TOWARDS FOUNDATIONS OF PROCESSING IMPRECISE DATA: FROM TRADITIONAL STATISTICAL TECHNIQUES OF PROCESSING CRISP DATA TO STATISTICAL PROCESSING OF FUZZY DATA

Hung T. Nguyen[1] Tonghui Wang[1] Vladik Kreinovich[2]

1. Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 88003, USA

2. Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA

Email: hunguyen@nmsu.edu, twang@nmsu.edu, vladik@cs.utep.edu

ABSTRACT: In traditional statistics, we process crisp data – usually, results of measurements and/or observations. Not all the knowledge comes from measurements and observations. In many real-life situations, in addition to the results of measurements and observations, we have *expert estimates*, estimates that are often formulated in terms of natural language, like "x is large". Before we analyze how to process these statements, we must be able to translate them in a language that a computer can understand. This translation of expert statements from natural language into a precise language of numbers is one of the main original objectives of fuzzy logic. It is therefore important to extend traditional statistical techniques from processing crisp data to processing fuzzy data. In this paper, we provide an overview of our related research.

Keywords: statistical processing, interval data, fuzzy data, random sets, possibility theory,

1 INTRODUCTION: IT IS NECESSARY TO GO BEYOND TRADITIONAL STATISTICS

In traditional statistics, we process crisp data – usually, results of measurements and/or observations:

- First, for each physical quantity, we observe and/or measure several values $x_1, \ldots, x_n$.
- Based on this data, we try to find the probability distribution that describes this data. Usually, we first guess the finite-parametric family of probability distributions that contains the desired one (e.g., Gaussian distribution), and then we use the measured values to *estimate* the values of the corresponding parameters.
- Based on the probability distributions describing the data, we can then use the known equations – i.e., the known relation between the measured quantities and the characteristics of the future events – to *predict* the results of future events.

Not all the knowledge comes from measurements and observations. In many real-life situations, in addition to the results of measurements and observations, we have *expert estimates*, estimates that are often formulated in terms of natural language, like "x is large". It is therefore necessary to extend traditional statistical techniques to the case when some of the processed values come from expert.

Statistical data processing often requires a lot of computations, and therefore, requires that we use computers. Thus, before we analyze how to process these statements, we must be able to translate them in a language that a computer can understand. This translation of expert statements from natural language into a precise language of numbers is one of the main original objectives of fuzzy logic (see, e.g., [7, 18]). It is therefore important to extend traditional statistical techniques from processing crisp data to processing fuzzy data.

In this paper, we provide an overview of our research in this direction, outline our main results and open problems.

2 ESTIMATING PARAMETERS OF A DISTRIBUTION BASED ON FUZZY DATA

When we have n results $x^{(1)},\ldots,x^{(n)}$ of repeated measurement of the same quantity, traditional statistical approach usually starts with computing their sample average

$$E = \frac{x^{(1)}+\ldots+x^{(n)}}{n}$$

and their sample variance

$$V = \frac{(x^{(1)}-E)^2+\ldots+(x^{(n)}-E)^2}{n-1}$$

(or, equivalently, the sample standard deviation $\sigma = \sqrt{V}$); see, e.g., [21]).

In some practical situations, we only have fuzzy estimates for $x^{(1)},\ldots,x^{(n)}$. In this case, E and V are also fuzzy values. How can we compute these values?

A convenient way of looking at the fuzzy number x is to consider it as a nested family of intervals $x(\alpha)$ – α-cuts of x corresponding to different values α. In this case, for each α, the corresponding α-cut $E(\alpha)$ of, say, E is equal to the range of the function $E(x^{(1)},\ldots,x^{(n)})$ when $x^{(i)}$ takes all possible values from the corresponding intervals $x^{(i)}(\alpha)$. So, in order to be able to compute E and V, it is sufficient to be able to compute the ranges $\mathbf{E}$ and $\mathbf{V}$ of the corresponding functions E and V on given intervals $\mathbf{x}^{(i)} = [\underline{x}^{(i)},\overline{x}^{(i)}]$.

The interval $\mathbf{E}$ for the sample average can be obtained by using straightforward interval computations, i.e., by replacing each elementary operation with numbers by the corresponding operation of interval arithmetic:

$$\mathbf{E} = \frac{\mathbf{x}^{(1)}+\ldots+\mathbf{x}^{(n)}}{n}.$$

What is the interval $[\underline{V},\overline{V}]$ of possible values for sample variance V?

