6-1-2022

Estimating Skewness and Higher Central Moments of an Interval-Valued Fuzzy Set

Juan Carlos Figueroa-Garcia
Universidad Distrital Bogata, jcfigueroag@udistrital.edu.co

Martine Ceberio
The University of Texas at El Paso, mceberio@utep.edu

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:

Technical Report: UTEP-CS-22-79

Estimating Skewness and Higher Central Moments of an Interval-Valued Fuzzy Set

Juan Carlos Figueroa Garcia, Martine Ceberio, Olga Kosheleva, and Vladik Kreinovich

Abstract A known relation between membership functions and probability density functions allows us to naturally extend statistical characteristics like central moments to the fuzzy case. In case of interval-valued fuzzy sets, we have several possible membership functions consistent with our knowledge. For different membership functions, in general, we have different values of the central moments. It is therefore desirable to compute, in the interval-valued fuzzy case, the range of possible values for each such moment. In this paper, we provide efficient algorithms for this computation.

1 Outline

From the purely mathematical viewpoint, the main difference between a membership function $\mu(x)$ (see, e.g., [1, 2, 3, 4, 5, 7]) and a probability density function $f(x)$ is that they are normalized differently:

- for $\mu(x)$, we require that its maximum is equal to 1, while
- for $f(x)$, we require that its integral is equal to 1.

Both functions can be re-normalized, so there is a natural probability density function assigned to each membership function, and vice versa. This assignments allow to naturally extend probabilistic notions to membership functions, and define mean (which turns out to be equivalent to a centroid), skewness, etc.

Juan Carlos Figueroa Garcia
Departamento de Ingenieria Industrial, Universidad Distrital
Bogota, Colombia, e-mail: jcfigueroag@udistrital.edu.co

Martine Ceberio, Olga Kosheleva, and Vladik Kreinovich
University of Texas at El Paso, 500 W. University, El Paso, Texas 79968, USA
e-mail: mceberio@utep.edu, olgak@utep.edu, vladik@utep.edu

These transformation implicitly assume that for each x, we can extract the exact value $\mu(x)$ from the expert. In practice, the experts can, at best, provide an interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$ of possible values of $\mu(x)$. This case – known as interval-valued fuzzy case – means that we have many different membership functions $\mu(x)$ as long as we have $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for all x. For different possible membership functions, we have different values of skewness and of other characteristics. It is therefore desirable to estimate the range of possible values of such a characteristic. In this paper, we provide efficient algorithms for such an estimation.

The structure of this paper is as follows. In Section 2, we formulate the problem in precise terms. In Section 3, we analyze the resulting computational problem for the case of skewness. In Section 4, we use this analysis to describe the resulting algorithm. In Section 5, we extend this algorithm to all central moments.

2 Formulation of the Problem

Known relation between probabilistic and fuzzy uncertainty. As Zadeh himself, the father of fuzzy logic, emphasized several times, the same data can be described by a probability density function $f(x)$ and a membership function $\mu(x)$. The main difference between these two descriptions is in the normalization:

- for a probability density function, the normalizing requirement is that the overall probability should be equal to 1:

$$\int f(x)\,dx = 1, \tag{1}$$

- while for a membership function, the normalizing requirement is that the largest value of this function must be equal to 1:

$$\max_x \mu(x) = 1. \tag{2}$$

Because of this relation, it is possible to re-normalized each of these functions by multiplying it by an appropriate constant. For example, to each membership function $\mu(x)$, we can assign the corresponding probability density function:

$$f(x) = \frac{\mu(x)}{\int \mu(y)\,dy}. \tag{3}$$

This allows us to naturally extend probabilistic characteristics to the fuzzy case. The transformation (3) enables us to naturally extend known characteristics of a probability distribution – such as its central moments (see, e.g., [6]) – to the fuzzy case. The first moment – the mean value – is defined as:

$$M_1 = \int x \cdot f(x)\, dx. \tag{4}$$

Substituting the expression (3) for $f(x)$ into this formula, we get the corresponding fuzzy expression:

$$M_1 = \frac{\int x \cdot \mu(x)\, dx}{\int \mu(x)\, dx}. \tag{5}$$

Interestingly, this is exactly what is called *centroid defuzzification*.

