On Probability of Making a Given Decision: A Theoretically Justified Transition From Interval to Fuzzy Uncertainty

Van Nam Huynh and Yoshiteru Nakamori
Japan Advanced Institute of Science and Technology (JAIST)
Tatsunokuchi, Ishikawa 923-1292, Japan
huynh@jaist.ac.jp

François Modave
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
fmodave@utep.edu

***Abstract*—In practice, it is often necessary to make a decision under uncertainty.**

In the case of *interval uncertainty*, for each alternative i, instead of the exact value v_i of the objective function, we only have an interval $\mathbf{v}_i = [\underline{v}_i, \overline{v}_i]$ of possible values. In this case, it is reasonable to assume that each value v_i is uniformly distributed on the corresponding interval $[\underline{v}_i, \overline{v}_i]$, and to take the probability that v_i is the largest as the probability of selecting the i-th alternative.

In some practical situations, we have *fuzzy uncertainty*, i.e., for every alternative i, we have a fuzzy number $\mathcal{V}_i$ describing the value of the objective function. Then, for every degree α, we have an interval $\mathcal{V}_i(\alpha)$, the α-cut of the corresponding fuzzy number. For each α, we can assume the uniform distributions on the corresponding α-cuts and get a probability $P_i(\alpha)$ that v_i will be selected for this α. From the practical viewpoint, it is desirable to combine these probabilities into a single probability corresponding to fuzzy uncertainty.

In deriving the appropriate combination, we use the fact that fuzzy values are not uniquely defined, different procedures can lead to differently scaled values. It turns out that the only scaling-invariant distribution on the set of all the degrees α is a uniform distribution. So, we justify the choice of $\int P_i(\alpha)\,d\alpha$ as the probability that under fuzzy uncertainty, an alternative i will be selected.

I. Making a Decision

Let us assume that we want to select an alternative with the largest possible value of a certain quantity. If for two alternatives a_1 and a_2, we know the exact values v_1 and v_2 of the corresponding quantity, then the question of which alternative to select is simple:

- we select a_1 if $v_1 > v_2$,
- we select a_2 if $v_2 > v_1$, and
- we can select any of these alternatives if $v_1 = v_2$.

II. Decision Making under Uncertainty

In many practical situations, we do not know the exact values of the desired quantity. In some situations, we only know the bounds $\underline{v}_i$ and $\overline{v}_i$ for the (unknown) actual value v_i, i.e., our only information about v_i is that v_i belongs to the *interval* $\mathbf{v}_i = [\underline{v}_i, \overline{v}_i]$.

In other situations, our only information about v_i comes from an expert estimate described by a term from natural language. In this case, a natural description of this information is by using a fuzzy number.

If we only know, e.g., intervals $[\underline{v}_1, \overline{v}_1]$ and $[\underline{v}_2, \overline{v}_2]$ of possible values of v_i, and these intervals share several common points, then:

- it may be that $v_1 > v_2$ and
- it may be that $v_2 > v_1$.

Thus, some decision makers will prefer v_1 and some may prefer v_2. In this case, it is reasonable to predict the *probability* of selecting v_1.

III. Decision Making under Interval Uncertainty: Formulas are Known

For decision making under interval uncertainty, there exist reasonable formulas for the probability of selecting v_1. For example, we can assume that v_1 is uniformly distributed within the interval $[\underline{v}_1, \overline{v}_1]$, v_2 is uniformly distributed within the interval $[\underline{v}_2, \overline{v}_2]$, and that v_1 and v_2 are independent random variables. Under these assumptions, we can compute the probability P_1 that $v_1 \geq v_2$ as

$$P_1 = \frac{I_1 + I_2 + I_3}{(\overline{v}_1 - \underline{v}_1) \cdot (\overline{v}_2 - \underline{v}_2)},$$

where

$$I_1 \stackrel{\text{def}}{=} \frac{1}{2} \cdot \max(0, \min(\overline{v}_1, \overline{v}_2) - \max(\underline{v}_1, \underline{v}_2))^2,$$

$$I_2 \stackrel{\text{def}}{=} (\overline{v}_2 - \underline{v}_2) \cdot \max(0, \overline{v}_1 - \overline{v}_2),$$

$$I_3 \stackrel{\text{def}}{=} (\overline{v}_1 - \underline{v}_1) \cdot \max(0, \underline{v}_1 - \underline{v}_2);$$

see, e.g., [6], [10], [12], [13].

