

# From Computation with Guaranteed Intervals to Computation with Confidence Intervals: A New Application of Fuzzy Techniques

Vladik Kreinovich<sup>1</sup>, Hung T. Nguyen<sup>2</sup>, Scott Ferson<sup>3</sup>, and Lev Ginzburg<sup>3</sup>

<sup>1</sup>Department of Computer Science, University of Texas at El Paso,  
El Paso, TX 79968, USA, vladik@cs.utep.edu

<sup>2</sup>Department of Mathematical Sciences,  
New Mexico State University, Las Cruces, NM 88003, USA

<sup>3</sup>Applied Biomathematics, 100 North Country Road  
Setauket, NY 11733, USA

## Abstract

Traditional interval computations provide an estimate for the result  $y = f(x_1, \dots, x_n)$  of data processing when we know intervals  $\mathbf{x}_1, \dots, \mathbf{x}_n$  that are guaranteed to contain the (unknown) actual values of the quantities  $x_1, \dots, x_n$ . Often, in addition to these *guaranteed* intervals, we have *confidence* intervals for these quantities, i.e., intervals  $\mathbf{x}_i$  that contain the corresponding values  $x_i$  with a certain probability. It is desirable, based on the confidence intervals for  $x_i$ , to produce the resulting confidence interval for  $y$ . It turns out that the formulas for computing such resulting confidence interval are closely related with the formulas for processing fuzzy numbers by using Zadeh's extension principle. Thus, known algorithms for processing fuzzy data can be used to process confidence intervals as well.

## 1 Why intervals?

Most information about real-life quantities come from measurements. Measurements are never 100% accurate; as a result, the measurement result  $\tilde{x}$  differs from the actual (unknown) value of the measured quantity  $x$ . In many real-life situations, the only information that we have about the measurement error  $\tilde{x} - x$  is the upper bound  $\Delta$  provided by the manufacturer. In such situations, after measurement, the only information that we have about  $x$  is that  $x$  belongs to the interval  $[\tilde{x} - \Delta, \tilde{x} + \Delta]$ .

## 2 Why process intervals?

Some quantities  $y$ , like a distance to a star or the amount of oil in a given area, are impossible (or difficult) to measure directly. To gauge  $y$ , we measure quantities  $x_1, \dots, x_n$  related to  $y$ , and then use the known relationship  $y = f(x_1, \dots, x_n)$  between  $x_i$  and  $y$  to reconstruct  $y$ .

When after each measurement, we only know the intervals  $\mathbf{x}_i$  of possible values of  $x_i$ , we must therefore find the interval of possible values of  $y = f(x_1, \dots, x_n)$  when  $x_1$  is in the interval  $\mathbf{x}_1$ ,  $\dots$ , and  $x_n$  is in the interval  $\mathbf{x}_n$ , i.e., we must transform the intervals  $\mathbf{x}_i$  into a new interval  $\mathbf{y}$ . Techniques for processing interval data – called “interval computations” – are successfully used in application areas ranging from space exploration to robotics to chemical engineering; see, e.g., [2, 3, 4].

## 3 From guaranteed intervals to confidence intervals

Interval computations handle cases when the upper bounds  $\Delta$  on the measurement errors are guaranteed and when, therefore, it is guaranteed that the actual value  $x$  of the measured quantity belongs to the corresponding interval  $[\tilde{x} - \Delta, \tilde{x} + \Delta]$ . In many practical situations, however, the bounds (and, correspondingly, the intervals) are not guaranteed. Instead of a single interval, we have several intervals that contain  $x$  with different confidence. We may have an interval that contains  $x$  with probability 99%, a slightly narrower interval that contains  $x$  with probability 98%, etc.

For indirect measurement, if we have such information about each of the variables  $x_1, \dots, x_n$ , what can we conclude about the value of the desired quantity  $y$ ?

The difficulty comes from the fact that, e.g., to get an interval that contains  $y$  with probability 99%, we cannot simply apply interval computations to the intervals  $\mathbf{x}_1$  and  $\mathbf{x}_2$  that contain, correspondingly,  $x_1$  and  $x_2$  with probability 99%: indeed, the resulting interval would include  $y$  with the probability that both  $x_1$  is in  $\mathbf{x}_1$  and  $x_2$  is in  $\mathbf{x}_2$ , and this probability of a joint event may be  $< 99\%$ .