Similarly, for any natural number $n \geq 2$, we can plug in the expression (3) into a formula for the n-th order moment:

$$M_n = \int (x - M_1)^n \cdot f(x)\, dx, \tag{6}$$

and get the corresponding fuzzy expression:

$$M_n = \frac{\int (x - M_1)^n \cdot \mu(x)\, dx}{\int \mu(x)\, dx}. \tag{7}$$

Need for interval-valued fuzzy. Traditional $[0, 1]$-based fuzzy approach implicitly assumes that for each x, we can extract the exact value $\mu(x)$ from the expert. In practice, the experts cannot very accurately gauge their degrees of confidence. At best, they provide an interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$ of possible values of $\mu(x)$.

This case – known as interval-valued fuzzy case – means that we have many different membership functions $\mu(x)$ as long as we have $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for all x.

For different possible membership functions, we have different values of skewness M_3 and of other central moments M_n. It is therefore desirable to estimate the range of possible values of each of these characteristics. In this paper, we provide efficient algorithms for such an estimation.

3 Case of Skewness: Analysis of the Problem

Case of skewness: explicit formula. Let us start with the case when $n = 3$. In this case, the formula (7) takes the form

$$M_3 = \frac{\int (x - M_1)^3 \cdot \mu(x)\, dx}{\int \mu(x)\, dx}. \tag{8}$$

Here,

$$(x - M_1)^2 = x^3 - 3 \cdot x^2 \cdot M_1 + 3 \cdot x \cdot M_1^2 - M_1^3, \tag{9}$$

thus the expression (9) takes the form

$$M_3 = \frac{\int x^3 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - 3 \cdot M_1 \cdot \frac{\int x^2 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} + 3 \cdot M_1^2 \cdot \frac{\int x \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - M_1^3. \quad (10)$$

Substituting the formula (5) into this expression, we get

$$M_3 = \frac{\int x^3 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - 3 \cdot \frac{\left(\int x^2 \cdot \mu(x)\,dx\right) \cdot \left(\int x \cdot \mu(x)\,dx\right)}{\left(\int \mu(x)\,dx\right)^2} +$$

$$3 \cdot \frac{\left(\int x \cdot \mu(x)\,dx\right)^3}{\left(\int \mu(x)\,dx\right)^3} - \frac{\left(\int x \cdot \mu(x)\,dx\right)^3}{\left(\int \mu(x)\,dx\right)^3} =$$

$$\frac{\int x^3 \cdot \mu(x)\,dx}{\int \mu(x)\,dx} - 3 \cdot \frac{\left(\int x^2 \cdot \mu(x)\,dx\right) \cdot \left(\int x \cdot \mu(x)\,dx\right)}{\left(\int \mu(x)\,dx\right)^2} + 2 \cdot \frac{\left(\int x \cdot \mu(x)\,dx\right)^3}{\left(\int \mu(x)\,dx\right)^3}. \quad (11)$$

Basic facts from calculus: reminder. In general, according to calculus, a function $F(v)$ attains its maximum with respect to the input $v \in [\underline{v}, \overline{v}]$ in one of the three cases:

- It can be that this maximum is attained inside the interval; in this case, at this point the derivative $F'(v)$ of the maximized function is equal to 0.
- It can be that this maximum is attained at the lower endpoint $\underline{v}$ of the given interval. In this case, the derivative $F'(v)$ at this point has to be non-positive: otherwise, if this derivative was positive, then for a sufficiently small $\varepsilon > 0$, at a point $\underline{v} + \varepsilon$ we would have larger values of $F(v)$ – which contradicts to our assumption that the largest value of the function $F(v)$ is attained for $v = \underline{v}$.
- It can also be that this maximum is attained at the upper endpoint $\overline{v}$ of the given interval. In this case, the derivative $F'(v)$ at this point has to be non-negative: otherwise, if this derivative was negative, then for a sufficiently small $\varepsilon > 0$, at a point $\overline{v} - \varepsilon$ we would have larger values of $F(v)$ – which contradicts to our assumption that the largest value of the function $F(v)$ is attained for $v = \overline{v}$.

