

Fuzzy Confidence Interval Estimation by Likelihood Ratio

Rédina Berkachy^a and Laurent Donzé^b

^{a,b}Applied Statistics And Modelling (ASAM) group, Department of Informatics,
Faculty of Economic and Social Sciences, University of Fribourg, Switzerland

^aRedina.Berkachy@UniFR.CH

^bLaurent.Donze@UniFR.CH

Abstract

We propose a fuzzy confidence interval estimation based on the likelihood ratio. This ratio, often used in hypotheses testing seems to be an efficient tool for calculating confidence intervals since it is known to be general, and thus can be applied on any parameter. The strength of the defended procedure is to use a wide range of estimators with any type of distribution for the estimation of confidence intervals when fuzziness occurs. The theoretical approach and the detailed steps of the calculation are given. This approach is illustrated by a classical problem: a fuzzy confidence interval for the fuzzy mean in the context of a normal distribution. Finally, a comparison between the interval by the defended approach and one calculated by a frequently used expression is made. Our results show that the support set of the fuzzy interval by the defended method is smaller than the one by the known expression.

Keywords: Estimation, Likelihood ratio, Fuzzy Confidence Interval, Fuzzy Statistical Inference.

1 Introduction and motivation

We use to postulate a null hypothesis and an alternative one for a given parameter, and to test it thanks to our data set. The aim in this case is to reject or not the null hypothesis at a given significance level. Another way to accomplish this same task is to construct a confidence interval in order to be able to position our parameter.

When fuzziness intervenes, the concept of confidence interval has to be re-defined in a way considering every

single value of the support set of the membership functions of the fuzzy numbers. A wide range of methods estimating fuzzy confidence intervals were given in the recent years. Kruse & Meyer [11] are from the first ones who proposed a clear theoretical definition of fuzzy confidence intervals. Viertl [13] and Viertl & Yeganeh [14] provided as well a definition of the concept of confidence regions for fuzzy data. They were particularly interested in intervals defined in a Bayesian context. In Kahraman & al. [10], the authors presented and analyzed the approaches of constructing fuzzy confidence intervals developed between 1980 and 2015. In addition, they closed their chapter by a method for confidence intervals with interval-valued intuitionistic fuzzy sets, and one with the so-called hesitant fuzzy confidence intervals. Moreover, it is interesting to mention the approach based on inner and outer approximations of confidence intervals in a fuzzy environment, given in Couso & Sánchez [5].

Even though a big panoply of methods is nowadays available, many of these methods were particularly introduced for specific parameters and for pre-defined distributions. As instance, these methods seem to be limited since no potential generalization of expressions have been proposed. It is then appealing to design and develop a practical method generalizing all possible cases of fuzzy confidence intervals. This method will be eventually useful in hypotheses testing. As such, we remember that hypotheses testing methods based on confidence intervals are widely used in different contexts in both the classical and the fuzzy environments (see Berkachy & Donzé [1], Berkachy & Donzé [2], Chachi & al. [3], Grzegorzewski [9] and others).

One way to fill in this gap is by using the so-called “likelihood ratio” often present in inference tests. For uncertainty modelling and based on Zadeh’s

probabilistic definition (see Zadeh [15]), Gil and Casals [8] described the extension of the likelihood ratio test to the case where data sets involve fuzzy imprecision. The likelihood function was used as well in some different contexts. Note Denoeux [6] which gave a maximum likelihood estimation from fuzzy data using the Expectation-Maximization algorithm. From another side, Viertl [13] presented in his study on Bayes' theorems, a generalized likelihood function for fuzzy data defined by its left and right α -cuts.

We hereby propose an original method to construct a fuzzy confidence interval based on the likelihood ratio. The main objective of this study is to provide a general way of computing a confidence interval when fuzziness occurs in the data set, and give us the possibility to extend our former results on fuzzy hypotheses tests by confidence intervals (see Berkachy & Donzé [1] and Berkachy & Donzé [2]). Our approach is adapted to a broad spectrum of parameters using any type of statistical distributions. An application followed by a comparison between our approach and a "traditional" one is given.

This paper is organized as follows. Section 2 is devoted to the theoretical approach. In Section 3, we give the steps of the computation related-procedure of the fuzzy confidence intervals. An illustrative application is afterwards given in Section 4. We close the paper by the Section 5 where we compare the application of the defended approach from one side and a known one from another side.

