Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (4) (2024), 1130 – 1140
DOI : 10.15672/hujms.1342344

Research Article

The t-test of a regression coefficient for imprecise data

Muhammad Aslam

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia

Abstract

The existing t-test for testing the significance of the regression coefficient is applied when cent percent observations of the data are precise, exact and certain. In practice, the measurement data or data recorded in an uncertain environment do not have all precise observations. The imprecise data cannot be analyzed using the existing t-test for testing the significance of the regression coefficient. In this paper, we will present the design of a t-test for testing the significance of the regression coefficient under neutrosophic statistics. The proposed t-test for testing the significance of the regression coefficient can be applied to imprecise data. The effect of the degree of uncertainty on the power of the test will be studied. The proposed t-test for testing the significance of the regression coefficient will be applied using the imprecise data. From the analysis, it is concluded that the proposed t-test for testing the significance of the regression coefficient will be informative, flexible and adequate to be applied to imprecise data.

Mathematics Subject Classification (2020). 97K80, 62J86, 62A86

Keywords. t-test, inference, power of the test, level of significance, simulation

1. Introduction

The regression line describes the relationship between the independent variable and the dependent variable.After establishing this relation, it is important to test the significance of the regression coefficient. For estimation and forecasting purposes, it is necessary to know the significance of the relationship between the independent variable and the dependent variable. For testing the significance of the relationship, the t-test is performed on the regression coefficient. During the implementation of the t-test for the regression coefficient, the null hypothesis that the regression coefficient has no significant effect vs. the alternative hypothesis that the regression coefficient has a significant effect in determining the relationship between the independent variable and dependent variable. Frank [13] studied the effect of the confounding variables on the coefficient of regression. Bewick et al. [3] discussed the misuse of the regression and correlations concepts and highlighted the reasons for the failure of the assumptions. Li and Yuan [16] used the regression analysis to evaluate the significance of educational variables. Nieminen [20] discussed the application of the regression using Meta-analysis. More applications can be seen in [12], [14], [17], [18] and [22]. As mentioned by [9] statistical regression analysis is one of the important

Email addresses: aslam_ravian@hotmail.com (M. Aslam)
Received: 13.08.2023; Accepted: 28.03.2024

statistical methods and has been widely applied to different scientific areas. Classical regression analysis models are limited to crisp data. In practice, however, data are usually imprecise because data are difficult to measure precisely or data are determined subjectively. When dealing with fuzzy data, using classical regression analysis method to test the regression coefficient would be improper and lead to an incorrect decision. Smarandache [25] provided the neutrosophic statistical analysis to analyze the imprecise data. Neutrosophic statistics was found to be more efficient than classical statistics, see [10] and [11]. More applications of neutrosophic logic can be seen in [4–8] and [23]. Smarandache [24] also proved that neutrosophic statistics is more efficient than classical statistics and interval-statistics.Neutrosophic statistics give additional information about the measure of indeterminacy that cannot be obtained from traditional statistics. The applications of neutrosophic statistics can be seen in [1], [2] and [21]. The existing t-test for testing the significance of the regression coefficient is applied when crisp data is available. The existing t-test for testing the significance of the regression coefficient has limitations in that it can be applied to imprecise data. By exploring the literature and according to the best of the authors knowledge, there is no work on t-test for testing the significance of regression coefficient under neutrosophic statistics. To fill this gap, in this paper, we will design a t-test for testing the significance of the regression coefficient under neutrosophic statistics. The testing statistic will be provided and applied using the simulated and real data. It is expected that the proposed t-test for testing the significance of the regression coefficient under classical statistics will be more efficient than the t-test for testing the significance of the regression coefficient under classical statistics.

2. Proposed test

Suppose that $(x_{1N}, y_{1N}), (x_{2N}, y_{2N}), , (x_{nN}, y_{nN})$; $(i = 1, 2, \ldots, n)$ be a neutrosophic pair of random variables and n is a sample size. The neutrosophic forms of pair data is $(x_{nN} = x_{nL} + x_{nU} I_{x_N})$;
$I_{x_N} \in [I_{x_L}, I_{x_U}], (y_{nN} = y_{nL} + y_{nU} I_{y_N}); I_{y_N} \in [I_{y_L}, I_{y_U}]$. Note that x_{nL}, y_{nL} are the lower values and $x_{nU} I_{x_N}, y_{nU} I_{y_N}$ are the indeterminate part $I_{x_N} \in [I_{x_L}, I_{x_U}] and I_{y_N} \in [I_{(y_L)}, I_{(y_U)}]$ are the measure of indeterminacy associated with both variables. Based on the information, the neutrosophic mean for variable x_{nN} is derived as:

$$\begin{aligned} \bar{x}_N &= \frac{((x_{1L} + x_{1U} I_{x_N}) + (x_{2L} + x_{2U} I_{x_N}) + \cdots + (x_{nL} + x_{nU} I_{x_N}))}{n} \\ &= \bar{x}_L + \bar{x}_U I_{\bar{x}_N}; I_{\bar{x}_N} \in [I_{\bar{x}_L}, I_{\bar{x}_U}] \end{aligned} \tag{2.1}$$

