On the approximation of a membership function by empirical quantile functions

Published Version:
Maria Letizia Guerra, Laerte Sorini, Luciano Stefanini (2020). On the approximation of a membership function by empirical quantile functions. INTERNATIONAL JOURNAL OF APPROXIMATE REASONING, 124(September), 133-146 [10.1016/j.ijar.2020.06.012].

Availability:
This version is available at: https://hdl.handle.net/11585/765625 since: 2020-09-16

Published:

DOI: http://doi.org/10.1016/j.ijar.2020.06.012

On the approximation of a membership function by empirical quantile functions

Maria Letizia Guerra

Department of Statistical Sciences "Paolo Fortunati" - University of Bologna, Italy

(Corresponding Author: *mletizia.guerra@unibo.it*)

Laerte Sorini, Luciano Stefanini

Department of Economics, Society, Politics - University of Urbino "Carlo Bo", Italy

(*laerte.sorini@uniurb.it, luciano.stefanini@uniurb.it*)

Abstract

The Average Cumulative representation of fuzzy intervals is connected with the possibility theory in the sense that the possibility and necessity functions are substituted by a pair of non decreasing functions defined as the positive and negative variations in the Jordan decomposition of a membership function. In this paper we motivate the crucial role of ACF in determining the membership function from experimental data; some examples and simulations are shown to state the robustness of the proposed construction.

KEYWORDS: Possibility distribution, Fuzzy intervals, Quantiles, Average cumulative function (ACF)

1 Introduction

Fuzzy numbers and possibility theory have been initiated by Zadeh in [31], as mathematical tools to formulate and reason about uncertainty in complex decision making events. During the recent years, they are receiving an increasing attention in research, as they constitute a complete and useful setting for modelling and managing uncertainty, in particular when information is characterized by partial or incomplete knowledge and when sources of empirical data sets are heterogeneous, as it is shown by Dubois in [8] and investigated in [16] concerning its potentialities arising in applications.

Following a similar approach, we introduced in [29] the Average Cumulative Function (ACF) representation for fuzzy intervals focusing on the properties it shares with the Cumulative Distribution Function (CDF) for probability distributions. We also showed that the ACF can be uniquely defined for any fuzzy interval and that any alfa-cut of a fuzzy interval can be directly deduced from the ACF. Regarding applications, ACF amounts to be a powerful instrument in order to deduce the membership function from the any kind of time series, a topic that has just been a matter of interest in [16] and [6]. An exhaustive overview of methods for building possibility distributions is in [17] where qualitative and quantitative possibility distributions are explored within order or similarity-based statistical methods. A wide scenario of real life decisions arising in company's management and based on possibility theory is analyzed in [3].

The possible connections between possibility and probability are studied in many contributions (see, e.g., [3], [4], [10], [14], [21], [22], [24], [25]); in [13] the problem of mutually transforming possibility measures into probability measures is handled with a deep attention to the philosophical nature of the two approaches that produces not equivalent representations of uncertainty. An additional class of transformations is extensively analyzed in [20] where the Arising Accumulation Transformation is applied to some decision making problems.

The construction of a possibility distribution when the probabilities are unknown and a data sample represented by a histogram is available is proposed in [22]; in [28] possibility distributions are deduced as families of upper and lower bounded probabilities identified by an informational distance function.

In this paper, we apply the properties of AC functions to (a) generating samples from fuzzy intervals (adopting the *insufficient reason principle*), and (b) estimating a membership function from empirical data.

The organization of the paper is based on four sections; after the Introduction, section 2 details the main properties of ACF. In section 3, subdivided into two subsections, examines simple ways to generate samples from possibility distributions associated to fuzzy intervals. Then (subsection 3.2), a general procedure is suggested to obtain the AC function from empirical data and to estimate a corresponding membership function under suitable assumptions on the location of its core; some computational experiments are included. Conclusions and future research lines are summarized in section 4.

2 Average cumulative functions associated to a fuzzy interval

We denote by $\mathbb{R}_{\mathcal{F}}$ the space of real fuzzy intervals, with normal, upper semi-continuous and quasi concave membership function $u : \mathbb{R} \rightarrow [0, 1]$ of bounded support (see [2], [9]). A fuzzy interval $u \in \mathbb{R}_{\mathcal{F}}$ is defined in terms of its membership function $u : \mathbb{R} \longrightarrow [0, 1]$ of the form

$$u(x) = \begin{cases} 0 & \text{if} \quad x < a \\ u^L(x) & \text{if} \quad a \leq x < c \\ 1 & \text{if} \quad c \leq x \leq d \\ u^R(x) & \text{if} \quad d < x \leq b \\ 0 & \text{if} \quad x > b. \end{cases} \tag{1}$$

We suppose a compact support $[a, b]$ and a compact non-empty core $[c, d] \subset [a, b]$ where it holds that $a < c \leq d < b \in \mathbb{R}$; $u^L : [a, c] \longrightarrow [0, 1[$ is the left side of the fuzzy interval, defined as a non-decreasing right-continuous function, $u^L(x) > 0$ for $x \in]a, c]$, and $u^R : [d, b] \longrightarrow [0, 1]$ is the right side, defined as a non-increasing left-continuous function, $u^R(x) > 0$ for $x \in [d, b[$.

The two functions $u^L(x)$ and $u^R(x)$ can be extended to the real domain by setting

$$u^L_{ext}(x) = \begin{cases} 0 & if \quad x < a \\ u^L(x) & if \quad a \leq x < c \\ 1 & if \quad x \geq c \end{cases} \tag{2}$$

$$u_{ext}^{R}(x) = \begin{cases} 1 & if \quad x \leq d \\ u^{R}(x) & if \quad d < x \leq b \\ 0 & if \quad x > b. \end{cases} \tag{3}$$

An extensive literature is devoted to the description of a fuzzy interval $u \in \mathbb{R}_{\mathcal{F}}$ as a possibility distribution on the real numbers such that a pair of cumulative distribution functions, called the lower and the upper CDFs of u, respectively, become the extended left side function $u_{ext}^{L}(x)$ and the extended right side function $u_{ext}^{R}(x)$. In particular, instead of the possibility and necessity functions (introduced in [31], [10] and extended in [3] and [15], [16], [18]); the main summarized results introduced in [29] follow, starting with the pair of functions $F_u^R, F_u^L : \mathbb{R} \longrightarrow [0,1]$:

$$F_u^R(x) = 1 - u_{ext}^{R}(x) = \begin{cases} 0 & if \quad x \leq d \\ 1 - u^{R}(x) & if \quad d < x \leq b \\ 1 & if \quad x > b. \end{cases} \tag{4}$$

$$F_u^L(x) = u_{ext}^{L}(x). \tag{5}$$

The two functions F_u^L and F_u^R are non decreasing; F_u^L is right continuous while F_u^R is left continuous because of the upper semi-continuity of u.It also holds that:

$$u\,(x) = F_u^L(x) - F_u^R(x) \;\; \forall x \in \mathbb{R}. \tag{6}$$

When equation (6) is viewed as the Jordan decomposition of u, then F_u^L and F_u^R are the positive and the negative variations of u and given the total variation function $V_u(x)$, $x \in \mathbb{R}$,

$$V_u(x) = \begin{cases} 0 & \text{if} \quad x \le a \\ \sup\limits_{\mathbb{P}_{x,n}} \left\{ \sum\limits_{j=1}^{n} |u(t_j) - u(t_{j-1})| \,; t_j \in \mathbb{P}_{x,n} \right\} & \text{if} \quad a < x \le b \\ V_u(b) & \text{if} \quad x > b \end{cases}$$

we can write

$$F_u^L(x) = \frac{V_u(x) + u(x)}{2} \tag{7}$$

$$F_u^R(x) = \frac{V_u(x) - u(x)}{2}. \tag{8}$$

where, for $x \in]a, b]$, $\mathbb{P}_{x,n} = \{a = t_0 < t_1 < ... < t_n = x\}$ is a finite decomposition of $[a, x]$ with n subintervals and the sup(...) is performed over all $\mathbb{P}_{x,n}$ with arbitrary $n \in \mathbb{N}$ and arbitrary points t_j, $j = 0, 1, ..., n$.

Definition 1 *For a fixed value of $\lambda \in [0, 1]$, the λ-Average Cumulative Function (λ-ACF for short) of u is defined to be the following convex combination of F_u^L and F_u^R, for all $x \in \mathbb{R}$,*

$$\begin{aligned} F_u^{(\lambda)}(x) &= (1-\lambda)F_u^L(x) + \lambda F_u^R(x) \\ &= \begin{cases} 0 & \text{if} \quad x < a \\ (1-\lambda)u^L(x) & \text{if} \quad a \le x < c \\ 1-\lambda & \text{if} \quad c \le x \le d \\ 1-\lambda u^R(x) & \text{if} \quad d < x \le b \\ 1 & \text{if} \quad x > b. \end{cases} \end{aligned} \tag{9}$$

$F_u^{(\lambda)}$ is non-decreasing, right continuous on $]-\infty, d[$ and left continuous on

$]c,+\infty[$. *For the value* $\lambda = \frac{1}{2}$ *we denote* $F_u^{(\frac{1}{2})}(x)$ *simply by* $F_u(x)$.