2 Theoretical approach

Consider a sequence $X_i, i = 1, \dots, n$ of random variables, independent identically distributed (i.i.d). Each variable X_i is supposed to be derived from a defined distribution with a probability density function (pdf) denoted by $f(x_i; \theta)$, where x_i is a realization of the variable X_i and θ is a vector of unknown parameters in the parameter space Θ . The likelihood function $L(\theta; x_i)$ is defined as follows:

$$L(\theta; x_i) = f(x_i; \theta). \quad (1)$$

We assume now each variable X_i to be imprecise and we would like to model every realization x_i by a fuzzy number. Consider the fuzzy realization \tilde{x}_i of the fuzzy random variable \tilde{X}_i , i.e. in this case \tilde{x}_i is the fuzzy version of x_i . The fuzzy number \tilde{x}_i is characterized by a Borel measurable membership function denoted by $\mu_{\tilde{x}_i}$, such that $\mu_{\tilde{x}_i} : x \rightarrow [0, 1]$. We are interested in extending the log-likelihood ratio to the fuzzy

context. For this purpose, some probability concepts are needed. Zadeh [15] has previously discussed the probability of a given fuzzy event in an uncertain environment. He provided the subsequent definitions.

Definition 1. Fuzzy event

Consider $(\mathbb{R}^n, \mathcal{A}, P)$ a probability space, where P is a probability measure over \mathbb{R}^n and \mathcal{A} is a σ -field of Borel sets on \mathbb{R}^n . A **fuzzy event** of \mathbb{R}^n is defined as a fuzzy subset \tilde{x} in \mathbb{R}^n such that its membership function $\mu_{\tilde{x}}$ is measurable in the Borel space.

Definition 2. Probability of a fuzzy event

The **probability of a fuzzy event** \tilde{x} is the expectation $E(\mu_{\tilde{x}})$ of $\mu_{\tilde{x}}$ with respect to the probability measure P as follows:

$$P(\tilde{x}) = \int_{\mathbb{R}^n} \mu_{\tilde{x}}(x) dP = E(\mu_{\tilde{x}}). \quad (2)$$

Definition 3. Independent fuzzy events

Let \tilde{x}_1 and \tilde{x}_2 be two fuzzy events in the probability space $(\mathbb{R}^n, \mathcal{A}, P)$, and $\tilde{x} = \tilde{x}_1 \tilde{x}_2$ a fuzzy subset of \mathbb{R}^n , representing the product between \tilde{x}_1 and \tilde{x}_2 . The respective membership function of \tilde{x} , denoted by $\mu_{\tilde{x}} = \mu_{\tilde{x}_1 \tilde{x}_2}(x)$, is defined as $\mu_{\tilde{x}_1}(x) \cdot \mu_{\tilde{x}_2}(x)$, $\forall x \in \mathbb{R}^n$. The fuzzy events \tilde{x}_1 and \tilde{x}_2 are said to be **independent** iff:

$$P(\tilde{x}) = P(\tilde{x}_1) \cdot P(\tilde{x}_2). \quad (3)$$

Furthermore, the fuzzy sample $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$ is seen as i.i.d. fuzzy realizations of the sequence of variables $X_i, i = 1, \dots, n$. It is a collection of n mutually independent fuzzy events $\tilde{x}_i, i = 1, \dots, n$. Its joint membership function $\mu_{\tilde{x}}$ can be written as:

$$\mu_{\tilde{x}}(\underline{x}) = \mu_{\tilde{x}_1}(x_1) \times \dots \times \mu_{\tilde{x}_n}(x_n). \quad (4)$$

In this case, we used the product t-norm in order to facilitate further integration operations.

Definition 4. Likelihood function of a fuzzy observation

Consider $\tilde{\theta}$ a vector of fuzzy parameters in the parameter space Θ . The **likelihood function** for a single fuzzy observation \tilde{x}_i can be expressed as:

$$L(\tilde{\theta}; \tilde{x}_i) = P(\tilde{x}_i; \tilde{\theta}) = \int_{\mathbb{R}} \mu_{\tilde{x}_i}(x) f(x_i; \tilde{\theta}) dx. \quad (5)$$

Remark. This probability can be analogously written in terms of the α -cuts of the modelling fuzzy numbers.