We must therefore generalize traditional interval methods to confidence intervals. This generalization is presented in the paper.

## 4 Relationship to fuzzy techniques

On the qualitative level, there is a clear relation between confidence intervals and fuzzy numbers:

- Confidence intervals are a “nested” collection of intervals corresponding to different levels of confidence probability.
- A fuzzy number can also be represented as a “nested” collection of  $\alpha$ -cuts, i.e., intervals corresponding to different thresholds of membership value (see, e.g., [5, 6]).

This seemingly natural qualitative relationships is known because it was sometimes used by statisticians in their past claims that problems solved by fuzzy logic can be solved even better by statistical techniques.

It turns out that this qualitative relationship can be actually upgraded to a quantitative one.

## 5 Input data for processing confidence intervals

Let us assume that a variable  $x$  is guaranteed to lie within an interval  $[x^-, x^+]$ . For every possible confidence level  $\beta \in (0, 1)$ , the corresponding confidence interval is defined as an interval that contains  $x$  with probability  $\geq \beta$ . In other words, a confidence interval  $[X^-, X^+]$  of confidence level  $\beta$  is defined as an interval for which the probability  $p_{\text{out}}$  to be outside  $[X^-, X^+]$  does not exceed  $\alpha \stackrel{\text{def}}{=} 1 - \beta$ .

There are two possibilities for  $x$  to be outside the interval  $[X^-, X^+]$ :

- when  $x$  is smaller than the lower bound  $X^-$ , and
- when  $x$  is larger than the upper bound  $X^+$ .

Thus, the probability  $p_{\text{out}}$  that  $x$  is outside the confidence interval is equal to the sum of the two probabilities:

- the probability that  $x$  is smaller than the lower bound  $X^-$  and
- the probability that  $x$  is larger than the upper bound  $X^+$ .

For the same probability distribution and for the same confidence level  $\beta$ , we can have different confidence intervals:

- we can set  $X^- = x^-$  and find the value  $X^+$  for which the probability

$$\text{Prob}[x \leq X^+]$$

that  $x \leq X^+$  is equal to  $\beta$ ;

- alternatively, we can set  $X^+ = x^+$  and find the value  $X^-$  for which the probability

$$\text{Prob}[x \geq X^-]$$

that  $x \geq X^-$  is equal to  $\beta$ ;

- we can also find  $X^-$  for which the probability that  $x < X^-$  is equal to  $\alpha/2 = (1 - \beta)/2$  and  $X^+$  for which the probability that  $x > X^+$  is equal to  $\alpha/2$ :

$$\text{Prob}[x < X^-] = \text{Prob}[x > X^+] = \alpha/2.$$

There are many other options.

In view of these various possibilities, before we start describing computations with such intervals, we must decide which confidence intervals we will consider. the above three examples provide us with the following three possibilities. For each confidence level  $\beta$ , we can define the corresponding confidence interval as:

- an interval  $[x^-, x^+(\beta)]$ , where  $x^+(\beta)$  is defined as the value for which  $\text{Prob}[x \leq x^+(\beta)] = \beta$ ;

- an interval  $[x^-(\beta), x^+]$ , where  $x^-(\beta)$  is defined as the value for which  $\text{Prob}[x \geq x^-(\beta)] = \beta$ ;
- an interval  $[x_{\text{sym}}^-(\beta), x_{\text{sym}}^+(\beta)]$ , where  $x^-(\beta)$  is defined as the value for which

$$\text{Prob}[x < x_{\text{sym}}^-(\beta)] = (1 - \beta)/2$$

and

$$\text{Prob}[x > x_{\text{sym}}^+(\beta)] = (1 - \beta)/2.$$

These three different possibilities express the same information, albeit in different form:

