# Fuzzy Confidence Intervals by Bootstrapped Fuzzy Distributions

Julien Rosset[1 [0000−0002−7883−1512]] and Laurent Donzé[1 [0000−0003−3522−4672]]

ASAM Group, Department of Informatics, University of Fribourg, Switzerland
Julien.Rosset@unifr.ch, Laurent.Donze@unifr.ch
https://www.unifr.ch/inf/asam

**Abstract.** We propose two practical procedures to construct fuzzy confidence intervals with fuzzy observations. These procedures use the fuzzy extensions of the quantile function and the concept of fuzzy empirical distribution. With bootstrapping techniques, we first compute fuzzy empirical distributions of the parameter of interest. We then build fuzzy confidence intervals by either finding the fuzzy quantiles of the distribution directly or using the fuzzy quantile function to find the relevant fuzzy quantiles. Our methods are illustrated through a numerical application. We construct fuzzy confidence intervals with the advocated methods and compare them to the fuzzy traditional way of creating such intervals and their construction using the likelihood ratio as in Berkachy and Donzé [2].

**Keywords:** Fuzzy Confidence Interval, Fuzzy Inference, Fuzzy Parameter, Fuzzy Distribution, Fuzzy Quantile

## 1 Introduction

The construction of fuzzy confidence intervals is nothing new. Indeed, for decades, various approaches have been proposed to build them. In the past, the suggested methods to find these so-called fuzzy confidence intervals were based on a predefined distribution around a specific parameter. Only recently have some more general approaches arisen, but they are often limited to a particular context. It thus would be very advantageous to generalise the previous constructions and have a practical tool to estimate fuzzy confidence intervals for any parameter with any distribution in any context. In the fuzzy world, most of the time, one cannot make any assumption on the fuzzy distribution of a fuzzy parameter. Therefore, we study the fuzzy extensions of the main concepts utilised to build classical confidence intervals, the fuzzy distribution and the fuzzy quantile function. We then provide two general methods to construct a fuzzy confidence interval. The idea behind these methods is first to use bootstrapping techniques adapted to the fuzzy world to generate a fuzzy empirical distribution of a fuzzy parameter, then find its fuzzy quantiles and use them to build fuzzy confidence intervals. We illustrated our methods with an empirical application. It allows comparing our findings with those obtained in Berkachy and Donzé [2].

Our work is organised as follows. We begin the study with a small literature review and define the notation in Sect. 2. The materials and methods for constructing fuzzy confidence intervals are provided in Sect. 3. The next section allows us to illustrate our two developed methods by comparing them to the traditional way of building fuzzy confidence intervals and their construction using the likelihood ratio. Details about the latter two can be found in Berkachy and Donzé [2]. Finally, Sect. 5 concludes our study.

## 2 Literatuve review and notation

In 1987, Kruse and Meyer [12] introduced a theoretical definition for the fuzzy confidence interval. Their theoretical definition has since then been used and refined by researchers in many computational procedures. For instance, Viertl and Yeganeh [15] generalised the concept of confidence regions for fuzzy data. Their work in this regard mainly concerns Bayesian statistics. Kahraman et al. [10], [11] studied interval-valued intuitionistic fuzzy sets (IVIFSs) and hesitant fuzzy sets (HFS). They then proposed two methods to construct fuzzy confidence intervals, one for interval-valued intuitionistic fuzzy sets and one for hesitant ones. Couso and Sanchez [7] provided an approach that considers the inner and outer approximations of confidence intervals in the context of fuzzy observations. Moreover, Wu [16] suggested a method to find fuzzy confidence intervals by considering unknown fuzzy parameters and fuzzy normal random variables through solving optimisation problems. In this direction, Chachi and Taheri. [6] also considered gaussian fuzzy random variables and built fuzzy confidence intervals for the mean of these gaussian fuzzy random variables. Based on Buckley's approach [5], Parchami and Mashinchi [13] used classical confidence intervals to model process capability indices as triangular fuzzy numbers. Unfortunately, these approaches are based on a pre-defined distribution around a specific parameter. To overcome this problem, Berkachy and Donzé [4] proposed a practical and general procedure to construct fuzzy confidence intervals with the use of the likelihood ratio method and a Bootstrap procedure extended to the fuzzy environment to estimate the distribution of this likelihood ratio.

Indeed, thanks to the work of Bradley Efron (See, e.g. [8]), Bootstrap methods became very popular among statisticians in the early 1980s. Then, various fuzzy Bootstrap approaches were found, extending the method to the fuzzy environment. To cite only a few relatively recent approaches, Berkachy and Donzé [4] provided two algorithms to constitute the bootstrapped samples mainly using the location and dispersion characteristics. On the other hand, Grzegorzewski and Romaniuk [9] described how to perform Bootstrap with epistemic fuzzy data.