Comment. Under interval uncertainty, similar formulas can be described for the case when we have several alternatives; see, e.g., [5]. For reader's convenience, these formulas are also given in the Appendix.

IV. Decision Making under Fuzzy Uncertainty: a Problem

In the fuzzy case, each value v_i is represented by a fuzzy number $\mathcal{V}_i$. A fuzzy number can be equivalently represented by a nested family of intervals $\mathcal{V}_i(\alpha)$ (α-cuts) corresponding to different values $\alpha \in [0,1]$.

For each α, we can use the corresponding intervals $\mathcal{V}_1(\alpha)$ and $\mathcal{V}_2(\alpha)$ to compute the probability $P_1(\alpha)$ that $v_1 \geq v_2$. The question is: how to combine these probabilities $P_1(\alpha)$ into a single probability for selecting v_1?

V. Idea

A fuzzy number means, crudely speaking, that we do not know which interval $\mathcal{V}_i(\alpha)$ actually describes the range. The actual value v_i is always within the range $\mathcal{V}_i(0)$, sometimes it is within the range $\mathcal{V}_i(1)$. We do not know the corresponding value α beforehand, but once we learn the actual value v_i, we can then find the largest α_m of all the values α for which this actual value is contained in the interval $\mathcal{V}_i(\alpha)$.

In principle, we can gather statistics of such values α_m. Once we know the probability corresponding to different values α_m, then we can estimate the desired probability of selecting v_1 as the expected value of the probability $P_1(\alpha)$ with respect to this probability distribution.

VI. What We Do in This Paper: Main Idea

In deriving the appropriate combination, we use the fact that fuzzy values are not uniquely defined, different procedures can lead to differently scaled values.

In this paper, we analyze these re-scalings and prove that the only scaling-invariant distribution on the set of all the degrees α is a uniform distribution. So, we justify the choice of $\int P_i(\alpha)\,d\alpha$ as the probability that under fuzzy uncertainty, an alternative i will be selected.

VII. Different Elicitation Methods Can Lead to Different Fuzzy Values

Polling: a natural way to assign fuzzy value. One of the natural methods to ascribe the degree of confidence $d(A)$ to a statement A is to take several (N) experts, and ask each of them whether he or she believes that A is true.

If $N(A)$ of them answer "yes", we take $d(A) = N(A)/N$ as the desired certainty value; see, e.g., [7], [8], [9].

Polling: examples.

- If all the experts believe in A, then this value is 1 (= 100%).
- If half of them believe in A, then $d(A) = 0.5$ (50%),
- etc.

To get more accurate polling results, we should ask as many experts as possible. Knowledge engineers want the system to include the knowledge of the entire scientific community, so they ask as many experts as possible.

Problem with asking too many experts. Asking too many experts leads to the following negative phenomenon: when the opinion of the most respected professors, Nobel-prize winners, etc., is known, some less self-confident experts will not be brave enough to express their own opinions, so they will:

- either say nothing,
- or follow the opinion of the majority.

The effect of additional experts on the degree of confidence. How does their presence influence the resulting uncertainty value? In line with the above description, let us consider three cases:

- adding shy experts (who do not answer anything);
- adding conformist experts; and
- adding experts of both type.

First case: adding shy experts. Let N denote the initial number of experts, $N(A)$ the number of those of them who believe in A, and M the number of shy experts added.

Initially, $d(A) = N(A)/N$. After we add M experts who do not answer anything when asked about A, the number of experts who believe in A is still $N(A)$, but the total number of experts is bigger ($M+N$). So the new value of the uncertainty ratio is

$$d'(A) = \frac{N(A)}{N+M} = c \cdot d(A),$$

where we denoted $c = N/(M+N)$.

Second case: adding adding conformist experts. When we add experts who give the same answers as the majority of N renowned experts, then, for the case when $d(A) > 1/2$, we get $N(A) + M$ experts saying that A is true.

