Analysis of Neutrosophic Multiple Regression

D. Nagarajan
S. Broumi
F. Smarandache
J. Kavikumar

Follow this and additional works at: https://digitalrepository.unm.edu/nss_journal

Recommended Citation
Nagarajan, D.; S. Broumi; F. Smarandache; and J. Kavikumar. "Analysis of Neutrosophic Multiple Regression." Neutrosophic Sets and Systems 43, 1 (). https://digitalrepository.unm.edu/nss_journal/vol43/iss1/4

This Article is brought to you for free and open access by UNM Digital Repository. It has been accepted for inclusion in Neutrosophic Sets and Systems by an authorized editor of UNM Digital Repository. For more information, please contact disc@unm.edu.
Analysis of Neutrosophic Multiple Regression

D. Nagarajan¹, S. Broumi²*, F. Smarandache³, J. Kavikumar⁴

¹Department of Mathematics, Hindustan Institute of Technology & Science, Chennai-603 103, India; E-mail: dnrmsu2002@yahoo.com,
²Laboratory of Information Processing, Faculty of Science Ben M’Sik, University Hassan II, B.P 7955, Sidi Othman, Casablanca, Morocco; E-mail: broumisaid78@gmail.com,
³Department of Mathematics, University of New Mexico, 705 Gurley Avenue, Gallup, NM 87301, USA; E-mail: fsmarandache@gmail.com
⁴Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Malaysia; E-mail: kavi@uthm.edu.my
* Correspondence: broumisaid78@gmail.com

Abstract: The idea of Neutrosophic statistics is utilized for the analysis of the uncertainty observation data. Neutrosophic multiple regression is one of a vital roles in the analysis of the impact between the dependent and independent variables. The Neutrosophic regression equation is useful to predict the future value of the dependent variable. This paper to predict the students' performance in campus interviews is based on aptitude and personality tests, which measures conscientiousness, and predict the future trend. Neutrosophic multiple regression is to authenticate the claim and examine the null hypothesis using the F-test. This study exhibits that Neutrosophic multiple regression is the most efficient model for uncertainty rather than the classical regression models.

Keywords: Neutrosophic multiple regression; Neutrosophic regression; Neutrosophic correlation

1. Introduction

The concept of fuzzy logic was introduced by Zadeh [1], the elements in the collections are represented by the membership value in the closed interval [0,1]. Atanassov [2,3,4] introduce the intuitionistic fuzzy set that is an extension of the fuzzy set. It is useful to examine the real-life circumstances by considering membership and non-membership grades but without indeterminate membership grades. Smarandache [5, 6] extend the idea of intuitionistic fuzzy sets with the account of indeterminate membership grades, which we called Neutrosophic sets. Aftermath, Salama et al., [7] introduced the operations on Neutrosophic sets and progressed Neutrosophic sets theory in [8, 9, 10, 11, 12].

The important role of analyzing the correlation of dependent and independent variables is to estimate the strength and relation between two variables. Hanafy et al., [13] introduced the concepts of Neutrosophic correlation and its coefficients for the case of finite spaces. The Neutrosophic regression analysis is a powerful method to identify the relationships between the dependent and independent variables and also forecasting the uncertainty observation data. Some of the applications of Neutrosophic regression can be seen in literature such as Karacoska [14], Cervigon, et al., [15], Kumar & Chong [16], and Abdul et al., [17]. Smarandach [18] introduced the
theory of Neutrosophic statistics that is the extension of classical statistics and also investigated Neutrosophic regression analysis. The real-time applications of Neutrosophic regression can be seen in Aslam [20], Salama et al., [21]. Prabhu et al., [22] analyzed the real-time multiple analysis. Some other contributions are in this domain have already been done by various researchers such as Tanaka & Ishibuchi [23] and Aslam [24].

