).

Neutrosophic logic was introduced by Smarandache (19), which is the generalization of fuzzy-logic. The former gives information about three measures (truth, false, and indeterminacy), while the latter gives information about two measures (truth and false). Smarandache (20) showed the efficiency of neutrosophic logic over fuzzy-logic and interval-based analysis. Neutrosophic logic has many applications in medical science. Ansari et al. (21) discussed a application of neutrosophic sets in artificial intelligence. Jafar et al. (22) used the neutrosophic logic in medical diagnosis. Basha et al. (23) applied neutrosophic logic in the classification of X-rays of the chest of patients with coronavirus disease 2019 (COVID-19). More information on applications of neutrosophic logic in medical science can be seen in Zhang et al. (24) and Zhang et al. (25). Neutrosophic statistics was developed by Smarandache (26) using the idea of neutrosophy in numbers. Chen et al. (27) and Chen et al. (28) discussed methods to analyze neutrosophic data. Aslam et al. (29) applied neutrosophic statistics on diabetics’ data. Ling et al. (30) analyzed neutrosophic numbers in medical waste treatment. More applications can be seen in Das et al.’s studies (31) and Saeed et al.’s studies (32).

Aslam (33) proposed a neutrosophic Z-test for two samples to investigate the relationship between metrological variables. By exploring the literature and to the best of our knowledge, there is still a gap to work on t-test for correlation under neutrosophic statistics for a single sample. We will present the design of a t-test for correlation using neutrosophic statistics in this study. The neutrosophic statistic will be given and applied using data obtained from patients with diabetes. We expect that the proposed t-test for correlation will beat the existing t-tests for correlations in terms of information, adequacy, and flexibility.

METHOD

The existing t-test of a correlation coefficient using classical statistics works only when decision-makers are sure about the parameters involved in the implementation of the test. In practice, it may not be possible to the level of significance, sample size, and observations obtained from a measurement or a complex process are always certain; see, for example, Doll and Carney (34). Modification of the existing t-test of a correlation coefficient is needed to investigate the significance of the correlation between variables in an indeterminate environment. Now, we will develop a t-test of a correlation coefficient using neutrosophic statistics with the expectation that the proposed test will be efficient and a general form of the existing t-test of a correlation coefficient. The procedure of the proposed t-test of a correlation coefficient using neutrosophic statistics will be explained as: Let $X_N = X_L + X_U I_{SN}; I_{SN} \in \{I_{L}, I_{U}\}$ and $Y_N = Y_L + Y_U I_{SN}; I_{SN} \in \{I_{L}, I_{U}\}$ be neutrosophic random variables, where the first values denote the determinate parts, the second values denote the indeterminate parts, and $I_{SN} \in \{I_{L}, I_{U}\}$. The neutrosophic correlation $r_{SN}$ can be written as:

$$
I = \frac{\sum r_{i}^{+} \cdot (X_{i} - \bar{X}) \cdot (Y_{i} - \bar{Y})}{\sqrt{\sum r_{i}^{+} \cdot (X_{i} - \bar{X})^{2}} \cdot \sqrt{\sum r_{i}^{+} \cdot (Y_{i} - \bar{Y})^{2}}}
$$

where $r_{IL} = r_{IU}$, and the neutrosophic correlation $r_{SN}$ can be written as:

$$
I = \frac{\sum r_{SN} \cdot (X_{iS} - \bar{X}_{S}) \cdot (Y_{iS} - \bar{Y}_{S})}{\sqrt{\sum r_{SN} \cdot (X_{iS} - \bar{X}_{S})^{2}} \cdot \sqrt{\sum r_{SN} \cdot (Y_{iS} - \bar{Y}_{S})^{2}}}
$$

Note that the neutrosophic correlation $r_{SN}$ is a generalization of the existing correlation under classical statistics. The neutrosophic correlation $r_{SN}$ reduces to correlation using classical statistics when $I_{SN} = 0$. To test the null hypothesis $H_{0}$ that there is no correlation between the variables vs. the alternative hypothesis $H_{1}$ that both variables are associated, the neutrosophic test statistic $t_{SN}$ is defined as:

$$
t_{SN} = t_{L} + t_{U} I_{SN} \in \{I_{L}, I_{U}\}
$$

The alternative form of $t_{SN}$ is defined as:

$$
t_{SN} = \frac{r_{IL} \times \sqrt{n_{L} - 2} + \frac{r_{IL}}{\sqrt{1 - r_{IL}^{2}}}}{\sqrt{n_{U} - 2} I_{SN} \in \{I_{L}, I_{U}\}}
$$

where $t_{IL} = t_{U}$ and the neutrosophic correlation $t_{SN}$ can be written as:

$$
t_{SN} = (1 + I_{SN}) \times \frac{r_{SN} \times \sqrt{n_{SN} - 2} + I_{SN} \in \{I_{SL}, I_{SU}\}}{\sqrt{1 - r_{SN}^{2}}}
$$
Note that $t_{N \in \mathcal{I}L, t_{U}}$ follows the neutrosophic $t$-distribution with the degree of freedom $n_{N} - 2$.

