New Trends in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume I: Foundations

Editors

Krassimir T. Atanassov
Michał Baczyński
Józef Drewniak
Janusz Kacprzyk
Maciej Krawczak
Eulalia Szmidt
Maciej Wygralak
Sławomir Zadrożny

SRI PAS IBS PAN

New Trends in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics Volume I: Foundations

iBS PAN

Systems Research Institute
Polish Academy of Sciences

New Trends in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume I: Foundations

Editors

Krassimir Atanassov
Michał Baczyński
Józef Drewniak
Janusz Kacprzyk
Maciej Krawczak
Eulalia Szmidt
Maciej Wygralak
Sławomir Zadrożny

IBS PAN **SRI PAS**

Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISBN 83-894-7546-4

On the Principal Component Analysis and the Correlation Coefficient

Beloslav Riečan
Faculty of Natural Sciences, Matej Bel University,
Department of Mathematics,
Tajovského 40, 97401 Banská Bastrica, Slovakia
Mathematical Institute, Slovak Academy of Sciences,
Štefánikova 49, 84101 Bratislava, Slovakia
beloslav.riecan@umb.sk

Abstract

The paper uses a new approach published in [4] by E. Szmidt, J. Kacprzyk and P. Bujnowski for data expressed in terms of Atanassov's intuitionistic fuzzy sets. It is shown that it can be expressed also in terms of the classical correlation coefficient.

1 Introduction

Consider the probabability space $(\Omega, \mathcal{S}, P)$ in the Kolmogorov sense, i.e. Ω is a nonempty set, $\mathcal{S}$ is a σ-algebra of subsets of Ω and $P : \Omega \to [0, 1]$ is a probability measure. Let $\xi : \Omega \to R$ be a random variable, $P_\xi : \mathcal{B}(R) \to [0, 1]$ its probability distribution defined by

$$P_\xi(A) = P(\xi^{-1}(A)), A \in \mathcal{B}(R),$$

New Trends in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume I: Foundations *(K.T. Atanassow, M. Baczyński, J. Drewniak, J. Kacprzyk, M. Krawczak, E. Szmidt, M. Wygralak, S. Zadrożny, Eds.), IBS PAN - SRI PAS, Warsaw, 2013.*

where $\mathcal{B}(R)$ is the σ-algebra of all Borel subsets of R. Then the mean value $E\xi)$ can be defined as the integral

$$E(\xi) = \int_\Omega \xi dP = \int_R t dP_\xi(t),$$

if the integral exists. The dispersion is defined as

$$\sigma^2(\xi) = \int_\Omega (\xi - E(\xi))^2 dP = \int_R (t - E(\xi))^2 dP_\xi(t),$$

if the function ξ is square integrable. If $\xi, \eta > \Omega \to R$ are two random variables then the correlation coefficient $r(\xi, \eta)$ is defined by the equality

$$r(\xi, \eta) = \frac{1}{\sigma(\xi)\sigma(\eta)} \int_\Omega (\xi - E(\xi)(\eta - E(\eta))dP =$$

$$= \frac{1}{\sigma(\xi)\sigma(\eta)} \int_R (u - E(\xi)((v - E(\eta))dP_T(u, v),$$

where $P_T : \mathcal{B}(R^2) \to [0,1]$ is defined by the equality

$$P_T(A) = P(T^{-1}(A)), A \in \mathcal{B}(R^2);$$

here $\mathcal{B}(R^2)$ is the σ-algebra of all Borel subsets of R^2, and $T : \Omega \to R^2$ is defined by the equality

$$T(\omega) = (\xi(\omega), \eta(\omega)).$$

Of course, the concept can be realizede not only on Boolean algenras but also in multilogic case of MV-algebras, and especially in Atanassov intuitionistic fuzzy sets.