- If we know the values  $x^+(\beta)$  for all  $\beta$ , then (provided that the distribution is non-atomic), we can reconstruct:
  - $x^-(\beta)$  as  $x^+(1 - \beta)$ , and
  - $x_{\text{sym}}^-(\beta)$  and  $x_{\text{sym}}^+(\beta)$  as, correspondingly,  $x^+((1 - \beta)/2)$  and  $x^+(1 - (1 - \beta)/2)$ .
- If we know the values  $x^-(\beta)$  for all  $\beta$ , then we can find:
  - $x^+(\beta)$  as  $x^-(1 - \beta)$ , and
  - $x_{\text{sym}}^-(\beta)$  and  $x_{\text{sym}}^+(\beta)$  as, correspondingly,  $x^-(1 - (1 - \beta)/2)$  and  $x^-(1 - \beta)/2$ .
- Finally, if we know the values  $x_{\text{sym}}^-(\beta)$  and  $x_{\text{sym}}^+(\beta)$  for all  $\beta$ , then:
  - when  $\beta \geq 0.5$ , we can reconstruct  $x^-(\beta)$  as  $x_{\text{sym}}^-(2\beta - 1)$  and  $x^+(\beta)$  as  $x_{\text{sym}}^+(2\beta - 1)$ ;
  - when  $\beta < 0.5$ , we can reconstruct  $x^-(\beta)$  as  $x_{\text{sym}}^+(1 - 2\beta)$  and  $x^+(\beta)$  as  $x_{\text{sym}}^-(1 - 2\beta)$ .

It is therefore sufficient to consider only one of these representations, e.g., the representation by intervals  $[x^-, x^+(\beta)]$ .

For simplicity, we will assume that the function  $f(x_1, \dots, x_n)$  is a non-decreasing function of all its variables.

## 6 Possible cases

In this paper, we will consider two possible cases:

- when the probabilities corresponding to different variables  $x_1, \dots, x_n$  are independent, and
- when we have no information about the possible correlation between these probabilities.

## 7 Main result for independent case

Let us first consider the case when the probabilities corresponding to different variables  $x_1, \dots, x_n$  are independent. Let us fix  $\beta$  and try to find, for  $y = f(x_1, \dots, x_n)$ , the value  $y^+(\beta)$  for which the probability  $\text{Prob}[y \leq y^+(\beta)]$  is equal to  $\beta$ .

For every tuple  $(\beta_1, \dots, \beta_n)$ , we have:

- $x_1 \in [x_1^-, x_1^+(\beta_1)]$  – i.e.,  $x_1 \leq x_1^+(\beta_1)$  – with probability  $\beta_1$ ;
- ...
- $x_n \in [x_n^-, x_n^+(\beta_n)]$  – i.e.,  $x_n \leq x_n^+(\beta_n)$  – with probability  $\beta_n$ .

Since the probabilities are independent, we can therefore conclude that with probability  $\beta_1 \cdot \dots \cdot \beta_n$ , all  $n$  inequalities hold, i.e.,  $x_1 \leq x_1^+(\beta_1)$ , ..., and  $x_n \leq x_n^+(\beta_n)$ . Since the function  $f(x_1, \dots, x_n)$  is non-decreasing, we conclude that with probability  $\beta_1 \cdot \dots \cdot \beta_n$ , we have

$$y = f(x_1, \dots, x_n) \leq f(x_1^+(\beta_1), \dots, x_n^+(\beta_n)).$$

Thus, we have a bound that bounds  $y$  from above with probability  $\beta_1 \cdot \dots \cdot \beta_n$ . In particular, if we select  $\beta_i$  for which the product is equal to  $\beta$ , we get the bound corresponding to the given  $\beta$ .

One such bound is easy to find: with probability  $\geq \beta$ , we have  $y \leq y^+$ . Our goal is to find the *smallest* of such bounds, so we can take

$$y^+(\beta) = \min f(x_1^+(\beta_1), \dots, x_n^+(\beta_n)), \quad (1)$$

where minimum is taken over all values  $\beta_1, \dots, \beta_n$  for which  $\beta_1 \cdot \dots \cdot \beta_n = \beta$ .

## 8 Relation between the main result for independent case and fuzzy data processing

As we have mentioned, for each variable  $x_i$ , we can combine the corresponding confidence intervals into a single fuzzy number  $X_i$ . Specifically, we must design a fuzzy number for which, for every  $\beta$ , the  $\beta$ -cut coincides with the interval  $[x_i^-, x_i^+(\beta)]$ .

