Uncertainty in Risk Analysis: Towards a General Second-Order Approach Combining Interval, Probabilistic, and Fuzzy Techniques

Scott Ferson[1], Lev Ginzburg[1], Vladik Kreinovich[2],
Hung T. Nguyen[3], and Scott A. Starks[2]

[1]Applied Biomathematics, 100 North Country Road,
Setauket, NY 11733, USA, {scott,lev}@ramas.com
[2]NASA Pan-American Center for Earth and Environmental Studies
University of Texas at El Paso, El Paso, TX 79968, USA
contact emails {vladik,sstarks}@utep.edu
[3]Department of Mathematical Sciences, New Mexico State University
Las Cruces, NM 88003, USA, hunguyen@nmsu.edu

***Abstract*— Uncertainty is very important in risk analysis. A natural way to describe this uncertainty is to describe a set of possible values of each unknown quantity (this set is usually an interval), plus any additional information that we may have about the probability of different values within this set. Traditional statistical techniques deal with the situations in which we have a complete information about the probabilities; in real life, however, we often have only partial information about them. We therefore need to describe methods of handling such partial information in risk analysis. Several such techniques have been presented, often on a heuristic basis. The main goal of this paper is to provide a justification for a general second-order formalism for handling different types of uncertainty.**

I. INTRODUCTION: UNCERTAINTY IN RISK ANALYSIS

Uncertainty in risk analysis: why. By definition, risk analysis deals with situations with uncertainty, i.e., with situations in which we do not have a complete and accurate knowledge about the state of the system. It is therefore very important that we be able to represent uncertainty in risk analysis as adequately as possible.

First component of uncertainty description: interval (set) uncertainty. In order to fully describe a system, we must know the exact values of all the physical quantities characterizing this system. For example, in environmental problems related to chemical pollution, a polluted system (e.g., a lake) can be fully described if we know the exact concentration of different pollutants in different parts of the lake.

Thus, to describe the uncertainty of our knowledge about a system, we must describe the uncertainty with which we know the values of each of the quantities (parameters) describing the system. Uncertainty means that we do not know the exact value of the quantity, several different values may be possible. For example, we may not know the exact value of the concentration but we may know that this concentration is between, say, 10^{-5} and 10^{-3}. In this case, any value from the interval $[10^{-5}, 10^{-3}]$ is possible; see, e.g., [7, 8, 18].

An important risk-related situation that leads to intervals is when a measurement does not detect any presence of a certain substance because its concentration x is below the detection limit D. In this case, the only information we have about x is that x belongs to the interval $[0, D]$.

In general, we usually known an interval $\mathbf{x}$ of possible values of the unknown quantity x – or, sometimes, a more general set X of possible values of x (different from an interval, e.g., the union of two intervals).

Second component of uncertainty description: probabilistic uncertainty. The set X of possible values describes which values of the analyzed quantity are possible and which values are not. In addition to this information, we often know which values are more probable and which are less probable. In other words, we often have some information about the probability of different values x from the interval (set) $\mathbf{x}$ of possible values.

Probabilistic uncertainty: traditional techniques. In some cases, we know the exact expression for this distribution. In these cases, we can use standard statistical techniques to represent, elicit, and aggregate uncertainty. A survey of the corresponding techniques as applied to risk analysis is given, e.g., in [1].

The need for techniques corresponding to partial information about probabilities. In many other real-life situations, however, we have only *partial* information about the probabilities. To handle such situations, it is necessary to expand known statistical techniques of representing, eliciting, and aggregating uncertainty to prob-

lems in which we only have partial information about the probabilities.

What we are planning to do. The main objective of this report is to introduce a new general technique dealing with partial information about probabilities. Several such techniques have been presented, often on a heuristic basis, without a proper justification typical for traditional statistical techniques. The main goal of this paper is to provide a justification for a general second-order formalism for handling different types of uncertainty.

For a survey with a detailed description of our approach see [6]; see also [2, 4, 5, 10, 11, 15].