Note that the neutrosophic statistics $t_{N \in \mathcal{I}L, t_{U}}$ is a generalization of the existing statistics under classical statistics. The neutrosophic statistics $t_{N \in \mathcal{I}L, t_{U}}$ reduces to statistic using classical statistics when $I_{r} = 0$.

The proposed $t$-test of a correlation coefficient will be carried out through the following steps:

**Step-1:** state $H_{0N}: r_{N \in \mathcal{I}L, t_{U}} = 0$ vs. $H_{1N}: r_{N \in \mathcal{I}L, t_{U}} \neq 0$;

**Step 2:** fix the level of significance $\alpha_{N} = \alpha_{L} + \alpha_{U}$ and select the tabulated value $t_{C}$ from (36);

**Step 3:** compute statistic $t_{N \in \mathcal{I}L, t_{U}}$ and compare with $t_{C}$;

**Step 4:** do not reject $H_{0N}: r_{N \in \mathcal{I}L, t_{U}} = 0$ if $t_{N \in \mathcal{I}L, t_{U}} < t_{C}$.

**APPLICATION USING DATA ON DIABETES**

To investigate the significance of the correlation between the sugar level and Drinking Glucose Solution about 237 ml contained 75 g of sugar, the data from 320 diabetics patients aged from 45 to 60 were collected from a hospital located in Hafizabad, Pakistan. A group of 20 patients at each age level is formed and the minimum and maximum blood sugar levels are recorded from 16 groups of patients. The schematic diagram to measure blood sugar level is depicted in Figure 1.

The data of blood sugar levels are reported in Table 1. The minimum and maximum levels of blood sugar (in mg/dl) after 8 h of fasting (G1) and 2 h after drinking, the glucose solution of about 237 ml and containing 75 g sugar (G2) are shown in Table 1. From Table 1, it can be seen that blood sugar level is expressed in intervals; therefore, investigation on the significance of correlation cannot be performed using the existing $t$-test for correlation. Decision-makers can apply the proposed $t$-test for correlation using neutrosophic statistics. To test the null hypothesis $H_{0N}$ that there is no correlation between G1 and G2 vs. $H_{1N}$ that G1 and G2 are associated, the neutrosophic correlation $r_{N \in \mathcal{I}L, t_{U}}$ for G1 and G2 is calculated as follows: $r_{N} = 0.9900 - 0.9899 I_{r} = 0.9900 - 0.9899 I_{r} = [0.9900, 0.9899]$.

The value of statistic $t_{N \in \mathcal{I}L, t_{U}}$ for G1 and G2 is calculated as: $t_{N} = 26.29 - 26.22 I_{t_{N}} = 26.29 - 26.22 I_{t_{N}} = [26.29, 26.22]$.

To investigate the relationship between G1 and G2, the following steps will be carried out:

**TABLE 1** Data of sugar levels in the blood.

| Age (years) | Data of “G1” (mg/dL) | Data of “G2” (mg/dL) |
|-------------|----------------------|----------------------|
| 45          | [159, 199]           | [166, 206]           |
| 46          | [150, 196]           | [156, 202]           |
| 47          | [139, 199]           | [147, 207]           |
| 48          | [142, 167]           | [148, 173]           |
| 49          | [152, 210]           | [160, 218]           |
| 50          | [143, 187]           | [150, 194]           |
| 51          | [151, 177]           | [159, 185]           |
| 52          | [140, 195]           | [147, 207]           |
| 53          | [154, 200]           | [160, 206]           |
| 54          | [142, 197]           | [149, 204]           |
| 55          | [150, 189]           | [157, 196]           |
| 56          | [160, 198]           | [168, 206]           |
| 57          | [162, 190]           | [170, 198]           |
| 58          | [146, 198]           | [152, 204]           |
| 59          | [149, 188]           | [156, 194]           |
| 60          | [177, 198]           | [179, 206]           |
Step 1: state \( H_{0N} \) : no correlation between G1 and G2 vs. \( H_{1N} \) : G1 and G2 are associated;
Step 2: fix the level of significance \( \alpha_N = 0.05 \) and the tabulated value is \( t_C = 1.76 \) at the degree of freedom 14 from (36);
Step 3: compute statistic \( t_N \) \( [t_L, t_U] \) = \( [26.29, 26.22] \) and compare with \( t_C = 1.76 \);
Step 4: As \( [26.29, 26.22] > 1.76 \), it is concluded that blood sugar levels between G1 and G2 are significant.