It is relatively natural to write the likelihood function $L(\tilde{\theta}; \tilde{x})$ of the fuzzy sample \tilde{x} as:

$$\begin{aligned} L(\tilde{\theta}; \tilde{x}) &= P(\tilde{x}; \tilde{\theta}) \\ &= \int_{\mathbb{R}} \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx \times \dots \times \int_{\mathbb{R}} \mu_{\tilde{x}_n}(x) f(x_n; \tilde{\theta}) dx \\ &= \prod_{i=1}^n \int_{\mathbb{R}} \mu_{\tilde{x}_i}(x) f(x_i; \tilde{\theta}) dx. \end{aligned} \quad (6)$$

Consequently, the log-likelihood function denoted by $l(\tilde{\theta}; \tilde{x})$ is expressed by:

$$l(\tilde{\theta}; \tilde{x}) = \log L(\tilde{\theta}; \tilde{x}) \quad (7)$$

$$\begin{aligned} &= \log \int_{\mathbb{R}} \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx + \dots \\ &\quad + \log \int_{\mathbb{R}} \mu_{\tilde{x}_n}(x) f(x_n; \tilde{\theta}) dx. \end{aligned} \quad (8)$$

Example. Consider a triangular fuzzy number \tilde{x}_1 given by the tuple $\tilde{x}_1 = (1, 2, 3)$ and taken from the normal distribution with a probability density function denoted by f . The membership function of \tilde{x}_1 is given by:

$$\mu_{\tilde{x}_1}(x) = \begin{cases} x - 1 & \text{if } 1 \leq x \leq 2, \\ 3 - x & \text{if } 2 < x \leq 3. \end{cases} \quad (9)$$

The likelihood function at \tilde{x}_1 is calculated as follows:

$$\begin{aligned} L(\tilde{\theta}; \tilde{x}_1) &= \int_{\mathbb{R}} \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx \\ &= \int_1^2 \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx + \int_2^3 \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx \\ &= \int_1^2 (x - 1) f(x_1; \tilde{\theta}) dx + \int_2^3 (3 - x) f(x_1; \tilde{\theta}) dx. \end{aligned}$$

Consider $\hat{\tilde{\theta}}$ the estimator by maximum likelihood of the fuzzy parameter $\tilde{\theta}$, and f the probability density function. The so-called likelihood ratio is the ratio of the likelihood function of the fuzzy parameter $L(\tilde{\theta}; \tilde{x})$ to the likelihood function of the fuzzy estimator $L(\hat{\tilde{\theta}}; \tilde{x})$, written as $L(\tilde{\theta}; \tilde{x})/L(\hat{\tilde{\theta}}; \tilde{x})$. Furthermore, the logarithm of this ratio is often given by the deviance, i.e. the difference between the log-likelihood functions of $\hat{\tilde{\theta}}$ and $\tilde{\theta}$. In the fuzzy context and analogously to the classical statistical theory, one can expect that LR given by

$$LR = -2 \log \frac{L(\tilde{\theta}; \tilde{x})}{L(\hat{\tilde{\theta}}; \tilde{x})} = 2 \left[l(\hat{\tilde{\theta}}; \tilde{x}) - l(\tilde{\theta}; \tilde{x}) \right] \quad (10)$$

is asymptotically Chi-squared distributed with k degrees of freedom, where k is determined by the constraints expressed by the null hypothesis. This result

can clearly lead to recalling the likelihood-ratio test statistic.

For the case of one constraint only ($k = 1$), a $100(1 - \delta)\%$ confidence interval constructed by a LR approach, is done by finding all possible values $\tilde{\theta}$ for which we will be at the point of accepting (not rejecting) the null hypothesis “ $H_0 : \tilde{\theta} = \tilde{\theta}_0$ ” at the significance level δ . As a consequence, the confidence interval can be found by writing the following:

$$2 \left[l(\hat{\tilde{\theta}}; \tilde{x}) - l(\tilde{\theta}; \tilde{x}) \right] \leq \chi_{(1; 1-\delta)}^2,$$

which can also be expressed as

$$l(\tilde{\theta}; \tilde{x}) \geq l(\hat{\tilde{\theta}}; \tilde{x}) - \frac{\chi_{(1; 1-\delta)}^2}{2}. \quad (11)$$

This last expression can be interpreted as calculating the $100(1 - \delta)\%$ confidence interval which includes all possible values of $\tilde{\theta}$ when the log-likelihood maximum variates for about no more than $\frac{\chi_{(1; 1-\delta)}^2}{2}$.

3 Procedure

A direct consequence of assuming that the data set is fuzzy is to get the log-likelihood as a function of fuzzy data. For this reason, we propose to update the method of computing the confidence intervals in a way of modelling the ambiguity and fuzziness. Under these circumstances, the estimator by maximum likelihood $\hat{\tilde{\theta}}$ is fuzzy as well. Therefore, instead of seeing it as a single value as in the classical theory, every value of the support set of this fuzzy parameter has to be considered. As the corresponding log-likelihood function is calculated for every value of the support set of the parameter, it is obvious that the computational burden becomes tedious.

Considering specific chosen values, we calculate threshold points in order to be able to find their intersections with the log-likelihood function. All these informations will be finally combined in the construction of the required fuzzy confidence intervals.

