From Type-2 Fuzzy to Type-2 Intervals and Type-2 Probabilities

Vladik Kreinovich[0000−0002−1244−1650], Olga Kosheleva[0000−0003−2587−4209], and Luc Longpré

University of Texas at El Paso, 500 W. University, El Paso, Texas 79968, USA,
vladik@utep.edu, olgak@utep.edu, longpre@utep.edu

Abstract. Our knowledge comes from observations, measurements, and expert opinions. Measurements and observations are never 100% accurate, there is always a difference between the measurement result and the actual value of the corresponding quantity. We gauge the resulting uncertainty either by an interval of possible values, or by a probability distribution on the set of possible values, or by a membership function that describes to what extent different values are possible. The information about uncertainty also comes either from measurements or from expert estimates and is, therefore, also uncertain. It is important to take such "type-2" uncertainty into account. This is a known idea in fuzzy, where type-2 fuzzy is a well-known effective technique. In this paper, we explain how a similar approach can be applied to type-2 intervals and type-2 probabilities.

Keywords: Type-2 fuzzy · Type-2 intervals · Type-2 probabilities.

1 Introduction

Uncertainty is ubiquitous. What are the main objectives of science and engineering?

- We want to know the current state of the world.
- We want to predict the future state of the world.
- We want to find out how to make the future state of the world better.

To describe the state of the world, we need to describe the values of the corresponding physical quantities. For example, in celestial mechanics, we need to know the positions and velocities of all celestial bodies.

Similarly, to describe appropriate actions, we need to describes the values of the parameters of these actions. For example, we need to know the initial velocity and orientation of the spaceship that will lead us to Mars.

Our information about the values of physical quantities come from measurements and from expert estimates.

- Measurements are never 100% accurate, there is always some uncertainty. For example, Vladik's height is 169.5 cm. This does not mean that it is exactly 169.5000, it means ± 0.5.

– Expert estimates are usually even less certain.

So, we always have uncertainty.

How can we describe this uncertainty. Uncertainty means that instead of a *single* value, we have the whole *set* of possible values of a quantity. This set is usually connected, i.e., it is an *interval*; see, e.g., [4, 6, 9, 11, 14].

In some cases, we also know the frequency with which, in similar situations, we encounter different values from this interval. This is known as *probabilistic* uncertainty.

In many practical cases, we do not know the probabilities. However, an expert can estimate the degree to which different values are possible. This corresponds to *fuzzy* uncertainty; see, e.g., [1, 5, 10, 12, 13, 18].

Need to take uncertainty into account when processing data. We do not just measure quantities, we perform some computations $y = f(x_1, \ldots, x_n)$ with the measurement results $x_1, \ldots, x_n$. Due to uncertainty, the actual values x_i are, in general, different from the measurement results $\widetilde{x}_i$. Because of this, the result $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ of processing measurement results is, in general, different from the value $y = f(x_1, \ldots, x_n)$ that we would have gotten if we knew the exact values x_i. It is important to gauge the resulting uncertainty in y.

For example, if $\widetilde{y} = 100$ million tons is the estimated amount of oil in a given region, then:

– if it is 100 ± 20, we should start exploiting, but
– it it is 100 ± 200, maybe there is no oil at all, so we better perform additional measurement before wasting resources.

For interval, probabilistic, and fuzzy uncertainty there are techniques for propagating uncertainty through computations.

Need for type-2 fuzzy. As we have mentioned, an expert cannot describe his/her estimate of the quantity of interest by a single number. Instead, the expert can produce an interval – or an interval with degrees assigned (= a fuzzy set).

Similarly, the same expert cannot describe his/her degree of confidence $m(x)$ that x is possible by a single number. For example:

– an expert can distinguish between degrees of confidence 0.7 and 0.8,
– but hardly anyone can distinguish between degrees 0.80 and 0.81.

It is more reasonable to expect that $m(x)$ is represented by an interval, or even by a fuzzy set.