When the intervals $\mathbf{x}^{(i)}$ intersect, then it is possible that all the actual (unknown) values $x^{(i)} \in \mathbf{x}^{(i)}$ are the same and hence, that the sample variance is 0. In other words, if the intervals have a non-empty intersection, then $\underline{V} = 0$. Conversely, if the intersection of $\mathbf{x}^{(i)}$ is empty, then V cannot be 0, hence $\underline{V} > 0$.

First, we design a *feasible* algorithm for computing the exact lower bound $\underline{V}$ of the sample variance. Specifically, our algorithm is *quadratic-time*, i.e., it requires $O(n^2)$ computational steps for n interval data points $\mathbf{x}^{(i)} = [\underline{x}^{(i)},\overline{x}^{(i)}]$. We have implemented this algorithm in C++, it works really fast. The algorithm is as follows (the proof that this algorithm is correct is provided in [1, 2]):

- First, we sort all $2n$ values $\underline{x}^{(i)}$, $\overline{x}^{(i)}$ into a sequence $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$. This sorting requires $O(n\cdot\log(n))$ steps.
- Second, we compute $\underline{E}$ and $\overline{E}$ and pick all "small intervals" $[x_{(k)},x_{(k+1)}]$ that intersect with $[\underline{E},\overline{E}]$.
- For each of remaining small intervals $[x_{(k)},x_{(k+1)}]$, we compute the ratio $r_k = S_k/N_k$, where

  $$S_k \stackrel{\text{def}}{=} \sum_{i:\underline{x}_i \ge x_{(k+1)}} \underline{x}_i + \sum_{j:\overline{x}_j \le x_{(k)}} \overline{x}_j,$$

  and N_k is the total number of such i's and j's (if $N_k = 0$, we take $V'_k \stackrel{\text{def}}{=} 0$). If $r_k \in [x_{(k)},x_{(k+1)}]$, then we compute

  $$V'_k \stackrel{\text{def}}{=} \frac{1}{n-1}\cdot\left(\sum_{i:\underline{x}_i \ge x_{(k+1)}} (\underline{x}_i - r_k)^2 + \sum_{j:\overline{x}_j \le x_{(k)}} (\overline{x}_j - r_k)^2\right).$$

- Finally, we return the smallest of the values V'_k as $\underline{V}$.

Our second result is that the general problem of computing $\overline{V}$ from given intervals $\mathbf{x}^{(i)}$ is NP-hard [1, 2].

NP-hard means, crudely speaking, that there are no general ways for solving all particular cases of this problem (i.e., computing $\overline{V}$) in reasonable time.

However, we show that there are algorithms for computing $\overline{V}$ for many reasonable situations. For example, we propose an efficient algorithm $\mathcal{A}$ that computes $\overline{V}$ for the case when the "narrowed" intervals $[\widetilde{x}^{(i)} - \Delta^{(i)}/n, \widetilde{x}^{(i)} + \Delta^{(i)}/n]$ – where $\widetilde{x}^{(i)} = (\underline{x}^{(i)} + \overline{x}^{(i)})/2$ is the interval's midpoint and $\Delta^{(i)} = (\underline{x}^{(i)} - \overline{x}^{(i)})/2$ is its half-width – do not intersect with each other. We also propose, for each positive integer k, an efficient algorithm $\mathcal{A}_k$ that works whenever no more than k "narrowed" intervals can have a common point [2].

3 USING FUZZY ESTIMATES FOR PARAMETERS OF PROBABILITY DISTRIBUTIONS FOR PREDICTIONS

Once we have the information about the measured quantities $x_1,\dots,x_n$, and we know that the desired quantity y is related to x_i by a known relation $y = f(x_1,\dots,x_n)$, we want to use the known information about x_i to make predictions about y. For example, once we know the current I and the resistance R, we can predict the voltage V by using the known relation (Ohm's law) $V = I \cdot R$.

For each of the quantities x_i, we have a fuzzy number representing its possible values, and from statistical data processing, we may also have some additional (fuzzy) information about the statistical characteristics of x_i. A natural case of such an information is when we have information about the average (first moment) of each variable. What can we then say about the fuzzy value corresponding to y and the fuzzy value corresponding to the average $E[y]$ of y?