II. WHAT IS A NATURAL WAY OF REPRESENTING PARTIAL INFORMATION ABOUT PROBABILITIES?

Which representation of probability distribution should we choose? In probability theory, there are many different ways of representing a probability distribution. For example, one can use a probability density function (pdf), or a cumulative distribution function (CDF), or a probability measure, i.e., a function which maps different sets into a probability that the corresponding random variable belongs to this set. The reason why there are many different representations is that in different problems, different representations turned out to be the most useful.

We would like to select a representation which is the most useful for problems related to risk analysis. To make this selection, we must recall where the information about probabilities provided by risk analysis is normally used.

How is the partial information about probabilities used in risk analysis? The main objective of risk analysis is to make decisions. A standard way of making a decision is to select the action a for which the expected utility (gain) is the largest possible. This is where probabilities are used: in computing, for every possible action a, the corresponding expected utility. To be more precise, we usually know, for each action a and for each actual value of the (unknown) quantity x, the corresponding value of the utility $u_a(x)$. We must use the probability distribution for x to compute the expected value $E[u_a(x)]$ of this utility.

In view of this application, the most useful characteristics of a probability distribution would be the ones which would enable us to compute the expected value $E[u_a(x)]$ of different functions $u_a(x)$.

Which representations are the most useful for this intended usage? General idea. Which characteristics of a probability distribution are the most useful for computing mathematical expectations of different functions $u_a(x)$? The answer to this question depends on the type of the function, i.e., on how the utility value u depends on the value x of the analyzed parameter.

Smooth utility functions naturally lead to moments. One natural case is when the utility function $u_a(x)$ is smooth. We have already mentioned, in Section I, that we usually know a (reasonably narrow) interval of possible values of x. So, to compute the expected value of $u_a(x)$, all we need to know is how the function $u_a(x)$ behaves on this narrow interval. Because the function is smooth, we can expand it into Taylor series. Because the interval is narrow, we can safely consider only linear and quadratic terms in this expansion and ignore higher-order terms:

$$u_a(x) \approx c_0 + c_1 \cdot (x - x_0) + c_2 \cdot (x - x_0)^2,$$

where x_0 is a point inside the interval. Thus, we can approximate the expectation of this function by the expectation of the corresponding quadratic expression:

$$E[u_a(x)] \approx E[c_0 + c_1 \cdot (x - x_0) + c_2 \cdot (x - x_0)^2],$$

i.e., by the following expression:

$$E[u_a(x)] \approx c_0 + c_1 \cdot E[x - x_0] + c_2 \cdot E[(x - x_0)^2].$$

So, to compute the expectations of such utility functions, it is sufficient to know the first and second moments of the probability distribution.

In particular, if we use, as the point x_0, the average $E[x]$, the second moment turns into the variance of the original probability distribution. So, instead of the first and the second moments, we can use the mean E and the variance V.

From numerical moments to interval-valued moments. When we know the exact probability distribution, we must use the *exact* values of the first and the second moment (or mean and variance).

If we only have a *partial* information about the probability distribution, then we cannot compute the exact value of these moments; instead, we have *intervals* of possible values of these moments. So, from this viewpoint, a natural representation of the partial information about the probability distribution is given by intervals $\mathbf{E}$ and $\mathbf{V}$ of possible values of mean E and variance V.

In risk analysis, non-smooth utility functions are common. In engineering applications, most functions are smooth, so usually the Taylor expansion works pretty well. In risk analysis, however, not all dependencies are smooth. There is often a threshold x_0 after which, say, a concentration of a certain chemical becomes dangerous.

This threshold sometimes comes from the detailed chemical and/or physical analysis. In this case, when we increase the value of this parameter, we see the drastic increase in effect and hence, the drastic change in utility value. Sometimes, this threshold simply comes from regulations. In this case, when we increase the value of this parameter past the threshold, there is no drastic increase in effects, but there is a drastic decrease of utility due

to the necessity to pay fines, change technology, etc. In both cases, we have a utility function which experiences an abrupt decrease at a certain threshold value x_0.

Non-smooth utility functions naturally lead to CDFs. We want to be able to compute the expected value $E[u_a(x)]$ of a function $u_a(x)$ which changes smoothly until a certain value x_0, then drops it value and continues smoothly for $x > x_0$. We usually know the (reasonably narrow) interval which contains all possible values of x. Because the interval is narrow and the dependence before and after the threshold is smooth, the resulting change in $u_a(x)$ before x_0 and after x_0 is much smaller than the change at x_0. Thus, with a reasonable accuracy, we can ignore the small changes before and after x_0, and assume that the function $u_a(x)$ is equal to a constant u^+ for $x < x_0$, and to some other constant $u^- < u^+$ for $x > x_0$.