Let us start by defining a so-called standardizing function:

Definition 5. Standardizing function

For a given value $\theta \in \mathbb{R}$ of the interval defining the membership function of $\tilde{\theta}$, consider the **standardizing function**

$$\begin{aligned} I_{stand} : \quad \mathbb{R} &\rightarrow \mathbb{R} \\ l(\theta, \tilde{x}) &\rightarrow I_{stand}(l(\theta, \tilde{x})) = \frac{l(\theta, \tilde{x}) - I_a}{I_b - I_a}, \end{aligned}$$

where $I_a, I_b \in \mathbb{R}$ and $I_a < I_b$, $I_{stand}(l(\theta, \tilde{x}))$ is bounded, $0 \leq I_{stand}(l(\theta, \tilde{x})) \leq 1$ and $I_a \leq l(\theta, \tilde{x}) \leq I_b$.

This step is introduced in the purpose of getting back to the concept of the interval $[0, 1]$ of the α -cuts.

Then, the following procedure of computations is adopted:

1. **Calculate the log-likelihood function $l(\tilde{\theta}; \tilde{x})$ (Equation 8).**

2. **Find the needed values for computations.**

Since the support set of the parameter contains an infinity of values, we propose to simply carefully reduce the number of values to only four. They are given by the lower and upper bounds of both the support and the core sets of the fuzzy parameter. If $\text{supp}(\hat{\theta})$ and $\text{core}(\hat{\theta})$ are respectively the support and the core sets of the fuzzy parameter $\hat{\theta}$, then the four needed values $p \leq q \leq r \leq s$ are found as follows:

$$p = \min(\text{supp}(\hat{\theta})), \quad (12)$$

$$q = \min(\text{core}(\hat{\theta})), \quad (13)$$

$$r = \max(\text{core}(\hat{\theta})), \quad (14)$$

$$s = \max(\text{supp}(\hat{\theta})). \quad (15)$$

Remark. For a symmetrical probability density function, this latter can be easily remarked since the left hand side of the log-likelihood function always increases and the right one always decreases (consequence of the shape of the pdf). Each part is monotonic.

3. **Calculate the threshold points.** We define the threshold values I_1, I_2, I_3 and I_4 , functions of p, q, r and s of Equations 12, 13, 14 and 15 respectively, as the right hand side of Equation 11 evaluated at $\theta = p, q, r$ and s , i.e.:

$$I_1 = l(p; \tilde{x}) - \frac{\chi_{(1;1-\delta)}^2}{2}, \quad (16)$$

$$I_2 = l(q; \tilde{x}) - \frac{\chi_{(1;1-\delta)}^2}{2}, \quad (17)$$

$$I_3 = l(r; \tilde{x}) - \frac{\chi_{(1;1-\delta)}^2}{2}, \quad (18)$$

$$I_4 = l(s; \tilde{x}) - \frac{\chi_{(1;1-\delta)}^2}{2}. \quad (19)$$

While the minimum threshold is

$$I_{\min} = \min(I_1, I_2, I_3, I_4), \quad (20)$$

the maximum one is

$$I_{\max} = \max(I_1, I_2, I_3, I_4). \quad (21)$$

4. **Find the intersection points.** The intersection abscisses $\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}, \theta_4^{*L}$ and $\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}, \theta_4^{*R}$ are calculated by solving the following equations

$$l^L(\theta_1^{*L}; \tilde{x}) = I_1 \quad \text{and} \quad l^R(\theta_1^{*R}; \tilde{x}) = I_1, \quad (22)$$

$$l^L(\theta_2^{*L}; \tilde{x}) = I_2 \quad \text{and} \quad l^R(\theta_2^{*R}; \tilde{x}) = I_2, \quad (23)$$

$$l^L(\theta_3^{*L}; \tilde{x}) = I_3 \quad \text{and} \quad l^R(\theta_3^{*R}; \tilde{x}) = I_3, \quad (24)$$

$$l^L(\theta_4^{*L}; \tilde{x}) = I_4 \quad \text{and} \quad l^R(\theta_4^{*R}; \tilde{x}) = I_4. \quad (25)$$

Note that the previous abscisses are of type values and not α -cuts. They will be used in the process of construction of the required α -cuts of the fuzzy confidence interval.

