Some Meta-theorems on Fuzzy Cardinalities and their Application

S. Janssens
Department of Applied Mathematics, Biometrics and Process Control,
Ghent University,
Coupure links 653,
B-9000 Gent, Belgium
Saskia.Janssens@rug.ac.be

B. De Baets
Department of Applied Mathematics, Biometrics and Process Control,
Ghent University,
Coupure links 653,
B-9000 Gent, Belgium
Bernard.DeBaets@rug.ac.be

H. De Meyer
Department of Applied Mathematics and Computer Science,
Ghent University,
Krijgslaan 281 (S9),
B-9000 Gent, Belgium
Hans.DeMeyer@rug.ac.be

Abstract

Several meta-theorems are proposed which state that certain inequalities satisfied by cardinalities of ordinary sets, are preserved under fuzzification.

Keywords: Bell inequalities, fuzzy cardinality, quasi-copula, similarity measure

1 Introduction

The task of measuring similarity occurs in a variety of disciplines such as numerical taxonomy, chemistry, information retrieval, ... The most frequently used similarity measures are the ones based on cardinalities of the sets involved. By far the most popular of these similarity measures is known as Jaccard's coefficient [2]. Another well-known similarity measure is the so-called simple matching coefficient [6].

Fuzzy relations can have certain properties, such as reflexivity, symmetry, T-transitivity, ... Similarity measures (resp. fuzzy similarity measures) can be seen as fuzzy relations on $\mathcal{P}(X)$ (resp. $\mathcal{F}(X)$) and since they are based on cardinalities (resp. fuzzy cardinalities), fulfilment of these properties translates into inequalities on these cardinalities. In this paper, we want to propose an intelligent way to verify when these properties, or equivalently, these inequalities on (fuzzy) cardinalities, are fulfilled.

We first recall some basic definitions in Section 2. In Section 3 we introduce the Bell inequalities, we rewrite them in terms of cardinalities in Section 4 and formulate necessary and sufficient conditions for the corresponding inequalities for fuzzy cardinalities in Section 5. Having expounded several meta-theorems in Section 6, we then present a parametrized family of cardinality-based similarity measures in Section 7 and investigate under what conditions the corresponding family of fuzzy similarity measures shows the same transitivity behaviour as in the crisp case.

2 Conjunctors and quasi-copulas

As usual, we define the intersection of two fuzzy sets A and B on a universe X pointwisely, i.e. $A \cap B(x) = I(A(x), B(x))$, by means of an appropriate function I that generalizes Boolean conjunction. Since in this paper we will intersect at most two fuzzy sets at the same time, it suffices to consider as suitable I a commutative conjunctor.

Definition 1 *A binary operation $I : [0,1]^2 \rightarrow [0,1]$ is called a conjunctor if it satisfies for any $x \in [0,1]$:*

(i) *Neutral element 1: $I(x,1) = I(1,x) = x$.*
(ii) *Monotonicity: I is increasing in each variable.*

Note that any conjunctor I coincides on $\{0,1\}^2$ with the Boolean conjunction and satisfies for any $x \in [0,1]$:

(i') Absorbing element 0: $I(x,0) = I(0,x) = 0$.

Moreover, any conjunctor I is bounded from above by $T_{\mathbf{M}}$, i.e. $I(x,y) \leq \min(x,y)$.

Definition 2 [1] *A binary operation $C : [0,1]^2 \rightarrow [0,1]$ is called a quasi-copula if it satisfies for any $x, x_1, x_2, y_1, y_2 \in [0,1]$:*

(i) *Neutral element 1: $C(x,1) = C(1,x) = x$.*
(i') *Absorbing element 0: $C(x,0) = C(x,0) = 0$.*
(ii) *Monotonicity: C is increasing in each variable.*

(iii) *1-Lipschitz property:*

$$|C(x_1, y_1) - C(x_2, y_2)| \leq |x_1 - x_2| + |y_1 - y_2|.$$

For any quasi-copula C it holds that $T_{\mathbf{L}} \leq C \leq T_{\mathbf{M}}$ (with $T_{\mathbf{L}}(x, y) = \max(x + y - 1, 0)$).

3 Bell inequalities

The Bell inequalities originate from quantum mechanics and are, from a mathematical point of view, strongly related to Kolmogorovian probability theory.