Since a fuzzy number can be represented as a nested family of intervals, it is sufficient, similarly to the previous section, to be able to consider this problem for interval uncertainty. In precise terms, we arrive at the following problem:

GIVEN: an algorithm computing a function $f(x_1,\dots,x_n)$ from R^n to R; n intervals $\mathbf{x}_1,\dots,\mathbf{x}_n$, and n intervals $\mathbf{E}_1,\dots,\mathbf{E}_n$,

TAKE: all possible joint probability distributions on R^n for which, for each i, $x_i \in \mathbf{x}_i$ with probability 1 and the mean E_i belongs to $\mathbf{E}_i$;

FIND: the set $\mathbf{Y}$ of all possible values of a random variable $y = f(x_1,\dots,x_n)$ and the set $\mathbf{E}$ of all possible values of $E[y]$ for all such distributions.

In this formulation, we have no information about the possible dependence between the random variables x_i. A similar problem can be formulated for the case when x_i are known to be independent, and for the cases when $n = 2$ and the values x_i are highly positively or highly negatively correlated (i.e., crudely speaking, when they are increasing or decreasing functions of each other).

If we can find the range for degenerate intervals $\mathbf{E}_i = [E_i, E_i]$, then we can use interval computation to extend these formulas to arbitrary intervals $\mathbf{E}_i$.

Similarly to interval computations, our main idea is to find the corresponding formulas for the cases when $n = 2$ and $f = \oplus$ is one of the basic arithmetic operations ($+$, $-$, $\cdot$, min, max). The algorithm for the general case is based on the fact that inside the computer, every algorithm consists of elementary operations (arithmetic operations, min, max, etc.). For each elementary operation $f(x,y)$, if we know the intervals $\mathbf{x}$ and $\mathbf{y}$ for x and y, we can compute the exact range $f(\mathbf{x},\mathbf{y})$; the corresponding formulas form the so-called *interval arithmetic*. We can therefore repeat the computations forming the program f step-by-step, replacing each operation with real numbers by the corresponding operation of interval arithmetic. It is known that, as a result, we get an enclosure for the desired range.

For example, if we know two "triples" $(\underline{x}_i, E_i, \overline{x}_i)$, $(i = 1,2)$, what are the possible triples $(\underline{y}, E, \overline{y})$ for $y = x_1 \cdot x_2$?

For all basic operations, the interval part $(\underline{y}, \overline{y})$ of the result is the same as for interval arithmetic.

We provide explicit formulas for the interval $\mathbf{E}$ of possible values of $E = E[y]$ [3]. For example, for multiplication, when we know nothing about the correlation,

$$\overline{E} = \min(p_1, p_2) \cdot \overline{x}_1 \cdot \overline{x}_2 + \max(p_1 - p_2, 0) \cdot \overline{x}_1 \cdot \underline{x}_2 + \max(p_2 - p_1, 0) \cdot \underline{x}_1 \cdot \overline{x}_2 +$$
$$\min(1 - p_1, 1 - p_2) \cdot \underline{x}_1 \cdot \underline{x}_2,$$

where $p_i \stackrel{\text{def}}{=} (E_i - \underline{x}_i)/(\overline{x}_i - \underline{x}_i)$.

This formula has a natural meaning. Indeed, the probability p_i can be interpreted as follows: if we only allow values $\underline{x}_i$ and $\overline{x}_i$, then there is only one probability distribution on x_i for which the average is exactly E_i. In this probability distribution, the probability $p[\overline{x}_i]$ of $\overline{x}_i$ is equal to p_i, and the probability $p[\underline{x}_i]$ of $\underline{x}_i$ is equal to $1 - p_i$.

