Neutrosophic Correlation and Simple Linear Regression

A. Salama
O. M. Khaled
K. M. Mahfouz

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

Recommended Citation

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 amywinter@unm.edu.
Neutrosophic Correlation and Simple Linear Regression

A. A. Salama¹, O. M. Khaled² and K. M. Mahfouz³

¹,²,³ Department of Mathematics and Computer Science, Faculty of Sciences, Port Said University, 23 December Street, Port Said 42522, Egypt.
Email: drsalama44@gmail.com

Abstract. Since the world is full of indeterminacy, the neutrosophics found their place into contemporary research. The fundamental concepts of neutrosophic set, introduced by Smarandache in [7, 8]. Recently, Salama et al. [14, 15, 16, 32] introduced the concept of correlation coefficient of neutrosophic data. In this paper, we introduce and study the concepts of correlation and correlation coefficient of neutrosophic data in probability spaces and study some of their properties. Also, we introduce and study the neutrosophic simple linear regression model. Possible applications to data processing are touched upon.

Keywords: Correlation Coefficient, Fuzzy Sets, Neutrosophic Sets, Intuitionistic Fuzzy Sets, Neutrosophic Data; Neutrosophic Simple Linear Regression

1 Introduction

In 1965 [13], Zadeh first introduced the concept of fuzzy sets. Fuzzy set is very much useful and in this one real value \( \mu_A(x) \in [0,1] \) is used to represent the grade of membership of a fuzzy set \( A \) defined on the crisp set \( X \). After two decades Atanassov [18, 19, 20] introduced another type of fuzzy sets that is called intuitionistic fuzzy set (IFS) which is more practical in real life situations. Intuitionistic fuzzy sets handle incomplete information i.e., the grade of membership function and non-membership function but not the indeterminate information and inconsistent information which exists obviously in belief system. Smarandache [7,8] introduced another concept of imprecise data called neutrosophic sets. Salama et al. [1] introduced and studied the operations on neutrosophic sets and developed neutrosophic sets theory in [25, 26, 27, 28, 29, 30, 31, 32]. In statistical analysis, the correlation coefficient plays an important role in measuring the strength of the linear relationship between two variables. As the correlation coefficients defined on crisp sets have been much discussed, it is also very common in the theory of fuzzy sets to find the correlation between fuzzy sets, which accounts for the relationship between the fuzzy sets. Salama et al. [15] introduced the concepts of correlation and correlation coefficient of neutrosophic in the case of finite spaces. In this paper we discuss and derived a formula for correlation coefficient, defined on the domain of neutrosophic sets in probability spaces.

2 Terminologies

Definition 2.1 [13]
Let \( X \) be a fixed set. A fuzzy set \( A \) of \( X \) is an object having the form \( A = \{(x, \mu_A(x)), x \in X \} \) where the function \( \mu_A : X \rightarrow [0,1] \) define the degree of membership of the element \( x \in X \) to the set \( A \). Let \( X \) be a fixed set. An intuitionistic fuzzy set \( A \) of \( X \) is an object having the form: \( A = \{(x, \mu_A(x), \gamma_A(x)), x \in X \} \), where the function: \( \mu_A : X \rightarrow [0,1] \) and \( \gamma_A : X \rightarrow [0,1] \) define respectively the degree of membership and degree of non-membership of the element \( x \in X \) to the set \( A \), which is a subset of \( X \) and for every \( x \in X \), \( 0 \leq \mu_A(x) + \gamma_A(x) \leq 1 \).

Let \( X \) be a non-empty fixed set. A neutrosophic set (NS) \( A \) is an object having the form: 
\[
A = \{(x, \mu_A(x), \gamma_A(x), \nu_A(x)), x \in X \}
\]
where \( \mu_A(x), \gamma_A(x) \) and \( \nu_A(x) \) represent the degree of membership function, the degree of indeterminacy, and the degree of non membership function respectively of each element \( x \in X \) to the set \( A \).