Note that the first value $\bar{x}_L$ presents the average of the lower values and $\bar{x}_U I_{\bar{x}_N}$ is the indeterminate part and $I_{\bar{x}_N} \in [I_{\bar{x}_L}, I_{\bar{x}_U}]$ is the degree of indeterminacy.

Based on the information, the neutrosophic mean for variable y_{nN} is derived as:

$$\begin{aligned} \overline{y}_N &= \frac{(y_{1L} + y_{1U} I_{y_N}) + (y_{2L} + y_{2U} I_{y_N}) + \ldots + (y_{nL} + y_{nU} I_{y_N})}{n} \\ &= \overline{y}_L + \overline{y}_U I_{\overline{y}_N}; I_{\overline{y}_N} \epsilon \, [I_{\overline{y}_L}, I_{\overline{y}_U}] \end{aligned} \tag{2.2}$$

Note that the first value $\overline{y}_L$ presents the average of the lower values and $\overline{y}_U I_{\overline{y}_N}$ is the indeterminate part and $I_{\overline{y}_N} \epsilon \, [I_{\overline{y}_L}, I_{\overline{y}_U}]$ is the degree of indeterminacy.

The neutrosophic variance for the variable x_{nN} is derived as:

$$s_{xN}^2 = \frac{\sum_{i=1}^{n} ((x_{nL} + x_{nU} I_N) - (\overline{x}_L + \overline{x}_U I_N))^2}{n-1} = s_{xL}^2 + s_{xU}^2 I_N; I_N \epsilon \, [I_L, I_U] \tag{2.3}$$

The neutrosophic variance for the variable y_{nN} is derived as:

$$s_{yN}^2 = \frac{\sum_{i=1}^{n} ((y_{nL} + y_{nU} I_N) - (\overline{y}_L + \overline{y}_U I_N))^2}{n-1} = s_{yL}^2 + s_{yU}^2 I_N; I_N \epsilon \, [I_L, I_U] \tag{2.4}$$

In Eq. (2.3) and Eq. (2.4), the first values presents the determinate part and the second values denote indeterminate part and $I_N \epsilon [I_L, I_U]$ is the degree of indeterminacy.

The neutrosophic linear regression for (x_{nN}, y_{nN}) with neutrosophic intercept $A_N = A_L + A_U I_N$ and neutrosophic rate of change $B_N = B_L + B_U I_N$ is defined as

$$y_{nN} = ((A_L + B_L (x_{nL} - \overline{x}_L)) + (A_U + B_U (x_{nU} - \overline{x}_U) I_N)) \,; I_N \epsilon [I_L, I_U] \tag{2.5}$$

Note that the initial expression in Eq. (2.5) represents the regression line within classical statistics, with the second part indicating the indeterminate component. The proposed regression line converges to the classical statistics regression line when I_L is set to zero. To find the estimates of $A_N \epsilon [A_L, A_U]$and $B_N \epsilon [B_L, B_U]$, the neutrosophic sum of squares of residual is defined by

$$S_N = \sum_{i=1}^{n} \left\{ \left((Y_{iL} - A_L - B_L (x_{nL} - \overline{x}_L))^2 \right) + \left((Y_{iU} - A_U - B_U (x_{nU} - \overline{x}_U))^2 \right) I_N \right\} \tag{2.6}$$

To find $A_N \epsilon [A_L, A_U]$, the partial derivative with respect to $A_N \epsilon [A_L, A_U]$ is given by

$$\frac{d_{S_N}}{dA_N} = \left[\sum_{i=1}^{n} \left\{ \left((Y_{iL} - A_L - B_L (x_{nL} - \overline{x}_L))^2 \right) + \left((Y_{iU} - A_U - B_U (x_{nU} - \overline{x}_U))^2 \right) I_N \right\} \right] \tag{2.7}$$

$$-2 \sum_{i=1}^{n} \left((Y_{iL} - A_L - B_L (x_{nL} - \overline{x}_L))^2 \right) - 2 \sum_{i=1}^{n} \left((Y_{iU} - A_U - B_U (x_{nU} - \overline{x}_U))^2 \right) I_N = 0$$