If we know the probabilities p_1 and p_2 of two events S_1 and S_2, then the probability of $S_1 \,\&\, S_2$ can take any value from $\max(p_1 + p_2 - 1, 0)$ to $\min(p_1, p_2)$. From this viewpoint, since $p_1 = p[\overline{x}_1]$ and $p_2 = p[\overline{x}_2]$, we can interpret $\min(p_1, p_2)$ as $\overline{p}[\overline{x}_1 \,\&\, \overline{x}_2]$. Similarly, we can interpret all other terms in the above formulas, so we can rewrite the formulas for $\underline{E}$ and $\overline{E}$ as follows:

$$\underline{E} = \underline{p}[\overline{x}_1 \,\&\, \overline{x}_2] \cdot \overline{x}_1 \cdot \overline{x}_2 + \overline{p}[\overline{x}_1 \,\&\, \underline{x}_2] \cdot \overline{x}_1 \cdot \underline{x}_2 + \overline{p}[\underline{x}_1 \,\&\, \overline{x}_2] \cdot \underline{x}_1 \cdot \overline{x}_2 + \underline{p}[\underline{x}_1 \,\&\, \underline{x}_2] \cdot \underline{x}_1 \cdot \underline{x}_2;$$

$$\overline{E} = \overline{p}[\overline{x}_1 \& \overline{x}_2] \cdot \overline{x}_1 \cdot \overline{x}_2 + \underline{p}[\overline{x}_1 \& \underline{x}_2] \cdot \overline{x}_1 \cdot \underline{x}_2 + \underline{p}[\underline{x}_1 \& \overline{x}_2] \cdot \underline{x}_1 \cdot \overline{x}_2 + \overline{p}[\underline{x}_1 \& \underline{x}_2] \cdot \underline{x}_1 \cdot \underline{x}_2.$$

4 A NATURAL WAY TO INTERPRET FUZZY TECHNIQUES IN PROBABILISTIC TERMS: RANDOM SETS

In the previous sections, we have considered fuzziness as an external feature explicitly added to the original probabilistic model. Fuzziness does not have to be added this way. There are known results showing that any consistent description of uncertainty can be, in principle, reformulated in probabilistic terms, and fuzzy description is no exception: from the purely mathematical viewpoint, we can always interpret fuzzy values as probabilities; this fact was first explicitly shown by one of the authors [14]. The corresponding probabilistic interpretation of fuzzy values is not only a mathematical equivalence, it makes sense from the viewpoint of knowledge representation as well.

Indeed, one of the main objectives of fuzzy logic is to interpret natural-language statements like "Mary is young". The problem with interpreting this statement lies in the ambiguity of the natural language. In the idealized situation when all experts could agree on what age corresponds to young and what age means that a person is no longer young, then whether a person is young or not is uniquely determined by the person's age. We can therefore describe this idealized situation by describing the set of all values of age that correspond to "young" – i.e., a certain subset of the set of all positive real numbers.

In reality, different experts may have (and actually have) different interpretations of the word "young". (Moreover, a single expert may have different interpretations of this word.) As a result, for different experts, we have different sets of values corresponding to "young". To capture the overall meaning of the word "young", we must therefore keep not just a *single* set, but *several* different sets – with an indication of which sets are more frequent and which are less frequent. In other words, to describe the meaning of the word "young", we must have a family of sets, with a frequency assigned to each of these sets. In mathematical terms, when we have several objects with probabilities (frequencies) assigned to each object, it is called a *random object*:

- if we have several numbers with different probabilities, it is called a *random number*;
- if we have several processes (functions of time) with different probabilities, it is called a *random process*; etc.

In our case, we have several sets with different probabilities. This structure is called a *random set*. From the mathematical viewpoint, a random set is a probability measure on the family of all sets.

One can show that this interpretation not only makes some sense, it actually explains some – otherwise heuristic and difficult to explain – empirically successful formulas and techniques of fuzzy logic; see, e.g., [6, 8, 13, 15].

Some applications will be described in detail in the talk.

5 CAN FUZZY RESEARCH HELP IN CRISP DATA PROCESSING? POSSIBILITY THEORY AS A TECHNIQUE FOR DESCRIBING ASYMPTOTIC PROPERTIES OF STATISTICAL DISTRIBUTIONS

In the previous section, we have shown that fuzzy techniques can be viewed as an important particular case of probabilistic techniques. From the purely mathematical viewpoint, this case is much more difficult to handle than traditional statistical techniques – because in this case, instead of a probability distribution on the set $\mathbb{R}$ of all real numbers, we have a much more complex object – a probability distribution on the set of all subsets of this set $\mathbb{R}$. However, since this case corresponds to commonsense reasoning, we have additional intuition that makes it easier to solve the corresponding problems. It is therefore natural to expect that in some cases, reformulating a purely probabilistic problem in fuzzy terms will help to solve it. In other words, we expect that not only crisp data processing can help in fuzzy research – by providing a crisp foundations for fuzzy case, but that also fuzzy research can help in crisp data processing.