Finally, let us mention two concepts closely linked to confidence intervals: the quantile function and the notion of distribution. Shvedov [14] extended the quantile function to the fuzzy world by defining the fuzzy quantile function, and Arefi et al. [1] gave different approaches to construct empirical fuzzy distributions.

We will use the following notation and conventions below. First, let us define by $\tilde{x}$ a fuzzy number and by $\mu_{\tilde{x}}(\cdot)$ its membership function. We consider also the $\alpha$-cuts of $\tilde{x}$ denoted by $\tilde{x}^{\alpha}$ or by its equivalent in interval form by $(x^{L,\alpha}, x^{R,\alpha})$. In practice, triangular fuzzy numbers are often used. We denote them by a triplet $\tilde{x} = (x^L, x, x^R)$ with $x^L \leq x \leq x^R \in \mathbb{R}$. Furthermore, assuming a random sample $X_1, \ldots, X_n$ with its realisations $x_1, \ldots, x_n$, we note by $\tilde{x}_1, \ldots, \tilde{x}_n$ the fuzzy equivalent of these quantities and by $\tilde{x}_i^{\alpha}$, $i = 1, \ldots, n$ their $\alpha$-cuts.

## 3 Fuzzy Confidence Intervals

This section will discuss two constructions of fuzzy confidence intervals for fuzzy parameters. Based on Kruse and Meyer [12], let us first define a fuzzy confidence interval (see, e.g. Berkachy and Donzé [2], Berkachy [3]). Recall that in a frequentist approach, the bounds of a confidence interval are functions of the sample. We note these functions by $\pi_i(\cdot), i = 1, 2$.

**Definition 1 (Fuzzy Confidence Interval).**
*Let $[\pi_1, \pi_2]$ be a symmetrical confidence interval for the parameter $\theta$ at significance level $\gamma$ such that $\pi_1$ and $\pi_2$ are functions of the observations, i.e. $\pi_1 := \pi_1(x_1, \ldots, x_n)$ and $\pi_2 := \pi_2(x_1, \ldots, x_n)$. A fuzzy confidence interval $\tilde{\sqcap}$ is a convex and normal fuzzy set, for which its respective left and right $\alpha$-cuts $[\sqcap^{L,\alpha}, \sqcap^{R,\alpha}]$ are defined as*

$$\sqcap^{L,\alpha} = \inf\{a \in \mathbb{R} : \exists x_i \in \tilde{x_i}^{\alpha}, \forall i = 1, ..., n, \text{ such that } \pi_1(x_1, ..., x_n) \leq a\},$$
$$\sqcap^{R,\alpha} = \sup\{a \in \mathbb{R} : \exists x_i \in \tilde{x_i}^{\alpha}, \forall i = 1, ..., n, \text{ such that } \pi_2(x_1, ..., x_n) \geq a\}.$$

*Its membership function $\mu_{\tilde{\sqcap}}(x)$ can be written as*

$$\mu_{\tilde{\sqcap}}(x) = \sup\{\alpha I_{[\sqcap^{L,\alpha}, \sqcap^{R,\alpha}]} : \alpha \in [0, 1]\}.$$

*This fuzzy confidence interval is said to belong to the $1 - \gamma$ confidence region such that for all parameter $\theta$, we have that*

$$\mathbb{P}(\sqcap^{L,\alpha} \leq \theta \leq \sqcap^{R,\alpha}) \geq 1 - \gamma, \forall \alpha \in [0, 1].$$

### 3.1 Fuzzy confidence interval: *method 1*

Let us consider a fuzzy estimator $\tilde{t}$ for the fuzzy parameter $\tilde{\theta}$. We can estimate $k$ new fuzzy parameters $\tilde{t^*}_1, \ldots, \tilde{t^*}_k$ by Bootstrap. This allows us to build an estimated empirical fuzzy distribution $\hat{\tilde{F}}$ with the $k$ new estimates plus the original one. We call this distribution "augmented" in the sense that one estimate comes from the original data, and the $k$ others are from the bootstrapped samples. Left and right parts of a fuzzy confidence interval at a given significance level $\gamma$ for each $\alpha$-cut are given by:

$$\hat{\theta}^{L,\alpha}_{\gamma/2} = t^{L,\alpha} - ({t^*}^{L,\alpha}_{(k+1)(1-\gamma/2)} - t^{L,\alpha}) = 2t^{L,\alpha} - {t^*}^{L,\alpha}_{(k+1)(1-\gamma/2)} \tag{1}$$