So, for each x from the original interval $[\underline{x}, \overline{x}]$:

– instead of a numerical degree of confidence $m(x)$,
– we have an interval $[\underline{m}(x), \overline{m}(x)]$ – or a fuzzy set.

This situation is known as *type-2 fuzzy*. Type-2 fuzzy sets – and especially interval-valued fuzzy sets – are well-studied and used in many applications; see, e.g., [10].

Need for type-2 intervals and probabilities, both subjective and objective. Intervals and probabilities also come from observations and measurements. Observations and measurements always have uncertainty. Thus, we know intervals and probabilities with uncertainty too. Similarly to type-2 fuzzy, we can call methods that take this uncertainty into account *type-2*.

Comment. There is an important difference between type-2 fuzzy and type-2 probabilities:

– fuzzy (and type-2 fuzzy) information is *subjective*, while
– interval and probabilistic uncertainty can be both *subjective* and *objective*.

2 Subjective and Objective Intervals – and Type-2 Intervals

Subjective vs. objective intervals. In some cases, there exists an exact value, but we do not know this value, we only know the interval containing this value. In this case, we have a *subjective* (epistemic) interval.

In other cases, we have a range of values. E.g., person's height and weight change during the day. So, the actual answer to the question "What is your height" is an interval. This is *objective* (*aleatory*) interval.

Need for type-2 intervals. Since measurements are not absolutely accurate, we do not know the exact endpoints ℓ and r of the objective interval. Instead, we have:

– an interval $[\underline{\ell}, \overline{\ell}]$ of possible values of ℓ, and
– an interval $[\underline{r}, \overline{r}]$ of possible values of r.

This type-2 situation can be described by two nested intervals:

– values from the inner interval $[\overline{\ell}, \underline{r}]$ *are* actual, while
– values from the outer interval $[\underline{\ell}, \overline{r}]$ *may be* actual, but we are not sure.

How can we process such type-2 intervals? If we know that $y = f(x_1, \ldots, x_n)$ and we have such type-2 interval about each x_i, what can we say about y? If all uncertainties are independent, i.e., if all combinations of possible values of x_i are possible, then:

– to find the inner interval for y, we apply interval computations to inner intervals for x_i, and
– to find the outer interval for y, we apply interval computations to outer intervals for x_i.

3 Subjective Type-2 Probabilities

What are subjective type-2 probabilities. Subjective type-2 probabilities mean that we only have partial information about the corresponding probabilities. This is known as *imprecise probabilities*.

Interval-valued probabilities: the basic case of subjective type-2 probabilities. In general, as we mentioned, the basic type of uncertainty is interval uncertainty. In line with this, the basic type of probabilistic uncertainty is interval-valued probabilities.

p-boxes: an important case of interval-valued probabilities. One of the main ways to describe a probability distribution is by a cumulative distribution function (cdf) $F(x) \stackrel{\text{def}}{=} \text{Prob}(X \le x)$. A natural idea is thus to consider, for each x, an interval $[\underline{F}(x), \overline{F}(x)]$ of possible values of $F(x)$. This is is known as *probability box*, or *p-box*, for short. p-boxes have been successfully used in many applications [2] (as well as fuzzy-valued probabilities).

4 Objective Type-2 Probabilities

Objective type-2 probabilities: what are they? How can we have objective uncertainty in probability values? To understand this, let us recall what is probability from the practical viewpoint.

In practice, probability p means, in effect, a frequency. We have a large number N of similar events (e.g., flipping a coin). These can be similar events occurring at different location and/or at different times. Probability p of a certain outcome means that this outcome is observed in $\approx p \cdot N$ cases.

An ideal case is when the event settings are absolutely identical. For example:

- we have a large set of identical atoms of a radioactive element, and
- we observe how many of them emit radiation during a given period of time.

In the usual quantum description, all the atoms are identical. However, the true quantum description is more complex; see, e.g., [3, 16]. In this sense, quantum physics is similar to fuzzy:

- one of the main ideas about fuzzy is Zadeh's statement that "everything is a matter of degree";
- in quantum physics, the main idea is that everything is a matter of probability.