Proposition 2 *For a given fuzzy interval* $u \in \mathbb{R}_{\mathcal{F}}$ *and a number* $\rho \in \mathbb{R}$, *the* λ-*ACF satisfies the translation property:*

$$\begin{aligned} F_{u+\rho}^{(\lambda)}(x) &= (1-\lambda)v_{ext}^L(x) + \lambda v_{ext}^R(x) \\ &= (1-\lambda)u_{ext}^L(x-\rho) + \lambda u_{ext}^R(x-\rho) \\ &= F_u^{(\lambda)}(x-\rho). \end{aligned} \tag{10}$$

where $v = u + \rho$ *is the translated fuzzy interval and* $v(x) = u(x-\rho)$ *is the membership function.*

An interesting relation holds:

Lemma 3 *Let* $u \in \mathbb{R}_{\mathcal{F}}$ *and let* $-u \in \mathbb{R}_{\mathcal{F}}$ *be its opposite interval; then, the following equality is true for all* $\lambda \in [0,1]$

$$F_u^{(\lambda)}(-x) + F_{-u}^{(1-\lambda)}(x) = 1, \text{ for all } x \in \mathbb{R}$$

where $F_{-u}^{(1-\lambda)}$ *is the* $(1-\lambda)$-*ACF of* $-u$.

For a given non-decreasing function $F : [a,b] \longrightarrow [0,1]$, the generalized inverse (also called the *quantile function* of F in probability theory, see, e.g. [19], when F is càdlàg) is defined to be the function $F^{-1} : [0,1] \longrightarrow [a,b]$ such that

$$F^{-1}(\alpha) = \inf\{x | F(x) \geq \alpha\} \text{ for all } \alpha \in]0,1] \text{ and } F^{-1}(0) = a. \tag{11}$$

The main theorem (proved in [29]) states that a partial càdlàg property of $F(x)$ is sufficient to determine the α-cuts $[u_\alpha^-, u_\alpha^+]$ of u for $\alpha \in]0,1]$. Recall that the fuzzy interval $-u$ has α-cuts given by $[-u_\alpha^+, -u_\alpha^-]$, so that, in particular, $u_\alpha^+ = -(-u)_\alpha^-$.

Theorem 4 *Let $u \in \mathbb{R}_\mathcal{F}$ and let $F_u^{(\lambda)}$, $F_{-u}^{(1-\lambda)}$ be the λ-ACF of u and the $(1-\lambda)$-ACF of $-u$, respectively, for any given value $\lambda \in]0,1[$. For all $\alpha \in]0,1]$, the α-cut $[u_\alpha^-, u_\alpha^+]$ of u is given by*

$$\begin{aligned} u_\alpha^-(\lambda) &= \inf\left\{x \in [a,c] | F_u^{(\lambda)}(x) \geq (1-\lambda)\alpha\right\} \qquad (12) \\ &= \left(F_u^{(\lambda)}|_*\right)^{-1}((1-\lambda)\alpha) \\ u_\alpha^+(\lambda) &= -(-u)_\alpha^- = -\inf\left\{x \in [-b,-d] | F_{-u}^{(1-\lambda)}(x) \geq \lambda\alpha\right\} \qquad (13) \\ &= -\left(F_{-u}^{(1-\lambda)}|_*\right)^{-1}(\lambda\alpha) \end{aligned}$$

where $\left(F_u^{(\lambda)}|_*\right)^{-1}$ and $\left(F_{-u}^{(1-\lambda)}|_*\right)^{-1}$ are the generalized inverses of the restrictions of $F_u^{(\lambda)}$ and $F_{-u}^{(1-\lambda)}$ to the subintervals $]-\infty, c]$ and $]-\infty, -d]$, respectively.

Assume for simplicity the membership function $u(x)$ (consequently $F(x)$) continuous.

In the particular case of $\lambda = \frac{1}{2}$, denoting $F_u = F_u^{(\frac{1}{2})}$ and $F_{-u} = F_{-u}^{(\frac{1}{2})}$, we have $F_{-u}(x) = 1 - F_u(-x)$ for all x, and we obtain, for $\alpha \in]0,1]$,

$$\begin{aligned} u_\alpha^- &= (F_u)^{-1}\left(\frac{\alpha}{2}\right) \\ u_\alpha^+ &= (F_u)^{-1}\left(1 - \frac{\alpha}{2}\right). \end{aligned} \tag{14}$$

Given any fixed value $\lambda \in \,]0,1[\,$, consider a non-decreasing function $F : \mathbb{R} \longrightarrow [0,1]$ satisfying the properties:

1) $a_F = \sup\{x | F(x) = 0\} \in \mathbb{R}$, $b_F = \inf\{x | F(x) = 1\} \in \mathbb{R}$ (clearly $a_F \leq b_F$);

2) $c_F = \inf\{x | F(x) \geq 1 - \lambda\} \in \mathbb{R}$, $d_F = \sup\{x | F(x) \leq 1 - \lambda\} \in \mathbb{R}$ (clearly $c_F \leq d_F$);

3) $a_F \leq c_F \leq d_F \leq b_F$ and F is right-continuous on $[a_F, c_F[$, left-continuous on $]d_F, b_F]$ and $F(x) = 1 - \lambda$ for all $x \in [c_F, d_F]$.

Then there exists a unique fuzzy interval $u_F \in \mathbb{R}_{\mathcal{F}}$ with λ-ACF, for $\lambda \in \,]0,1[$ given by F. Indeed, the membership function of u_F is given by (compare with Definition 1)

$$u_F(x) = \begin{cases} 0 & \text{if} \quad x < a_F \\ \frac{1}{1-\lambda} F(x) & \text{if} \quad a_F \leq x < c_F \\ 1 & \text{if} \quad c_F \leq x \leq d_F \\ \frac{1}{\lambda}(1 - F(x)) & \text{if} \quad d_F < x \leq b_F \\ 0 & \text{if} \quad x > b_F \end{cases} \tag{15}$$

and, from the assumptions 1), 2) and 3) on F, u_F is a fuzzy interval (the proof is immediate by directly verifying that $u_F \in \mathbb{R}_{\mathcal{F}}$).

We denote by $\mathbb{F}_\lambda(\mathbb{R})$ the family of all functions $F : \mathbb{R} \longrightarrow [0,1]$ satisfying properties 1)-2)-3).

In the particular case of $\lambda = \frac{1}{2}$, the family $\mathbb{F}_\lambda(\mathbb{R})$ will be simply denoted by $\mathbb{F}(\mathbb{R})$; the $\frac{1}{2}-$AC function of $u \in \mathbb{R}_\mathcal{F}$ is a non-decreasing function $F_u : \mathbb{R} \longrightarrow [0,1]$ such that $F_u(x) = \frac{1}{2}u^L(x)$ on $[a,c[$, $F_u(x) = 1 - \frac{1}{2}u^R(x)$ on $]d,b]$ and $F_u(x) = \frac{1}{2}$ on the core $[c,d]$ of u, i.e., $u(x) = 2\min\{F(x), 1-F(x)\}$.

In [29] we have formalized a bijection between the space $\mathbb{R}_\mathcal{F}$ of real fuzzy intervals and the spaces $\mathbb{F}_\lambda(\mathbb{R})$ of non-decreasing functions $F : \mathbb{R} \longrightarrow [0,1]$ such that, for a fixed $\lambda \in]0,1[$, $F(x)$ is right-continuous on $[a_F, c_F]$ and left-continuous on $[d_F, b_F]$ and $F(x) = 1-\lambda$ on $[c_F, d_F]$ where

$$\begin{cases} a_F = \sup\{x|F(x) = 0\} \in \mathbb{R}, \\ b_F = \inf\{x|F(x) = 1\} \in \mathbb{R}, \\ c_F = \inf\{x|F(x) \geq 1-\lambda\} \in \mathbb{R}, \\ d_F = \sup\{x|F(x) \leq 1-\lambda\} \in \mathbb{R}. \end{cases} \tag{16}$$

Each $F \in \mathbb{F}_\lambda(\mathbb{R})$ is called a λ-AC function.