$$\hat{\theta}^{R,\alpha}_{1-\gamma/2} = t^{R,\alpha} - ({t^*}^{R,\alpha}_{(k+1)(\gamma/2)} - t^{R,\alpha}) = 2t^{R,\alpha} - {t^*}^{R,\alpha}_{(k+1)(\gamma/2)} \tag{2}$$

where, for the left $\alpha$-cut, ${t^*}^{L,\alpha}_{(k+1)(1-\gamma/2)}$ is the $(1-\gamma/2)$-quantile, and for the right $\alpha$-cut, ${t^*}^{R,\alpha}_{(k+1)(\gamma/2)}$ is the $\gamma/2$-quantile of the augmented distribution. The bootstrapped fuzzy confidence interval at confidence level $1-\gamma$ is given by $\tilde{\sqcap}^{\alpha} = [\sqcap^{L,\alpha}, \sqcap^{R,\alpha}] = [\hat{\theta}^{L,\alpha}_{\gamma/2}, \hat{\theta}^{R,\alpha}_{1-\gamma/2}]$. Let us prove that the latter is a fuzzy confidence interval.

*Proof.*

1. By construction, we have,

$$[(\hat{\theta}_{\gamma/2})^{L,\alpha_1}, (\hat{\theta}_{1-\gamma/2})^{R,\alpha_1}] \subset [(\hat{\theta}_{\gamma/2})^{L,\alpha_2}, (\hat{\theta}_{1-\gamma/2})^{R,\alpha_2}] \tag{3}$$

for $\alpha_2 < \alpha_1$, $\alpha_1, \alpha_2 \in (0,1]$. Let $\mu_{\tilde{\sqcap}}$ be the membership function of $\tilde{\sqcap}$. Choose $x_1 \in [(\hat{\theta}_{\gamma/2})^{L,\alpha_1}, (\hat{\theta}_{1-\gamma/2})^{R,\alpha_1}] = \tilde{\sqcap}^{\alpha_1}$ and $x_2 \in [(\hat{\theta}_{\gamma/2})^{L,\alpha_2}, (\hat{\theta}_{1-\gamma/2})^{R,\alpha_2}] = \tilde{\sqcap}^{\alpha_2}$. As $\alpha_2 < \alpha_1$, $\mu_{\tilde{\sqcap}}(x_1) \geq \min(\mu_{\tilde{\sqcap}}(x_2))$. Let $x_3 = tx_1 + (1-t)x_2$, $t \in [0,1]$. Then, $\mu_{\tilde{\sqcap}}(x_3) \geq \min(\mu_{\tilde{\sqcap}}(x_1), \mu_{\tilde{\sqcap}}(x_2))$, which proves the convexity of $\tilde{\sqcap}$.
2. First, observe that (3) implies that the left and right membership functions are respectively monotonically non-decreasing and non-increasing. Then, let $x \in [(\hat{\theta}_{\gamma/2})^{L,\alpha}, (\hat{\theta}_{1-\gamma/2})^{R,\alpha}], \quad \forall\alpha \in (0,1]$. For $\alpha = 1$, $\mu_{\tilde{\sqcap}}(x) = 1$. Thus, $\sup(\mu_{\tilde{\sqcap}}(x)) = 1$, which proves the normality;
3. Finally, by construction, it is evident that $\mathbb{P}(\sqcap^{L,\alpha} \leq \theta \leq \sqcap^{R,\alpha}) \geq 1-\gamma, \forall\alpha \in [0,1]$. □

### 3.2 Fuzzy confidence interval: *method 2*

The second method we propose consists of building quantile functions of a fuzzy random variable. Shvedov [14, p. 478] gives us a detailed definition.

Let us again consider a fuzzy estimator $\tilde{t}$ for the fuzzy parameter $\tilde{\theta}$. Suppose that $\tilde{\hat{F}}$ gives its fuzzy empirical distribution. Similar to method 1, a bootstrapping technique generates this one, augmented by the estimate based on the initial dataset. At a given $\alpha$-cut, we write that $\tilde{t}^{\alpha} = (t^{L,\alpha}, t, t^{R,\alpha}) \sim \tilde{\hat{F}}^{\alpha} = (\hat{F}^{L,\alpha}, \hat{F}, \hat{F}^{R,\alpha})$. The fuzzy $p$ quantiles of this fuzzy empirical distribution are the quantiles of the left, centred and right part of $\tilde{\hat{F}}$, i.e.

$$\tilde{Q}(p, \tilde{\hat{F}}) = \Big(Q(p, \hat{F}^{L,\alpha}), Q(p, \hat{F}), Q(p, \hat{F}^{R,\alpha})\Big), \quad \alpha \in (0,1],$$

with

$$Q(p, \hat{F}^{L,\alpha}) \leq Q(p, \hat{F}) \leq Q(p, \hat{F}^{R,\alpha}) \tag{4}$$

and where $Q(\cdot)$ is the quantile function.