5. **Define the fuzzy confidence interval.** The fuzzy confidence interval by the likelihood ratio $\tilde{\Pi}_{LR}$ can be written by its α -cuts $(\tilde{\Pi}_{LR})_\alpha = [(\tilde{\Pi}_{LR})_\alpha^L, (\tilde{\Pi}_{LR})_\alpha^R]$. The left and right α -cuts $(\tilde{\Pi}_{LR})_\alpha^L$ and $(\tilde{\Pi}_{LR})_\alpha^R$ are defined in the following manner:

$$(\tilde{\Pi}_{LR})_\alpha^L = \left\{ \theta \in \mathbb{R} \mid \theta_{\inf}^{*L} \leq \theta \leq \theta_{\sup}^{*L} \text{ and } \alpha = I_{stand}(l(\theta, \tilde{x})) = \frac{l(\theta, \tilde{x}) - I_{\min}}{I_{\max} - I_{\min}} \right\}, \quad (26)$$

where $\theta_{\inf}^{*L} = \inf(\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}, \theta_4^{*L})$ and $\theta_{\sup}^{*L} = \sup(\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}, \theta_4^{*L})$ and

$$(\tilde{\Pi}_{LR})_\alpha^R = \left\{ \theta \in \mathbb{R} \mid \theta_{\inf}^{*R} \leq \theta \leq \theta_{\sup}^{*R} \text{ and } \alpha = I_{stand}(l(\theta, \tilde{x})) = \frac{l(\theta, \tilde{x}) - I_{\min}}{I_{\max} - I_{\min}} \right\}, \quad (27)$$

where $\theta_{\inf}^{*R} = \inf(\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}, \theta_4^{*R})$ and $\theta_{\sup}^{*R} = \sup(\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}, \theta_4^{*R})$.

We highlight that this fuzzy confidence interval is constructed intentionally in the perspective of integrating it in the approach of the fuzzy inference tests by confidence intervals described in Berkachy & Donzé [1] and [2].

4 Numerical application

Consider a random sample composed by $n = 10$ observations $x_l, l = 1, \dots, 10$ derived from a normal distribution with a standard deviation of 0.79. We

suppose that the observations are fuzzy and will be modelled by triangular fuzzy numbers. The data set and the corresponding fuzzy numbers are shown in Table 1.

The aim is to calculate the fuzzy confidence interval for the fuzzy mean given by $\tilde{x} = (0.8, 1.8, 2.8)$ at the significance level $\delta = 0.05$. The aforementioned procedure is considered and will be applied in the following manner:

1. Calculate the log-likelihood function:

$$\begin{aligned} l(\tilde{\theta}; \tilde{x}) &= \log \int_{\mathbb{R}} \mu_{\tilde{x}_1}(x) f(x_1; \tilde{\theta}) dx + \dots + \\ &\quad \log \int_{\mathbb{R}} \mu_{\tilde{x}_{10}}(x) f(x_{10}; \tilde{\theta}) dx \\ &= \int_0^1 x f(x_1; \tilde{\theta}) dx + \\ &\quad \int_1^2 (2-x) f(x_1; \tilde{\theta}) dx + \dots + \\ &\quad \int_1^2 (x-1) f(x_{10}; \tilde{\theta}) dx + \\ &\quad \int_2^3 (3-x) f(x_{10}; \tilde{\theta}) dx, \end{aligned}$$

where $f(x_l; \tilde{\theta})$, $l = 1, \dots, 10$, is the probability density function of the normal distribution with a standard deviation of 0.79.

2. By analogy to the classical theory, the concerned estimator denoted by $\hat{\theta}$, is nothing but the mean. This estimator $\hat{\theta}$ is fuzzy, consequence of the fuzziness of the data. The support set of the fuzzy mean \tilde{x} is the interval $[p, s]$ where $p = 0.8$ and $s = 2.8$. Since the fuzzy mean is triangular alike the data, the core set is reduced to one value only, $q = r = 1.8$.

3. Calculate after I_1, I_2, I_3 and I_4 . We have that $\frac{\chi^2_{(1;1-0.05)}}{2} = \frac{3.84}{2} = 1.92$, and we get the following:

$$I_1 = l(0.8; \tilde{x}) - 1.92 = -19.8, \quad (28)$$

$$I_2 = l(1.8; \tilde{x}) - 1.92 = -13.47, \quad (29)$$

$$I_3 = l(1.8; \tilde{x}) - 1.92 = -13.47, \quad (30)$$

$$I_4 = l(2.8; \tilde{x}) - 1.92 = -19.8. \quad (31)$$

In addition, the minimum and maximum thresholds are respectively $I_{\min} = \min(I_1, I_2, I_3, I_4) = -19.8$ and $I_{\max} = \max(I_1, I_2, I_3, I_4) = -13.47$. These points are seen in Figure 1.