For a given $u \in \mathbb{R}_\mathcal{F}$ with membership

$$u(x) = \begin{cases} 0 & \text{if} \quad x < a \text{ or } x > b \\ u^L(x) & \text{if} \quad a \leq x < c \\ 1 & \text{if} \quad c \leq x \leq d \\ u^R(x) & \text{if} \quad d < x \leq b \end{cases} \tag{17}$$

the λ-AC function $F_u^{(\lambda)} \in \mathbb{F}_\lambda(\mathbb{R})$ corresponding to u is

$$F_u^{(\lambda)}(x) = \begin{cases} 0 & \text{if} \quad x < a \\ (1-\lambda)u^L(x) & \text{if} \quad a \le x < c \\ 1-\lambda & \text{if} \quad c \le x \le d \\ 1-\lambda u^R(x) & \text{if} \quad d < x \le b \\ 1 & \text{if} \quad x > b \end{cases} \tag{18}$$

Vice versa, if $F \in \mathbb{F}_\lambda(\mathbb{R})$ is given and $\lambda \in]0,1[$ is fixed, with $a_F \le c_F \le d_F \le b_F$, the corresponding $u \in \mathbb{R}_\mathcal{F}$ is

$$u_{F,\lambda}(x) = \begin{cases} 0 & \text{if} \quad x < a_F \text{ or } x > b_F \\ \frac{1}{1-\lambda}F(x) & \text{if} \quad a_F \le x < c_F \\ 1-\lambda & \text{if} \quad c_F \le x \le d_F \\ \frac{1}{\lambda}(1-F(x)) & \text{if} \quad d_F < x \le b_F \end{cases}. \tag{19}$$

Remark that an equivalent compact for of (19) is the following (the explicit computation of $a_F, ..., b_F$ is not required)

$$u_{F,\lambda}(x) = \min\left(\frac{1}{1-\lambda}F(x), \frac{1}{\lambda}(1-F(x))\right), \text{ for all } x. \tag{20}$$

For a real random variable X with CDF F_X, a *quantile* of order $p \in]0,1[$ is a real value κ_p where F_X crosses or jumps over p, i.e., such that

$$\lim_{x \uparrow \kappa_p} F_X(x) \le p \quad \text{and} \quad F_X(\kappa_p) \ge p.$$

Given a simple sample $x_1, x_2, ..., x_N$ from a real random variable X, the (empirical) p-quantile $\widehat{\kappa}_p(N)$ is obtained by minimizing, with respect to κ, the function

$$S_{p,N}(\kappa) = (1-p)\sum_{\substack{i=1\\ x_i<\kappa}}^{N}(\kappa - x_i) + p\sum_{\substack{i=1\\ x_i>\kappa}}^{N}(x_i - \kappa); \tag{21}$$

furthermore,

$$\widehat{\kappa}_p(N) = \arg\min_{\kappa} S_{p,N}(\kappa)$$

is an unbiased estimate of κ_p.

According to Theorem 4 it is easy to deduce the following proposition.

Proposition 5 *Let $u \in \mathbb{R}_{\mathcal{F}}$ have continuous membership function (1); let $F_u(x)$, $x \in \mathbb{R}$ be its $\frac{1}{2}$-ACF. Then for all $\alpha \in]0,1]$, the α-cuts $[u_\alpha^-, u_\alpha^+]$ of u are such that u_α^- is the $\frac{\alpha}{2}$-quantile of $F_u(x)$ and u_α^+ is the $\frac{\alpha}{2}$-quantile of $F_{-u}(x)$.*

Proof. We have

$$F_u(x) = \frac{1}{2}u_{ext}^L(x) + \frac{1}{2}\left(1 - u_{ext}^R(x)\right)$$

and, from equality $F_{-u}(x) = 1 - F_u(-x)$,

$$F_{-u}(x) = \frac{1}{2}u_{ext}^R(-x) + \frac{1}{2}\left(1 - u_{ext}^L(-x)\right);$$

From the continuity of $u_{ext}^L(x)$ and $u_{ext}^R(x)$ it follows that both F_u and F_{-u} are continuous and their inverses are quantile functions. ■

Let us consider the case where the membership function is given at a finite number of points, i.e. suppose that the fuzzy number $u \in \mathbb{R}_{\mathcal{F}}$ is "measured" at N (independent) observations $(t_i, u(t_i))$; this is equivalent to consider a set of independent variables $X_1, X_2, ..., X_N$ identically distributed on the support $[a, b]$ and to extract a simple sample of N distinct values t_i from each X_i, $i = 1, 2, ..., N$.

Consider the decomposition $\mathbb{P}_N = \{x_1 < x_2 < ... < x_N\}$ of the support $[a, b]$, obtained by ordering the t_i such that $t_{(1)} < t_{(2)} ... < t_{(N)}$ and defining $x_i = t_{(i)}$ for $i = 1, 2..., N$. The corresponding *empirical AC function* is

$$\widehat{F}_{\mathbb{P}_N}(x) = \frac{1}{N} \sum_{i=1}^{N} \widehat{I}(x \geq x_i) \tag{22}$$

where

$$\widehat{I}(x \geq x_i) = \begin{cases} 1 & \text{if} \quad x \geq x_i \\ 0 & \text{if} \quad x < x_i \end{cases} . \tag{23}$$

For $\alpha \in]0, 1]$, the α-cuts of u can be estimated by computing the empirical $\frac{\alpha}{2}$-quantile of the sample data $\{x_i | i : 1, ..., N\}$ and the empirical $\frac{\alpha}{2}$-quantile of the data $\{-x_i | i : 1, ..., N\}$. To this issue, we have to minimize the two empirical functions, as in eq. (21). The obtained values

$$\kappa_\alpha^-(N) = \arg\min_\kappa S_{\frac{\alpha}{2}, N}(\kappa) \tag{24}$$

$$\kappa_\alpha^+(N) = \arg\min_\kappa S_{1-\frac{\alpha}{2}, N}(\kappa) \tag{25}$$

give an estimate $[\kappa_\alpha^-(N), \kappa_\alpha^+(N)]$ of the α-cut $[u_\alpha^-, u_\alpha^+]$ of u and are obtained without computing directly the (empirical) AC function from the data.

The Glivenko-Cantelli theorem ensures the convergence, for $N \to \infty$, of interval $[\kappa_\alpha^-(N), \kappa_\alpha^+(N)]$ to the $\alpha-$cut $[u_\alpha^-, u_\alpha^+]$.

In practical applications the assumption of continuity of F_u is not restrictive and it is a standard approach in statistics to estimate the quantiles and the cumulative distribution function from empirical data.

It is also interesting to observe that, for fuzzy numbers, the same empirical AC function can be associated to possibly several membership functions, according to the value $\lambda \in]0, 1[$ appearing in equations (19), (20). For example, consider the linear ACF $F(x) = \frac{x-a}{b-a}$ if $x \in [a, b]$, $F(x) = 0$ if $x < a$, $F(x) = 1$ if $x > b$ (we can consider F as a "uniform" ACF). If we take $\lambda = \frac{1}{2}$, then the associated λ-membership function $u_{0.5}(x)$ gives the linear-shaped triangular fuzzy number, usually represented by the triplet $(a, \frac{a+b}{2}, b)$, with core at $c_{0.5} = \frac{a+b}{2}$ and support $[a, b]$; if we take, e.g., $\lambda = \frac{3}{4}$, then the associated $u_{0.75}(x)$ gives the triangular fuzzy number $(a, \frac{a+b}{2}, b)$, with the same support but core at $c_{0.75} = 0.75a + 0.25b$. In general, then, before associating the membership to a uniform AC function, we have to chose the value of λ; this is equivalent to chose the position of the core $c_\lambda = \lambda a + (1 - \lambda)b$ between a and b.

A similar reasoning can be applied to the case of a non-linear AC function F on $[a, b]$ by choosing the position of the core $c \in]a, b[$ and by computing the membership function according to (20) with the value $\lambda = F(c)$.

3 Membership functions from empirical quantiles

Two questions are relevant to managing a membership function u empirically: (a) how to generate (independent) observations from u, and (b) how to estimate u from data.

3.1 Generating data from a given AC function

A natural way to generate observations from a fuzzy interval $u \in \mathbb{R}_{\mathcal{F}}$ with α-cuts $[u_\alpha^-, u_\alpha^+]$, $\alpha \in [0,1]$, is based on the application of the so called and well known *insufficient reason principle*, (*IRP* for short, see Dubois-Prade [11] and [12]): select a uniform sample of m α-cuts, corresponding to levels $\alpha_1 < \alpha_2 < ... < \alpha_m \in [0,1]$ (eventually $\alpha_k = \frac{k-1}{m-1}$, $k = 1, 2, ..., m$ equispaced between 0 and 1) and choose uniformly from each $[u]_{\alpha_k} = [u_{\alpha_k}^-, u_{\alpha_k}^+]$ an equal number of values $t_{k,j} \in [u]_{\alpha_k}$, $j = 1, 2, ..., n \geq 2$, so that a matrix of $N = mn$ observations $\{t_{k,j} | k = 1, ..., m; j = 1, ..., n\}$ is obtained (eventually values $t_{k,j} = u_{\alpha_k}^- + \frac{j-1}{n-1}(u_{\alpha_k}^+ - u_{\alpha_k}^-)$, equally spaced on each $[u]_{\alpha_k}$).

Suppose for simplicity that the values $t_{k,j}$ are all distinct. Let us organize the matrix $t_{k,j}$ (for example by rows) into a vector $\widehat{t}_i$, $i = 1, 2, ..., N$ and denote by $\widehat{t}_{(i)}$ the ascending ordered element in i-th position with $\widehat{t}_{(i+1)} > \widehat{t}_{(i)}$, $i = 1, ..., N-1$.

The corresponding empirical AC function of the data set is obtained by setting $x_i = \widehat{t}_{(i)}$, $i = 1, ..., N$ and, for $x \in \mathbb{R}$,

$$\widehat{F}_{\{t_{k,j}\}}(x) = \frac{1}{N} \sum_{i=1}^{N} \widehat{I}(x \geq x_i). \tag{26}$$

We have the following obvious property.