$$- \sum_{i=1}^{n} Y_{iL} + nA_L + B_L \sum_{i=1}^{n} (x_{nL} - \overline{x}_L) - I_N \sum_{i=1}^{n} Y_{iL} + nI_N A_U + B_U I_N \sum_{i=1}^{n} (x_{nU} - \overline{x}_U) = 0$$

$$- \sum_{i=1}^{n} Y_{iL} - I_N \sum_{i=1}^{n} Y_{iL} + n (A_N) = 0$$

After simplification, $A_N \epsilon [A_L, A_U]$ is given by

$$A_N = \overline{y}_L + \overline{y}_U I_N \tag{2.8}$$

where $A_N \epsilon [A_L, A_U]$

To find $B_N \epsilon [B_L, B_U]$, the partial derivative with respect to $B_N \epsilon [B_L, B_U]$ is given by

$$\frac{d_{S_N}}{dB_N} = \left[\sum_{i=1}^{n} \left\{ \left((Y_{iL} - A_L - B_L (x_{nL} - \overline{x}_L))^2 \right) + \left((Y_{iU} - A_U - B_U (x_{nU} - \overline{x}_U))^2 \right) I_N \right\} \right] \tag{2.9}$$

Let $A_L = \overline{y}_L$, $A_U = \overline{y}_U$, Eq. (2.9) can be rewritten as:

$$\frac{d_{S_N}}{dB_N} = \sum_{i=1}^{n} \left\{ \left(((Y_{iL} - \overline{y}_L) - B_L (x_{nL} - \overline{x}_L))^2 \right) + \left(((Y_{iU} - \overline{y}_U) - B_U (x_{nU} - \overline{x}_U))^2 \right) I_N \right\}$$

$$\sum_{i=1}^{n} (Y_{iL} - \overline{y}_L) (x_{nL} - \overline{x}_L) - B_L \sum_{i=1}^{n} (x_{nL} - \overline{x}_L)^2 + \sum_{i=1}^{n} (Y_{iU} - \overline{y}_U) (x_{nU} - \overline{x}_U) I_N$$
$$-B_U \sum_{i=1}^{n} (x_{nU} - \overline{x}_U)^2 I_N = 0$$

$$\begin{aligned} B_L \sum_{i=1}^{n} (x_{nL} - \overline{x}_L)^2 + B_U \sum_{i=1}^{n} (x_{nU} - \overline{x}_U)^2 I_N = & \sum_{i=1}^{n} (Y_{iL} - \overline{y}_L) (x_{nL} - \overline{x}_L) \\ & + \sum_{i=1}^{n} (Y_{iU} - \overline{y}_U) (x_{nU} - \overline{x}_U) I_N; \\ & I_N \in [I_L, I_U] \end{aligned} \tag{2.10}$$

Let

$$\begin{aligned}\sum_{i=1}^{n}\left(Y_{iN}-\overline{y}_N\right)\left(x_{nN}-\overline{x}_N\right) &= \sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right) \\ &\quad +\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_N;I_N\epsilon\left[I_L,I_U\right]\end{aligned} \tag{2.11}$$

The solution of $\sum_{i=1}^{n}\left(Y_{iN}-\overline{y}_N\right)\left(x_{nN}-\overline{x}_N\right)$ for lower value I_L and upper value I_U is given by

$$\begin{aligned}\sum_{i=1}^{n}\left(Y_{iN}-\overline{y}_N\right)\left(x_{nN}-\overline{x}_N\right)\in\Bigg[&\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)+\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_L, \\ &\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)+\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_U\Bigg]\end{aligned} \tag{2.12}$$

Let

$$B_N\sum_{i=1}^{n}\left(x_{nN}-\overline{x}_N\right)^2=B_L\sum_{i=1}^{n}\left(x_{nL}-\overline{x}_L\right)^2+B_U\sum_{i=1}^{n}\left(x_{nU}-\overline{x}_U\right)^2I_N;I_N\epsilon\left[I_L,I_U\right]$$

The Eq. (2.10) can be written as:

$$B_N\sum_{i=1}^{n}\left(x_{nN}-\overline{x}_N\right)^2=\sum_{i=1}^{n}\left(Y_{iN}-\overline{y}_N\right)\left(x_{nN}-\overline{x}_N\right)$$

or

$$\begin{aligned}B_N &= \frac{\sum_{i=1}^{n}\left(Y_{iN}-\overline{y}_N\right)\left(x_{nN}-\overline{x}_N\right)}{\sum_{i=1}^{n}\left(x_{nN}-\overline{x}_N\right)^2} \\ &= \frac{\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)+\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_N}{\sum_{i=1}^{n}\left(x_{nN}-\overline{x}_N\right)^2};\ I_N\epsilon\left[I_L,I_U\right]\end{aligned}$$