Broumi & Smarandache [25] studied the weighted correlation and correlation coefficient between two interval Neutrosophic sets that were defined by Wang et al., [26]. Zhang et al., [27] explained the correlation coefficient measures and their entropy for interval Neutrosophic sets. Ye [28] proposed the two correlation coefficients between normal Neutrosophic numbers (NNSs) based on the score functions of normal Neutrosophic numbers (NNNs) and investigated their properties. He also developed a MADM method with NNSs under normal Neutrosophic numbers. Ye [29] presented a new correlation coefficient measure between dynamic single-valued Neutrosophic multisets. Karaaslan [30] studied the measures between two Neutrosophic sets; two interval-Neutrosophic sets; two Neutrosophic-refined sets and their applications of these methods are utilized in multi-criteria decision-making problems. Broumi and Smarandache [31] also proposed the correlation coefficient between interval Neutrosophic sets. Rajarajeswari and Uma [32] put forward the correlation measure for IFMS. Recently, Broumi and Smarandache [reference] defined the Hausdorff distance between Neutrosophic sets and some similarity measures based on the distance such as the set-theoretic approach and matching function to calculate the similarity degree between Neutrosophic sets. Broumi [32] explained the concept of correlation measure of Neutrosophic-refined sets that is the extension of the correlation measure of Neutrosophic sets and intuitionistic fuzzy multi-sets. Le [33] established the fuzzy decision-making method based on the weighted correlation coefficient under the intuitionistic fuzzy environment. Le [34] explained the cosine similarity measures for intuitionistic fuzzy sets and their applications. Gerstenkorn [35] studied the concept of correlation under the environment of intuitionistic fuzzy sets. Further, Hung [36] defined the correlation for intuitionistic fuzzy sets based on the centroid method. Ye [37] introduced the multicriteria decision-making method by the use of the correlation coefficient under a single-valued Neutrosophic environment. Deli [38] studied the concept of Neutrosophic-refined sets and their applications in medical diagnosis. Sahin [39] explained the correlation coefficient of single-valued Neutrosophic hesitant fuzzy sets and applied them in decision-making problems. Pramanik et al., [40] studied the multicriteria decision-making problems by applying a rough Neutrosophic correlation coefficient. Nagarajan et al., [41] explained Neutrosophic interval valued graphs. Lathamaheswari et al., [42] explained type 2 fuzzy in bio medicine. Ye [43] explained the improved correlation coefficients of single-valued Neutrosophic sets and interval Neutrosophic sets for multiple attribute decision-making problems. Liu et al., [44] established a correlation coefficient for the interval-valued Neutrosophic hesitant fuzzy sets and applied them in multiple attribute decision-making. Ye [45] studied the multi-criteria decision-making method using the correlation coefficient under a single-valued Neutrosophic environment. González-Rodríguez et al.,[46] explained ANOVA test for Fuzzy data. Jiryaei A et al.,[47] studied fuzzy random variables.
2. Preliminaries:

Regression line with dependent and one independent equation is

\[ Y = a + bX + e. \] (1)

When \( Y \) is the output value on dependent, variable \( X \) is the input value of the independent variable, \( b \) is the slope, \( a \) is the intercept and \( e \) is the residual.

More than one independent variable equation as:

\[ Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n + e \] (2)

Here \( n \) number of independent variables and \( b_1, b_2, \ldots, b_n \) are number of slopes for each. \( e \) is the standard error. The estimation of \( a \) and \( b \) for to minimize the error of prediction equation

\[ Y' = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n \] (3)

The equation for \( a \) with two independent variables is:

\[ a = Y - b_1X_1 - b_2X_2 \] (4)

For the two-variable case:

\[ b_1 = \frac{\sum x_1^2 \sum y - \sum x_1 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \] (5)

\[ b_2 = \frac{\sum x_1^2 \sum x_2 y - \sum x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \] (6)

From the above equations 5 & 6 only for two variables \( x_1 \) and \( x_2 \).

Smarandache [18] is given the Neutrosophic extended for classical statistics operation. The operations are as follows.

Let’s \( S_1 \) and \( S_2 \) be two sets of numbers.