These expectations are indeed correct. Let us cite two examples. The first example is described in detail in [11]. This application is based on the fact that in statistics, there is a situation that is similar to nested intervals that form a fuzzy set – a nested set of confidence intervals. It turns out that indeed, we can use techniques and algorithms developed for processing fuzzy data to design new efficient algorithms for processing confidence intervals.

Another – less trivial – example comes from the necessity to consider rare events. Rare events – such as unusually large deviations – are extremely important, because they account for catastrophic failures of technical systems, for natural disasters such as earthquakes and floods, etc. Since they are rare, we do not have a large number of observed events of this type and therefore, we cannot use traditional engineering statistical techniques for processing such events. As an alternative, statisticians have developed *asymptotic* techniques, in which instead of describing the probability $p(L)$ of a specific large deviation L, we try to describe an asymptotic expression $p_a(L)$ that describes how the probability $p(L)$ of a deviation of size $\geq L$ depends on L. By definition of asymptotic, when L is large, the actual probability is close to this asymptotic expression, and the larger L (i.e., the more important the deviation), the closer this asymptotic estimate $p_a(L)$ to the actual value $p(L)$.

In many cases, for large deviations L, the dependence of p on L is *scale-invariant*, i.e., crudely speaking, leads to $p(L) \sim C \cdot L^{-\alpha}$ for some real number α; for details, see [16, 17].

In this case, we have two parameters to characterize this dependence: C and α. If we want to characterized the rarity of an event by a single parameter, then which of these two parameters should we choose? A small change in α leads to a much faster decrease in $p(L)$ than a small change in C, so it is natural to select α as a measure of rarity.

In this case, if we have two rare events with rarities $\alpha(A)$ and $\alpha(B)$, what is the rarity of $A \vee B$? In other words, how can we estimate the probability of the event that either A or B will lead to a large deviation $\geq L$? One can easily see that if, say, $\alpha(A) < \alpha(B)$, then $P_A(L) \gg P_B(L)$, moreover, $P_B(L)/P_A(L) \to 0$ and therefore, asymptotically, the total probability $P_{A\,\&\,B}(L)$ is equal to $p_A(L)$. In other words, in contrast to traditional probability theory in which – provided A and B are incompatible – the probability of $A \vee B$ is equal to the sum $p(A \vee B) = p(A) + p(B)$, for our newly defined measure of rarity, $\alpha(A \vee B) = \max(\alpha(A), \alpha(B))$. This is exactly the formula used in fuzzy logic – and in a related formalism of possibility theory. It is therefore not surprising that a theory of "idempotent probabilities" – a mathematical theory which is very similar to possibility theory – turned out to be very helpful in the analysis of asymptotic properties of large deviations [15, 20].

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC2-1232, by 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, by NSF grants EAR-0112968, EAR-0225670, and 9710940 Mexico/Conacyt, by IEEE/ACM SC2001 and SC2002 Minority Serving Institutions Participation Grants, and by a research grant from Sandia National Laboratories as part of the Department of Energy Accelerated Strategic Computing Initiative (ASCI).

This research was partly done when V.K. was a Visiting Faculty Member at the Fields Institute for Research in Mathematical Sciences (Toronto, Canada) and Visiting Researcher at the Euler International Mathematical Institute (St. Petersburg, Russia).

REFERENCE

[1] S. Ferson, L. Ginzburg, V. Kreinovich, and M. Aviles (2002) "Exact Bounds on Sample Variance of Interval Data", Extended Abstracts of the 2002 SIAM Workshop on Validated Computing, Toronto, Canada, May 23–25, 2002: 67–69.

[2] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré, and M. Aviles (2002) "Computing Variance for Interval Data is NP-Hard", ACM SIGACT News, 33(2):108–118.

[3] S. Ferson, L. Ginzburg, V. Kreinovich, and J. Lopez (2002) "Absolute Bounds on the Mean of Sum, Product, etc.: A Probabilistic Extension of Interval Arithmetic", Extended Abstracts of the 2002 SIAM Workshop on Validated Computing, Toronto, Canada, May 23–25, 2002, pp. 70–72.