In 1991, Gerstenkorn and Manko [24] defined the correlation of intuitionistic fuzzy sets \( A \) and \( B \) in a finite set \( X = \{x_1, x_2, \ldots, x_n \} \) as follows:
\[
C_{GM}(A, B) = \sum_{i=1}^{n} (\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i))
\]
and the correlation coefficient of fuzzy numbers \( A, B \) was given by:
\[
\rho_{GM} = \frac{C_{GM}(A, B)}{\sqrt{T(A) \cdot T(B)}}
\]
where \[ T(A) = \sum_{i=1}^{n} (\mu_A^2(x_i) + \nu_A^2(x_i)). \] (2.3)

Yu [4] defined the correlation of \( A \) and \( B \) in the collection \( F([a,b]) \) of all fuzzy numbers whose supports are included in a closed interval \([a,b]\) as follows:

\[ C_Y(A,B) = \frac{1}{b-a} \int_a^b \mu_A(x)\mu_B(x) + \nu_A(x)\nu_B(x) \, dx. \] (2.4)

where \( \mu_A(x) + \nu_A(x) = 1 \) and the correlation coefficient of fuzzy numbers \( A,B \) was defined by

\[ \rho_Y = \frac{C_Y(A,B)}{\sqrt{C_Y(A,A) \cdot C_Y(B,B)}}. \] (2.5)

In 1995, Hong and Hwang [5] defined the correlation of intuitionistic fuzzy sets \( A \) and \( B \) in a probability space \((X, B,P)\) as follows:

\[ C_{HHH}(A,B) = \int_X (\mu_A \mu_B + \nu_A \nu_B) \, dP \] (2.6)

and the correlation coefficient of intuitionistic fuzzy numbers \( A,B \) was given by

\[ \rho_{HHH} = \frac{C_{HHH}(A,B)}{\sqrt{C_{HHH}(A,A) \cdot C_{HHH}(B,B)}}. \] (2.7)

Salama et al. [15] defined the correlation of neutrosophic data in a finite set \( X = \{x_1, x_2, \ldots, x_n\} \) as follows:

\[ C_{HS}(A,B) = \sum_{i=1}^{n} (\mu_A(x_i) + \nu_A(x_i) + \gamma_A(x_i)\gamma_B(x_i)) \] (2.10)

and the correlation coefficient of fuzzy numbers \( A,B \) was given by:

\[ \rho_{HS} = \frac{C_{HS}(A,B)}{\sqrt{T(A) \cdot T(B)}} \] (2.11)

where

\[ T(A) = \sum_{i=1}^{n} (\mu_A^2(x_i) + \nu_A^2(x_i) + \gamma_A(x_i)) \] (2.12)

\[ T(B) = \sum_{i=1}^{n} (\mu_B^2(x_i) + \nu_B^2(x_i) + \gamma_B(x_i)) \] (2.13)

Salama et al. [16] introduce the concept of positively and negatively correlated and used the concept of centroid to define the correlation coefficient of neutrosophic sets which lies in the interval \([-1, 1]\) was given by:

\[ \rho_{HS} = \frac{C_{HS}(A,B)}{\sqrt{C_{HS}(A,A) \cdot C_{HS}(B,B)}} \] (2.14)

where

\[ C_{HS} = m(\mu_A)m(\mu_B) + m(\nu_A)m(\nu_B) + m(\gamma_A)m(\gamma_B) \] (2.15)

\[ m(\mu_A) = \int x \mu_A(x) \, dx \] \[ \int \mu_A(x) \, dx \]

\[ m(\nu_A) = \int x \nu_A(x) \, dx \] \[ \int \nu_A(x) \, dx \]

A. A. Salama, O. M. Khaled and K. M. Mahfouz, Neutrosophic Correlation and Simple Linear Regression
\[ m(v_A) = \int x v_A(x) dx \quad m(v_B) = \int x v_B(x) dx \]
\[ m(\gamma_A) = \int x \gamma_A(x) dx \quad m(\gamma_B) = \int x \gamma_B(x) dx \]

3. Correlation Coefficient of Neutrosophic Sets

Let \((X, B, P)\) be a probability space and \(A\) be a neutrosophic set in a probability space \(X\),
\[ A = \{ (x, \mu_A(x), \gamma_A(x), v_A(x)) \ | x \in X \} \]
where \(\mu_A(x), \gamma_A(x), v_A(x): X \to [0,1]\) are, respectively, Borel measurable functions satisfying
\[-1 \leq \mu_A(x) + \gamma_A(x) + v_A(x) \leq 1^+ \]
\[-1 \leq \mu_A(x) + \gamma_A(x) + v_A(x) \leq 1^+ \]
where \([-1, 1^+]\) is non-standard unit interval [3].