Based on these quantities, assuming $p < (1-p)$, three intervals can be built:

$$I_p^{U,\alpha} = [Q(p, \hat{F}^{L,\alpha}); Q(1-p, \hat{F}^{R,\alpha})] \tag{5}$$

$$I_p = [Q(p, \hat{F}); Q(1-p, \hat{F})] \tag{6}$$

$$I_p^{L,\alpha} = \begin{cases} [Q(p, \hat{F}^{R,\alpha}); Q(1-p, \hat{F}^{L,\alpha})] & \text{if} \quad Q(p, \hat{F}^{R,\alpha}) < Q(1-p, \hat{F}^{L,\alpha}) \\ \emptyset \quad \text{otherwise.} & \end{cases} \tag{7}$$

One can verify that

$$I_p^{U,\alpha} \supseteq I_p \supseteq I_p^{L,\alpha}.$$

The interval $I_p$ gives us a classical crisp interval and can be interpreted as the core of a fuzzy interval. The limits of this fuzzy interval are those of $I_p^{U,\alpha}$. The last interval, $I_p^{L,\alpha}$, has no interest because it is included in the core.

*Proof.* Analogue to the latter one. □

## 4 Numerical Application

Let us discuss and compare the resulting fuzzy confidence intervals for the mean $\tilde{\bar{x}} = (1.8, 2.8, 3.8)$ from Table 2 arising from our two methods described above, the fuzzy traditional way of constructing such intervals and by the likelihood ratio. The latter two are explained thoroughly in Berkachy [3, p. 101–112] and Berkachy and Donzé [2]. The set of observations shown in Table 1 used to build such confidence intervals for the mean is from Berkachy and Donzé [2].

### 4.1 Discussion and comparison

First, notice how both the cores and left and right parts of the fuzzy confidence intervals from *methods 1* and *2* encapsulate the ones arising from the fuzzy traditional method and the likelihood ratio approach. Thus, in this example, *methods 1* and *2* generate larger, i.e. fuzzier confidence intervals. Moreover, we can also say that the core of the confidence interval found by the likelihood ratio approach is different than the cores found with the three other methods. This effect could come from the fuzzy modelling choice of the maximum likelihood estimator needed in the likelihood ratio approach. When comparing the results of *methods 1* and *2*, we can see that the left and right parts of the fuzzy confidence interval given by *method 1* are the left and right parts found by using *method 2* minus a factor $0.0\bar{6}$. The reason behind this effect can be understood by looking at (1) and (2). Indeed, one observes that to build confidence intervals, *method 1* starts from the value of the original estimate and then uses the quantiles of the augmented distribution of the mean. This thus yields a slightly different result than *method 2*, which only considers the quantiles of the mean's augmented distribution when constructing such confidence intervals, as described by (6) and (7).

## 5 Conclusion

Based on the work of Berkachy and Donzé [2] and Shvedov's fuzzy quantile function [14], we developed two methods to find fuzzy confidence intervals for any parameter and any distribution. The idea behind the two methods consisted of constructing an augmented fuzzy empirical distribution of a fuzzy parameter with bootstrapping techniques adapted to the fuzzy environment and then finding the fuzzy quantiles of this distribution. The main difference between the two methods is that the first considers the original estimate when constructing the fuzzy intervals. In contrast, the second method directly picks the quantiles of the augmented distribution to build fuzzy confidence intervals. We also provided an empirical application. We discuss the results and compare them with other methods. This comparison showed that our developed methods give larger fuzzy confidence intervals. Lastly, given that our methods are simple, easy to implement, and present no particular computational difficulties, they seem to be a promising way to build fuzzy statistical inference procedures.