4. Compute the values $\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}$ and θ_4^{*L} as seen in Equations 22 to 25. They are given by:

$$\theta_1^{*L} = \theta_4^{*L} = 0.66, \quad (32)$$

$$\theta_2^{*L} = \theta_3^{*L} = 1.25. \quad (33)$$

In the same way, we compute $\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}$ and θ_4^{*R} and we get:

$$\theta_1^{*R} = \theta_4^{*R} = 2.94, \quad (34)$$

$$\theta_2^{*R} = \theta_3^{*R} = 2.35. \quad (35)$$

5. The fuzzy confidence interval $\tilde{\Pi}_{LR}$ given by its α -cuts as $(\tilde{\Pi}_{LR})_{\alpha} = [(\tilde{\Pi}_{LR})_{\alpha}^L, (\tilde{\Pi}_{LR})_{\alpha}^R]$ is seen in Figure 2. The left α -cut can be written as follows:

$$\begin{aligned} (\tilde{\Pi}_{LR})_{\alpha}^L &= \left\{ \theta \in \mathbb{R} \mid \theta_{\inf}^{*L} \leq \theta \leq \theta_{\sup}^{*L} \text{ and} \right. \\ &\quad \left. \alpha = \frac{l(\theta, \tilde{x}) + 19.8}{6.33} \right\}, \end{aligned}$$

where $\theta_{\inf}^{*L} = \inf(\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}, \theta_4^{*L}) = \theta_1^{*L} = 0.66$, and $\theta_{\sup}^{*L} = \sup(\theta_1^{*L}, \theta_2^{*L}, \theta_3^{*L}, \theta_4^{*L}) = \theta_2^{*L} = 1.25$.

Similarly, the right α -cut can be given by:

$$\begin{aligned} (\tilde{\Pi}_{LR})_{\alpha}^R &= \left\{ \theta \in \mathbb{R} \mid \theta_{\inf}^{*R} \leq \theta \leq \theta_{\sup}^{*R} \text{ and} \right. \\ &\quad \left. \alpha = \frac{l(\theta, \tilde{x}) + 19.8}{6.33} \right\}, \end{aligned}$$

where $\theta_{\inf}^{*R} = \inf(\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}, \theta_4^{*R}) = \theta_2^{*R} = 2.35$ and $\theta_{\sup}^{*R} = \sup(\theta_1^{*R}, \theta_2^{*R}, \theta_3^{*R}, \theta_4^{*R}) = \theta_1^{*R} = 2.94$.

Index	x_l	Triangular Fuzzy Number
1	1	(0, 1, 2)
2	2	(1, 2, 3)
3	3	(2, 3, 4)
4	2	(1, 2, 3)
5	2	(1, 2, 3)
6	1	(0, 1, 2)
7	1	(0, 1, 2)
8	3	(2, 3, 4)
9	1	(0, 1, 2)
10	2	(1, 2, 3)

Table 1: The data set and the corresponding fuzzy number of each observation - Example of Section 4

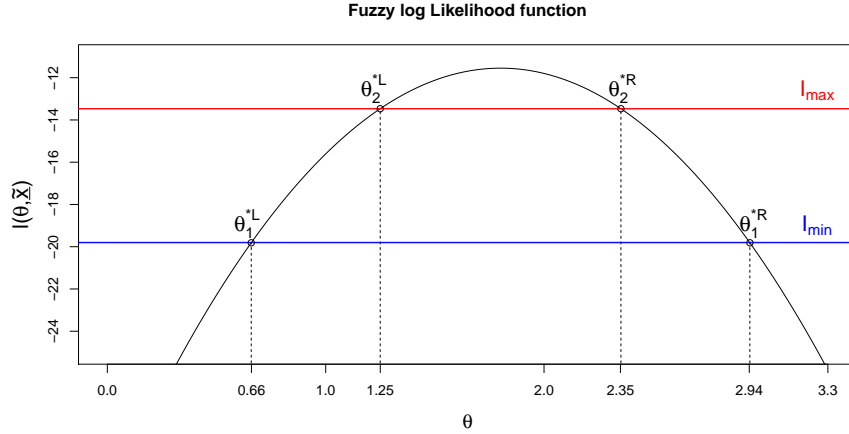


Figure 1: Fuzzy log-likelihood function for the mean and the intersection with the upper and lower bounds of the fuzzy parameter \tilde{x} .