2 Intuitionistic fuzzy sets

An IF-set is a couple $A = (\mu_A, \nu_A)$ of two fuzzy sets such that

$$\mu_A : \Omega \to [0,1], \nu_A : \Omega \to [0,1],$$

and

$$\mu_A + \nu_A \leq 1.$$

We shall call $\mu_A : \Omega \to [0,1]$ the membership function, $\nu_A : \Omega \to [0,1]$ tho non-membership function, and $\pi_A : \Omega \to [0,1]$ the hesitation margin. Let $\mathcal{F}$ be the family of all IF-sets on Ω. We write

$$A \leq B \Longleftrightarrow \mu_A \leq \mu_B, \nu_A \geq \mu_B.$$

Then $(1_\Omega, 0_\Omega)$ is the greatest element of $\mathcal{F}$, $(0_\Omega, 1_\Omega)$ is the smallest element of $\mathcal{F}$. We shall use three binary operations on $\mathcal{F}$: the union of IF-sets A, B (the disjunction of corresponding assertions)

$$A \oplus B = ((\mu_A + \mu_B) \wedge 1, (\nu_A + \nu_B - 1) \vee 0),$$

the intersection of A, B (the conjunction of corresponding asertions)

$$A \odot B = (\mu_A + \mu_B - 1) \vee 0, (\nu_A + \nu_B) \wedge 1),$$

and the product of A, B

$$A.B = (\mu_A.\mu_B, 1 - (1 - \nu_A).(1 - \nu_B)) =$$

$$= (\mu_A.\mu_B, \nu_A + \nu_B - \nu_A.\nu_B).$$

Instead of a probability measure we consider a state $m : \mathcal{F} \to [0, 1]$ satisfying the following properties:

(i) $m(1_\Omega, 0_\Omega) = 1, m(0_\Omega, 1_\Omega)) = 0,$

(ii) $A \odot B = 0_\Omega, 1_\Omega) \Longrightarrow m(A \oplus B) = m(A) + m(B),$

(iii) $A_n \nearrow A(i.e.\mu_{A_n} \nearrow \mu_A, \nu_{A_n} \searrow \nu_A) \Longrightarrow m(A_n) \nearrow m(A).$

Instead of a random variable we shall consider an observable what is a mapping $x : \mathcal{B}(R) \to \mathcal{F}$ satisfying the following properties:

(i) $x(R) = (1_\Omega, 0_\Omega), x(\emptyset) = (0_\Omega, 1_\Omega),$

(ii) $A \cap B = \emptyset \Longrightarrow x(A) \odot x(B) = (0_\Omega, 1_\Omega), x(A \cup B) = x(A) + x(B),$

(iii) $A_n \nearrow A \Longrightarrow x(A_n) \nearrow x(B).$

If $\xi : \Omega \to R$ is a random variable, then $x : \mathcal{B}(R) \to \mathcal{S}$ defined by $x(A) = \xi^{-1}(A)$ has the properties stated above.

Theorem 1. If $x : \mathcal{B}(R) \to \mathcal{F}$ is an observable and $m : \mathcal{F} \to [0, 1]$ is a state, then $m_x : \mathcal{B}(R) \to [0, 1]$ defined by $m_x(A) = m(x(A))$ is a probability measure.

Proof is straightforward.

Theorem 1 gives a possibility to define moments. We asume that there is given a fixed state $m : \mathcal{F} \to [0, 1]$.

Definition 1. If $x : \mathcal{B}(R) \to \mathcal{F}$ is an observable, then the mean value $E(x)$ is defined by the formula

$$E(x) = \int_R t dm_x(t),$$

if the integral exists. If there exists $\int_R t^2 dm_x(t)$, then we define the dispersion

$$\sigma^2(x) = \int_r (t - E(x))^2 dm_x(t).$$

For defining the correlation coefficient we need the notion of the joint observable. The notion corresponds to the notion of a random vector.

Theorem 2. For any observables $x, y : \mathcal{B}(R) \to \mathcal{F}$ there exists their joint observable $h : \mathcal{B}(R^2) \to \mathcal{F}$, hence the following properties are satisfied:

(i) $h(R^2) = (1_\Omega, 0_\Omega), x(\emptyset) = (0_\Omega, 1_\Omega)$,

(ii) $A \cap B = \emptyset \Longrightarrow h(A) \odot h(B) = (0_\Omega, 1_\Omega), h(A \cup B) = h(A) + h(B)$,

(iii) $A_n \nearrow A \Longrightarrow h(A_n) \nearrow h(B)$.

(iv) $C, D \in \mathcal{B}(R) \Longrightarrow h(C \times D) = x(C).y(D)$.

Proof. [2], Theorem 2.

Theorem 3. Let $h : \mathcal{B}(R^2) \to \mathcal{F}$ be the joint observable of observables $x, y : \mathcal{B}(R) \to \mathcal{F}$. Define $m_h : \mathcal{B}(R^2) \to [0, 1]$ by $m_h(A) = m(h(A))$. Then m_h is a probability measure.