Based on the analysis, it can be concluded that there is a significant correlation between 8-h fasting sugar level and 2 h after drinking, the glucose solution of about 237 ml and containing 75 g of sugar.

**ADVANTAGES**

The proposed test for correlation is a generalization of \( t \)-test for correlation using classical statistics, interval-based statistics, and fuzzy logic. Now, the efficiency of the proposed \( t \)-test for correlation will be discussed over these tests in terms of flexibility and information. For comparisons, we will consider the neutrosophic form of the statistic \( t_N \) \( [t_L, t_U] \) that is \( t_N = 26.29 = 26.22 \) \( t_N \) \( [0, 0.0027] \). This neutrosophic form has two parts of information: the first one is about the statistic of classical statistics, and the second one is about the indeterminate part of the proposed test. For example, when \( t_L = 0 \), the value 26.29 presents the value of test statistic using classical statistics. According to the proposed test, the value of \( t_N \) \( [t_L, t_U] \) will in the interval from 26.29 to 26.22 rather than the exact value. The proposed test also indicates the measure of indeterminacy associated with the interval that is 0.0027. From this comparison, it is clear that the \( t \)-test using neutrosophic statistics has an edge over the existing \( t \)-test for correlation. The \( t \)-test using interval statistics and fuzzy-based logic gives the values of the test statistic in intervals without giving any information about the measure of indeterminacy. For example, for testing the hypothesis \( H_{0N} \) : no correlation between G1 and G2, the probability of committing a type-1 error is 0.05 (false), the probability of accepting \( H_{0N} \) : no correlation between G1 and G2 is 0.95 (true), and the measure of indeterminacy is 0.0027. The \( t \)-test using fuzzy logic will give information only about the measures of falseness and truth. Based on the analysis, it is concluded that the proposed \( t \)-test for correlation is better than the existing tests.

**DISCUSSIONS**

As the data is collected from a group having 20 people at the fasting time and then after two hours of drinking glucose solution about 237 milliliters contained 75-gram. The neutrosophic form of the correlation between G1 and G2 is \( r_N = 0.99 - 0.9899 \) \( t_N \) \( [0, 0.0001] \). It is interesting to note that the correlation between the two groups, G1 and G2, varies from 0.99 to 0.9899, with the measure of indeterminacy \( I_N = 0.001 \). From this correlation analysis, it can be seen that there is a strong positive correlation between fasting of 8 h and after 2 h of drinking the glucose solution. It means that if an 8-h fasting blood sugar level is high, then the blood sugar level after 2 h of drinking the glucose solution is also high and vice versa. It is important to note that after the 8-h fasting, the minimum blood sugar level of those aged 45 is 159. The value 159 indicates that these patients should take some energy drink before 2 h before sleeping, so that blood sugar can be utilized properly by the body. In addition, with an increase in 8-h fasting sugar, patients aged 45 to 60 should avoid taking carbohydrate or glucose items.

**CONCLUDING REMARKS**

The \( t \)-test of a correlation coefficient under neutrosophic statistics was presented in the article. The proposed \( t \)-test of a correlation coefficient was a generalization of the existing \( t \)-test of a correlation coefficient under classical statistics. From the real example, the proposed \( t \)-test of a correlation coefficient was found to be effective for investigating the significance of correlation in an indeterminate environment. The simulation study showed that measures of indeterminacy affect the decision on the significance of correlation. The proposed test can be applied to investigate correlations in the fields of economics, business, medicine, and industry. The proposed \( t \)-test of a correlation coefficient using a double sampling scheme can be considered as future research. Further statistical properties can be studied in future research. The proposed study can be extended for blood sugar measurement under different conditions and validation methods as future research. In addition, some disturbances can also be considered for blood glucose measurement in future studies.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

**AUTHOR CONTRIBUTIONS**

Both authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

**FUNDING**

This work was funded by the Deanship of Scientific Research at King Abdulaziz University.