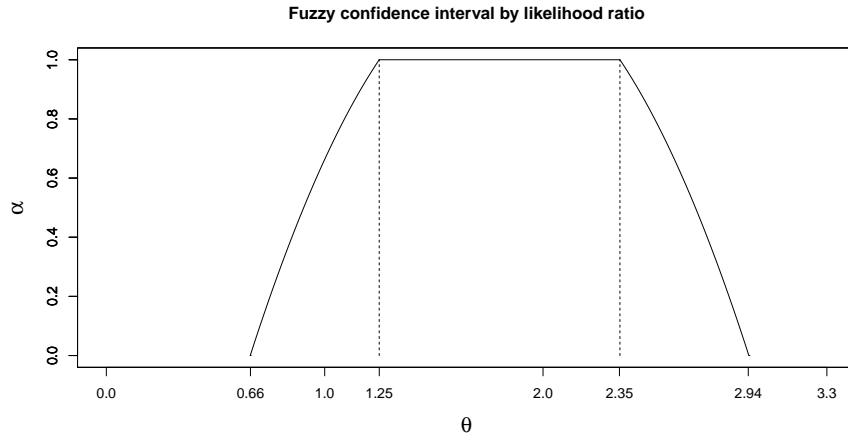


Figure 2: Fuzzy confidence interval by likelihood ratio $\tilde{\Pi}_{LR}$.

5 Comparison of computations of the fuzzy confidence intervals

This section is dedicated to compare the “traditional” way of estimating a fuzzy confidence interval with the one by our defended approach. For this purpose, let us first calculate the “traditional” two-sided fuzzy confidence interval for the mean given the normality of the distribution, using the following well-known expression

$$(\tilde{\Pi})_\alpha = \left[\bar{X}_\alpha^L - u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X}_\alpha^R + u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}} \right], \quad (36)$$

where \bar{X}_α^L and \bar{X}_α^R are the left and right α -cuts of the fuzzy mean, n is the sample size, σ is the standard deviation and $u_{1-\frac{\delta}{2}}$ is the $1 - \frac{\delta}{2}$ ordered quantile of the normal distribution for the confidence interval at

the δ significance level.

We apply this formula on the data set of example of Section 4, and we get the following result:

$$\begin{aligned} (\tilde{\Pi})_\alpha &= \left[\bar{X}_\alpha^L - u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X}_\alpha^R + u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}} \right], \\ &= \left[0.08 + \alpha - u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}}, 2.8 - \alpha + u_{1-\frac{\delta}{2}} \frac{\sigma}{\sqrt{n}} \right], \\ &= [0.31 + \alpha, 3.29 - \alpha], \end{aligned}$$

with $\bar{X}_\alpha^L = 0.8 + \alpha$ and $\bar{X}_\alpha^R = 2.8 - \alpha$, $u_{1-\frac{\delta}{2}} = u_{1-\frac{0.05}{2}} = 1.96$, $\sigma = 0.79$ and $n = 10$. This interval is shown in Figure 3. From another side, we remind that the fuzzy confidence interval calculated by our approach and denoted by $\tilde{\Pi}_{LR}$, is given in Step 5 of Section 4.

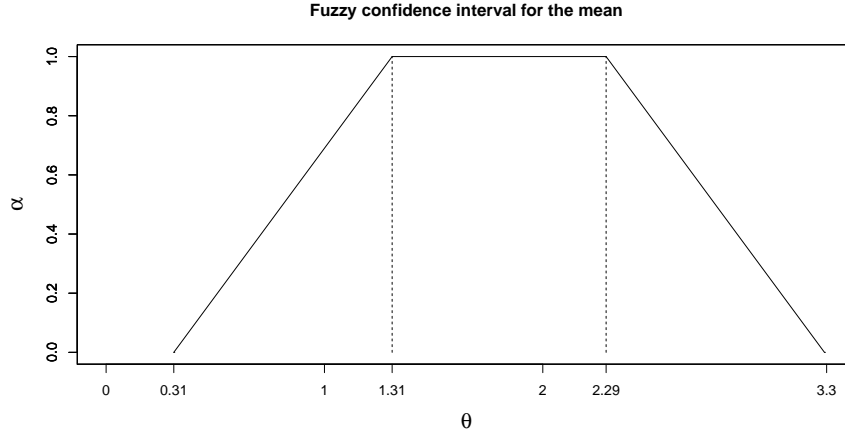


Figure 3: Fuzzy confidence interval for the mean $\tilde{\Pi}$.

For the comparison between the two methods of computation, it is clear that the shape of the fuzzy interval $\tilde{\Pi}$ depends strongly on the shape of the membership function of the fuzzy mean; this latter depending on the shape of the membership functions modelling the fuzzy observations. Contrariwise, the fuzzy confidence interval $\tilde{\Pi}_{LR}$ depends mainly on the shape of the probability density function pre-defined from the beginning.

We also have to note that the support set of the fuzzy confidence interval for the mean $\tilde{\Pi}$ (i.e. Figure 3) is larger than the one based on our approach $\tilde{\Pi}_{LR}$ (i.e. Figure 2), i.e. $\text{supp}(\tilde{\Pi}_{LR}) \subset \text{supp}(\tilde{\Pi})$. As a consequence, the defended method is seen to be more restrictive: we tend to reject more often when a given hypothesis is not included in the interval.