**Definition 3.1**

For a neutrosophic sets \(A, B\), we define the correlation of neutrosophic sets \(A\) and \(B\) as follows:
\[ C(A, B) = \int (\mu_A \mu_B + \gamma_A \gamma_B + v_A v_B) dP \quad (3.1) \]
Where \(P\) is the probability measure over \(X\). Furthermore, we define the correlation coefficient of neutrosophic sets \(A\) and \(B\) as follows:
\[ \rho(A, B) = \frac{C(A, B)}{\sqrt{T(A) T(B)}} \quad (3.2) \]
where \(T(A) = C(A, A) = \int (\mu_A^2 + \gamma_A^2 + v_A^2) dP \)
\[ T(B) = C(B, B) = \int (\mu_B^2 + \gamma_B^2 + v_B^2) dP \]

The following proposition is immediate from the definitions.

**Proposition 3.1**

For neutrosophic sets \(A\) and \(B\) in \(X\), we have
i. \( C(A, B) = C(B, A) \), \( \rho(A, B) = \rho(B, A) \).
ii. If \( A=B \), then \( \rho(A, B) = 1 \).

The following theorem generalizes both Theorem 1 [24], Proposition 2.3 [4144] and Theorem 1[15] of which the proof is remarkably simple.

**Theorem 3.1**

For neutrosophic sets \(A\) and \(B\) in \(X\), we have
\[ 0 \leq \rho(A, B) \leq 1 \quad (3.3) \]

**Proof**

The inequality \(\rho(A, B) \geq 0\) is evident since \(C(A, B) \geq 0\) and \(T(A) T(B) \geq 0\). Thus, we need only to show that \(\rho(A, B) \leq 1\), or \(C(A, B) \leq \frac{1}{2} \sqrt{T(A) T(B)}\).
For an arbitrary real number \(k\), we have
\[ 0 \leq \int \left( (\mu_A - k \mu_B)^2 + (\gamma_A - k \gamma_B)^2 + (v_A - k v_B)^2 \right) dP \]
\[ = \int \left( \mu_A^2 + \gamma_A^2 + v_A^2 \right) dP - 2k \int (\mu_A \mu_B + \gamma_A \gamma_B + v_A v_B) dP \]
\[ + k^2 \int (\mu_B^2 + \gamma_B^2 + v_B^2) dP \]

Thus, we can get:
\[ \left( \int (\mu_A \mu_B + \gamma_A \gamma_B + v_A v_B) dP \right)^2 \leq \left( \int (\mu_A^2 + \gamma_A^2 + v_A^2) dP \right) \left( \int (\mu_B^2 + \gamma_B^2 + v_B^2) dP \right) \]
\[ C(A, B)^2 \leq T(A) T(B) \]
Therefore, we have \(\rho(A, B) \leq 1\).

**Theorem 3.2**

\(\rho(A, B) = 1\) if and only if \(A = cB\) for some \(c \in IR\).

**Proof**

Considering the inequality in the proof of Theorem 3.1, then the equality holds if and only if
\[ P[\mu_A = c \mu_B] = P[\gamma_A = c \gamma_B] = P[v_A = c v_B] = 1 \]
for some \(c \in IR\).

which completes the proof.