Similarly, a function $F(v)$ attains its minimum with respect to the input $v \in [\underline{v}, \overline{v}]$ in one of the three cases:

- It can be that this minimum is attained inside the interval; in this case, at this point the derivative $F'(v)$ of the maximized function is equal to 0.
- It can be that this minimum is attained at the lower endpoint $\underline{v}$ of the given interval. In this case, the derivative $F'(v)$ at this point has to be non-negative: otherwise, if this derivative was negative, then for a sufficiently small $\varepsilon > 0$, at a point $\underline{v} + \varepsilon$ we would have smaller values of $F(v)$ – which contradicts to our assumption that the smallest value of the function $F(v)$ is attained for $v = \underline{v}$.

- It can also be that this minimum is attained at the upper endpoint $\overline{v}$ of the given interval. In this case, the derivative $F'(v)$ at this point has to be non-positive: otherwise, if this derivative was positive, then for a sufficiently small $\varepsilon > 0$, at a point $\overline{v} - \varepsilon$ we would have smaller values of $F(v)$ – which contradicts to our assumption that the smallest value of the function $F(v)$ is attained for $v = \overline{v}$.

Let us apply this idea to our case. In our case, unknown are values $\mu(x)$. For each expression

$$I_k \stackrel{\text{def}}{=} \int x^k \cdot \mu(x)\, dx, \tag{12}$$

the derivative with respect to $\mu(x)$ is equal to

$$\frac{\partial I_k}{\partial \mu(x)} = x^k. \tag{13}$$

Thus, for the expression (11) – which, in terms of the notations I_k has the form

$$M_3 = \frac{I_3}{I_0} - 3 \cdot \frac{I_2 \cdot I_1}{I_0^2} + 2 \cdot \frac{I_1^3}{I_0^3}, \quad (14$$

the derivative $d(x)$ with respect to $\mu(x)$ takes the following form:

$$d(x) = \frac{x^3 \cdot I_0 - I_3}{I_0^2} - 2 \cdot \frac{x^2 \cdot I_1 \cdot I_0^2 + I_2 \cdot x \cdot I_0^2 - 2I_0}{I_0^4} +$$

$$2 \frac{3I_1 \cdot x \cdot I_0^3 - I_3 \cdot 3I_0^2}{I_0^6}. \tag{15}$$

In other words, this derivative is a cubic polynomial

$$d(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3, \tag{16}$$

with a positive coefficient $a_3 = 1/I_0$ at x^3.

It is known that a cubic function has either 1 or 3 roots, i.e., in general, we have values $x_1 \le x_2 \le x_3$ at which $d(x) = 0$. Since the coefficient at x^3 is positive:

- we have $d(x) < 0$ for $x < x_1$,
- we have $d(x) > 0$ for $x_1 < x < x_2$,
- we have $d(x) < 0$ for $x_2 < x < x_3$, and
- we have $d(x) > 0$ for $x > x_3$.

When $d(x) > 0$, then, as we have mentioned:

- maximum cannot be attained inside the interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$, and it cannot be attained at the lower endpoint of this interval, so maximum has to be attained for $\mu(x) = \overline{\mu}(x)$;

- similarly, minimum cannot be attained inside the interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$, and it cannot be attained at the upper endpoint of this interval, so minimum has to be attained for $\mu(x) = \underline{\mu}(x)$.

When $d(x) < 0$:

- maximum cannot be attained inside the interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$, and it cannot be attained at the upper endpoint of this interval, so maximum has to be attained for $\mu(x) = \underline{\mu}(x)$;
- similarly, minimum cannot be attained inside the interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$, and it cannot be attained at the lower endpoint of this interval, so minimum has to be attained for $\mu(x) = \overline{\mu}(x)$.

Thus, once we know the values x_i, we can uniquely determine the membership functions $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ at which the skewness attains, for these x_i, its largest and smallest values. So, to find the overall largest and smallest values of the skewness, it is sufficient to find the values x_i for which the resulting skewness is the largest and for which it is the smallest. Hence, we arriove at the following algorithm for computing the range $\left[\underline{M}_3, \overline{M}_3\right]$ of possible values of skewness.