Proposition 6 *The empirical AC function $\widehat{F}_{\{t_{k,j}\}}(x)$ is such that*

$$\lim_{n,m \to \infty} \sup \left| \widehat{F}_{\{t_{k,j}\}}(x) - F_u(x) \right| = 0$$

where $F_u(x)$ is the AC function of u.

Proof. It is sufficient to apply Proposition 25 in [29]. ■

By the uniform construction of $\{t_{k,j}\}$ from the m α-cuts $[u]_{\alpha_k}$ we can see that the probability of $x_i \in [u]_{\alpha_k}$ does not depend on n, the number of elements taken from each α-cut. Indeed, letting $[a,b]$ the support and $[c,d]$ the core of u, denote the lengths of the intervals $\delta_1 = u^+_{\alpha_1} - u^-_{\alpha_1} = b - a$, $\delta_k = u^+_{\alpha_k} - u^-_{\alpha_k}$ and $\delta_m = u^+_{\alpha_m} - u^-_{\alpha_m} = d - c$, the nesting property of the α-cuts ensures that $x_i \in [u]_{\alpha_{k+1}} \Longrightarrow x_i \in [u]_{\alpha_k}$ and all x_i belong to $[u]_{\alpha_1}$ (the support); on the other hand, by uniformity, the number of elements x_i in $[u]_{\alpha_j}$ and not in $[u]_{\alpha_{j+1}}$, for fixed j, is the proportion $\frac{\delta_j - \delta_{j+1}}{\delta_j}$ of the difference of lengths of the two intervals. We then have

$$\begin{aligned}
\Pr\{x_i \in [u]_{\alpha_k}\} &= \frac{1}{N}\left(n(m-k+1) + n\sum_{j=1}^{k-1} \frac{\delta_j - \delta_{j+1}}{\delta_j} \right) \\
&= \frac{1}{m}\left(m-k+1+k-1+\sum_{j=1}^{k-1} \frac{\delta_{j+1}}{\delta_j} \right) = \frac{1}{m}\left(m + \sum_{j=1}^{k-1} \frac{\delta_{j+1}}{\delta_j} \right) \\
&= 1 - \frac{1}{m}\sum_{j=1}^{k-1} \frac{\delta_{j+1}}{\delta_j}.
\end{aligned}$$

The last expression does not depend on n but only on the number m of α-cuts and on the lengths of the intervals $[u]_{\alpha_k} = [u^-_{\alpha_k}, u^+_{\alpha_k}]$.

It is also possible to see that the statistical *modal value* belongs (or coincides with it, if $c = d$) to the core of u; furthermore, the modal value and the *median* of the data set, under the IRP condition, will also coincide.

Assume that the core $[c, d]$ is not a singleton, i.e. that $\delta_m > 0$; eventually, if $c = d$ the n data values corresponding to the core are generated uniformly from a neighbor $[c-\varepsilon, c+\varepsilon]$ with a small $\varepsilon > 0$ and, in this case, $\delta_m = 2\varepsilon > 0$. Lets now consider a histogram with p uniform classes $C_1, C_2, ..., C_p$, obtained from the data set $\{x_i | i = 1, ..., N\}$; each class has length $\frac{b-a}{p}$ and is an interval $C_l = \left[a + \frac{l-1}{p}(b-a), a + \frac{l}{p}(b-a)\right]$, $l = 1, 2, ..., p$. It is not difficult to determine the composition of each class C_l, i.e. to compute the probability that $x_i \in C_l$. Clearly, it will depend on the position of the class with respect to each α-cut $[u]_{\alpha_k}$, $k = 1, ..., m$. Denote $I_{l,k} = C_l \cap [u]_{\alpha_k}$ and let $\widehat{\delta}_{l,k} = length(I_{l,k}) \le \frac{b-a}{p}$ (set $\widehat{\delta}_{l,k} = 0$ if $I_{l,k}$ is empty); then, the proportion of elements $t_{k,j}$, $j = 1, ..., n$, from the interval $[u]_{\alpha_k}$ that belong to C_l will be the ratio $\frac{\widehat{\delta}_{l,k}}{\delta_k}$ and we have

$$\Pr\{x_i \in C_l\} = \frac{n}{N} \sum_{k=1}^{m} \frac{\widehat{\delta}_{l,k}}{\delta_k} = \frac{1}{m} \sum_{k=1}^{m} \frac{\widehat{\delta}_{l,k}}{\delta_k} \le \frac{b-a}{p} \frac{1}{m} \sum_{\substack{k=1 \\ I_{l,k} \neq \emptyset}}^{m} \frac{1}{\delta_k}$$

i.e., each class has a number of elements proportional to the average ratio of the length of its intersection with each α-cut to the length of the α-cut itself. As a consequence, the highest probability corresponds to a class which intersects the core of u, denote it by $C_{\widehat{l}}$; by the nesting property of the α-cuts and by the uniformity of the $t_{k,j}$ with respect to each $[u]_{\alpha_k}$, it also follows that the proportion of elements x_i located on the left (before) the modal class is the same as after the modal class, with the consequence that the modal value coincides with the median of the data set $\{x_i | i = 1, ..., N\}$.

In the special case of $n = 2$, i.e., to each α-cut two elements $t_{i,1}, t_{i,2}$ are generated for all $i = 1, ..., m$, we can deduce a procedure to generate a (random) sample from a fuzzy number u with a singleton core $\{c\}$, support $[a,b]$ and *continuous* AC function F_u: let α_i be m (independent) random numbers generated between 0 and 1, determine $t_{k,1} = F_u^{-1}(\frac{\alpha_k}{2})$, $t_{k,2} = F_u^{-1}(1-\frac{\alpha_k}{2})$ and set $x_{2k-1} = t_{k,1}$, $x_{2k} = t_{k,2}$ for $k = 1,...,m$ (a set of $N = 2m$ values), where $F^{-1}(\alpha) = \inf\{a|F(x) \geq \alpha\}$ is the generalized inverse of F.

If we chose $n = 1$, i.e., to each α-cut only one element x_i is selected for all $i = 1,...,m$, we can generate a (random) sample of values $x_1, x_2, ...x_m \in [a,b]$ from u: let α_i be (independent) random numbers generated between 0 and 1 and determine x_i as

$$x_i = \begin{cases} F_u^{-1}(\frac{\alpha_i}{2}) & \text{with probability } \frac{1}{2} \\ F_u^{-1}(1-\frac{\alpha_i}{2}) & \text{with probability } \frac{1}{2} \end{cases}, \; i = 1,...,m. \tag{27}$$

A Glivenko-Cantelli-like theorem (see [29]) ensures that, for $m \to \infty$, the empirical cumulative distribution function of the data set $\{x_i|i = 1,...,m\}$, given in (26) with $n = 1$, converges almost surely and uniformly to F_u.

In equation (27), the quantile function is used under the assumption that the core of u corresponds to $\lambda = \frac{1}{2}$, i.e., if F_u is the $\frac{1}{2}$-ACF of u. A similar result is also valid if F_u is assumed to be the λ-ACF of u with $\lambda \in]0,1[$; in this case, we substitute equation (27) with the following

$$x_i = \begin{cases} F_u^{-1}(\lambda\alpha_i) & \text{with probability } \frac{1}{2} \\ F_u^{-1}(1-\lambda\alpha_i) & \text{with probability } \frac{1}{2} \end{cases}, \; i = 1,...,m. \tag{28}$$

Example 1. Consider two fuzzy numbers X and Y (X with linear shaped membership function), as in Figure 1. We have generated, according to the rule expressed in (27), i.e., assuming $\lambda = 0.5$, two samples of size

$m = 100$ and $m = 500$; in this case we know exactly the position of the core of X and Y (1.0 and 2.0, respectively), so we do not face the problem with their location.

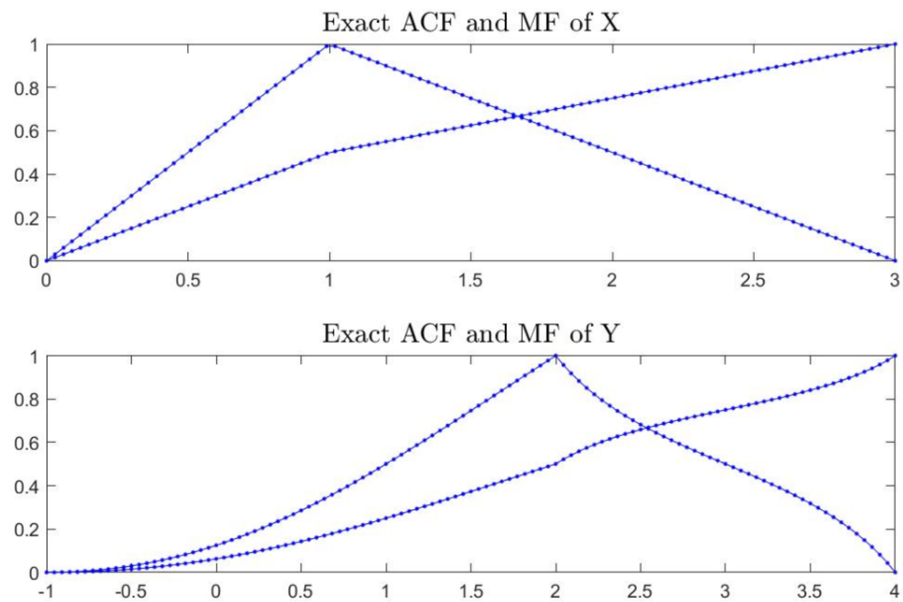

Figure 1: Membership function (MF) and $\frac{1}{2}$-Average Cumulative function (ACF) of fuzzy numbers X and Y as in Example 1.