**Theorem 3.3**

\(\rho(A, B) = 0\) if and only if \(A\) and \(B\) are non-fuzzy sets and they satisfy the condition: \(\mu_A + \mu_B = 1\) or
\[ \gamma_A + \gamma_B = 1 \]
\[ v_A + v_B = 1 \]

**Proof**

Suppose that \(\rho(A, B) = 0\), then \(C(A, B) = 0\). Since \(\mu_A \mu_B + \gamma_A \gamma_B + v_A v_B \geq 0\), then \(C(A, B) = 0\) implies
\[ P[\mu_A \mu_B = 0] = P[\gamma_A \gamma_B = 0] = P[v_A v_B = 0] = 1 \]
which means that
\[ P[\mu_A = 0] = 1, P[\gamma_A = 0] = 1 \text{ and } P[v_A = 0] = 1 \]
If \(\mu_A(x) = 1\), then we can get \(\mu_B(x) = 0\) and \(\gamma_A(x) = v_A(x) = 0\). At the same time, if \(\mu_B(x) = 1\), then we can get \(\mu_B(x) = 0\) and \(\gamma_B(x) = v_B(x) = 0\), hence, we have \(\mu_A + \mu_B = 1\). Conversely, if \(A\) and \(B\) are non-fuzzy sets and \(\mu_A + \mu_B = 1\). If \(\mu_A(x) = 1\), then we can have \(\mu_B(x) = 0\) and \(\gamma_B(x) = v_B(x) = 0\), which implies \(C(A, B) = 0\). Similarly
we can give the proof when $\gamma_A + \gamma_B = 1$ or $\nu_A + \nu_B = 1$.

**Theorem 3.4**

If $A$ is a non-fuzzy set, then $T(A) = 1$.

**The proof** is obvious.

**Example**

For a continuous universal set $X = [1, 2]$, if two neutrosophic sets are written, respectively, $A = \{(x, \mu_A(x), \nu_A(x), \gamma_A(x)) | x \in [1, 2]\}$, $B = \{(x, \mu_B(x), \nu_B(x), \gamma_B(x)) | x \in [1, 2]\}$, where

\[
\mu_A(x) = 0.5(x - 1), \quad 1 \leq x \leq 2, \\
\mu_B(x) = 0.3(x - 1), \quad 1 \leq x \leq 2, \\
\nu_A(x) = 1.9 - 0.9x, \quad 1 \leq x \leq 2, \\
\nu_B(x) = 1.4 - 0.4x, \quad 1 \leq x \leq 2, \\
\gamma_A(x) = (5 - x)/6, \quad 1 \leq x \leq 2, \\
\gamma_B(x) = 0.5x - 0.3, \quad 1 \leq x \leq 2.
\]

Thus, we have $C(A, B) = 0.79556, \ T(A) = 0.79593$ and $T(B) = 0.93656$. Then we get $\rho(A, B) = 0.936506$.

It shows that neutrosophic sets $A$ and $B$ have a good positively correlated.

**4. Neutrosophic linear regression model**

Linear regression models are widely used today in business administration, economics, and engineering as well as in many other traditionally non-quantitative fields including social, health and biological sciences. Regression analysis is a methodology for analyzing phenomena in which a variable (output or response) depends on other variables called input (independent or explanatory) variables. Function is fitted to a set of given data to predict the value of dependent variable for a specified value of the independent variable. However, the phenomena in the real world cannot be analyzed exactly, because they depend on some uncertain factors and in some cases, it may be appropriate to use neutrosophic regression. Tanaka et al. (1982) [8] proposed the first linear regression analysis with a fuzzy model. According to this method, the regression coefficients are fuzzy numbers, which can be expressed as interval numbers with membership values. Since the regression coefficients are fuzzy numbers, the estimated dependent variable is also a fuzzy number. A collection of recent papers dealing with several approaches to fuzzy regression analysis can be found in Kacprzyk and Fedrizzi (1992)[17]. Other contributions in this area are by Diamond (1988)[22], Tanaka and Ishibuchi (1991)[11], Savic and Pedrycz (1991)[16] and Ishibuchi (1992) [9]. Yen et al. (1999) [12] extended the results of a fuzzy linear regression model that uses symmetric triangular coefficient to one with non-symmetric fuzzy triangular coefficients.

In this section we will define the simple linear regression in neutrosophic set.