The solution of B_N for lower value I_L and upper value I_U is given by

$$\begin{aligned}B_N\epsilon\Bigg[&\frac{\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)+\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_L}{\sum_{i=1}^{n}\left(x_{nL}-\overline{x}_L\right)^2}, \\ &\frac{\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)+\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)I_U}{\sum_{i=1}^{n}\left(x_{nU}-\overline{x}_U\right)^2}\Bigg]\end{aligned}$$

The neutrosophic form of $B_N\epsilon\left[B_L,B_U\right]$ is given by

$$B_N=\frac{\left(\sum_{i=1}^{n}\left(Y_{iL}-\overline{y}_L\right)\left(x_{nL}-\overline{x}_L\right)\right)}{\sum_{i=1}^{n}\left(x_{nL}-\overline{x}_L\right)^2}+\frac{\left(\sum_{i=1}^{n}\left(Y_{iU}-\overline{y}_U\right)\left(x_{nU}-\overline{x}_U\right)\right)}{\sum_{i=1}^{n}\left(x_{nU}-\overline{x}_U\right)^2}I_N \tag{2.13}$$

After some simplification, $B_N\epsilon\left[B_L,B_U\right]$ in Eq. (2.13) can be given by

$$B_N=\frac{\sum x_Ly_L-\frac{1}{n}\sum x_L\sum y_L}{\sum x_L^2-\frac{1}{n}\left(\sum x_L\right)^2}+\frac{\sum x_Uy_U-\frac{1}{n}\sum x_U\sum y_U}{\sum x_U^2-\frac{1}{n}\left(\sum x_U\right)^2}I_N \tag{2.14}$$

To test the significance of the regression coefficient, the neutrosophic test statistic is given by

$$t_N=\frac{b_LS_{xL}}{S_{y_L.x_L}}(n-1)^{-\frac{1}{2}}+\frac{b_US_{xU}}{S_{y_U.x_U}}(n-1)^{-\frac{1}{2}}I_N;I_N\epsilon\left[I_L,I_U\right] \tag{2.15}$$

where

$$S_{y_N.x_N}\epsilon\left[\sqrt{\frac{\sum\left\{y_L-\overline{y}_L-b_L\left(x_L-\overline{x}_L\right)\right\}^2}{n-2}},\sqrt{\frac{\sum\left\{y_U-\overline{y}_U-b_U\left(x_U-\overline{x}_U\right)\right\}^2}{n-2}}\right]$$

Note that the neutrosophic test statistic is given in Eq. (2.13) is the generalization of the test statistic mentioned in [15]. The proposed neutrosophic test reduces to the test statistic is given in [15] when $I_N = 0$. Note that the neutrosophic test statistic t_N follows the neutrosophic t-test distribution with $(n-2)$ degree of freedom. Note here that a t_N^2 critical value (for n degrees of freedom) is equal to the critical value (for 1 and n degrees of freedom) of F_N-test and, thus, this method can be considered as the one-dimensional version of the FN-test proposed by [19].

3. Application

Now, the application of the proposed t-test for regression coefficient will be discussed in this section. The data is about the students' performance interviews based on the aptitude tests is selected from [19]. Conscientiousness is a dependent variable and it is denoted by y_{nN} and aptitude test is represented by x_{nN}. The data is shown in Table 1. The necessary computations to carry out the proposed test are given as:

The neutrosophic mean for variable x_{nN} and y_{nN} are calculated as:

$\overline{x}_N = 3.3 + 3I_{\overline{x}_N}$ and $\overline{y}_N = 4.7 + 6I_{\overline{y}_N}$

The neutrosophic variance for the variable x_{nN} is calculated as:

$s_{xN}^2 = 4.9 + 4I_N$

The neutrosophic linear regression for (x_{nN}, y_{nN}) is given by

$y_{nN} = (4.7 + 1.4263\,(x_{nL} - 3.3)) + (6 + 1.9722\,(x_{nU} - 3)\,I_N)$

where $B_N \epsilon\,[B_L, B_U]$ is calculated using Eq. (2.14) and given by

$B_N = 1.4263 + 1.9722 I_N; I_N \epsilon [0, 0.2768]$

The upper value I_U can be computed as:

$I_U = (1.9722 - 1.4263)/1.9722 = 0.2768.$

To test the significance of the regression coefficient, the neutrosophic test statistic is calculated by using Eq. (2.15) and is given by

$t_N = 0.3406 + 0.6577 I_N; I_N \epsilon [0, 0.4821]$

where

$S_{y_L.x_L} \epsilon [3.0900, 1.9991]$ and $s_{xN} \epsilon [2.2136, 2]$

The upper value I_U can be computed as:

$I_U = (0.6577 - 0.3406)/0.6577 = 0.4821.$

The proposed t-test for the regression coefficient is implemented in the following steps

Step-1: Null hypothesis H_0 : regression coefficient is insignificant vs. alternative hypothesis

H_1 : regression coefficient is significant

Step-2: Let the level of significance α=0.05 and the tabulated value is 2.306.