In Figure 2, a sample of $m = 100$ data is extracted from the two fuzzy numbers and the membership function is reconstructed from the empirical AC function with $\lambda = \frac{1}{2}$. The resulting percentage relative error between exact and estimated AC functions is 0.40% for X and 0.59% for Y. Remark that in this case a small portion on left and right of the support of X and Y is not covered completely (in particular, the right side of X).

In Figure 3, the sample has $m = 500$ points and the AC function is estimated empirically. The resulting percentage relative error between exact

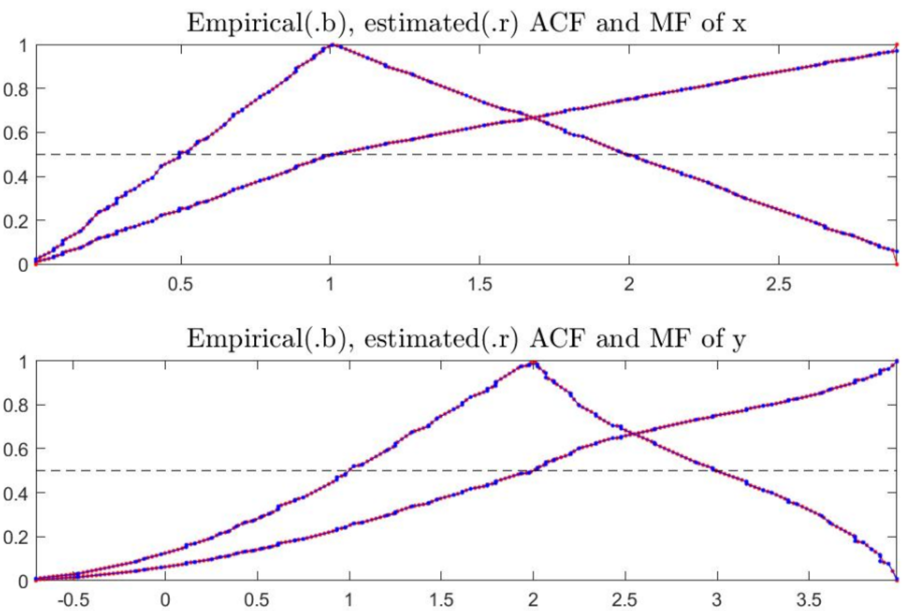

Figure 2: Empirical (m = 100) and estimated Membership function (MF) and Average Cumulative function (ACF) of fuzzy numbers X and Y, as in Example 2.

and estimated AC functions is 0.34% for X and 0.52% for Y. In this case, the covering of the supports of X and Y is more satisfactory than for the previous smaller sample.

3.2 Estimating a membership function from empirical data

As we have mentioned, the λ-AC function (ACF) of a fuzzy number $u \in \mathbb{R}_F$, in the continuous case, has the same properties of a cumulative distribution function (CDF) of a (real) random variable X, defined on the same (support) domain. This gives an interesting (formal) similarity between probability and possibility as, at least in principle, we can examine the same function

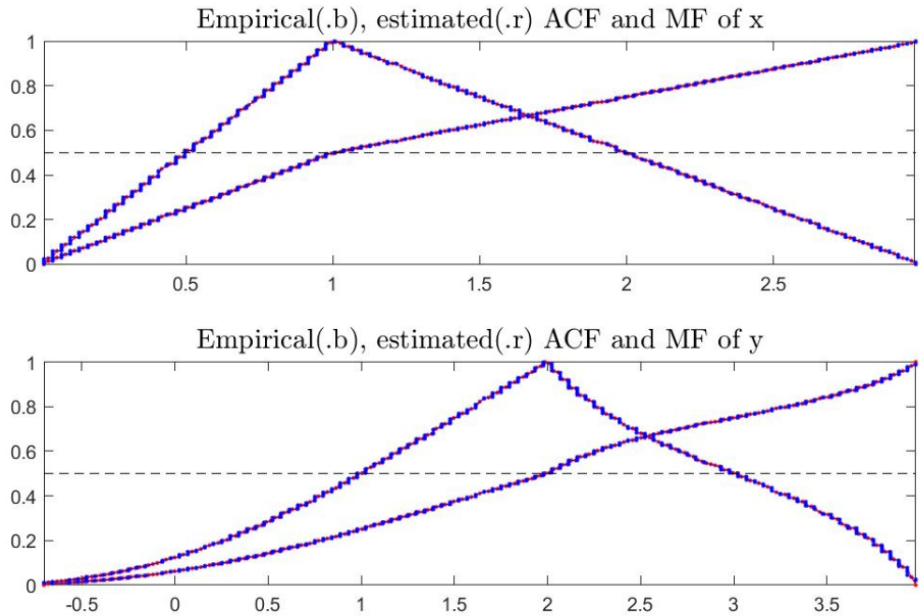

Figure 3: Empirical (m = 500) and estimated Membership function (MF) and Average Cumulative function (ACF) of fuzzy numbers X and Y, as in Example 2.

by combining the two settings, as we did in section 3.1, to simulate data from the possibility distribution. If we work with fuzzy numbers, we can perform all the operations in terms of the corresponding λ-AC functions (see also papers by G. De Cooman, in particular [7]).

Clearly, the information about how the sample is extracted, becomes crucial when considering the empirical F_u: the connection between F_u, interpreted as a quantile function, and the underlying fuzzy membership u corresponds to the value $\lambda = \frac{1}{2}$. On the other hand, starting with a given ACF, we can apply Theorem 4 to reconstruct the membership function u (or its α-cuts) only if the value of $\lambda \in]0, 1[$ is preliminarily decided; we know indeed that, in general, $\lambda = \frac{1}{2}$ corresponds to the core of u only when it is coincident with the empirical median.

In the cases where we do not have such precise information, we face a situation where other values of $\lambda \in]0, 1[$ can be used, according to equation (15), and we have to chose an appropriate value of λ, which corresponds to a particular choice of the position of the core of u. By choosing $\lambda = \frac{1}{2}$, we continue to assume that the core coincides with the median; on the other hand, if we assume that the core $c = c^*$ corresponds to a different value of $\lambda = \lambda^*$, the following equality holds: $F(c^*) = 1 - \lambda^*$, equivalently $c^* = F^{-1}(1 - \lambda^*)$ and $\lambda^* = 1 - F(c^*)$.

This is the rule (coherent with Theorem 4) to determine from the AC function either the location c^* of the core if $\lambda = \lambda^*$ is fixed, or the value of λ if the core $c = c^*$ is fixed.

We suggest the following procedure to obtain a membership function $\widehat{u}(x)$ from an empirical AC function $\widehat{F}(x)$.

Procedure eACMF: *Estimation of AC and membership functions from a data set.*

Step 1 (Initialization): Let $X = \{x_1, x_2, ..., x_m\}$ be the available sample of values with the corresponding relative frequencies $\{p_1, p_2, ..., p_m\}$ (eventually, $p_j = \frac{1}{m}$ for all j); let $a = \min\{x_j\} - \varepsilon$, $b = \max\{x_j\} + \varepsilon$, for some positive ε, be the support of $\widehat{u}(x)$.

Step 2: Construct the empirical AC function as $F_X(x) = \sum_{j=1}^{m} p_j \widehat{I}(x \geq x_j)$ where $\widehat{I}(x \geq x_j) = 1$ if $x \geq x_j$; $= 0$ otherwise.

Step 3: Determine the central value $c^* \in]a, b[$ to be considered as the core of the fuzzy number $\widehat{u}$, e.g., $c^* = median\{x_j\}$, or $c^* = mode\{x_j\}$ (if the data are uni-modal), or $c^* = mean\{x_j\}$.

Step 4: From the empirical AC function $F_X(x)$ compute $F_X(c^*)$, e.g. by interpolation, and set $\lambda^* = 1 - F_X(c^*)$; if $c^* = median\{x_j\}$, then set $\lambda^* = 0.5$; the empirical membership function of $\widehat{u}$ at the points $x_j \in X$ is $u_X(x_j) = \min\left\{\frac{1}{1-\lambda}F_X(x_j), \frac{1}{\lambda}\left(1 - F_X(x_j)\right)\right\}$.

Step 5: From the empirical AC function construct, by some approximation, an estimated AC function $\widehat{F}(x)$, $x \in [a, b]$.

Step 6: Use the estimated AC function $\widehat{F}(x)$ obtained in Step 5 to compute the membership function $\widehat{u}$ as $\widehat{u}(x) = \min\left\{\frac{1}{1-\lambda^*}\widehat{F}(x_j), \frac{1}{\lambda^*}\left(1 - \widehat{F}(x)\right)\right\}$, $x \in [a, b]$.