Definition 2. If $x, y : \mathcal{B}(R) \to \mathcal{F}$ are observables, and h is their joint observable, then we define the correlation coefficient $r(x, y)$ by the formula

$$r(x,y) = \frac{1}{\sigma(x)\sigma(y)} \int\int_{R^2} (u - E(x))(v - E(y)) dm_h(u, v).$$

3 PCA and Correlation Coefficient

According to [4] consider a finite set $\Omega = \{x_1, x_2, ..., x_n\}$ and two Atanassov intuitionistic fuzzy sets

$$A = (\mu_A, \nu_A), B = (\mu_B, \nu_B),$$

where

$$\mu_A : \Omega \to [0, 1], \nu_A : \Omega \to [0, 1], \mu_B : \Omega \to [0, 1], \nu_B : \Omega \to [0, 1],$$

and

$$\mu_A + \nu_A \le 1, \mu_B + \nu_B \le 1.$$

Consider further

$$\pi_A = 1 - \mu_A - \nu_A, \pi_B = 1 - \mu_B - \nu_B.$$

In [4] the correlation coefficient $r_{A-IFS}(A, B)$ between A and B in Ω is

$$r_{A-IFS}(A, B) = \frac{1}{3}(r_1(A, B) + r_2(A, B) + rt_3(A, B)),$$

where

$$r_1(A, B) = \frac{\Sigma_{i=1}^n (\mu_A(x_i) - \overline{\mu_A})(\mu_B(x_i) - \overline{\mu_B})}{(\Sigma_{i=1}^n (\mu_A(x_i) - \overline{\mu_A})^2)^{0.5} (\Sigma_{i=1}^n (\mu_B(x_i) - \overline{\mu_B})^2)^{0.5}},$$

$$r_2(A,B) = \frac{\Sigma_{i=1}^n (\nu_A(x_i) - \overline{\nu_A})(\nu_B(x_i) - \overline{\nu_B})}{(\Sigma_{i=1}^n (\nu_A(x_i) - \overline{\nu_A})^2)^{0.5} (\Sigma_{i=1}^n (\nu_B(x_i) - \overline{\nu_B})^2)^{0.5}},$$

$$r_3(A,B) = \frac{\Sigma_{i=1}^n (\pi_A(x_i) - \overline{\pi_A})(\pi_B(x_i) - \overline{\pi_B})}{(\Sigma_{i=1}^n (\pi_A(x_i) - \overline{\pi_A})^2)^{0.5} (\Sigma_{i=1}^n (\pi_B(x_i) - \overline{\pi_B})^2)^{0.5}}.$$

The main result of our paper is a presentation of all three coefficients r_1, r_2, r_3 by terms of correlations in the sense of Definition 2. We shall find for any $A, B \in \mathcal{F}$ such observables $x, y : \mathcal{B}(R) \to \mathcal{F}$ that

$$r(x,y) = r_1(A,B),$$

and similarly for $r_2(A,B)$, and $r_3(A,B)$.

The main instrument in our investigations will be the state representation theorem from [3].

Theorem 4. To any state $m : \mathcal{F} \to [0,1]$ there exist $\alpha \in R$ and probability measures $P, Q : \mathcal{S} \to [0,1]$ such that

$$m(A) = \int_\Omega \mu_A dP + \alpha(1 - \int_\Omega (\mu_A + \nu_A) dQ).$$

for any $A \in \mathcal{F}$.

4 Membership correlation coefficient

Theorem 5. To any Atanassov intuitionistic fuzzy sets $A, B \in \mathcal{F}$ there exist observables $x, y : \mathcal{B}(R) \to \mathcal{F}$ such that

$$r_1(A,B) = r(x,y).$$

Proof. In Theorem 4 put $\alpha = 0$ and define $P : 2^\Omega \to [0,1]$ by the equality

$$P(K) = \frac{1}{n} card K,$$

hence

$$P(\{x_i\}) = \frac{1}{n}, i = 1, 2, ..., n$$

We want to define an observable $x : \mathcal{B}(R) \to \mathcal{F}$. Let $C \in \mathcal{B}(R)$. Then we put

$$\mu_{x(C)} = \frac{1}{n} card\{i; \mu_A(x_i) \in C\},$$

and

$$\nu_{x(C)} = 1 - \mu_{x(C)}.$$

Then

$$m(x(C)) = \int_\Omega \mu_{x(C)} dP = \frac{1}{n} card\{i; \mu_A(x_i) \in C\},$$

hence

$$m(x(\{\mu_A(x_i)\}) = \frac{1}{n}.$$

Therefore

$$E(x) = \int_R t dm_x(t) = \Sigma_{i=1}^n \int_{\{\mu_A(x_i)\}} t dm_x(t) =$$

$$= \Sigma_{i=1}^n \mu_A(x_i) \frac{1}{n} = \frac{1}{n} \Sigma_{i=1}^n \mu_A(x_i) = \overline{\mu_A}.$$