4 Resulting Algorithm for Computing the Range $\left[\underline{M}_3, \overline{M}_3\right]$ of Possible Values of Skewness

What is given and what we want. For each x, we have an interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$. We want to find the range $\left[\underline{M}_3, \overline{M}_3\right]$ of all possible values of the skewness M_3 – as defined by the formulas (5) and (8) – over all functions $\mu(x)$ for which $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for all x.

Computing $\overline{M}_3$. For each triple of real numbers $x_1 \le x_2 \le x_3$, we compute the skewness $\overline{M}_3(x_1, x_2, x_3)$ corresponding to the following membership function:

- when $x < x_1$ or $x_2 < x < x_3$, we take $\mu(x) = \underline{\mu}(x)$; and
- we take $\mu(x) = \overline{\mu}(x)$ when $x_1 < x < x_2$ or $x > x_2$.

Then, we use an optimization algorithm to compute

$$\overline{M}_3 = \max_{x_1, x_2, x_3} \overline{M}_3(x_1, x_2, x_3). \tag{17}$$

Computing $\underline{M}_3$. For each triple of real numbers $x_1 \le x_2 \le x_3$, we compute the skewness $\underline{M}_3(x_1, x_2, x_3)$ corresponding to the following membership function:

- when $x < x_1$ or $x_2 < x < x_3$, we take $\mu(x) = \overline{\mu}(x)$; and
- we take $\mu(x) = \underline{\mu}(x)$ when $x_1 < x < x_2$ or $x > x_2$.

Then, we use an optimization algorithm to compute

$$\underline{M}_3 = \min_{x_1,x_2,x_3} \underline{M}_3(x_1, x_2, x_3). \tag{18}$$

5 General Case of Arbitrary Central Moments

What is given and what we want. For each x, we have an interval $\left[\underline{\mu}(x), \overline{\mu}(x)\right]$. We are also given a natural number n. We want to find the range $\left[\underline{M}_n, \overline{M}_n\right]$ of all possible values of the n-th central moment M_n – as defined by the formulas (5) and (7) – over all functions $\mu(x)$ for which $\mu(x) \in \left[\underline{\mu}(x), \overline{\mu}(x)\right]$ for all x.

Discussion. In the general case of the n-th order central moment, similar arguments leads to an n-th order polynomial $d(x)$ with a positive coefficient at x^n. Thus, similar arguments lead to the following algorithm.

Computing $\overline{M}_n$. For each n-tuple of real numbers $x_1 \le x_2 \le \ldots \le x_n$, we take $x_0 \stackrel{\text{def}}{=} -\infty$ and $x_{n+1} \stackrel{\text{def}}{=} +\infty$. Thus, the whole real line is divided into $n+1$ intervals (x_0, x_1), $(x_1, x_2), \ldots, (x_n, x_{n+1})$. Then, we compute the n-th central moment $\overline{M}_n(x_1, \ldots, x_n)$ corresponding to the following membership function:

- for $x \in (x_k, x_{k+1})$ for which $n - k$ is odd, we take $\mu(x) = \underline{\mu}(x)$ and
- when $n - k$ is even, we take $\mu(x) = \overline{\mu}(x)$.

Then, we use an optimization algorithm to compute

$$\overline{M}_n = \max_{x_1,\ldots,x_n} \overline{M}_n(x_1, \ldots, x_n). \tag{19}$$

Computing $\underline{M}_n$. For each n-tuple of real numbers $x_1 \le x_2 \le \ldots \le x_n$, we take $x_0 \stackrel{\text{def}}{=} -\infty$ and $x_{n+1} \stackrel{\text{def}}{=} +\infty$. Thus, the whole real line is divided into $n+1$ intervals (x_0, x_1), $(x_1, x_2), \ldots, (x_n, x_{n+1})$. Then, we compute the n-th central moment $\underline{M}_n(x_1, \ldots, x_n)$ corresponding to the following membership function:

- for $x \in (x_k, x_{k+1})$ for which $n - k$ is odd, we take $\mu(x) = \overline{\mu}(x)$ and
- when $n - k$ is even, we take $\mu(x) = \underline{\mu}(x)$.

Then, we use an optimization algorithm to compute

$$\underline{M}_n = \min_{x_1,\ldots,x_n} \underline{M}_n(x_1, \ldots, x_n). \tag{20}$$

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
2. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
3. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
4. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
5. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
6. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
7. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.