In the first approximation – traditional quantum mechanics:

- particle locations and velocities are only known with probabilities, they can fluctuate around their classical values,
- however, the forces between particles are described by the usual formulas, e.g., by the Coulomb law
$$F = -c \cdot \frac{q_1 \cdot q_2}{r^2}.$$

In secondary quantization, we take into account that the forces can also fluctuate around the classical values. In other words, the fields – that describe these forces – are also quantum objects whose values are only known with some probabilities.

In general, no matter what kind of events we consider, these events are not identical. There are always quantum fluctuations because of which, for each event, the probability p_i is slightly different from p. Here, the values p_i are randomly fluctuating around the classical value p. In other words, here, we have objective type-2 probabilities.

What does this mean in terms of observations? Can we experimentally detect the difference between type-1 and type-2 probabilities? To answer this question, let us recall what randomness means in terms of observations.

What does randomness mean in terms of observations: reminder. What does randomness mean in terms of observations?

Randomness means more than frequency. For example, according to Central Limit Theorem (see, e.g., [15]), differences between frequency and probability should be normally distributed.

The general idea is that if a sequence is random, it must satisfy all the probability laws.

A probability law is something that happens with probability 1. In mathematical terms, it is a set of probability measure 1 – so that its complement has measure 0.

So, a sequence is random if:

- it does not belong to any definable set of probability measure 0,
- or, equivalently, it does not belong to the union of all definable sets of measure 0.

This is Kolmogorov's definition of a random sequence; see, e.g., [7]:

Definition 1. *An object x is called* definable *if there is a formula $P(y)$ – built from the basic constants, objects, functions, and predicates of the corresponding theory – that is satisfied if and only if $y = x$.*

Example. For example, the formula $y \cdot y = 1 + 1 \,\&\, y > 0$ uniquely defines the value $\sqrt{2}$.

Comment. In this sense, every element, every function, every set that we can define is definable.

Definition 2. *Let μ be a probability measure on a set X. An element $x \in X$ is called μ-*random *if it does not belong to any definable set of μ-measure 0.*

Comment. The original Kolmogorov's definition was proposed for the case when the measure itself is definable. In this paper, we will be also interested in measures which are not definable. For such measures, we need to make a small modification of this definition.

Definition 3. *Let z be an object. An object x is called z-*definable *if there is a formula $P(z, y)$ that is satisfied if and only if $y = x$.*

Definition 4. *Let μ be a probability measure on a set X. An element $x \in X$ is called* μ-random *if it does not belong to any μ-definable set of μ-measure 0.*

Comment. We have already mentioned that every element whose definition we can write down is definable. From this viewpoint, it is not possible to provide a specific example of an element which is not definable – whatever example will give will be, by definition, definable. So, a natural question is: are there elements which are not definable? It turns out that not only there are such element, but also almost all elements are not definable.

Proposition 1. *For every probability measure μ on a set X, almost all elements are not μ-definable.*

Proof. Every definable set is described by a finite text – its definition. There are only countably many texts, so there are only countably many definable sets. The union of countably many sets of measure 0 still has measure 0. So, almost all sequence are indeed random. The proposition is proven.

So can we experimentally detect the difference between type-1 and type-2 probabilities? We are interested in a sequence of events. Let $n_i = 1$ if the selected outcome occurred and $n_i = 0$ if it did not.

We compare two cases:

- type-1 case when each n_i occurs with probability p, and
- type-2 case when each n_i occurs with probability p_i.

Here:

- we select some distribution on the set of all probabilities with mean p, and
- then, we take, as p_i, a random sequence of independent values corresponding to these probabilities.

Definition 5.