The probability of a single random event A_i is denoted by $p_i = P(A_i)$ and the probability of the intersection of a pair of random events is denoted by $p_{ij} = P(A_i \cap A_j)$. Since p_i and p_{ij} are probabilities, the following properties should hold:

$$\begin{aligned} p_{ij} &\leq \min(p_i, p_j)\,, \\ p_i + p_j - p_{ij} &\leq 1\,. \end{aligned}$$

Pitowsky also suggested a geometrical interpretation [5]. The above inequalities represent the faces of a closed convex polytope with vertices $(a, b, \min(a, b))$, where $(a, b) \in \{0, 1\}^2$.

In the case of random experiments concerning three events whereby at most two events are intersected at the same time, Pitowsky found the following set of inequalities, which he called the Bell–Wigner inequalities:

$$\begin{aligned} 0 \leq p_i - p_{ij} - p_{ik} + p_{jk}\,, \\ p_i + p_j + p_k - p_{ij} - p_{ik} - p_{jk} \leq 1\,, \end{aligned}$$

for every $i \neq j \neq k \neq i$.

4 Bell inequalities in terms of cardinalities

Since the Bell inequalities apply in particular to classical probabilities, we can rewrite them also in terms of cardinalities. Consider a finite universe X with dimension n, then the classical probability $P(A)$ is given by $|A|/n$ and the Bell inequalities can be rewritten in the following form:

$$|A| + |B| - n \leq |A \cap B| \leq \min(|A|, |B|)\,,$$
$$0 \leq |A| - |A \cap B| - |A \cap C| + |B \cap C|\,,$$
$$|A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| \leq n\,,$$

for any $A, B, C \in \mathcal{P}(X)$. Note that the vertices of the polytope now correspond to the extreme cases $A, B, C \in \{\emptyset, X\}$. We can prove the following theorem.

Theorem 1 *All linear inequalities with real coefficients of the form:*

$$a_1|A| + a_2|B| + b_{12}|A \cap B| + c \geq 0$$

$$\begin{aligned} &a_1|A| + a_2|B| + a_3|C| \\ &+ b_{12}|A \cap B| + b_{13}|A \cap C| + b_{23}|B \cap C| + c \geq 0 \end{aligned}$$

hold for all ordinary sets A, B, C in a finite universe X if and only if these inequalities hold for all $A, B, C \in \{\emptyset, X\}$.

5 Bell-type inequalities in fuzzy logic

We can rewrite the above Bell inequalities in the following way. Let A_i and A_j be fuzzy sets in a finite universe X with arbitrary dimension n. The fuzzy probability $P(A_i)$ is usually defined as $\sum_u A_i(u)/n$. Starting from the inequality

$$0 \leq P(A_i) + P(A_j) - P(A_i \cap A_j) \leq 1,$$

we can write the following, with $A_i \cap A_j$ modelled by means of a commutative conjunctor I

$$0 \leq \frac{\sum A_i(u)}{n} + \frac{\sum A_j(u)}{n} - \frac{\sum I(A_i(u), A_j(u))}{n} \leq 1\,.$$

This inequality is fulfilled when

$$0 \leq A_i(u) + A_j(u) - I(A_i(u), A_j(u)) \leq 1$$

for any $u \in X$, which is in turn equivalent with

$$0 \leq x + y - I(x, y) \leq 1$$

for any $(x, y) \in [0, 1]^2$.

The Bell-type inequalities for commutative conjunctors are necessary and sufficient conditions for the corresponding Bell-type inequalities for fuzzy probabilities to hold for any dimension n. In this way, we can write the Bell-type inequalities as follows:

$$\begin{aligned} &I_2^1 : T_{\mathbf{L}} \leq I \leq T_{\mathbf{M}} \\ &I_3^2 : 0 \leq x - I(x, y) - I(x, z) + I(y, z) \\ &I_3^3 : x + y + z - I(x, y) - I(x, z) - I(y, z) \leq 1 \end{aligned}$$

for any $x, y, z \in [0, 1]$.

Remark that the double Bell-type inequality I_2^1 is fulfilled for any commutative conjunctor which is bounded from below by $T_{\mathbf{L}}$, in particular for any commutative quasi-copula. Inequality I_3^2 is fulfilled for any commutative quasi-copula, while inequality I_3^3 holds only for certain t-norms [4].

6 Meta-theorems

In this section we present several theorems which state that certain inequalities satisfied by cardinalities of ordinary sets, are preserved under fuzzification.

Theorem 2 *Consider a finite universe X with dimension n. Let a, b and u denote the fuzzy cardinalities of respectively A, B and $A \cap B$. Consider an inequality of the form*

$$\mathcal{H}(a, b, u, n) \geq 0\,, \tag{1}$$

where $\mathcal{H}$ denotes a continuous function which is homogeneous in its arguments. Then, if this inequality holds for all n and all crisp sets A and B, the same inequality also holds for all n and all fuzzy sets A and B if I_2^1 is satisfied.