\( S_1+S_2 = \{ x_1+x_2 \mid x_1 \in S_1 \text{ and } x_2 \in S_2 \} \)

\( S_1-S_2 = \{ x_1-x_2 \mid x_1 \in S_1 \text{ and } x_2 \in S_2 \} \)

\( S_1 \cdot S_2 = \{ x_1 \cdot x_2 \mid x_1 \in S_1 \text{ and } x_2 \in S_2 \} \)

\( a \cdot S_1 = \{ a \cdot x \mid x \in S_1 \} \)

\( a + S_1 = S_1 + a = \{ a+x \mid x \in S_1 \} \)

\( a - S_1 = \{ a-x \mid x \in S_1 \} \)

\( S_1 - a = \{ x-a \mid x \in S_1 \} \)

\( S_1 \cdot S_2 = \{ x_1 \cdot x_2 \mid x_1 \in S_1 \text{ and } x_2 \neq 0 \} \)

\( S_1 \cap S_2 = \{ x \mid x \in S_1 \text{ and } x \in S_2 \} \)
$S_1a = \{x: x \in S_1, a \neq 0\}$

$aS_1 = \{ax_1 : x_1 \in S_1, \ a \neq 0\}$

$\sqrt[S_1]{n} = \{\sqrt[x_1]{n} : x_1 \in S_1\}$

3. Numerical example

In table 1 shows the student performance campus interview based on aptitude and personality test, that measure the conscientiousness

$Y$ is the dependent variable conscientiousness $x_1$ is the aptitude test and personality test as shown in the following table 1.

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x_1$</th>
<th>$x_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,3]</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2,1</td>
</tr>
<tr>
<td>[2,4]</td>
<td>[1,2]</td>
<td>[3,2]</td>
</tr>
<tr>
<td>4</td>
<td>[2,3]</td>
<td>4</td>
</tr>
<tr>
<td>[1,4]</td>
<td>[2,1]</td>
<td>[4,4]</td>
</tr>
<tr>
<td>6</td>
<td>[2,3]</td>
<td>[4,5]</td>
</tr>
<tr>
<td>[2,4]</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>[10,13]</td>
<td>[5,6]</td>
<td>[6,7]</td>
</tr>
<tr>
<td>[14,15]</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>[7,1]</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
\sum x_1y = \sum X_1Y - \frac{\sum X_1 \sum Y}{N} \quad \quad (7)
\]

\[
\sum x_2y = \sum X_2Y - \frac{\sum X_2 \sum Y}{N} \quad \quad (8)
\]

\[
\sum x_1x_2 = \sum X_1X_2 - \frac{\sum X_1 \sum X_2}{N} \quad \quad (9)
\]

Using the equation 7, 8, and 9

\[
\sum x_1y = [38, 91.9] \quad \sum x_2y = [23, 136.1] \quad \sum x_1x_2 = [35, 19.9]
\]

Matrix form of the values is corresponding to the correlation, sum of square, and cross product of the variables as shown in the following table 2.
Using equation 5 and 6 the value of the regression coefficient
\[ b_1 = [-2.34988, 1.650734], \quad b_2 = [0.965172, 1.093934] \]

from equation 4 the value of the intercept is
\[ a = [-4.29976, 10.18347] \]

Therefore the Neutrosophic regression equation is
\[ Y = [-4.29976, 10.18347] + [-2.34988, 1.650734] x_1 + [0.965172, 1.093934] x_2 \]

The proportion of variance is in the set of independent variables is R square value. The Neutrosophic R square value is

A Neutrosophic residual sum of squares is \( NRSS = \sum (y - \hat{y})^2 \)  
(9)

\[ NRSS = \sum (y - \hat{y})^2 = [183,267.7] \]

A Neutrosophic total sum of squares \( NTSS = \sum (y - \bar{y})^2 \)  
(10)

\[ NTSS = \sum (y - \bar{y})^2 = [22682,1875] \]

A Neutrosophic coefficient of determination is \( NCD = 1 - \frac{NRSS}{NTSS} \)  
(11)

\[ NCD = 1 - \frac{NRSS}{NTSS} = [0.097,0.129] \]