Similarly

$$\sigma^2(x) = \int_R (t - E(x))^2 dm_x(t) = \frac{1}{n} \Sigma_{i=1}^n (\mu_A(x_i) - \overline{\mu_A})^2.$$

Similarly it can be defined an observable $y : \mathcal{B}(RE) \to \mathcal{F}$ suh that

$$m(y(\{\mu_B(x_i)\})) = \frac{1}{n},$$

$$E(y) = \overline{\mu_B}, \sigma(y)^2 - \frac{1}{n} \Sigma_{i=1}^n (\mu_B(x_i) - \overline{\mu_B})^2.$$

Let $h : \mathcal{B}(R^2) \to \mathcal{F}$ be the joint observable of observables x, y. Then

$$m_h(\{(\mu_A(x_i), \mu_B(x_i))\}) = m(h(\{(\mu_A(x_i), \mu_B(x_i))\})) =$$

$$= m(h(\{\mu_A(x_i)\} \times R)) - m(h(\{\mu_A(x_i)\} \times (R \setminus \{\mu_B(x_i)\}))) =$$

$$= m_x(\{\mu_A(x_i)\}) - 0 = \frac{1}{n}.$$

Put $Q_i = (\mu_A(x_i), \mu_B(x_i))$, and compute

$$\int\int_{R^2} (u - E(x))(v - E(y)) dm_h(u, v) =$$

$$= \Sigma_{i=1}^n \int\int_{\{Q_i\}} (u - \overline{\mu_A})(v - \overline{\mu_B}) dm_h(u, v) =$$

$$= \Sigma_{i=1}^n (\mu_A(x_i) - \overline{\mu_A})(\mu_B(x_i) - \overline{\mu_B}) \frac{1}{n} =$$

$$= \frac{1}{n} \Sigma_{i=1}^n (\mu_A(x_i) - \overline{\mu_A})(\mu_B(x_i) - \overline{\mu_B}).$$

Therefore

$$r(x,y) = \frac{1}{\sigma(x)\sigma(y)}\frac{1}{n}\Sigma_{i=1}^{n}(\mu_A(x_i) - \overline{\mu_A})(\mu_B(x_i) - \overline{\mu_B})$$

$$= \frac{\frac{1}{n}\Sigma_{i=1}^{n}(\mu_A(x_i) - \overline{\mu_A})(\mu_B(x_i) - \overline{\mu_B})}{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(\mu_A(x_i) - \overline{\mu_A})^2}\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(\mu_B(x_i) - \overline{\mu_B})^2}} = r_1(A,B)$$

5 Non-membership correlation coefficient

Theorem 6. To any Atanassov intuitionistic fuzzy sets $A, B \in \mathcal{F}$ there exist observables $x, y : \mathcal{B}(R) \to \mathcal{F}$ such that

$$r_2(A,B) = r(x,y).$$

Proof. In Theorem 4 put $\alpha = 1$, and $P = Q, P(K) = \frac{1}{n} card K, K \subset \Omega$. Let $C \in \mathcal{B}(R)$. Define $x(C) \in \mathcal{F}$ by the formulas

$$\nu_{x(C)} = 1 - card\{i; \nu_A(x_i) \in C\},$$

$$\mu_{x(C)} - 1 - \nu_{x(C)}.$$

Then by Theorem 4

$$m(x(C))) = \int_\Omega \mu_{x(C)} dP + 1 - \int_\Omega (\mu_{x(C)} + \nu_{x(C)}) dP =$$

$$= \int_\Omega (1 - \nu_{x(C)}) dP = \frac{1}{n} card\{i; \nu_A(x_i) \in C\},$$

hence

$$m_x(\{\nu_A(x_i)\}) = \frac{1}{n}, i = 1, 2, ..., n.$$

Therefore

$$E(x) = \int_R dm_x(t) = \Sigma_{i=1}^{n} \int_{\{\nu_A(x_i)\}} t dm_x(t) = \Sigma_{i=1}^{n} \nu_A(x_i)\frac{1}{n} = \overline{\nu_A}.$$

The rest of the proof can be realized similarly as in Theorem 5.