[4] S. Ferson, L. Ginzburg, V. Kreinovich, H.T. Nguyen, and S.A. Starks (2002) "Uncertainty in Risk Analysis: Towards a General Second-Order Approach Combining Interval, Probabilistic, and Fuzzy Techniques", Proceedings of FUZZ-IEEE'2002, Honolulu, Hawaii, May 12–17, 2002, 2:1342–1347.

[5] S. Ferson, L. Ginzburg, V. Kreinovich, and H. Schulte (2001) "Interval Computations as a Particular Case of a General Scheme Involving Classes of Probability Distributions", In: J. Wolff von Gudenberg and W. Kraemer (eds.), Scientific Computing, Validated Numerics, Interval Methods, Kluwer, Dordrecht, 355–366.

[6] J. Goutsias, R.P.S. Mahler, and H. T. Nguyen, eds. (1997) Random Sets: Theory and Applications, Springer-Verlag, N.Y.

[7] G.J. Klir and B. Yuan (1995) Fuzzy Sets and Fuzzy Logic, Prentice Hall, NJ.

[8] V. Kreinovich (1997) "Random Sets Unify, Explain, and Aid Known Uncertainty Methods in Expert Systems", In: J. Goutsias, R.P.S. Mahler, and H. T. Nguyen, eds., Random Sets: Theory and Applications, Springer-Verlag, N.Y.: 321–345.

[9] V. Kreinovich, S. Ferson, L. Ginzburg, H. Schulte, M.R. Barry, and H.T. Nguyen (2000) "From Interval Methods of Representing Uncertainty To A General Description of Uncertainty", In: H. Mohanty and C. Baral (eds.), Trends in Information Technology, Proceedings of the International Conference on Information Technology ICIT'99, Bhubaneswar, India, December 20–22, 1999, Tata McGraw-Hill, New Delhi, 161–166.

[10] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl (1997) Computational Complexity and Feasibility of Data Processing and Interval Computations, Kluwer, Dordrecht.

[11] V. Kreinovich, H.T. Nguyen, S. Ferson, and L. Ginzburg (2002) "From Computation with Guaranteed Intervals to Computation with Confidence Intervals: A New Application of Fuzzy Techniques", Proceedings of the 21st International Conference of the North American Fuzzy Information Processing Society NAFIPS'2002, New Orleans, Louisiana, June 27–29, 2002: 418–422.

[12] V.P. Kuznetsov (1991) Interval Statistical Models, Moscow, Radio i Svyaz Publ. (in Russian).

[13] S. Li, Y. Ogura, and V. Kreinovich (2002) Limit Theorems and Applications of Set Valued and Fuzzy Valued Random Variables, Kluwer, Dordrecht.

[14] H.T. Nguyen (1979) "Some Mathematical Tools for Linguistic Probabilities", Fuzzy Sets and Systems, 2:53–65.

[15] H.T. Nguyen and B. Bouchon-Meunier (2003) "Random Sets and Large Deviations Principle as a Foundation for Possibility Measures", Soft Computing (to appear).

[16] H.T. Nguyen and V. Kreinovich (1997) Applications of Continuous Mathematics to Computer Science, Kluwer, Dordrecht.

[17] H.T. Nguyen, V. Kreinovich, and B. Wu (1999) "Fuzzy/Probability∼Fractal/Smooth", International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS), 7:363–370.

[18] H.T. Nguyen and E. A. Walker (1999) A First Course in Fuzzy Logic, CRC Press, Boca Raton, Florida.

[19] R. Osegueda, V. Kreinovich, L. Potluri, and R. Aló (2002) "Non-Destructive Testing of Aerospace Structures: Granularity and Data Mining Approach", Proceedings of FUZZ-IEEE'2002, Honolulu, Hawaii, May 12–17, 2002, 1:685–689.

[20] A. Puhalskii (2001) Large Deviations and Idempotent Probability, Chapman and Hall/CRC, Boca Raton.

[21] S. Rabinovich (1993) Measurement Errors: Theory and Practice, American Institute of Physics, New York.

[22] S.A. Vavasis (1991) Nonlinear Optimization: Complexity Issues, Oxford University Press, N.Y.

[23] P. Walley (1991) Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, N.Y.