The simplest case is when $u^+ = 1$ and $u^- = 0$. In this case, the desired expected value $E[u_a^{(0)}(x)]$ coincides with the probability that $x < x_0$, i.e., with the corresponding value $F(x_0)$ of the cumulative distribution function (CDF). A generic function $u_a(x)$ of this type, with arbitrary values u^- and u^+, can be easily reduced to this simplest case, because, as one can easily check, $u_a(x) = u^- + (u^+ - u^-) \cdot u^{(0)}(x)$ and hence, $E[u_a(x)] = u^- + (u^+ - u^-) \cdot F(x_0)$.

Thus, to be able to easily compute the expected values of all possible non-smooth utility functions, it is sufficient to know the values of the CDF $F(x_0)$ for all possible x_0.

From CDF to interval-valued CDF: the notion of a p-bound. When we know the exact probability distribution, we must use the exact values $F(x)$ of the CDF. If we only have a *partial* information about the probability distribution, then we cannot compute the exact values $F(x)$ of the CDF. Instead, for every x, we have an *interval* $[F^-(x), F^+(x)]$ of possible values of the probability $F(x)$. Such a pair of two CDFs $F^-(x)$ and $F^+(x)$ which bounds the (unknown) actual CDF is called a *probability bound*, or a *p-bound*, for short.

So, in risk analysis, a natural representation of the partial information about the probability distribution is given by a p-bound.

p-bounds or moments? We have shown that for decision problems with smooth utility functions, the best representation is by interval mean and interval variance, and for decision problems with discontinuous utility functions, the best representation of partial information is a p-bound.

Of the two corresponding representations of a probability distribution, CDF is much more informative: if we know CDF, we can compute the moments, but if we only know the moments, we can have many different CDFs. Thus, because we want to make our representation as informative as possible, it makes sense to use CDFs and their interval analogues – p-bounds.

Real numbers, intervals, and probability distributions are particular cases of p-bounds It is worth mentioning that several other types of uncertainty can be viewed as particular cases of p-bounds.

For example, the case of complete certainty, when we know the exact value x_0 of the desired quantity, can be represented as a p-bound in which

$$F^-(x) = F^+(x) = \begin{cases} 0 & \text{if } x \le x_0, \\ 1 & \text{otherwise} \end{cases}$$

The case when our only information about x is that x belongs to the interval $[x^-, x^+]$ can be represented by the following p-bound:

$$F^-(x) = \begin{cases} 0 & \text{if } x \le x^+, \\ 1 & \text{otherwise} \end{cases}$$

$$F^+(x) = \begin{cases} 0 & \text{if } x \le x^-, \\ 1 & \text{otherwise} \end{cases}$$

Finally, a probability distribution with a CDF $F(x)$ can be represented as a p-bound with $F^-(x) = F^+(x) = F(x)$.

Information about moments can also be represented in terms of p-bounds; see, e.g., [4, 15]. For example, if we know the interval $[x^-, x^+]$ on which the distribution is located, and if we know its mean E, then we can conclude that $F(x) \in [F^-(x), F^+(x)]$, where, e.g.,

$$F^+(x) = \min\left(1, \frac{x^+ - E}{x^+ - x}\right).$$

p-bounds have been successfully used in practice. We have shown that in risk analysis, a natural way to represent risk-related partial information about probabilities is by using a second-order probability distribution, namely, a *p-bound* – a pair of CDFs $F^-(x)$ and $F^+(x)$ for which $F^-(x) \le F^+(x)$. In particular, a real number, an interval, and a probability distribution are all particular cases of p-bounds.

p-bounds have been successfully used in different risk analysis problems ranging from problems related to pollution and environment to risk analysis for nuclear engineering; see, e.g., [4, 15].