So, the new uncertainty value is

$$d'(A) = \frac{N(A)+M}{N+M} = \frac{N \cdot d(A) + M}{N+M}.$$

General case: adding both shy and conformist experts. If we add M "silent" experts and M' "conformists" (who vote as the majority), then we get a transformation

$$d(A) \to \frac{N \cdot d(A) + M'}{N+M+M'}.$$

Mathematical observation. In all these cases, the transformation from an old scale $d(A)$ to a new scale $d'(A)$ is a linear function $d(A) \to a \cdot d(A) b$ for some constants a and b; in the most general case

$$a = \frac{N}{N+M+M'}$$

and

$$b = \frac{M'}{N+M+M'}.$$

By selecting appropriate values of N, M, and M', we can get arbitrary linear functions with positive linear coefficients. Thus, we arrive at the following conclusion.

Conclusion. Fuzzy degree of confidence $d(A)$ is defined modulo an arbitrary linear re-scaling transformation.

VIII. Probability Distributions on $[0,1]$ Which Are Conditionally Invariant under Arbitrary Re-scalings

Formulation of the problem. Since fuzzy values are defined modulo a re-scaling (linear transformation), it is reasonable to require that the corresponding probability measure on the interval $[0,1]$ be conditionally invariant with respect to these re-scalings.

In other words, we require that after each linear re-scaling, the conditional probabilities should not change. Since a conditional probability $P(A\,|\,B)$ is defined as ratio of two probabilities

$$P(A\,|\,B) = \frac{P(A\,\&\,B)}{P(B)},$$

this means that the ratios of probabilities must be preserved – i.e., that the probabilities must be invariant modulo some additive constant.

Towards a description of probability measures. In order to describe all probability measures which are conditionally invariant under re-scalings, let us first recall how probability measures can be described.

A continuous probability distribution can be described by its probability density function $\rho(x)$.

Comment. More general distributions can also be described in similar terms, if we allow "generalized" functions (distributions) $\rho(x)$ – such as the delta-function $\delta(x)$ which is only equal to 0 for $x=0$ and for which $\int \delta(x)\,dx = 1$.

To avoid mathematical complications, we can simply consider such functions as limits of "normal" functions. For example, the delta-function can be viewed as a limit of functions $\delta_\varepsilon(x)$ for which:

- $\delta_\varepsilon(s) = \dfrac{1}{2\varepsilon}$ for $x \in [-\varepsilon, \varepsilon]$, and
- $\delta_\varepsilon(s) = 0$ for all other x.

How to describe probability measures which are conditionally invariant under re-scalings. Conditional invariance means that the probabilities may change by a multiplicative constant. If probabilities change by a multiplicative constant, this means that the corresponding probability densities also change by a constant. Thus, we arrive at the following definitions.

Definition. *We say that a function $\rho(x) \ge 0$ is* conditionally invariant under re-scalings *if for every two real numbers $\lambda > 0$ and s, there exists a constant $C(\lambda, s)$ such that*

$$\rho(\lambda \cdot x + s) = C(\lambda, s) \cdot \rho(x)$$

for all x.

Comment. One can easily check that a constant function $\rho(x) = \text{const}$ (corresponding to the uniform distribution) is conditionally invariant under re-scalings. It turns out that constant functions are the only functions with this invariance property:

Main result. *If a function $\rho(x)$ is conditionally invariant under re-scalings, then it is a constant function.*

Proof.

1°. Let us first prove that the function $\rho(x)$ is either always equal to 0, or always positive. In other words, we prove that if $\rho(x) = 0$ for some x, then $\rho(x') = 0$ for all real values x'.

Indeed, let us assume that for some x, we have $\rho(x) = 0$. Every real number x' can be represented as $x + (x' - x)$, i.e., as $\lambda \cdot x + s$, where $\lambda = -1$ and $s = x' - x$. Thus, we conclude that

$$\rho(x') = C(1, x' - x) \cdot \rho(x).$$

Since $\rho(x) = 0$, we get $\rho(x') = 0$.

2°. If $\rho(x) = 0$ for all x, then the function $\rho(x)$ is clearly a constant.

So, to complete the proof, in the remaining part of the proof, we will consider only the remaining case, when all the values of the function $\rho(x)$ are positive.