Step-3: Compare $t_N = 0.3406 + 0.6577I_N; I_N \epsilon [0, 0.4821]$ with the tabulated value.

Step-4: Do not reject H_0 and it is concluded that the regression coefficient is insignificant.

Table 1. Conscientiousness and aptitude test data .

[1,3]	[2,2]	[2,4]	[4,4]	[1,4]	[6,6]	[2,4]	[10,13]	[14,15]	[5,5]
[3,3]	[2,2]	[1,2]	[2,3]	[2,1]	[2,3]	[2,2]	[5,6]	[7,7]	[7,1]

4. Simulation study

This section discusses the evaluation of test power using neutrosophic simulated data. Test power, defined as the probability of rejecting H_0 when it is false, is examined across various sample sizes (ranging from 10 to 60) with neutrosophic data featuring lower and upper values. Through a simulation study involving 100 neutrosophic samples, comparisons are made with tabulated values at α =0.05 and α =0.10. The resulting test power (1- β) is calculated and presented in Table 2. Table 2 reveals that the proposed t-test for assessing the significance of the regression coefficient yields power values within indeterminate intervals. Lower values represent the power of the existing t-test for this purpose, while upper values indicate indeterminate power values. Notably, there is a decreasing trend in test power for the same sample size (n). Furthermore, an increase in α from 0.05 to 0.10 is associated with a higher likelihood of Type I errors, leading to a reduction in test power. To visually illustrate these findings, neutrosophic power curves are presented in Figures 1 and 2. Figure 1 depicts the curve for α =0.05, while Figure 2 shows the curve for α =0.10. It is evident from these figures that the power curve for the existing test surpasses that of the proposed t-test for the regression coefficient.

Table 2. The values of the power of the test for various n.

	$\alpha = 0.05$	$\alpha = 10$
n	$(1 - \beta)$	$(1 - \beta)$
10	[0.9505, 0.9394]	[0.8992, 0.8869]
15	[0.9503, 0.9392]	[0.9004, 0.8870]
20	[0.9500, 0.9379]	[0.9001, 0.8867]
25	[0.9498, 0.9397]	[0.9000, 0.8820]
30	[0.9507, 0.9400]	[0.9002, 0.8850]
40	[0.9512, 0.9388]	[0.8999, 0.8828]
50	[0.9505, 0.9391]	[0.9002, 0.8836]
60	[0.9495, 0.9380]	[0.9000, 0.8855]

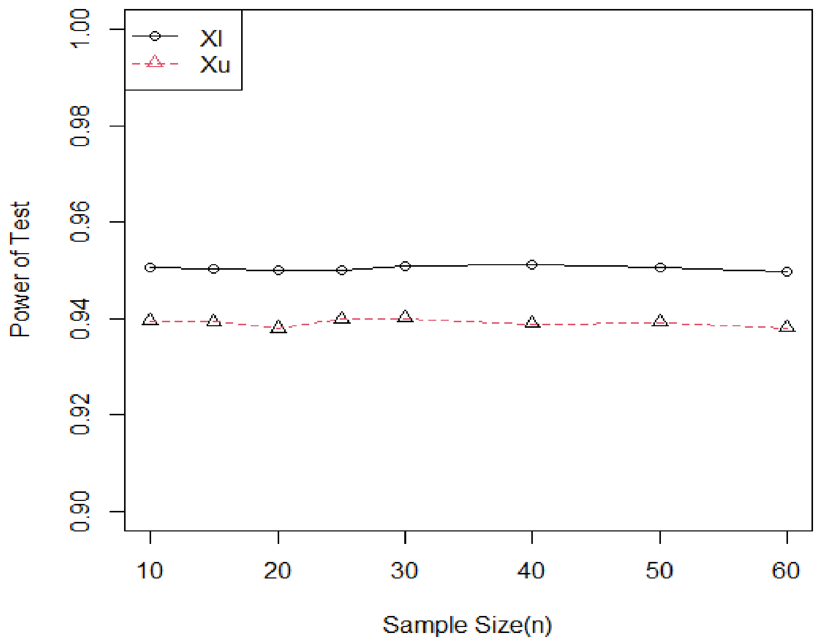

Figure 1. Power curves for various sample size when $\alpha = 0.05$.