III. P-BOUNDS FROM EXPERT ESTIMATES: CASE OF NATURAL-LANGUAGE ESTIMATES

Fuzzy properties: a general reminder. Words from natural language are not precise ("fuzzy"). Let us take the word "small" as an example. When the value of, say, concentration, is really small, everyone would 100% agree that this value is small indeed. When the value is really large, everyone would agree that this value is not small.

For intermediate values, however, we bound to have a disagreement.

The need to translate expert knowledge from natural language to a computer-understandable language of numbers was recognized as early as the 1960s, when the first expert systems started to be designed. A special formalism called *fuzzy logic* was designed to help us capture the meaning of words. In this formalism, to represent a meaning of a word like "small", we assign, to every possible value x, a degree $\mu_{\text{small}}(x)$ to which x is small. The dependence of this degree of x is called a *membership function*, or a *fuzzy set*.

Where do the values $\mu(x)$ come from? There are several dozen different techniques for eliciting these values; see, e.g., [9, 14]. Sometimes, the experts can present these numbers ("subjective probabilities") directly. If they cannot, then for every x, we can poll several (N) experts on whether they believe that this particular values x is, say, small, and if M out of N experts answer "yes", we take $\mu(x) = M/N$. What is a natural way to translate these "subjective probabilities" into p-bounds? We will answer this question on the example of membership functions of three most frequent types.

Fuzzy properties like "large". The first type is a function which describes words like "large", for which $\mu(x)$ is increasing from 0 at $x = 0$ to 1 for $x \to \infty$. Let us give a simple example of such function:

$$\mu_{\text{large}}(x) = \begin{cases} 0 & \text{if } x \le 1, \\ x - 1 & \text{if } 1 \le x \le 2, \\ 1 & \text{if } x > 2 \end{cases}$$

Suppose that the expert tells us that the actual value of some quantity X is large. What does it say about the possible values of the probability $F(x)$ (that $X \le x$) for different x?

Let us start with a value $x \le 1$. For this value, $\mu_{\text{large}}(x) = 0$. This means that the values below x cannot be large, so it is reasonable to take $F(x) = 0$.

Let us now take a value $x \ge 2$. For this value, $\mu_{\text{large}}(x) = 1$, which means that the value x is definitely large. Based on the expert opinion, we only know that the actual value X is large. It may be below x with probability 1 – in which case $F(x) = 1$; it may be above X with probability 1 – in which case $F(x) = 0$. So, here, the corresponding value of the p-bound – i.e., the interval of possible values of $F(x)$ – is $\mathbf{F}(x) = [0, 1]$.

What if x is in between 1 and 2, e.g., $x = 1.6$? In this case, the probability $\mu(x)$ that x is large is equal to 0.6. Since the function $\mu(x)$ is increasing, the probability $\mu(X)$ that X is large even smaller for $X < x$. Thus, out of all large values, values ≤ 0.6 should have a frequency ≤ 0.6. So, since we know that actual value X is large, we conclude that the probability $F(x)$ cannot exceed 0.6.

In general, the value $F(x)$ cannot exceed the probability $\mu(x)$, i.e., $\mu(x)$ serves as the upper part $F^+(x)$ of the p-bound. The lower part $F^-(x)$ should be 0, because we may have X so large than it is much larger than 2.

Combining these three cases, we conclude that for increasing membership functions $\mu(x)$ like "large", a natural translation of the membership function is a p-bound $[0, \mu(x)]$.

Fuzzy properties like "small". The second type of membership functions that we will consider is a function which describes words like "small", for which $\mu(x)$ is decreasing from 1 at $x = 0$ to 0 for $x \to \infty$. Let us give a simple example of such function:

$$\mu_{\text{small}}(x) = \begin{cases} 1 - x & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise} \end{cases}$$

Suppose that the expert tells us that the actual value of some quantity X is small. What does it say about the possible values of the probability $F(x)$ (that $X \le x$) for different x?

Let us start with a value $x \ge 1$. For this value, $\mu_{\text{small}}(x) = 0$, which means that the value x is definitely not small. Based on the expert opinion, we only know that the actual value X is small. All values X which can be small (i.e., for which $\mu(X) > 0$) are below 1, so they are all below x. Thus, all values of X are below x with probability 1, and $F(x) = 1$.