By definition, the  $\beta$ -cut consists of all the values  $x$  for which the values of the corresponding membership function  $\mu_i(x)$  is greater than or equal to  $\beta$ . Thus, the right-hand side  $x_i^+(\beta)$  of the  $\beta$ -cut interval is the value where the inequality  $\mu_i(x) \geq \beta$  changes to  $\mu_i(x) < \beta$ , i.e. – for continuous membership functions – the value for which  $\mu_i(x) = \beta$ . Therefore,  $x_i^+(\beta) = x$  if and only if  $\mu_i(x) = \beta$ .

Let us use this equivalence to reformulate the above formula (1) in terms of the corresponding membership functions  $\mu_i(x_i)$  and  $\mu(y)$ . Indeed, according to (1), for every value  $y$ ,  $\beta = \mu(y)$  is the smallest possible value of  $\beta_1 \dots \beta_n$  among all the tuples for which  $f(x_1^+(\beta_1), \dots, x_n^+(\beta_n)) = y$ , i.e., in terms of  $\mu_i(x_i)$ , the smallest possible value of  $\mu_1(x_1) \dots \mu_n(x_n)$  among all tuples for which  $f(x_1, \dots, x_n) = y$ .

In other words, if we form:

- a fuzzy number  $X_1$  (with membership function  $\mu_1(x_1)$ ) from the confidence intervals corresponding to  $x_1$ ,
- another fuzzy number  $X_2$  (with membership function  $\mu_2(x_2)$ ) from the confidence intervals corresponding to  $x_2$ ,
- etc.

then the fuzzy number  $Y$  formed by the resulting confidence intervals for  $y$  can be obtained from  $X_1, \dots, X_n$  by applying Zadeh's extension principle:

$$\mu(y) = \max t(\mu_1(x_1), \dots, \mu_n(x_n)), \quad (2)$$

where  $\max$  is taken over all tuples  $(x_1, \dots, x_n)$  for which  $f(x_1, \dots, x_n) = y$ , and  $t(a, b) = a \cdot b$  is an appropriate t-norm.

## 9 Main result for possibly dependent case and its relation to fuzzy data processing

Let us now consider the case when we have no information about the correlation between  $x_i$ . Let us fix  $\beta$  and try to find, for  $y = f(x_1, \dots, x_n)$ , the value  $y^+(\beta)$  for which the probability  $\text{Prob}[y \leq y^+(\beta)]$  is equal to  $\beta$ .

Similarly to the independent case, for every tuple  $(\beta_1, \dots, \beta_n)$ , we have:

- $x_1 \in [x_1^-, x_1^+(\beta_1)]$  – i.e.,  $x_1 \leq x_1^+(\beta_1)$  – with probability  $\beta_1$ ;
- ...
- $x_n \in [x_n^-, x_n^+(\beta_n)]$  – i.e.,  $x_n \leq x_n^+(\beta_n)$  – with probability  $\beta_n$ .

The probability that  $i$ -th inequality fails is equal to  $1 - \beta_i$ , hence the probability that one of  $n$  inequalities fail cannot exceed the sum of these probabilities, i.e.,  $\leq (1 - \beta_1) + \dots + (1 - \beta_n)$ . Thus, with probability

$$\geq 1 - ((1 - \beta_1) + \dots + (1 - \beta_n)) = \beta_1 + \dots + \beta_n - (n - 1),$$

all  $n$  inequalities hold, i.e.,  $x_1 \leq x_1^+(\beta_1)$ , ..., and  $x_n \leq x_n^+(\beta_n)$ . Since the function  $f(x_1, \dots, x_n)$  is non-decreasing, we conclude that with probability  $\max(0, \beta_1 + \dots + \beta_n - (n - 1))$ , we have

$$y = f(x_1, \dots, x_n) \leq f(x_1^+(\beta_1), \dots, x_n^+(\beta_n)).$$

Thus, we have a bound that bound  $y$  from above with probability  $\beta_1 \dots \beta_n$ . In particular, if we select  $\beta_i$  for which the product is equal to  $\beta$ , we get the bound corresponding to the given  $\beta$ .