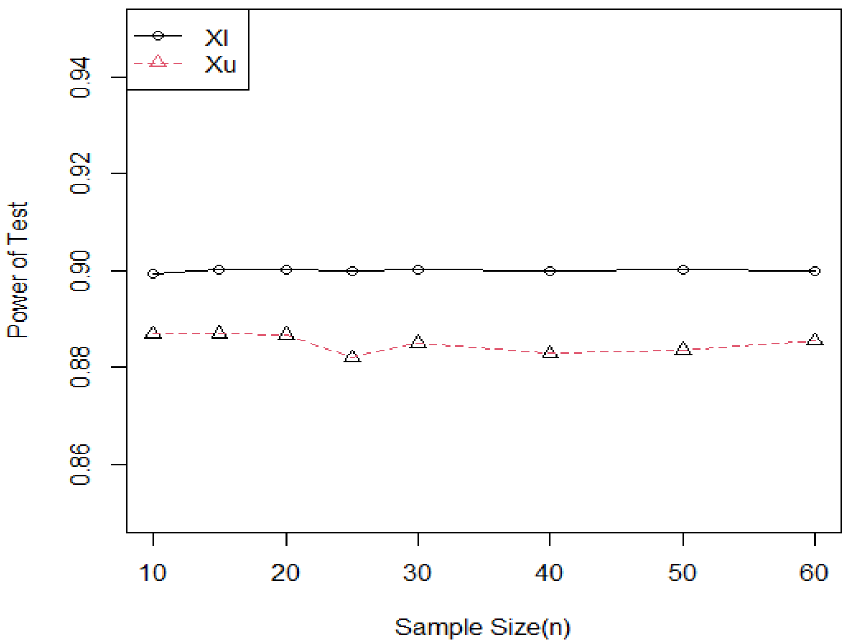

Figure 2. Power curves for various sample size when $\alpha = 0.10$.

5. Performance comparison

Now, the performance comparisons in terms of the power of the test will be discussed in this section. The effect of the measure of indeterminacy on the power of the test will be evaluated. As mentioned earlier, the proposed t-test of the regression coefficient is a generalization of the t-test of the regression coefficient under classical statistics. The proposed t-test of regression coefficient reduces to the t-test of regression coefficient under classical statistics when all observations in the data are determinate. Table 3 and Table 4 depict the values of power of the test for the proposed t-test of regression coefficient and t-test of regression coefficient under classical statistics ($I_N = 0$). Table 3 shows the values of power of the test for the t-test of regression coefficient and t-test of regression coefficient under classical statistics when α=0.05 and n=10. Table 4 shows the values of power of the test for the t-test of regression coefficient and t-test of regression coefficient under classical statistics when α=0.10 and n=10. From Table 3, it can be noted that as the measure of indeterminacy I_N increases from 0 to 0.20, the power of the test decreases. It is important to note that for the same value of sample size, there is a decreasing trend in the power of the test as the level of significance α increases. The values of the power of the test for these parameters are also shown in Figures 3-4. From these figures, it can be noted that the power curve for the existing test is higher than the other values of measure of indeterminacy. From these figures, it can be concluded that the test of regression under an uncertain environment loses the power to reject the null hypothesis when it is false. From the study, it is concluded that decisions-makers should be very careful in applying the t-test of the regression coefficient. In addition, there is a significant effect on the power of the test when the level of significance or measure of indeterminacy changes. Therefore, the proposed test is helpful in making decisions about the testing of the hypothesis in uncertain situations.

Table 3. the power of the test when α=0.05 and n=10.

Sample number	$I_N = 0.00$	$I_N = 0.02$	$I_N = 0.05$	$I_N = 0.10$	$I_N = 0.15$	$I_N = 0.20$
1	0.9488	0.9441	0.937	0.9277	0.916	0.9052
2	0.9463	0.9422	0.9354	0.9236	0.912	0.9003
3	0.9501	0.9462	0.9398	0.9325	0.9198	0.9092
4	0.9476	0.9443	0.937	0.9278	0.9167	0.9064
5	0.9487	0.9451	0.9399	0.9286	0.9159	0.9052
6	0.9523	0.9495	0.9444	0.9314	0.9222	0.9101
7	0.9484	0.9447	0.9398	0.9293	0.9186	0.9064
8	0.9504	0.9458	0.9382	0.9276	0.9158	0.9036
9	0.9524	0.9471	0.9412	0.9324	0.9214	0.9096
10	0.9483	0.9441	0.9377	0.927	0.9151	0.9034

6. Concluding remarks

The t-test for the regression coefficient is applied to test the significance of the regression coefficient in the presence of imprecise data. The test statistic under classical statistics is modified in this paper so that it can be applied when imprecise data is obtained from complex processes or under an uncertain environment. The proposed test was found to be more efficient than the existing t-test for regression coefficient. It is concluded that the use of the existing t-test for the regression coefficient under uncertainty may mislead the decision-makers. Based on the study, it is recommended to use the proposed test in the imprecise data obtained from education, metrology, business, dam temperature or level of water and in poetical science. The proposed test for testing the significance of the regression coefficient for multiple regressions can be extended as future research.