On the other hand and regarding the shape, one can get a more elaborated interval, and then more accurate. In addition, we note that the core set of our fuzzy interval is larger than the known one, $\text{core}(\tilde{\Pi}) \subset \text{core}(\tilde{\Pi}_{LR})$.

Despite the computational burden that can be faced due to the calculations of our interval, by the proposed procedure we were able to manage this complexity and thus, provide an affordable one for further sophisticated use.

A final remark is that our defended approach is in some sense general. Hence, it can be applied in several contexts with a widespread of estimators, without the need of specifying a given distribution. However, in this numerical example, we showed also the

“traditional” fuzzy confidence interval, where the expression for such calculations was used exclusively for the mean and in the case of the normal distribution only. Any further interval using the same “traditional” expression should be adapted to the type of estimator and the distribution chosen.

6 Conclusion and further research

A theoretical approach of computing the fuzzy confidence interval based on the log-likelihood ratio is presented. An application illustrating this approach is given. In parallel, in the same setups, the interval by the “traditional” well-known expression was shown. For this calculation, specifying the parameter and the distribution were both needed. The obtained two intervals are then compared. We highlight that the defended approach is more flexible in terms of the possibility to adapt to a wider range of estimators without the need of specifying a given type of distributions. However, regardless of the computational complexity that the calculations through such methods might procure, our procedure gives a light way to get an elaborated fuzzy confidence interval. Indeed, our approach paves the way for a more accurate but severe one. Finally, for further researches, we would like to understand the influence of the use of such confidence intervals on decision making in fuzzy inference. A study on different parameters in multiple fuzzy inference methods can eventually be considered.

References

- [1] R. Berkachy, L. Donzé, Defuzzification of a fuzzy hypothesis decision by the signed distance method, in: Proceedings of the 61st World Statistics Congress, Marrakech, Morocco, 2017.

- [2] R. Berkachy, L. Donzé, Testing Hypotheses by Fuzzy Methods: A Comparison with the Classical Approach, in *Applying Fuzzy Logic for the Digital Economy and Society*, Springer International Publishing, Cham, 2019, pp. 1–22.
- [3] J. Chachi, S. M. Taheri, R. Viertl, Testing statistical hypotheses based on fuzzy confidence intervals, *Austrian Journal of Statistics* 41 (4) (2012) 267–286.
- [4] I. Couso, D. Dubois, Statistical reasoning with set-valued information: Ontic vs. epistemic views, *International Journal of Approximate Reasoning* 55 (7) (2014) 1502–1518.
- [5] I. Couso, L. Sánchez, Inner and outer fuzzy approximations of confidence intervals, *Fuzzy Sets and Systems* 184 (1) (2011) 68–83.
- [6] T. Denoeux, Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets and Systems* 183 (1) (2011) 72–91.
- [7] J. Gebhardt, M. A. Gil, R. Kruse, Fuzzy Set-Theoretic Methods in Statistics, in *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, Springer US, Boston, MA, 1998, pp. 311–347.
- [8] M. A. Gil, M. R. Casals, An operative extension of the likelihood ratio test from fuzzy data, *Statistical Papers* 29 (1) (1988) 191–203.
- [9] P. Grzegorzewski, Testing statistical hypotheses with vague data, *Fuzzy Sets and Systems* 112 (3) (2000) 501–510.
- [10] C. Kahraman, I. Otay, B. Öztayşi, Fuzzy Extensions of Confidence Intervals: Estimation for μ , σ^2 , and p , in *Fuzzy Statistical Decision-Making: Theory and Applications*, Springer International Publishing, Cham, 2016, pp. 129–154.
- [11] R. Kruse, K. D. Meyer, *Statistics with vague data*, Vol. 6, Springer Netherlands, 1987.
- [12] M. L. Puri, D. A. Ralescu, Fuzzy random variables, *Journal of Mathematical Analysis and Applications* 114 (2) (1986) 409–422.
- [13] R. Viertl, *Statistical Methods for Fuzzy Data*, John Wiley and Sons, Ltd, 2011.
- [14] R. Viertl, S. M. Yeganeh, Fuzzy Confidence Regions, in *Fuzzy Statistical Decision-Making: Theory and Applications*, Springer International Publishing, Cham, 2016, pp. 119–127.
- [15] L. Zadeh, Probability measures of fuzzy events, *Journal of Mathematical Analysis and Applications* 23 (2) (1968) 421–427.