One such bound is easy to find: with probability  $\geq \beta$ , we have  $y \leq y^+$ . Our goal is to find the *smallest* of such bounds, so we can take

$$y^+(\beta) = \min f(x_1^+(\beta_1), \dots, x_n^+(\beta_n)), \quad (3)$$

where minimum is taken over all values  $\beta_1, \dots, \beta_n$  for which  $\max(0, \beta_1 + \dots + \beta_n - (n - 1)) = \beta$ .

Similar to the independent case, we can thus conclude that if we form:

- a fuzzy number  $X_1$  (with membership function  $\mu_1(x_1)$ ) from the confidence intervals corresponding to  $x_1$ ,
- another fuzzy number  $X_2$  (with membership function  $\mu_2(x_2)$ ) from the confidence intervals corresponding to  $x_2$ ,
- etc.

then the fuzzy number  $Y$  formed by the resulting confidence intervals for  $y$  can be obtained from  $X_1, \dots, X_n$  by applying Zadeh's extension principle (2) – albeit with a different t-norm  $t(a, b) = \max(a + b - 1, 0)$ . In short:

- if the probabilities are independent, then we should use the algebraic product

$$t(a, b) = a \cdot b;$$

- if we have no information about the possible correlation, then we should use

$$t(a, b) = \max(a + b - 1, 0).$$

## 10 The relationship between confidence intervals and fuzzy data processing is very useful

The above relationship between nested confidence intervals and fuzzy numbers is, in our opinion, methodologically interesting, because they provide one more case when the formulas of fuzzy logic, formulas originally designed to handle human uncertainty, can be also used to solve well formulated and mathematically well-defined practical problems.

This relationship is also practically useful because we can now apply the multi-decade experience of computing extension principle formulas to solve problems with purely statistical uncertainty.

It is worth mentioning that a similar example of usefulness of fuzzy formulas in solving purely statistical problems (planning, including robotic planning) was presented in our paper [8].

## 11 From intervals to p-boxes

Confidence intervals appear not only when we measure a physical quantity, they also appear when we reconstruct the value of the probability (or of any other physical characteristic) from the sample; see, e.g., [9].

For example, when we reconstruct a cumulative distribution function (CDF) from the sample, then Kolmogorov-Smirnov inequalities enable us to provide bounds for the actual CDF (called *probability boxes*, or *p-boxes*) with a certain confidence interval. Thus, we have a nested collection of bounds corresponding to different confidence probabilities. Such a nested collection is called *hybrid number* in [1].

There are methods which enables us, knowing p-boxes for  $x_1$  and  $x_2$ , to find the p-box containing

the CDF for  $f(x_1, x_2)$ ; see, e.g., [7]. We must therefore extend these methods to the case when, instead of p-boxes, we have a collection of p-boxes corresponding to different confidence probabilities. Here too formulas based on extension principle work.

## Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC2-1232, by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grants numbers F49620-95-1-0518 and F49620-00-1-0365, and by NSF grants CDA-9522207, ERA-0112968 and 9710940 Mexico/Conacyt.

The authors are very thankful to Lotfi A. Zadeh for his attention to our problem and for his inspiring advice, and to the anonymous referees for their useful comments.

## References

- [1] S. Ferson and L. Ginzburg, "Hybrid arithmetic", In: B. M. Ayyub (ed.), *Proceedings of the ISUMA-NAFIPS'95*, IEEE Computer Society Press, Los Alamitos, CA, 1995, pp. 619–623.
- [2] Interval Computations webpage <http://www.cs.utep.edu/interval-comp>, 2002.
- [3] L. Jaulin, M. Keiffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*, Springer-Verlag, 2001.
- [4] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.
- [5] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, 1995.
- [6] H. T. Nguyen and E. A. Walker, *First Course in Fuzzy Logic*, CRC Press, Boca Raton, FL, 1999.
- [7] Ramas website <http://www.ramas.com>, 2002

- [8] R. Trejo, V. Kreinovich, and C. Baral, "Towards Feasible Approach to Plan Checking Under Probabilistic Uncertainty: Interval Methods", *Proc. of the 17th National Conference on Artificial Intelligence AAAI'2000*, Austin, TX, July 30–August 3, 2000, pp. 545–550.
- [9] H. M. Wadsworth, Jr. (eds.), *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York, 1990.