What if x is in between 0 and 1, e.g., $x = 0.2$? In this case, the probability $\mu(x)$ that x is small is equal to 0.8. Hence, the probability that any larger value $X > x$ is "small" also does not exceed 0.8. This means that if $F(x)$ is smaller than $1 - 0.8 = 0.2$ – e.g., equal to 0.1 – then there will be more than > 0.8 of values which are $\ge x$ – and thus, some values $X > x$ cannot be reasonably called small, in contradiction to the expert's opinion. So, if the actual value X is small, the probability $F(x)$ cannot exceed 0.2. In general, the value $F(x)$ cannot be smaller than $1 - \mu(x)$, i.e., $1 - \mu(x)$ serves as the upper part $F^-(x)$ of the p-bound. The upper part $F^+(x)$ should be 1, because we may have $X = 0$ with probability 1.

Combining these two cases, we conclude that for increasing membership functions $\mu(x)$ like "small", a natural translation of the membership function is a p-bound $[1 - \mu(x), 1]$.

Fuzzy properties like "around 1". Finally, we can consider membership functions describing terms like "around x_0", which increase from 0 to 1 until they reach a certain value x_0, and then decrease from 1 to 0. For such membership functions, possible values (i.e., values for which the degree $\mu(x)$ is large enough) are concentrated around the number x_0, that is why such membership functions are called *fuzzy numbers*.

As an example, we will consider the following function corresponding to "around 1":

$$\mu_{\approx 1}(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ 2 - x & \text{if } 1 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

For a membership function of this type, with a maximum at some value x_0, similar arguments lead to the following p-bound $[F^-(x), F^+(x)]$:

$$F^-(x) = \begin{cases} 0 & \text{if } x \le x_0, \\ 1 - \mu(x) & \text{if } x > x_0 \end{cases}$$

$$F^+(x) = \begin{cases} \mu(x) & \text{if } x \le x_0, \\ 1 & \text{if } x > x_0 \end{cases}$$

In particular, for the above membership function "around 1", the corresponding p-bound has the following form:

$$F^-(x) = \begin{cases} 0 & \text{if } x \le 1, \\ x - 1 & \text{if } 1 \le x \le 2, \\ 1 & \text{if } x > 2 \end{cases}$$

$$F^+(x) = \begin{cases} x & \text{if } x \le 1, \\ 1 & \text{if } x > 1 \end{cases}$$

These three cases can be described in a way which is similar to our transformation of measurements into p-bounds. Indeed, how can we describe a fuzzy set that corresponds to a certain property like "around 1"? A natural way to characterize a fuzzy set is to describe, for every level α, the set $X_\alpha = \{x \,|\, \mu(x) \ge \alpha\}$ of all the values which have this property with degree at least α. Such sets are called *α-cuts* – because on the graph, they really correspond to horizontal cuts. For example, for the above membership function "around 1", the α-cuts are $X_\alpha = [\alpha, 2 - \alpha]$.

If we, e.g., have α-cuts $X_{0.1}$, $X_{0.2}$, etc., corresponding to $\alpha = 0.1$, $\alpha = 0.2$, etc., this means, crudely speaking, that all experts agree that $x \in X_0$, that 90% of them agree that $x \in X_{0.1}$, that 80% of experts agree that $x \in X_{0.2}$, etc., until we reach we level $X_{0.9}$ in which only 10% of the experts agree; see, e.g., [13]. So, we have a natural subdivision of experts into 10 groups: 10% believe that x is somewhere on the interval $X_{0.1}$ – and no narrower bounds are possible; 10% believe that x is somewhere on the interval $X_{0.2}$ – and no narrower bounds are possible, etc. One can easily see that if we use the above algorithm to transform this knowledge base into a p-bound, we get exactly the p-bound that we came up with.

Second order fuzzy sets. p-bounds can be viewed as second-order probability distributions. We have shown that fuzzy descriptions can be naturally interpreted in terms of p-bounds. It can also be shown that second-order fuzzy descriptions (see, e.g., [12]) can also be naturally translated into the corresponding p-bounds.