The Neutrosophic mean of Y is [46.50]. The Neutrosophic r square is [0.09,0.12] from the above results shows that the variation between independent and dependent variables is 9 % and 12 %. That means the student performance campus interview variation based on aptitude and personality test is between 9 % and 12 %. Hence, it is revealed that these variables are also affected by the student performance on-campus interview.
4. Significance test of R square

Using the F test for significance of R square is

$$F = \frac{R^2 / K}{(1 - R^2)(N - K - 1)}$$  \hspace{1cm} (12)

Which is distributed as F with K and N-K-1 degrees of freedom when the null hypothesis is true. Now $R^2$ represents the multiple correlations rather than the single correlation.

The null hypothesis: R square value is not zero population with degrees of freedom is N-K-1

Using (12), the Neutrosophic F value is [0.007904,0093]

Comparing the tabulated value using degrees of freedom and the calculated value. It shows that the null hypothesis is accepted.

5. Regression with beta weights

Comparison of correlation and regression equation is

$$Z_Y' = r_{xy}Z_x$$  \hspace{1cm} (13)

But $\beta$ means a b weight when X and Y are in standard scores, so for the simple regression case, $r = \beta$, and we have:

$$Z_Y' = \beta Z_x$$  \hspace{1cm} (14)

The bottom line on this is we can estimate $\beta$ weights using a correlation matrix.

$$\beta_1 = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{1 - r_{x_1x_2}}$$  \hspace{1cm} (15)

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{1 - r_{x_1x_2}^2}$$  \hspace{1cm} (16)

where $r_{yx_1}$ is the correlation of y with $X_1$, $r_{yx_2}$ is the correlation of y with $X_2$ and $r_{12}$ is the correlation of $x_1$ with $x_2$. Note that the two formulas are nearly identical and the correlation matrix shows in table:3
Using the equation 15 and 16 calculate the Neutrosophic beta coefficients. That is

\[ \beta_1 = [-0.50399, -1.3679], \beta_2 = [-1.23030, 0.329977] \]

Note that there is a surprisingly large difference in beta weights given the magnitude of correlations.

**6. The limitations on statistics**

In table 4 shows that limitation on different category statistics.

### Table 4: Limitation on Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical statistics</td>
<td>The analysis only for the determined parameter. Testing the analysis of variance and significance under classical statistics only for determined observation.</td>
</tr>
<tr>
<td>Fuzzy statistics</td>
<td>It will be applied for observations in Fuzzy. Under fuzzy statistics testing the analysis of variance and significance only for the observations are fuzzy and uncertain.</td>
</tr>
<tr>
<td>Intuitionistic fuzzy statistics</td>
<td>It will apply only intervals belongs to membership and non-membership. Under Intuitionistic statistics testing the analysis of variance and significance only for the observation are membership and non-membership that belongs to the real unit interval.</td>
</tr>
<tr>
<td>Neutrosophic statistics</td>
<td>It is applied to an uncertain environment. Under Neutrosophic statistics testing the analysis of...</td>
</tr>
</tbody>
</table>
extension of intuitionistic fuzzy sets. variance and significance when the observations are not fuzzy in the interval and it is an extension of classical and fuzzy statistics.

7. Conclusion

In this paper, we introduce the multiple regression method under the environment of Neutrosophic sets. Moreover, we proposed a method to compute the correlation coefficient of Neutrosophic sets which is given us information about the degree of the relationships between the variables based on Neutrosophic sets. Further, the method is applied to predict the students' performance in campus interviews based on aptitude and personality tests. Based on the above method the result shows that the variation between independent and dependent variables is 9% and 12%, which means that the students' performance variations based on aptitude and personality tests are between 9% and 12%. Thus, it is revealed that aptitude and personality tests are affected students' performance in campus interviews. Future work will be focused on the concept of interval Neutrosophic multiple regression analysis.

Acknowledgment

This research was supported by the Ministry of Higher Education Malaysia, Malaysia, under FRGS Grant No: K179.

References


Received: March 3, 2021. Accepted: Jun 4, 2021