- *For each value $x \in (0,1)$, let $\mu(x)$ denote a probability measure that is located at 1 with probability x and at 0 with the remaining probability $1-x$.*
- *Let $p \in (0,1)$ be a real number.*
- *By $\nu(p)$, we mean a probability measure $\mu(p) \times \mu(p) \times \ldots$ that describes an infinite sequence of independent random variables $n_i^{(1)} \in \{0,1\}$ each of which is equal to 1 with probability p.*
- *By a* p-random sequence*, we mean a sequence which is random with respect to $\nu(p)$.*

Definition 6. *We say that the probability measure on the interval $[0,1]$ is* trivial *if it is concentrated at some value $p \in [0,1]$ with probability 1.*

Definition 7.

- *Let μ_0 be a non-trivial probability measure on the interval $[0,1]$.*
- *Let p denote the mean value of the corresponding random variable.*

- *Let* $\mu_\infty = \mu_0 \times \mu_0 \times \ldots$ *be a probability measure describing an infinite sequence of independent identically distributed random variables distributed according to the probability* μ_0.
- *Let* $p_1, p_2, \ldots$ *be a* μ_∞*-random sequence.*
- *Let* $\nu^{(2)}(p) = \mu(p_1) \times \mu(p_2) \times \ldots$ *be a probability measure describing an infinite sequence of independent random variables* $n_i^{(2)} \in \{0,1\}$ *each of which is equal to 1 with probability* p_i.
- *By a* type-2 p-random sequence*, we means a sequence which is random with respect to* $\nu^{(2)}(p)$.

Discussion. How can we distinguish between p-random sequences and type-2 p-random sequences? Often, random variables can be distinguished by their moments. For both sequences, we can compare moments, i.e., averages over i from 1 to N of products $n_i^{a_0} \cdot n_{i+i_1}^{a_1} \cdot \ldots \cdot n_{i+i_m}^{a_m}$. For example:

- mean is the average of n_i,
- covariance with next neighbor is the average of $n_i \cdot n_{i+1}$, etc.

Our *first result* is that for both sequences, each moment tends to the same limit: e.g., the mean tends to p.

Definition 8. *By a* moment*, we mean the expression*

$$m \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N} \cdot \sum_{i=1}^{N} n_i^{a_0} \cdot n_{i+i_1}^{a_1} \cdot \ldots \cdot n_{i+i_m}^{a_m}$$

corresponding to some values $a_j > 0$ *and* i_j.

Proposition 2. *For any type-2 p-random sequence, the moment is the same as for any p-random sequence.*

Proof. Since $n_i \in \{0,1\}$, for each power $a_j > 0$, we have $n_{i+i_j}^{a_j} = n_{i+i_j}$. Thus, the expression for the moment takes the following simplified form"

$$m = \lim_{N\to\infty} \frac{1}{N} \cdot \sum_{i=1}^{N} n_i \cdot n_{i+i_1} \cdot \ldots \cdot n_{i+i_m}.$$

Due to the Large Numbers Theorem, the limit value m is equal to the expected value $E[\cdot]$ of the corresponding expression with respect to the measure $\nu^{(2)}(p)$, i.e., to the value

$$m = \lim_{N\to\infty} \frac{1}{N} \cdot \sum_{i=1}^{N} E^{(2)}[n_i \cdot n_{i+i_1} \cdot \ldots \cdot n_{i+i_m}],$$

where $E^{(2)}[\cdot]$ denotes the corresponding expected value.

Since all the variables n_i are independent, the expected values of the product is equal to the product of the expected values, so

$$m = \lim_{N\to\infty} \frac{1}{N} \cdot \sum_{i=1}^{N} E^{(2)}[n_i] \cdot E^{(2)}[n_{i+i_1}] \cdot \ldots \cdot E^{(2)}[n_{i+i_m}],$$

i.e.,

$$m=\lim_{N\to\infty}\frac{1}{N}\cdot\sum_{i=1}^{N}p_i\cdot p_{i+i_1}\cdot\ldots\cdot p_{i+i_m}.$$

The values p_i are themselves random with respect to the probability measure μ_0. Thus, again due to the Large Numbers Theorem, m is equal to the expected value with respect to the measure μ_0:

$$m=\lim_{N\to\infty}\frac{1}{N}\cdot\sum_{i=1}^{N}E_0[p_i\cdot p_{i+i_1}\cdot\ldots\cdot p_{i+i_m}],$$

where E_0 denoted the corresponding expected value.