In the next two examples, we apply formula (27) to a pair of fuzzy sets X, Y by comparing the construction of the membership functions corresponding to three cases: 1) the core c coincides with the median of the distribution; 2) the core c is the mode (assuming uni-modality); 3) the core c is the mean of the data.

Example 2: A sample of $m = 250$ data for two fuzzy numbers X, Y are generated from the same AC functions as in Example 1, but this time the x_i and y_i are perturbed by adding a normal random number generated from $N(0, \sigma^2)$ with $\sigma = 0.3$. The obtained AC functions are pictured in Figure 4; remark that the median of X and Y are essentially preserved.

Considering that the AC functions are monotone, we have used the method in [5] to obtain their best (smooth) monotonic approximation.

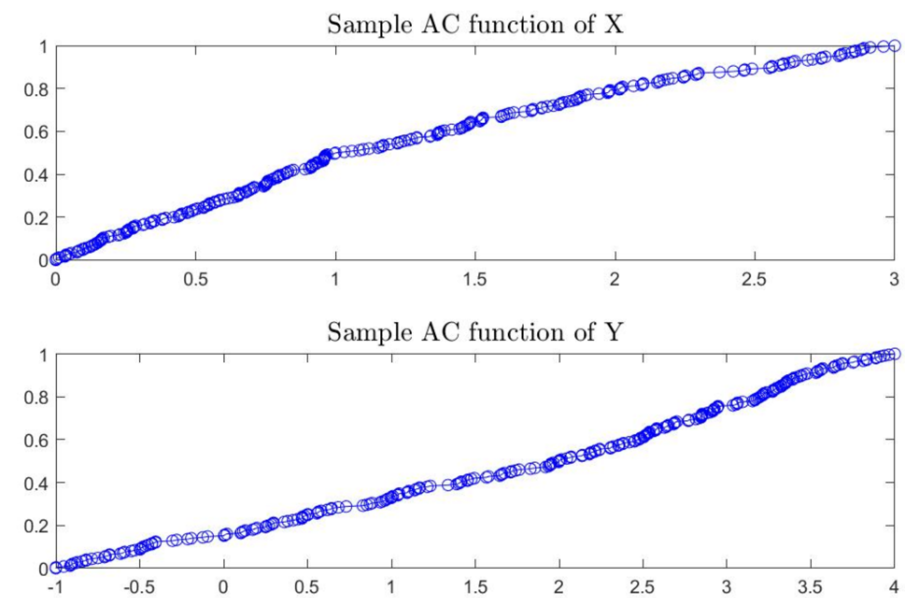

Figure 4: The ACFs of numbers X ad Y are represented, following hypothesis in Example 2.

From the data set, the cores c_X^* and c_Y^* are estimated from the empirical median, mode and mean of X and Y; they are reported in Table 1, with the corresponding values of λ^*, obtained from the empirical AC functions by linear interpolation.

Table 1: *Median, mode, mean and corresponding λ^* values for Example 2*

Data	*1) Median*	*2) Mode*	*3) Mean*
X	$c_X^* = 1.01$, $\lambda^* = 0.5$	$c_X^* = 0.69$, $\lambda^* = 0.68$	$c_X^* = 1.22$, $\lambda^* = 0.45$
Y	$c_Y^* = 2.01$, $\lambda^* = 0.5$	$c_Y^* = 2.83$, $\lambda^* = 0.30$	$c_Y^* = 1.75$, $\lambda^* = 0.54$

The empirical and estimated AC functions of X and Y do not change for the three cases; instead, the membership functions, estimated according to Procedure *eACMF* corresponding to the appropriate values of λ^* (see Table 1) are quite different in the three cases.

Figure 5 reproduces the situation where the core of X and Y coincide with the median of ACFs; they are similar, considering that the data are perturbed, to the ACFs and to the membership functions reproduced in Figure 4.

When the core is made coincident with the mode (Figure 6) or the mean (Figure 7), then the membership functions change significantly their position and shape (in particular, when the empirical ACFs exhibit some nonlinearity).

Example 3: In this case, the two variables X, Y represent two properties (variables) from a sub-sample of the **quakes** data set, available in the *R-language Package datasets (version 3.6.0)*. X and Y are the (rescaled) latitude and longitude from a cluster of $m = 795$ seismic events having loc-

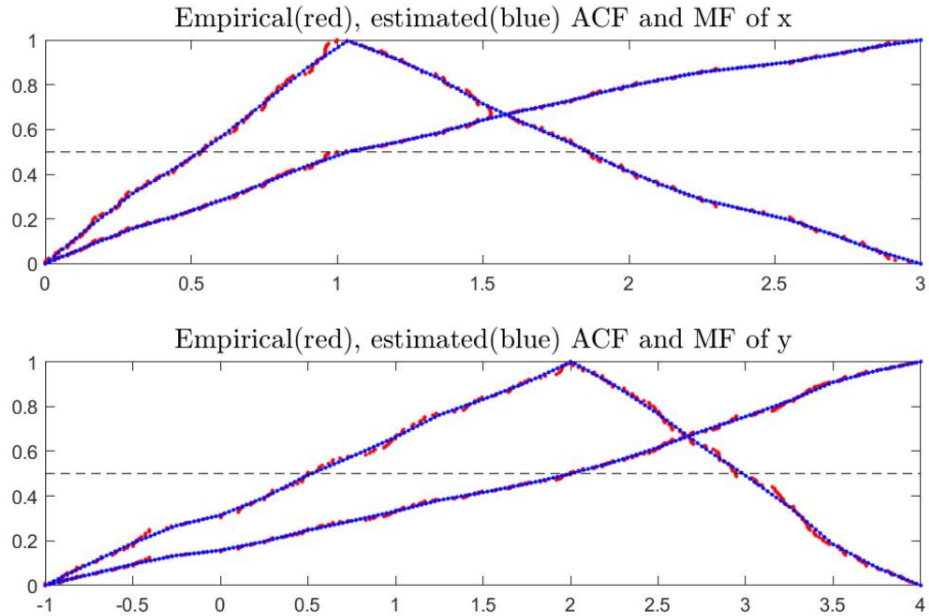

Figure 5: Empirical (m = 250) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 2, assuming the core coincident with the **median**; the horizontal dashed lines correspond to values of λ^*.

ations in a cube near Fiji (frequently used to test clustering procedures). The empirical median, mode and mean of X and Y are given in Table 2.

Table 2: *Median, mode, mean and corresponding* λ^* *values for Example 3*

Data	*1) Median*	*2) Mode*	*3) Mean*
X	$c_X^* = -0.33$, $\lambda^* = 0.5$	$c_X^* = 0.64$, $\lambda^* = 0.38$	$c_X^* = -1.23$, $\lambda^* = 0.59$
Y	$c_Y^* = 2.42$, $\lambda^* = 0.5$	$c_Y^* = 2.07$, $\lambda^* = 0.59$	$c_Y^* = 2.89$, $\lambda^* = 0.42$

The empirical AC functions of X and Y are pictured in Figure 8.

The estimated ACFs and the corresponding membership functions, obtained according to Procedure *eACMF* for the three cases, are shown in Figure 9 for the median, Figure 10 for the mode and Figure 11 for the

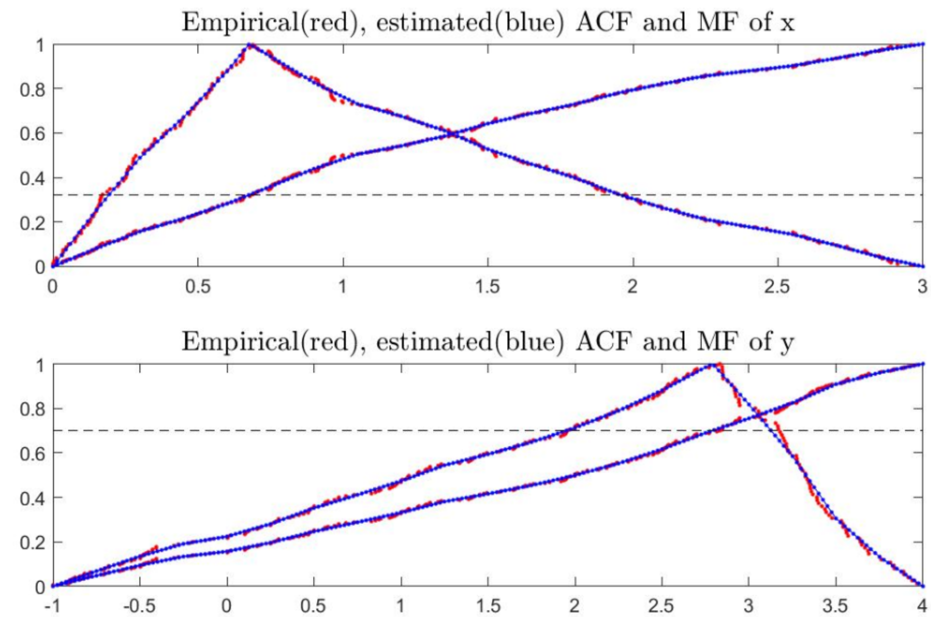

Figure 6: Empirical (m = 250) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 2, assuming the core coincident with the **mode**; the horizontal dashed lines correspond to values of λ^*.

mean.