**ACKNOWLEDGMENTS**

We are thankful to the editor, reviewers, and Usama Afzal from the University of the Education Lahore for their valuable suggestions to improve the quality and presentation of the article.
REFERENCES

1. Bartroff J, Song J. Sequential tests of multiple hypotheses controlling type I and II familywise error rates. J Stat Plan Infer. (2014) 153:100–14. doi: 10.1016/j.jspi.2014.05.010

2. Aleixandre-Benavent R, Moreno-Solano LM, Sapena AF, Pérez EAS. Correlation between impact factor and public availability of published research data in Information Science and Library Science journals. Scientometrics. (2016) 107:1–13. doi: 10.1007/s11192-016-1868-7

3. McGillivray B, Estell M. The relationship between usage and citations in an open access mega-journal. Scientometrics. (2019) 121:817–38. doi: 10.1007/s11192-019-03288-3

4. Pang J, Landes RD. Some t-tests for N-of-1 trials with serial correlation. PLoS ONE. (2020) 15:e0228077. doi: 10.1371/journal.pone.0228077

5. Silva LO, Maldonado G, Brigham T, Mullan AF, Utengen A, Cabrera D. Evaluating scholars’ impact and influence: cross-sectional study of the correlation between a novel social media–based score and an author-level citation metric. J Med Internet Res. (2021) 23:e28859. doi: 10.2196/28859

6. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. (2018) 126:1763–8. doi: 10.1213/ANE.0000000000002864

7. Najmi A, Balakrishnan Sadasivam AR. How to choose and interpret a statistical test? An update for budding researchers. J Fam Med Prim Care. (2021) 10:2763. doi: 10.4103/jfmpc.jfmpc_433_21

8. Khan HM, Mende S, Rafiq A, Gabbidon K, Reddy PH. Needed to measure predictive accuracy: a study of diabetic patients. Biochem Biophys Acta Mol Basis Dis. (2017) 1863:1046–53. doi: 10.1016/j.bbadis.2017.01.007

9. Liu S, Gao Y, Shen Y, Zhang M, Li J, Sun P. Application of three statistical models for predicting the risk of diabetes. BMC Endor Disord. (2019) 19:1–10. doi: 10.1186/s12902-019-01456-2

10. Wani HA, Majid S, Khan MS, Bhat AA, Wani RA, Bhat SA, et al. Scope of honey in diabetes and metabolic disorders. In: Therapeutic Applications of Honey and its Phytochemicals. Springer. (2020). p. 195–217. doi: 10.1007/978-981-15-7305-7_9

11. Nedjalkova M, Madurga S, Ballabio D, Bobeva R, Romanova J, Kichev I, et al. Diabetes mellitus type 2: exploratory data analysis based on clinical reading. Open Chem. (2020) 18:1041–1053. doi: 10.1515/chem-2020-0086

12. Adaobi OO, Iwueze IS, Biu EO, Arimie CO. On the analysis of blood glucose levels of diabetic patients. Fortune J Health Sci. (2021) 4:257–83. doi: 10.26502/fjhs021

13. Mukasheva A, Saparkhoyayev N, Akanov Z, Apon A, Kalra S. Predicting the scope of honey in diabetes mellitus using emogondometric models. Diab Ther. (2019) 10:2079–93. doi: 10.1007/s13380-019-00684-1

14. Eyniadeh Z, Ameli Z, Sahranavard M, Daneshparvar M, Dolag MA, Rooykosh M, et al. Biostatistical Investigation of Correlation Between COVID-19 and Diabetes Mellitus. medRxiv. (2020). doi: 10.1101/2020.11.21.20235853

15. Balamurugan SAA, Saranya K, Sasiakla S, Chinthama G. Statistical and machine learning approaches for clinical decision on drug usage in diabetes with reference to competence and safeness. Int J Comput Intell Syst. (2021) 14:859–68. doi: 10.2991/ijcis.d.210212.002

16. Alsag AM. Remarks on the use of Pearson’s and Spearman’s correlation coefficients in assessing relationships in ophthalmic data. African Vision Eye Health. (2021) 8:10. doi: 10.4102/aveh.v8i1.612

17. Janse RJ, Hoekstra T, Jager KJ, Zoellic C, Tripepi G, Dekker FW, et al. Conducting correlation analysis: important limitations and pitfalls. Clin Kidney J. (2021) 14:2336–7. doi: 10.1093/ckj/sfab085

18. Sun Q, Tang L, Zeng Q, Gu M. Assessment for the correlation between diabetic retinopathy and metabolic syndrome: a cross-sectional study. Diaib Metabol Syndr Obesity. (2021) 14:1773. doi: 10.2147/DMSO.S265214

19. Smarandache F. Neutrosophy neurostrophic probability, set, and logic, proquest information and learning. Ann Arbor.

20. Smarandache F. Introduction to Neutrosophic Measure, Neutrosophic Integral, and Neutrosophic Probability, Infinite Study. (2013).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Aslam and Albassam. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.