3°. Let us now prove that in this remaining case, the value $C(\lambda, s)$ does not depend on λ, i.e., that $C(\lambda, s) = C(s)$ for some function $C(s)$.

Indeed, for $x = 0$, the formula that describes conditional invariance takes the form

$$\rho(s) = C(\lambda, s) \cdot \rho(0).$$

Since $\rho(x) > 0$ for all x, we have $\rho(0) > 0$. Dividing both sides of the above equality by a positive number $\rho(0)$, we conclude that

$$C(\lambda, s) = \frac{\rho(s)}{\rho(0)}.$$

The right-hand side of this equality does not depend on λ, so we indeed conclude that $C(\lambda, s)$ only depends on s.

4°. In view of the statement from Part 3 of this proof, we have

$$\rho(\lambda \cdot x + s) = C(s) \cdot \rho(x)$$

for all λ, s, and x. Let us prove that under this condition, $\rho(x)$ is a constant function.

Indeed, let us take $x = 1$ and $s = 0$. Then, the above formula means that

$$\rho(\lambda) = C(0) \cdot \rho(1)$$

for all real values λ. The right-hand side of this equality does not depend on λ, so the function $\rho(x)$ is indeed constant.

The proposition is proven.

Comment. Solutions to similar functional equations are well-known; see, e.g., [1]; the above derivation is similar to the one from [4].

IX. Conclusion

We conclude that such a distribution should be uniform on the interval $[0,1]$, so the resulting probability of selecting the alternative v_1 under fuzzy uncertainty is $\int_0^1 P_1(\alpha)\,d\alpha$.

ACKNOWLEDGMENTS

This work was partially supported by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006–08.

The authors are thankful to the anonymous referees for valuable suggestions.

REFERENCES

[1] J. Aczel, *Lectures on functional equations and their applications*, Dover, New York, 2006.

[2] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré, and M. Aviles, "Computing Variance for Interval Data is NP-Hard", *ACM SIGACT News*, 2002, Vol. 33, No. 2, pp. 108–118.

[3] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré, and M. Aviles, "Exact Bounds on Finite Populations of Interval Data", *Reliable Computing*, 2005, Vol. 11, No. 3, pp. 207–233.

[4] V. N. Huynh, C. Hu, Y. Nakamori, and V. Kreinovich, *On Decision Making under Interval Uncertainty: A New Justification of Hurwicz Optimism-Pessimism Approach and Its Use in Group Decision Making*, University of Texas at El Paso, Department of Computer Science, Technical Report UTEP-CS-07-13, March 2007.

[5] V. N. Huynh, V. Kreinovich, Y. Nakamori, and H. T. Nguyen, *Towards Efficient Prediction of Decisions under Interval Uncertainty*, University of Texas at El Paso, Department of Computer Science, Technical Report, May 2007.

[6] V. N. Huynh, Y. Nakamori, and J. Lawry, "A probability-based approach to comparison of fuzzy numbers and applications to target-oriented decision making", *IEEE Transactions on Fuzzy Systems*, to appear.

[7] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*. Prentice Hall, Upper Saddle River, New Jersey, 1995.

[8] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

[9] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

[10] H. T. Nguyen, V. Kreinovich, and L. Longpré, "Dirty Pages of Logarithm Tables, Lifetime of the Universe, and (Subjective) Probabilities on Finite and Infinite Intervals", *Reliable Computing*, 2004, Vol. 10, No. 2, pp. 83–106.

[11] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 2004.

[12] D. Wagman, M. Schneider, and E. Schnaider, "On the use of interval mathematics in fuzzy expert systems", *International Journal of Intelligent Systems*, 1994, Vol. 9, pp. 241–259.

[13] R. R. Yager, M. Detyniecki, and B. Bouchon-Meunier, "A context-dependent method for ordering fuzzy numbers using probabilities", *Information Sciences*, 2001, Vol. 138, pp. 237–255.

APPENDIX

Formulation of the problem. We have n alternatives. For each i from 1 to n, we assume that the (unknown) actual value v_i is uniformly distributed in the known interval $[\underline{v}_i, \overline{v}_i]$, and that different values v_i are independent random variables.

Based on these intervals and corresponding distributions, we want to compute the probability P_i that v_i is the largest of n values $v_1, \ldots, v_n$.