**Table 1.** Fuzzy observations (See Berkachy and Donzé [2])

| **Index** | $x_i$ | $\tilde{x}_i$ |
|---|---|---|
| **1** | 4 | $(3, 4, 5)$ |
| **2** | 1 | $(0, 1, 2)$ |
| **3** | 3 | $(2, 3, 4)$ |
| **4** | 2 | $(1, 2, 3)$ |
| **5** | 3 | $(2, 3, 4)$ |
| **6** | 2 | $(1, 2, 3)$ |
| **7** | 5 | $(4, 5, 6)$ |
| **8** | 2 | $(1, 2, 3)$ |
| **9** | 3 | $(2, 3, 4)$ |
| **10** | 3 | $(2, 3, 4)$ |
| | $\bar{x} = 2.8$ | $\tilde{\bar{x}} = (1.8, 2.8, 3.8)$ |

**Table 2.** Fuzzy confidence intervals for the fuzzy mean $(1.8, 2.8, 3.8)$ at confidence level $1 - \gamma = 1 - 0.05$ obtained by the fuzzy traditional method, the likelihood ratio approach, the *methods 1* and *2*.

| **Methods** | **Fuzzy Confidence Interval** |
|---|---|
| Fuzzy traditional method | (1.0965, 2.0965, 3.5034, 4.5034 ) |
| Likelihood ratio approach | (1.3674, 2.2175, 3.3838, 4.2332 ) |
| *Method 1* | (0.9333, 1.9333, 3.6000, 4.6000 ) |
| *Method 2* | (1.0000, 2.0000, 3.6667, 4.6667 ) |

## References


1. Arefi, M., Viertl, R., Taheri, S.M.: Fuzzy density estimation. Metrika **75**(1), 5–22 (2012). https://doi.org/10.1007/s00184-010-0311-y
2. Berkachy, R., Donzé, L.: Fuzzy Confidence Intervals by the Likelihood Ratio with Bootstrapped Distribution. Proceedings of the 12th International Joint Conference on Computational Intelligence pp. 231–242 (2020). https://doi.org/10.5220/0010023602310242
3. Berkachy, R.: The Signed Distance Measure in Fuzzy Statistical Analysis. Some theoretical, empirical and programming advances. Ph.D. thesis, University of Fribourg (2020)
4. Berkachy, R., Donzé, L.: Fuzzy Confidence Intervals by the Likelihood Ratio: Testing Equality of Means—Application on Swiss SILC Data. SN Computer Science **3**(5), 374 (2022). https://doi.org/10.1007/s42979-022-01257-z
5. Buckley, J.J.: Fuzzy Statistics. Springer (2004). https://doi.org/10.1007/978-3-540-39919-3
6. Chachi, J., Taheri, S.M.: Fuzzy confidence intervals for mean of Gaussian fuzzy random variables. Expert Systems with Applications **38**(5), 5240–5244 (2011). https://doi.org/10.1016/j.eswa.2010.10.034
7. Couso, I., Sánchez, L.: Inner and outer fuzzy approximations of confidence intervals. Fuzzy Sets and Systems **184**(1), 68–83 (2011). https://doi.org/10.1016/j.fss.2010.11.004
8. Efron, B.: Bootstrap methods. another look at the jacknife. The Annals of Statistics **7**(1), 1–26 (1979), https://www.jstor.org/stable/2958830
9. Grzegorzewski, P., Romaniuk, M.: Epistemic Bootstrap for Fuzzy Data. Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and the 11th International Summer School on Aggregation Operators (AGOP) pp. 538–545 (2021). https://doi.org/10.2991/asum.k.210827.071
10. Kahraman, C., Otay, I., Öztayşi, B.: Fuzzy Extensions of Confidence Intervals: Estimation for $\mu$, $\sigma^2$, and $p$, pp. 129–154. Springer International Publishing, Cham (2016), https://doi.org/10.1007/978-3-319-39014-7_9
11. Kahraman, C., Oztaysi, B., Onar, S.C.: Interval-Valued Intuitionistic Fuzzy Confidence Intervals. Journal of Intelligent Systems **28**(2), 307–319 (2019). https://doi.org/10.1515/jisys-2017-0139
12. Kruse, R., Meyer, K.D.: Statistics with Vague Data. D. Reidel Publishing Company (1987). https://doi.org/10.1007/978-94-009-3943-1
13. Parchami, A., Mashinchi, M.: Fuzzy estimation for process capability indices. Information Sciences **177**(6), 1452–1462 (2007). https://doi.org/10.1016/j.ins.2006.08.016
14. Shvedov, A.S.: Quantile function of a fuzzy random variable and an expression for expectations. Mathematical Notes **100**(3-4), 477–481 (2016). https://doi.org/10.1134/s0001434616090157
15. Viertl, R., Yeganeh, S.M.: Fuzzy Confidence Regions, pp. 119–127. Springer International Publishing, Cham (2016), https://doi.org/10.1007/978-3-319-39014-7_8
16. Wu, H.C.: Statistical confidence intervals for fuzzy data. Expert Systems with Applications **36**(2), 2670–2676 (2009). https://doi.org/10.1016/j.eswa.2008.01.022