Also for Example 3, the smooth monotonic approximations of the ACFs are obtained by the method described in [5]. We can remark that, in this example, the membership functions for the three cases exhibit similar forms (in particular for Y), due essentially to the fact that, with respect to the relatively large supports, the values of median, mode and mean are not so different as it is in Example 2.

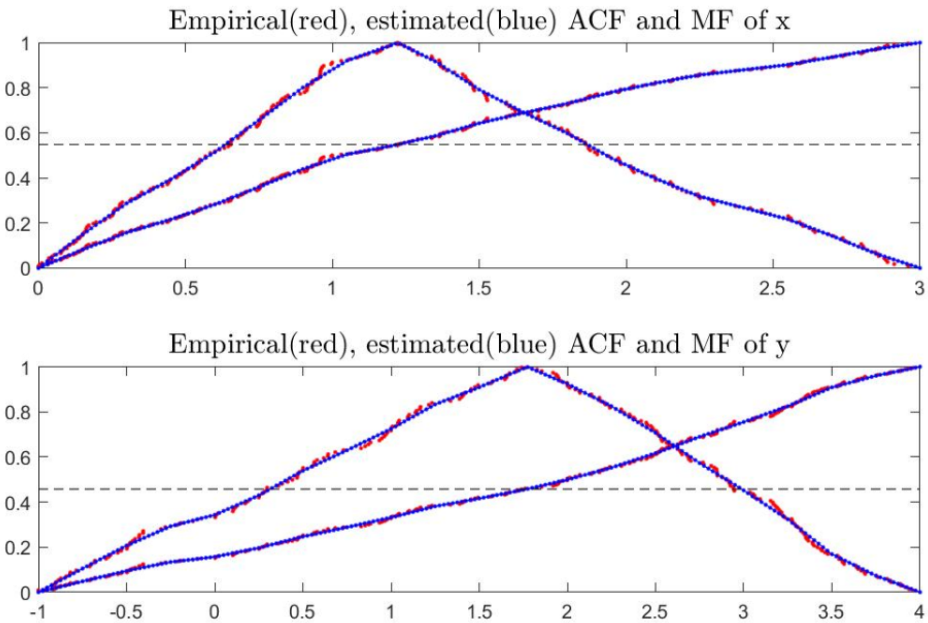

Figure 7: Empirical (m = 250) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 2, assuming the core coincident with the **mean**; the horizontal dashed lines correspond to values of λ^*.

4 Conclusions and further research

The AC function F_u associated to a (normal, convex) fuzzy number $u \in \mathbb{R}_{\mathcal{F}}$ creates a bridge with possibility theory and allows a setting to work with empirical observations. As we have seen in section 3, under the assumptions that

(a) data are generated according to an underlining membership function, *and*

(b) we have a precise information about the location of the core,

we can apply Theorem 4 to obtain an estimated function $\widehat{u} \in \mathbb{R}_{\mathcal{F}}$ from the empirical (monotonic) AC function $\widehat{F}$, assuming that the value $\lambda^* \in]0,1[$ is

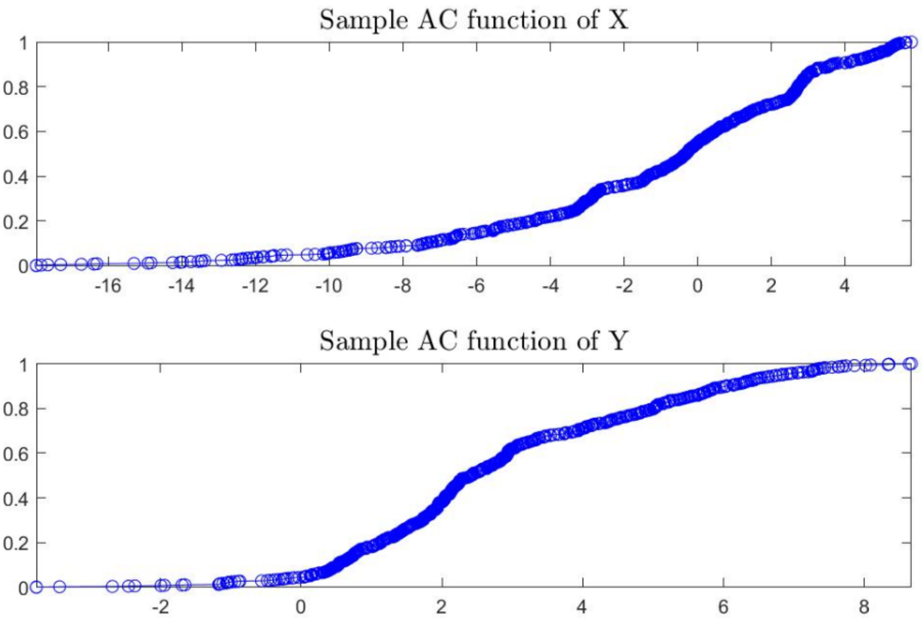

Figure 8: Empirical (m = 795) Average Cumulative functions of X and Y from the *quakes* dataset, as in Example 3.

fixed. On the other hand, the selection of λ^* cannot be deduced directly from the data set without a precise (at least qualitative) assumption.

A natural choice, having a valid statistical justification, is to determine λ^* such that the core of u coincides with the median of observed values. In this case, we know that the α-cuts of u coincide with the $\frac{\alpha}{2}$-quantile intervals; but other possible values can be determined, e.g., by the help of depth functions (see, [26], [1], [23]), or by analyzing the order structure of the data set.

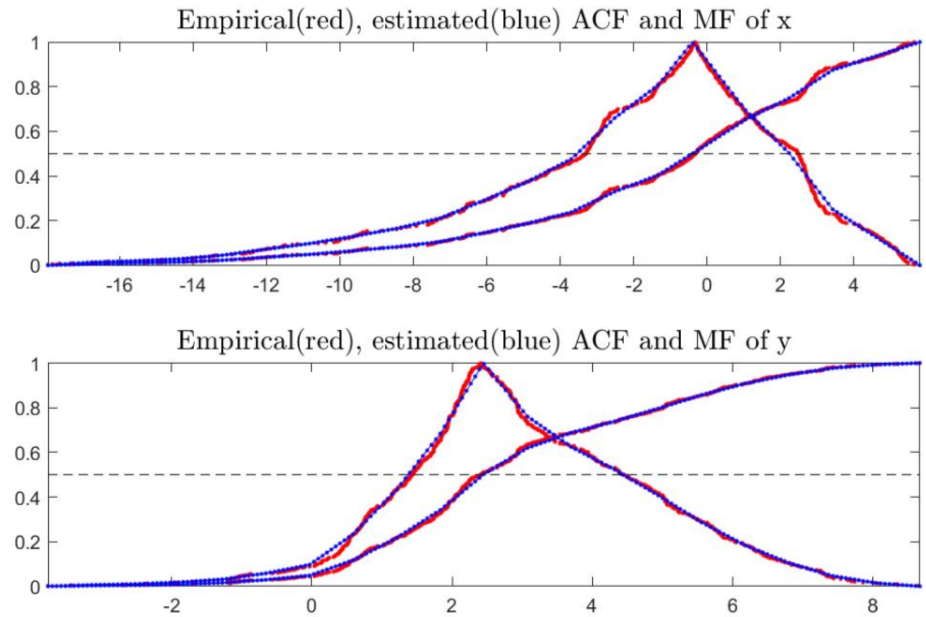

Figure 9: Empirical (m = 795) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 3 when the core coincides with the **median**; horizontal dashed lines correspond to the values of λ^*.

4.1 Open problems

The choice of an appropriate meaningful value of λ^* still remains an essentially open question, even if median, mode or mean values seems to merit special attention.

A second open question concerns the cases when (ordered) data are not able to evidence a unique *central* value, for example if the observations can be clustered into subsets around reasonably identified centroids; in these situations, we may have distinct fuzzy sets to identify and the use of clustering techniques can be of help to preliminarily subdivide the data into a number of sub-samples each giving a membership function with reduced

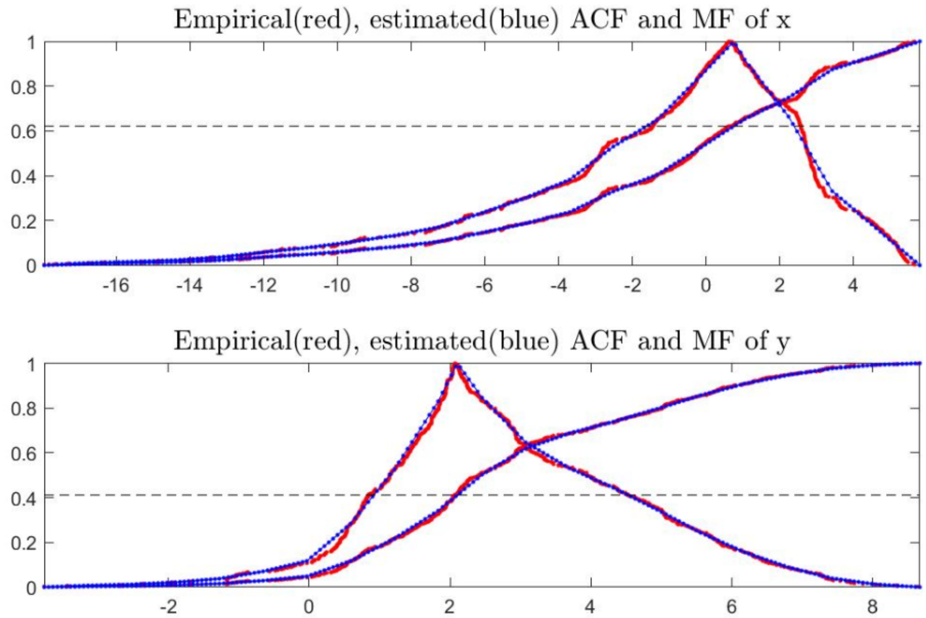

Figure 10: Empirical (m = 795) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 3 when the core coincides with the **modal value**; horizontal dashed lines correspond to the values of λ^*.

supports.