6 Hesitation margin correlation coefficient

Theorem 7. To any Atanassov intuitionistic fuzzy sets $A, B \in \mathcal{F}$ there exist observables $x, y : \mathcal{B}(R) \to \mathcal{F}$ such that

$$r_3(A, B) = r(x, y).$$

Proof. In Theorem 4 put $\alpha = 1$, and $P = Q, P(K) = \frac{1}{n} card K, K \subset \Omega$. Let $C \in \mathcal{B}(R)$. Define $x(C) \in \mathcal{F}$ by the formulas

$$\mu_{x(C)} = 0,$$

$$nu_{x(C)} = \frac{1}{n} card\{i; \mu_A(x_i) \in C\}.$$

Then

$$m(x(C)) = 1 - \int_\Omega \nu_x(C) dP = \frac{1}{n} card\{i; \pi_A(x_i) \in C\},$$

hence

$$m_x(\{\pi_A(x_i)\}) = \frac{1}{n}, i = 1, 2, ..., n.$$

Therefore

$$E(x) = \int_R t dm_x(t) = \Sigma_{i=1}^n \int_{\{\pi_A(x_i)\}} t dm_x(t) =$$

$$= \Sigma_{i=1}^n \pi_A(x_i) frac1n = \overline{\pi_A}.$$

The rest of the proof can be realized similarly as in Theorem 5.

7 Conclusions

We have shown that tehe three correlation components r_1, r_2, r_3 can be expressed by the general correlation coefficient. It could have two possible applications., Firstly some general results could be applied to the case studied in [4]. Secondly, in the concept presented in [4] it could be studied not only discrete case, but e.g. the continuous case of A-IFS.

References

[1] Atanassov K. (1999) Intuitionistic Fuzzy Sets: Theory and Applications. Springer-Verlag.

[2] Atanassov K. Riečan B.: On some properties of IF observables. Noteson IFS, to appear.

[3] Ciungu L., Riečan B. (2009) General form of probabilities. In: Fuzzy Logic and Applications. Proc. WILF Palermo 2009, 101 - 107.

[4] Szmidt E., Kacprzyk J., Bujnowski P.: Advances in principal Component Analysis fo Intuitionistic Fuzzy Data Sets. 2012 IEEE 6th Int. Conf Intelligent Systems, 194 - 199.

The papers presented in this Volume 1 constitute a collection of contributions, both of a foundational and applied type, by both well-known experts and young researchers in various fields of broadly perceived intelligent systems.
It may be viewed as a result of fruitful discussions held during the Eleventh International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2012) organized in Warsaw on October 12, 2012 by the Systems Research Institute, Polish Academy of Sciences, in Warsaw, Poland, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences in Sofia, Bulgaria, and WIT - Warsaw School of Information Technology in Warsaw, Poland, and co-organized by: the Matej Bel University, Banska Bystrica, Slovakia, Universidad Publica de Navarra, Pamplona, Spain, Universidade de Tras-Os-Montes e Alto Douro, Vila Real, Portugal, Prof. Asen Zlatarov University, Burgas, Bulgaria, and the University of Westminster, Harrow, UK:

Http://www.ibspan.waw.pl/ifs2012

The Workshop has also been in part technically supported by COST Action IC0806 ``Intelligent Monitoring, Control and Security of Critical Infrastructure Systems" (INTELLICIS).

The consecutive International Workshops on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGNs) have been meant to provide a forum for the presentation of new results and for scientific discussion on new developments in foundations and applications of intuitionistic fuzzy sets and generalized nets pioneered by Professor Krassimir T. Atanassov. Other topics related to broadly perceived representation and processing of uncertain and imprecise information and intelligent systems have also been included. The Eleventh International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2012) is a continuation of this undertaking, and provides many new ideas and results in the areas concerned.

We hope that a collection of main contributions presented at the Workshop, completed with many papers by leading experts who have not been able to participate, will provide a source of much needed information on recent trends in the topics considered.

ISBN-13 9788389475466