From the proof of Theorem 2, which will be published elsewhere, we obtain

Theorem 3 *Under the assumptions of Theorem 2: if (1) holds for all n and all crisp sets A and B, and if $\mathcal{H}$ does not depend explicitly upon n, then (1) also holds for all n and all fuzzy sets A and B.*

We can use Theorem 3 to conclude that any inequality of the following form :

$$a_1 x + a_2 y + b_{12} I(x, y) \geq 0$$

which holds for the extreme cases $(x, y) \in \{0, 1\}^2$, als holds for any $(x, y) \in [0, 1]^2$ and any commutative conjunctor I. Furthermore, we can prove the following theorem.

Theorem 4 *Consider an inequality of the following form which holds for the extreme cases $(x, y) \in \{0, 1\}^2$:*

$$a_1 x + a_2 y + b_{12} I(x, y) + c \geq 0\,.$$

Then, only in the case:

$$a_1 < 0 \ \wedge \ a_2 < 0 \ \wedge \ b_{12} > 0\,,$$

the condition $T_{\mathbf{L}} \leq I$ is explicitly needed in order to guarantee the above inequality.

We can repeat the same reasoning for three fuzzy sets.

Theorem 5 *Consider a finite universe X with dimension n. Let a, b, c, u, v and w denote the fuzzy cardinalities of respectively $A, B, C, A \cap B, A \cap C$ and $B \cap C$. Consider an inequality of the form*

$$\mathcal{H}(a, b, c, u, v, w, n) \geq 0\,, \tag{2}$$

where $\mathcal{H}$ denotes a continuous function which is homogeneous in its arguments. Then, if this inequality holds for all n and all crisp sets A, B and C, the same inequality also holds for all n and all fuzzy sets A, B and C if I_2^1, I_3^2 and I_3^3 are satisfied.

Again, from the proof of Theorem 5, which will also be published elsewhere, we obtain

Theorem 6 *Under the assumptions of Theorem 5: if (2) holds for all n and all crisp sets A, B and C, and if $\mathcal{H}$ does not depend explicitly upon n, then (2) also holds for all n and all fuzzy sets A, B and C if I_2^1 and I_3^2 are satisfied.*

We can use Theorem 6 to conclude that any inequality of the form:

$$\begin{aligned} & a_1 x + a_2 y + a_3 z \\ & + b_{12} I(x, y) + b_{13} I(x, z) + b_{23} I(y, z) \geq 0 \end{aligned}$$

which holds for the extreme cases $(x, y, z) \in \{0, 1\}^3$, holds for any $(x, y, z) \in [0, 1]^3$ and any commutative conjunctor I that satisfies inequalities I_2^1 and I_3^2. Furthermore, we can prove the following theorem

Theorem 7 *Consider an inequality of the following form which holds for the extreme cases $(x, y, z) \in \{0, 1\}^3$:*

$$\begin{aligned} & a_1 x + a_2 y + a_3 z \\ & + b_{12} I(x, y) + b_{13} I(x, z) + b_{23} I(y, z) + c \geq 0\,. \end{aligned}$$

Then, only in the case:

$$\begin{aligned} & (a_1 < 0 \ \wedge \ a_2 < 0 \ \wedge \ a_3 < 0) \\ \wedge \ & (b_{12} > 0 \ \wedge \ b_{13} > 0 \ \wedge \ b_{23} > 0) \\ \wedge \ & [\,(-a_1 \geq b_{12} \ \wedge \ -a_2 \geq b_{12} \ \wedge \ -a_3 > c - b_{12}) \\ & \vee (-a_1 \geq b_{13} \ \wedge \ -a_2 > c - b_{13} \ \wedge \ -a_3 \geq b_{13}) \\ & \vee (-a_1 > c - b_{23} \ \wedge \ -a_2 \geq b_{23} \ \wedge \ -a_3 \geq b_{23})]\,, \end{aligned}$$

inequality I_3^3 is explicitly needed in order to guarantee the above inequality.