**Definition 4.1**

Assume that there is a random sample $(x_1, x_2, \ldots, x_n) \in X$, alone with the sequence of data;

\[
((\mu_A(x_1), \gamma_A(x_1), \nu_A(x_1)), (\mu_B(x_1), \gamma_B(x_1), \nu_B(x_1)), \ldots, \\
(\mu_A(x_n), \gamma_A(x_n), \nu_A(x_n)), (\mu_B(x_n), \gamma_B(x_n), \nu_B(x_n)), \text{ as defined in} \\
\text{Definition 2.3, } \mu_A(x), \gamma_A(x), \nu_A(x) \text{ represent the degree of membership function (namely } \mu_A(x) \text{), the degree of non-membership (namely } \nu_A(x) \text{), and the degree of indeterminacy (namely } \gamma_A(x) \text{) respectively of each element } x \in X \text{ to the set A. Also } \mu_B(x), \gamma_B(x), \nu_B(x) \text{ represent the degree of membership function (namely } \mu_B(x) \text{), the degree of non-membership (namely } \nu_B(x) \text{), and the degree of indeterminacy (namely } \gamma_B(x) \text{) respectively of each element } x \in X \text{ to the set B.}
\]

Consider the following simple neutrosophic linear regression model:

\[
A_i = aB_i + \beta,
\]

where $X_i$ denotes the independent variables, $b$ the estimated neutrosophic intercept coefficient, $a$ the estimated neutrosophic slope coefficients and $Y_i$ the estimated neutrosophic output. As classical statistics linear regression we will define the neutrosophic coefficients $a$ and $b$

\[
\alpha = \frac{C(A, B)}{C(A, A)} \text{ and } \beta = E(Y) - \alpha E(X),
\]

where $C(A, B)$ define in [15] as follows

\[
C(A, B) = \frac{1}{n} \sum_{i=1}^{n} \mu_A(x_i) \mu_B(x_i) + \nu_A(x_i) \nu_B(x_i) + \gamma_A(x_i) \gamma_B(x_i)) \ldots
\]

$E(Y) = \frac{1}{3} (\bar{\mu}_B(x_i) + \bar{\gamma}_B(x_i) + \bar{\nu}_B(x_i))$, and

$E(X) = \frac{1}{3} (\bar{\mu}_A(x_i) + \bar{\gamma}_A(x_i) + \bar{\nu}_A(x_i))$.

$\bar{\mu}_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \mu_A(x_i), \bar{\mu}_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \mu_B(x_i) \ldots$

$\bar{\gamma}_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \gamma_A(x_i), \bar{\gamma}_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \gamma_B(x_i) \ldots$

$\bar{\nu}_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \nu_A(x_i), \bar{\nu}_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \nu_B(x_i) \ldots$

**Example 4.1**

In example [15], we compute that

\[
C(A, B) = 0.88, \quad C(A, A) = T(A) = 0.83,
\]
\[ \mu_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \mu_A(x_i) = 0.4 , \]
\[ \mu_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \mu_B(x_i) = 0.3 , \]
\[ \nu_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \nu_A(x_i) = 0.3 , \]
\[ \nu_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \nu_B(x_i) = 0.35 , \]
\[ \psi_A(x_i) = \frac{1}{n} \sum_{i=1}^{n} \psi_A(x_i) = 0.35 , \]
\[ \psi_B(x_i) = \frac{1}{n} \sum_{i=1}^{n} \psi_B(x_i) = 0.6 . \]
\[ E(\bar{Y}) = \frac{1}{3}(\mu_B(x_i) + \nu_A(x_i) + \psi_B(x_i)) = 0.42 , \]
\[ E(\bar{X}) = \frac{1}{3}(\mu_A(x_i) + \psi_A(x_i) + \nu_A(x_i)) = 0.35 , \]
then \( \alpha = 1.06 \) and \( \beta = 0.49 . \)

Then neutrosophic linear regression model is given by
\[ A_i = 1.06B_i + 0.49 . \]

**Conclusion**

Our main goal of this work is propose a method to calculate the correlation coefficient of neutrosophic sets which lies in \([0,1]\), give us information for the degree of the relationship between the neutrosophic sets. Further, we discuss some of their properties and give example to illustrate our proposed method reasonable. Also we get the simple linear regression on neutrosophic sets.

**References**


Received: May 23, 2014. Accepted: June 30, 2014.