Problem: how to compute P_i for large n? As we have mentioned, for $n = 2$, there are explicit formulas for P_i.

In general, since the distribution is uniform, the desired probability P_i is equal to the ratio V_i/V, where

$$V = (\overline{v}_1 - \underline{v}_1) \cdot \ldots \cdot (\overline{v}_n - \underline{v}_n)$$

is the (n-dimensional) volume of the box, and V_i is the volume of the part of which box for which v_i is larger than the values of all other values v_j.

In principle, we can compute the volume V_i by computing the corresponding n-dimensional integral. However, computing n-dimensional integrals with a given accuracy $\varepsilon > 0$ means that we have to consider a grid of size $\sim \varepsilon$ along each axis – i.e., consider $\sim \dfrac{1}{\varepsilon}$ points along each axis and $\sim \dfrac{1}{\varepsilon^n}$ points overall.

For large n, this computation time is too high to be practically useful. It is therefore desirable to come up with more efficient algorithms for computing P_i.

First idea: Monte-Carlo Simulations. A natural idea is to use Monte-Carlo simulations; see, e.g., [11]. Specifically, we select a number N, and then N times, we simulate each v_i as a uniformly distributed random variable. After that, we take N_i/N as an estimate for P_i, where N_i is the number of simulations in which v_i was the largest value.

It is known that the accuracy of the Monte-Carlo simulation is $1/\sqrt{N}$. So, to get 10% accuracy in computing P_i, it is sufficient to take $N \approx 100$ simulations.

Limitations of Monte-Carlo simulations. The main limitation of this approach is that if we want accurate estimates, with accuracy $\varepsilon \ll 1$, we need a large number of simulations $N \approx \dfrac{1}{\varepsilon^2}$. This number is not impossible (as for direct integration) but still large. It is therefore desirable to design an algorithm for computing P_i exactly.

Towards efficient algorithm for exact computations. We will describe an efficient ($O(n^2)$) algorithm for computing P_i. Without losing generality, we can assume that $i = 1$, i.e., that we need to compute the probability P_1 that v_1 is the largest of n values v_i. The outline of this section is as follows:

- First, we will describe the main idea behind this algorithm.
- Then, we will show how this idea translates into an actual $O(n^2)$ algorithm.
- Finally, we will explicitly describe the resulting algorithm.

Main idea behind the new algorithm. Our idea is to first describe, for each given v_1, the conditional probability $p_1(v_1)$ that this v_1 is the largest – under the condition that v_1 is the actual value. Then, due to the Bayes formula, the overall probability P_1 that v_1 is the largest can be obtained by integrating this conditional probability $p_1(v_1)$ times the probability density of v_1:

$$\mathrm{Prob}(v_1 \text{ is the largest}) =$$

$$\int \mathrm{Prob}(v_1 \text{ is the largest} \,|\, v_1 \text{ is actual}) \cdot \rho_1(v_1)\, dv_1.$$

The distribution of v_1 is uniform on the interval $[\underline{v}_1, \overline{v}_1]$, hence

$$P_1 = \frac{1}{\overline{v}_1 - \underline{v}_1} \cdot \int p_1(v_1)\, dv_1.$$

How can we describe the expression for $p_1(v_1)$? Once v_1 is fixed, the fact that v_1 is the largest means that $v_2 \le v_1$,

$v_3 \le v_1$, etc. Since all the variables v_i are independent, this probability is equal to the product of $n-1$ probabilities: the probability that $v_2 \le v_1$, the probability that $v_3 \le v_1$, etc.

For each i, the probability that $v_i < v_1$ can be determined as follows:

- If $\overline{v}_i \le v_1$, then $v_i \le v_1$ with probability 1. This probability does not change the product and can thus simply be omitted.
- If $v_1 < \underline{v}_i$, this means that $v_i \le v_1$ cannot happen at all. The resulting probability is 0, so such terms can be completely ignored.
- Finally, if $\underline{v}_i \le v_1 < \overline{v}_i$, then, since the distribution of v_i is uniform on the interval $[\underline{v}_i, \overline{v}_i]$, the probability that $v_i \le v_1$ is equal to $\dfrac{v_1 - \underline{v}_i}{\overline{v}_i - \underline{v}_i}$.