7 A parametrized family of similarity measures

The similarity measures considered so far do not cover the wide variety of comparison coefficients encountered in practice. Also, it may be useful to have at ones disposal parametrized families of similarity measures, so that one can progress continuously between two specific measures and thereby investigate the variation of certain properties. Previously we have introduced a family of similarity measures for crisp sets [3] in the form of a rational expression solely based on the cardinalities of the sets involved:

$$S(A,B) = \frac{x(\Delta_{A,B}) + y\delta_{A,B} + z\nu_{A,B}}{x'(\Delta_{A,B}) + y\delta_{A,B} + z\nu_{A,B}}, \qquad (3)$$

in which $\Delta_{A,B} = |A\Delta B|$, $\delta_{A,B} = |A \cap B|$ and $\nu_{A,B} = |(A \cup B)^c|$. The parameters x, y, z and x' are positive and real. In particular, for $x = z = 0$, $x' = y = 1$, we obtain Jaccard's coefficient. The choice $x = 0$, $x' = y = z = 1$, yields the simple matching coefficient, and for $x = z = 0$, $x' = 1$ and $y = 2$ we retrieve the Dice coefficient.

In order to use our meta-theorems we have to rewrite the expression (3) only in terms of $|A|, |B|$ and $|A \cap B|$. The resulting expressions are then fuzzified leading to a family of fuzzy similarity measures.

Next we want to investigate what are the conditions to be imposed upon the parameters x, y, z and x' of family (3), in order to obtain respectively $T_{\mathbf{L}}$-transitive and $T_{\mathbf{P}}$-transitive similarity measures for fuzzy sets. The $T_{\mathbf{L}}$-transitivity condition for (fuzzy) similarity measures S is given by

$$1 + S(A,C) - S(A,B) - S(B,C) \geq 0\,. \qquad (4)$$

Note that the $T_{\mathbf{P}}$-transitivity condition for (fuzzy) similarity measures S is given by

$$S(A,C) - S(A,B)\,S(B,C) \geq 0\,. \qquad (5)$$

Theorem 8 *The $T_{\mathbf{L}}$-transitive members of the family of fuzzy similarity measures are for any commutative quasi-copula characterized by:*

$$(x' = x) \ \vee \ (x' > x \wedge x' \geq \max(y, z))\,. \qquad (6)$$

First, we have proven that (4) holds identically for all n and all crisp sets A, B, C, if and only if the parameters satisfy (6). For fuzzy sets A, B, C, the left-hand side of (4) is a homogeneous function of n, a, b, c, u, v, w. If all Bell-type inequalities are fulfilled, we can use Theorem 5 to conclude that inequality (4) also holds for all fuzzy sets A, B, C under the same parameter conditions (6). When $z = 0$, the homogeneous function on the left-hand side of (4) is then independent of n and using Theorem 6 we can conclude that inequality (4) also holds for all fuzzy sets A, B, C under the same parameter conditions (6), even if I_3^3 is not fulfilled. For the remaining case where the quasi-copula does not satisfy inequality I_3^3 and the parameter z is not zero, we have established a direct algebraic proof. Details of the complete proof will be published elsewhere.

Theorem 9 *The $T_{\mathbf{P}}$-transitive members of the family of fuzzy similarity measures are for any commutative quasi-copula characterized by:*

$$(x' = x) \ \vee \ (x' > x \wedge x\,x' \geq \max(y^2, z^2))\,. \qquad (7)$$

8 Conclusions

In this paper, we have proposed several meta-theorems, which find their application in proving transitivity properties of a parametrized family of fuzzy similarity measures.

Acknowledgements

This work is supported in part by the Bilateral Scientific and Technological Cooperation Flanders-Hungary BIL00/51.

References

[1] C. Genest, L. Molina, L. Lallena and C. Sempi, A characterization of quasi-copulas, *Journal of Multivariate Analysis* **69** (1999) 193–205.

[2] P. Jaccard, Nouvelles recherches sur la distribution florale, *Bulletin de la Sociétée Vaudoise des Sciences Naturelles* **44** (1908) 223–270.

[3] S. Janssens, B. De Baets and H. De Meyer, On the transitivity of a parametrized family of similarity measures, in: *Proc. EUROFUSE Workshop on Information Systems, Varenna, Italy* (2002) pp. 239–243.

[4] S. Janssens, B. De Baets and H. De Meyer, Inequalities in fuzzy probability calculus, in: *Proc. IFSA Congress, Istanbul, Turkey* (2003,) to appear.

[5] I. Pitowsky, *Quantum Probability – Quantum Logic*, Lecture Notes in Physics, Vol. 321 (Springer–Verlag, 1989).

[6] R. Sokal and C. Michener, A statistical method for evaluating systematic relationships, *Univ. Kansas Sci. Bull.* **38** (1958) 1409–1438.