IV. RESULTS: A BRIEF OVERVIEW

Aggregation problem. In many risk-related situations, we have several estimates for the same quantity: e.g., several measurements and/or several different expert estimates. As a result of each measurement, we get a p-bound. It is desirable to *aggregate* these p-bounds, i.e., to combine them into a single p-bound.

There are several ways to do that. In many applications (such as nuclear engineering) in which the error can be disastrous, it is desirable to be as cautious as possible. The most cautious way to combine the p-bounds $\mathbf{F}(x) = [F^-(x), F^+(x)]$ and $\mathbf{G}(x) = [G^-(x), G^+(x)]$ is to consider CDFs from both p-bounds as possible, and thus, to take the *union* $\mathbf{F}(x) \cup \mathbf{G}(x)$ of these p-bounds, i.e., a p-bound $[\min(F^-(x), G^-(x)), \max(F^+(x), G^+(x))]$.

In other applications, we may use some averaging techniques instead; see [6] for an overview.

Indirect measurements. In many risk-related applications, it is difficult (or even impossible) to directly measure the desired quantity y. Instead, we measure y *indirectly*, i.e., we measure some directly measurable quantities $x_1, \ldots, x_n$ which are related to y in a known way $y = f(x_1, \ldots, x_n)$, and then reconstruct y from the results of measuring x_i. So, we know p-bounds for each x_i, and we have some information (maybe none) about the dependence between x_i (e.g., that they are independent, or that they correlated in a known way, etc.), and we must estimate a p-bound for y.

When p-bounds are intervals, the resulting p-bound is also an interval. The problem of computing this interval is called the problem of *interval computations* [7, 8]. The original straightforward approach to solving interval computation problems takes into consideration the fact that inside a computer, each computation consists of elementary steps $\odot$ like $+$, $-$, $\cdot$, etc.

Therefore, first, for each elementary operation $a \odot b$, we find out how to perform the corresponding interval operation, i.e., how to transform intervals $\mathbf{a}$ and $\mathbf{b}$ of possible values of operands a and b into an (enclosure for) the interval of possible values of $a \odot b$. Then, we replace each elementary operation with numbers by the corresponding interval operation with intervals and thus, get, at the end, the enclosure for the desired interval $\mathbf{y}$ of possible values of $y = f(x_1, \ldots, x_n)$.

A similar approach can be applied to p-bounds. For p-bounds, the corresponding elementary operations are summarized in [4, 15].

V. FUTURE WORK

Combining moments with p-bounds. In Section II, we have argued that for decision problems with smooth utility functions, the best representation is by interval mean and interval variance, and for decision problems with discontinuous utility functions, the best representation of partial information is a p-bound. Therefore, ideally, if we want to be able to use our partial information about the probabilities in all possible decision problems, we should represent this partial information by keeping *both* p-bounds and interval moments.

One of the ultimate objectives of our research is to be able to handle such combined data. We are working on it, but as of now, we only have preliminary results.

Hybrid numbers. In the above text, we assumed that, when we are given an interval $\mathbf{x}$ of possible values of x, we are thus 100% guaranteed that the actual value x belongs to this given interval. In reality, often, there is a possibility that x is outside this interval. For example, the interval $\mathbf{x}$ may come from statistical analysis, when it arises as confidence interval corresponding to a certain confidence level.

In this case, we know, e.g., that x belongs to $\mathbf{x}$ with confidence 99%. In statistics, the probability of an error is usually used as a numerical characteristic of confidence; so, e.g., the case when $x \in \mathbf{x}$ in 99% of the cases is described by confidence level $\alpha = 0.01$.

In principle, most statistical methods enable us to make conclusions of different levels of confidence. Therefore, in addition to the original confidence interval that corresponds to a confidence level α, we can get another confidence level corresponding to a smaller confidence level $\alpha' < \alpha$. However, to decrease the confidence level, i.e., to increase our belief that the actual (unknown) value x belongs to the interval $\mathbf{x}$, we must widen the interval. So, the interval $\mathbf{x}_{\alpha'}$ correspond to the new confidence level $\alpha' < \alpha$ must contain the interval $\mathbf{x}_\alpha$ corresponding to the original confidence level α.