Thus, the conditional probability $p_1(v_1)$ is equal to

$$p_1(v_1) = \prod_{i: v_1 \le \overline{v}_i} \frac{v_1 - \underline{v}_i}{\overline{v}_i - \underline{v}_i},$$

if $v_1 \ge \underline{v}_i$ for all i, and to 0 otherwise.

Transforming this idea into the actual algorithm. As we see, the expression for $p_1(v_1)$ depends on the relation between v_1 and the endpoints $\underline{v}_i$ and $\overline{v}_i$ of the intervals $[\underline{v}_i, \overline{v}_i]$. So, if we sort these endpoints into an increasing sequence

$$v_{(1)} \le v_{(2)} \le \ldots \le v_{(2n)},$$

then, in each of the resulting $2n+1$ zones

$$z_0 = (-\infty, v_{(1)}), z_1 = [v_{(1)}, v_{(2)}), \ldots,$$

$$z_j = [v_{(j)}, v_{(j+1)}), \ldots, z_{2n} = [v_{(2n)}, \infty),$$

we will have the same analytical expression for $p_1(v_1)$.

For each zone, the corresponding expression is a product of $\le n$ linear terms. Multiplying these terms one by one, we get a polynomial of degree $\le n$ in $\le n$ computational steps.

The integral $\int p_1(v_1)\,dv_1$ can be computed as the sum of integrals p_{1j} over all the zones z_j, $j = 0, \ldots, 2n$. An integral of a polynomial

$$a_0 + a_1 \cdot v_1 + \ldots + a_k \cdot v_1^k$$

is equal to

$$a_0 \cdot v_1 + \frac{a_2}{2} \cdot v_1^2 + \ldots + \frac{a_k}{k+1} \cdot v_1^{k+1},$$

i.e., it can be also computed coefficient-by-coefficient in linear time. Since we have $2n$ zones, we thus need $(2n+1) \cdot O(n) = O(n^2)$ time to compute all $2n+1$ sub-integrals, and then $2n = O(n)$ operations to add them and get $\int p_1(v_1)\,dv_1$. Dividing this integral by $\overline{v}_1 - \underline{v}_1$, we get P_1. Thus, overall, we indeed need quadratic time.

Resulting algorithm. At the first step of this algorithm, we order all $2n$ endpoints $\underline{v}_i$ and $\overline{v}_i$ into an increasing sequence $v_{(1)} \le v_{(2)} \le \ldots \le v_{(2n)}$. As a result, we divide the real line into $2n+1$ zones

$$z_0 = (-\infty, v_{(1)}), z_1 = [v_{(1)}, v_{(2)}), \ldots,$$

$$z_j = [v_{(j)}, v_{(j+1)}), \ldots, z_{2n} = [v_{(2n)}, \infty).$$

For the zones z_j for which $v_{(j)} < \underline{v}_1$, $v_{(j+1)} > \overline{v}_1$, or $v_{(j+1)} < \underline{v}_i$ for some i, the integral p_{1j} is equal to 0.

For every other zone, we form the expression

$$p_1(v_1) = \prod_{i: v_{(j+1)} \le \overline{v}_i} \frac{v_1 - \underline{v}_i}{\overline{v}_i - \underline{v}_i}.$$

This expression is a product of $\le n$ linear functions of the unknown v_1. By multiplying by these functions one by one, we get an explicit expression for a polynomial in v_1. By processing the coefficients of this polynomial one by one, we can provide the explicit analytical expression for the (indefinite) integral $P_{1j}(v_1)$ of this polynomial. The desired integral p_{1j} can then be computed as the difference $P_{1j}(v_{(j+1)}) - P_{1j}(v_{(j)})$.

Finally, the desired probability p_1 is computed as

$$P_1 = \frac{1}{\overline{v}_1 - \underline{v}_1} \cdot \sum_{j=0}^{2n} p_{1j}.$$

Comment. The idea of dividing the real line into zones corresponding to sorted endpoints of the given intervals comes from another situation where we need to combine probabilities and intervals: namely, from the algorithms for algorithms for computing population variance under interval uncertainty [2], [3].