Table 4. The power of the test when α=0.10 and n=10.

Sample number	$I_N = 0.00$	$I_N = 0.02$	$I_N = 0.05$	$I_N = 0.10$	$I_N = 0.15$	$I_N = 0.20$
1	0.8954	0.8899	0.8814	0.8641	0.8485	0.8326
2	0.8994	0.8941	0.8844	0.8689	0.8536	0.8385
3	0.9063	0.9012	0.8907	0.8749	0.8589	0.8426
4	0.8992	0.8928	0.8831	0.8676	0.8542	0.8377
5	0.9008	0.8945	0.8868	0.8698	0.8534	0.8362
6	0.8994	0.8937	0.8846	0.8708	0.8561	0.8413
7	0.8944	0.8889	0.8813	0.8671	0.8512	0.8374
8	0.8989	0.8929	0.8845	0.8701	0.8563	0.8385
9	0.9083	0.9024	0.8923	0.8775	0.8618	0.8452
10	0.9034	0.8966	0.8869	0.8728	0.8573	0.8423

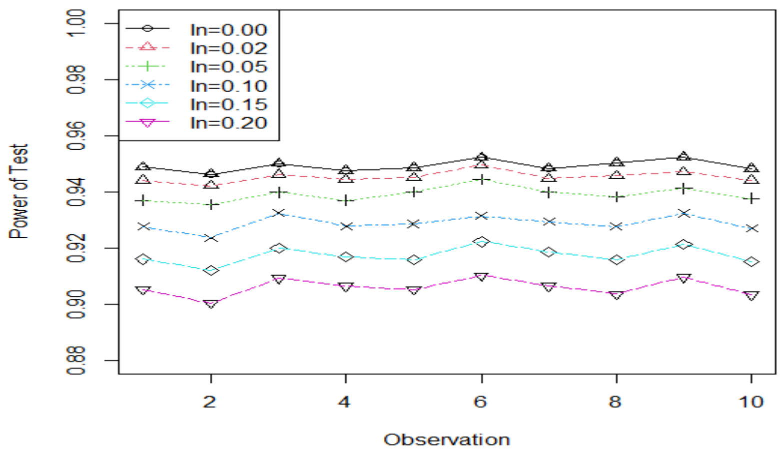

Figure 3. Power curves for various sample size when $\alpha = 0.05$.

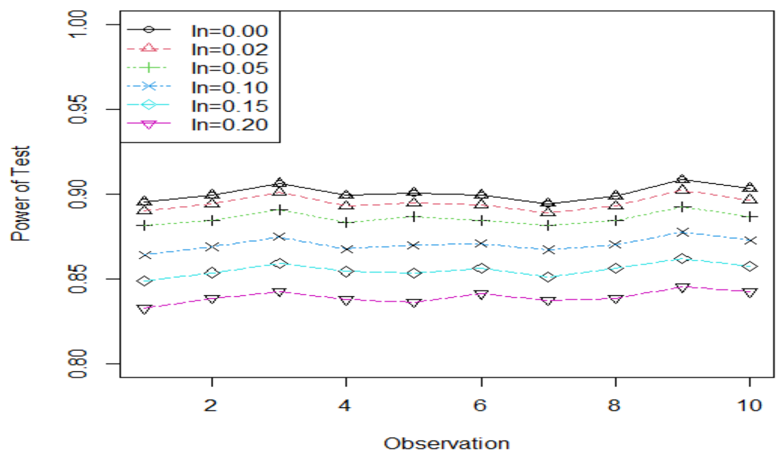

Figure 4. Power curves for various sample size when $\alpha = 0.10$.

Acknowledgements

The author expresses sincere gratitude to the editor and reviewers for their valuable suggestions aimed at enhancing the quality and presentation of the paper.

Author contributions. M.A wrote the paper.

Conflict of interest statement. No conflict of interest regarding the paper.

Funding. No funding was obtained for the current study.

Data availability. The data is given in the paper.

References

[1] R. Alhabib and A. Salama, *The neutrosophic time series-study its models (linear-logarithmic) and test the coefficients significance of its linear model*, Neutrosophic Sets Syst. **33**, 105-115, 2020.