To describe uncertainty, it therefore makes sense to keep not just a *single* interval corresponding to a single confidence level, but to keep *several* intervals corresponding to different confidence levels. In more precise terms, we have several confidence values $0 < \alpha_1 < \ldots < \alpha_n < 1$, and for each of these values, we have intervals $\mathbf{x}_{\alpha_i}$ which are *nested* (i.e., when $\alpha_i < \alpha_j$, then $\mathbf{x}_{\alpha_i} \supseteq \mathbf{x}_{\alpha_j}$).

In particular, this conclusion can be applied to intervals $[F^-(x), F^+(x)]$ of possible values of CDF $F(x)$. So, instead of a single p-bound, we get several nested p-bounds corresponding to different confidence levels. This construction is called a *hybrid number*; see, e.g., [2]. For every x, $F(x)$ is a fuzzy set, so a hybrid number can be viewed as a second-order fuzzy set. For aggregation and indirect measurement with hybrid numbers, we also have only preliminary results.

Acknowledgments

This work was supported in part by NASA grants NCC5-209 and NCC 2-1232, by the Air Force Office of Scientific Research grants F49620-95-1-0518 and F49620-00-1-0365, by NSF grants CDA-9522207, ERA-0112968, and 9710940 Mexico/Conacyt, and by Small Business Innovation Research grant 9R44CA81741 to Applied Biomathematics from the National Cancer Institute (NCI), a component of the National Institutes of Health (NIH).

The authors are thankful to Jerry M. Mendel and Robert John for their encouragement.

References

[1] R. T. Clemen and R. L. Winkler, "Combining probability distributions from experts in risk analysis", *Risk Analysis*, 1999, Vol. 19, No. 2, pp. 187–203.

[2] J. A. Cooper, S. Ferson, and L. R. Ginzburg, "Hybrid processing of stochastic and subjective uncertainty data". Risk Analysis, 1996, Vol. 16, pp. 785–791.

[3] G. de Cooman, T. L. Fine, et al. *ISIPTA'01; 2nd Int'l Symposium on Imprecise Probabilities and Their Applications*, Cornell University, Ithaca, NY, 2001.

[4] S. Ferson, L. Ginzburg, and R. Akçakaya, "Whereof one cannot speak: when input distributions are unknown", *Risk analysis* (to appear).

[5] S. Ferson, L. Ginzburg, V. Kreinovich, and H. Schulte, "Interval Computations as a Particular Case of a General Scheme Involving Classes of Probability Distributions", In: [18], pp. 355–366.

[6] S. Ferson and V. Kreinovich, *Representation, Elicitation, and Aggregation of Uncertainty in Risk Analysis – From Traditional Probabilistic Techniques to More General, More Realistic Approaches: A Survey*, Technical Report, Applied Biomathematics, 2001, 135 pp.

[7] Interval computations website http://www.cs.utep.edu/interval-comp

[8] R. B. Kearfott and V. Kreinovich, *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

[9] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, 1995.

[10] V. Kreinovich et al., "From Interval Methods of Representing Uncertainty To A General Description of Uncertainty", In: H. Mohanty and C. Baral (eds.), *Trends in Information Technology*, Tata McGraw-Hill, New Delhi, 2000, pp. 161–166.

[11] V. Kreinovich, C. Langrand, and H. T. Nguyen, "Combining Fuzzy and Probabilistic Knowledge Using Belief Functions", *Proc. 2nd Vietnam-Japan Bilateral Symposium on Fuzzy Systems and Applications VJFUZZY'2001*, Hanoi, Vietnam, December 7–8, 2001, pp. 191–198.

[12] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, 2001.

[13] H. T. Nguyen and V. Kreinovich, "Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications", In: [8], pp. 245–290.

[14] H. T. Nguyen and E. A. Walker, *First Course in Fuzzy Logic*, CRC Press, Boca Raton, FL, 1999.

[15] Ramas website http://www.ramas.com

[16] P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, N.Y., 1991.

[17] T. Whalen, "Interval probabilities induced by decision problems", In: *Advances in the Dempster-Shafer Theory of Evidence*, John Wiley & Sons, 1994.

[18] J. Wolff von Gudenberg and W. Kraemer (eds.), *Scientific Computing, Validated Numerics, Interval Methods*, Kluwer, Dordrecht, 2001.