Since all the variables p_i are independent, the expected values of the product is equal to the product of the expected values, so

$$m=\lim_{N\to\infty}\frac{1}{N}\cdot\sum_{i=1}^{N}E_0[p_i]\cdot E_0[p_{i+i_1}]\cdot\ldots\cdot E_0[p_{i+i_m}].$$

By the choice of p_i, we have $E_0[p_i]=p$ for all i, thus

$$m=\lim_{N\to\infty}\frac{1}{N}\cdot\sum_{i=1}^{N}p\cdot p\cdot\ldots\cdot p=p^{m+1}.$$

For p-random sequences, we get the exact same moment – so these moments are indeed equal to each other. The proposition is proven.

Discussion. So does this result mean that we cannot distinguish between p-random and type-2 p-random sequences? No, our *second result* is that no sequence can be random with respect to both type-1 and type-2 distributions.

Proposition 3.

– *No p-random sequence is also type-2 p-random.*
– *No type-2 p-random sequence is also p-random.*

Proof. According to the main result from [17] – which is also mentioned in [7] and in [8] – for two sequences of probabilities p_i and q_i, the following two conditions are equivalent to each other:

– no sequence is random with respect both to probabilities p_i and to probabilities q_i, and
– $\sum\limits_{i=1}^{\infty}\left[\left(\sqrt{p_i}-\sqrt{q_i}\right)^2+\left(\sqrt{1-p_i}-\sqrt{1-q_i}\right)^2\right]=+\infty.$

In our case, $q_i=p$ for all i. So, to prove the proposition, it is sufficient to prove that

$$\sum_{i=1}^{\infty}\left[(\sqrt{p_i}-\sqrt{p})^2+\left(\sqrt{1-p_i}-\sqrt{1-p}\right)^2\right]=+\infty.$$

Due to the Large Numbers Theorem, for $N \to \infty$, we have

$$\lim_{N\to\infty} \frac{1}{N} \cdot \sum_{i=1}^{N} \left[(\sqrt{p_i} - \sqrt{p})^2 + \left(\sqrt{1-p_i} - \sqrt{1-p} \right)^2 \right] =$$

$$v \stackrel{\text{def}}{=} E_0 \left[(\sqrt{p_i} - \sqrt{p})^2 + \left(\sqrt{1-p_i} - \sqrt{1-p} \right)^2 \right] > 0.$$

So, asymptotically, we have

$$\sum_{i=1}^{N} \left[(\sqrt{p_i} - \sqrt{p})^2 + \left(\sqrt{1-p_i} - \sqrt{1-p} \right)^2 \right] \sim N \cdot v,$$

and thus, the limit of this sum is indeed equal to infinity. The proposition is proven.

Discussion. This means that there are probability laws that are only true for type-1 sequences but not for type-2 sequences, and vice versa. So, it *is* possible to experimentally detect the difference between type-1 and type-2 random sequences!

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), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), 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).

The authors are greatly thankful to Janusz Kacprzyk for his encouragement.

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. S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz, *Constructing Probability Boxes and Dempster-Shafer Structures*, Sandia National Laboratories, Report SAND2002-4015, January 2003.
3. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.
4. L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2012.

5. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
6. B. J. Kubica, *Interval Methods for Solving Nonlinear Constraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.
7. M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer, Berlin, Heidelberg, New York, 2008.
8. L. Longpré and V. Kreinovich, "When are two wave functions distinguishable: a new answer to pauli's question, with potential application to quantum cosmology", *International Journal of Theoretical Physics*, 2008, Vol. 47, No. 3, pp. 814–831.
9. G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.
10. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
11. R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
12. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
13. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
14. S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, New York, 2005.
15. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
16. K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2021.
17. V. G. Vovk, "On a randomness criterion", *Soviet Mathematical Doklady*, 1987, Vol. 35, pp. 656–660.
18. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.