[2] M. Aslam, *Neutrosophic F-Test for two counts of data from the Poisson distribution with application in climatology*, Stats **5** (3), 773-783, 2022.

[3] V. Bewick, L. Cheek and J. Ball, *Statistics review 7: Correlation and regression*, Crit. Care **7**, 1-9, 2003.

[4] S. Broumi, S. Krishna Prabha and V. Uluçay, *Interval-valued Fermatean neutrosophic shortest path problem via score function*, Neutrosophic syst. appl. **11**, 1-10, 2023.

[5] S. Broumi, S. Mohanaselvi, T. Witczak, M. Talea, A. Bakali and F. Smarandache, *Complex fermatean neutrosophic graph and application to decision making*, Decis. Mak. Appl. Manag. Eng. **6** (1), 474-501, 2023.

[6] S. Broumi, P.K. Raut and S.P. Behera, *Solving shortest path problems using an ant colony algorithm with triangular neutrosophic arc weights*, Int. J. Neutrosophic Sci. **20** (4), 128-128-137, 2023.

[7] S. Broumi, R. Sundareswaran, M. Shanmugapriya, A. Bakali and M. Talea, *Theory and applications of Fermatean neutrosophic graphs*, Neutrosophic Sets Syst. **50**, 248-286, 2022.

[8] S. Broumi, R. Sundareswaran, M. Shanmugapriya, P.K. Singh, M. Voskoglou and M. Talea, *Faculty performance evaluation through multi-criteria decision analysis using interval-valued Fermatean neutrosophic sets*, Mathematics **11** (18), 3817, 2023.

[9] C.C. Chen, C.M. Lai and W.C. Sun, *Fuzzy testing for regression coefficient of Fuzzy numbers*, J. Test. Eval. **41** (1), 116-121, 2013.

[10] J. Chen, J. Ye and S. Du, S*cale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics*, Symmetry **9** (10), 208, 2017.

[11] J. Chen, J. Ye, S. Du and R. Yong, *Expressions of rock joint roughness coefficient using neutrosophic interval statistical numbers*, Symmetry **9** (7), 123, 2017.

[12] M. Fischer, B.T. West, M.R. Elliott and F. Kreuter, *The impact of interviewer effects on regression coefficients*, J. Surv. Stat. Methodol. **7** (2), 250-274, 2019.

[13] K.A. Frank, *Impact of a confounding variable on a regression coefficient,* Sociol. Methods Res. **29** (2), 147-194, 2000.

[14] A.R. Ives, *For testing the significance of regression coefficients, go ahead and logtransform count data*, Methods Ecol. Evol. **6** (7), 828-835, 2015.

[15] G.K. Kanji, 100 Statistical Tests, Sage, 2006.
[16] S. Li and X. Yuan, *Application of linear regression mathematical model in the evaluation of teachers informatization quality.* Complexity **1**, 1-10, 2021.
[17] K. Mu, Q. Shi, Y. Ma and J. Tan, *Exploration of entrepreneurship education by linear regression and psychological factor analysis*, Front. Psychol. **11**, 2045, 2020.
[18] J. Mulder and A. Olsson-Collentine, *Simple Bayesian testing of scientific expectations in linear regression models*, Behav. Res. Methods **51**, 1117-1130, 2019.
[19] D. Nagarajan, S. Broumi, F. Smarandache and J. Kavikumar, *Analysis of neutrosophic multiple regression*, Neutrosophic Sets Syst. **43**, 44-53, 2021.
[20] P. Nieminen, *Application of standardized regression coefficient in meta-analysis*, BioMedInformatics **2** (3), 434-458, 2022.
[21] A. Polymenis, *A neutrosophic Studentsttype of statistic for AR (1) random processes*, J. Fuzzy. Ext. Appl **2** (4), 388-393, 2021.
[22] T. Ramanathan and M. Rajarshi, *Rank tests for testing randomness of a regression coefficient in a linear regression model*, Metrika **39**, 113-124, 1992.
[23] F. Smarandache, New types of Soft Sets: HyperSoft Set, IndetermSoft Set, IndetermHyperSoft Set, and TreeSoft Set, Int. J. Neutrosophic Sci. **20** (4), 58-64, 2023.
[24] F. Smarandache, Neutrosophic Statistics is an extension of Interval Statistics, while Plithogenic Statistics is the most general form of statistics (second version), Int. J. Neutrosophic Sci. **19** (1), 148-165, 2022.
[25] F. Smarandache, *Introduction to Neutrosophic Statistics, Sitech and Education*, Craiova, Romania-Educational Publisher, Columbus, Ohio, USA, 123, 2014.