4.2 Further research

Some additional research will also involve computational issues, in order to improve the empirical applicability of the proposed ideas. In particular, we have planned some work to address

- the generation of random fuzzy intervals (or random possibility distributions) via AC functions and the search for possible metrics on ACFs that focus on useful topological structures (see for example [27] and [30]);

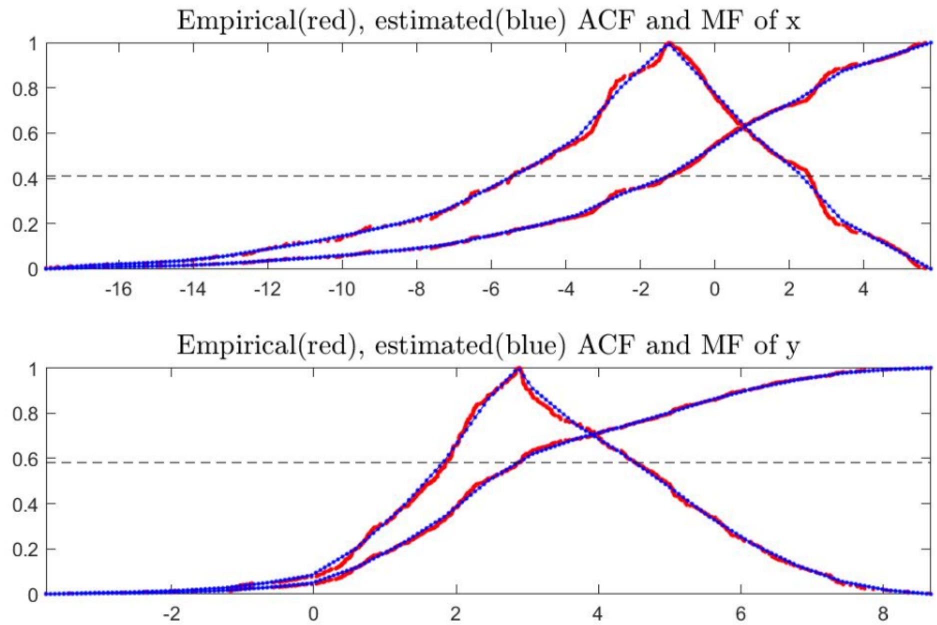

Figure 11: Empirical (m = 795) and estimated MF and ACF of fuzzy numbers X and Y, as in Example 3 when the core coincides with the **mean**; horizontal dashed lines correspond to the values of λ^*.

- the search for general and flexible procedures for robust approximation of ACFs from empirical data: indeed, the ACF-representation based on monotonic functions eases the search of approximation methods and algorithms, as is done in [5] by the use of F-transform. We remark that a similar procedure can be applied to the estimation of density and distribution functions of random variables;

- the identification of multiple membership functions with possibly overlapping supports, as in the case where observed data are obtained from mixtures of possibility distributions, e.g., as unions of fuzzy sets;

- the extension of a similar construction to the multi-dimensional case, using ideas from statistical or geometrical cluster analysis (including fuzzy clustering as, e.g., in), related to multi-dimensional depth functions, copulas (or more general functions) to represent multidimensional fuzzy quantities and associated AC functions, with or without the convexity requirement.

References

[1] M. Aments, M. Bern, D. Eppstein, S.H. Teng, Regression depth and center points. *Discrete Computational Geometry*, 23 (2000) 305-323.

[2] B. Bede, *Mathematics of Fuzzy Sets and Fuzzy Logic*, Studies in Fuzziness and Soft Computing n. 295, Springer, 2013.

[3] C. Carlsson, R. Fullér, *Possibility for Decision: A Possibilistic Approach to Real Life Decisions*, Studies in Fuzziness and Soft Computing, Volume 270, Springer, 2011.

[4] G. Coletti, D. Petturiti, B. Vantaggi, Fuzzy memberships as likelihood functions in a possibilistic framework, *International Journal of Approximate Reasoning*, 88 (2017) 547-566.

[5] L. Coroianu, L. Stefanini, General approximation of fuzzy numbers by F-transform, *Fuzzy Sets and Systems*, 288 (2016) 46-74.

[6] L. De Campos, J. Huete, Measurement of possibility distributions, *International Journal of General Systems*, 30, 3 (2001) 309-346.

[7] G. De Cooman, Possibility theory I: the measure-and integral-theoretic groundwork, *International journal of general systems* 25/4 (1997), 291-323.

[8] D. Dubois, Possibility theory and statistical reasoning, *Computational Statistics & Data Analysis,* 51 (2006) 47-69.

[9] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications.* Academic Press, New York, 1980.

[10] D. Dubois, H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York (1988).

[11] D. Dubois, H. Prade, On several representations of an uncertain body of evidence, in M.M. Gupta ,E. Sanchez, eds. *Fuzzy Information and Decision Processes*, North-Holland, Amsterdam, (1982) 167-181.

[12] Dubois D., Prade H. (1983) Unfair coins and necessity measures : towards a possibilistic interpretation of histograms, *Fuzzy Sets and Systems*, 10 (1983)15-20.

[13] D. Dubois, H. Prade, S. Sandri, On Possibility/Probability Transformations, in R. Lowen, M. Roubens (eds) *Fuzzy Logic, state of the art*, Theory and Decision Library book series TDLD, n. 12, Kluwer Academic Publisher, 1993, 103-112.

[14] D. Dubois, H. Prade, Fuzzy sets in approximate reasoning, Part 1: Inference with possibility distributions, *Fuzzy Sets and Systems*, 100-1 (1999) 73-132.

[15] D. Dubois, H. Prade, Qualitative possibility functions and integrals, Chapter 36 in A. Pap (Ed.) *Handbook of Measure Theory*, North Holland, 2002.

[16] D. Dubois, H. Prade, Possibility Theory and its Applications: where do we stand?, Chapter A/3 in J. Kacprzyk, W. Pedrycz (Eds.), *Handbook of Computational Intelligence*, Springer (2015).

[17] D. Dubois, H. Prade, Practical methods for constructing possibility distributions, *International Journal of Intelligent Systems*, 31 (2016) 215-239.

[18] D. Dubois, E. Kerre, R. Mesiar, H. Prade, Fuzzy Interval Analysis, chapter 10 in D. Dubois, H. Prade (Eds), *Fundamentals of Fuzzy sets*, Kluwer, 2000, pp 483-581.

[19] R. Durrett, *Probability: Theory and Examples*, IV edition, Cambridge University Press, 2010.

[20] L. Jin, M. Kalina, R. Mesiar, S. Borkotokey, Characterizations of the possibility-probability transformations and some applications, *Information Sciences*, 477 (2019) 281-290.

[21] G.J. Klir, On fuzzy-set interpretation of possibility theory, *Fuzzy Sets and Systems*, 108 (1999) 263-273.

[22] M.H. Masson, T. Denoeux, Inferring a possibility distribution from empirical data, *Fuzzy Sets and Systems* 157 (2006) 319 – 340.

[23] I. Mizera, On depth and deep points: A calculus, *The Annals of Statistics*, 30/6 (2002) 1682-1736.

[24] M. Pota, M. Esposito, G. De Pietro, Transforming probability distributions into membership functions of fuzzy classes: A hypothesis test approach, *Fuzzy Sets and Systems*, 233 (2013) 52-73.

[25] M.L. Puri, D.A Ralescu, Fuzzy random variables, *J. Math. Anal. Appl.*, 114 (1986) 409-422.

[26] P.J. Rousseeuw, M. Hubert, Regression depth, J. Amer. Statist. Assoc., 94 (1999) 388-402.

[27] B. Sinova, M.A. Gil, A. Colubi, A. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach, *Fuzzy Sets and Systems*, 200 (2012) 99-115.

[28] M. Serrurier, H. Prade, An informational distance for estimating the faithfulness of a possibility distribution, viewed as a family of probability distributions, with respect to data, *International Journal of Approximate Reasoning*, 54/7 (2013) 919-933.

[29] L. Stefanini, M.L. Guerra, On possibilistic representations of fuzzy intervals, *Information Sciences*, 405 (2017) 33-54.

[30] W. Trutsching, G. Gonzales-Rodriguez, A.Colubi, M.A. Gil, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread, *Information Sciences*, 179 (2009) 3964-3972.

[31] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems,* 1 (1978) 3–28.