From Interval Methods of Representing Uncertainty To A General Description of Uncertainty

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email vladik@cs.utep.edu

Scott Ferson
Applied Biomathematics
100 North Country Road
Setauket, NY 11733, USA
email scott@ramas.com

Lev Ginzburg
Dept. of Ecology and Evolution
State University of New York
Stony Brook, NY 11794, USA

Harry Schulte
NASA NTRS
University of Texas at El Paso
El Paso, TX 79968, USA
email hschulte@utep.edu

Matthew R. Barry
Advanced Technology
United Space Alliance
600 Gemini
Houston TX 77058-2783, USA

Hung T. Nguyen
Dept. of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA
email hunguyen@nmsu.edu

Abstract

Measurements do not result in an exact value of the measured quantity; even after the most accurate measurement, there is still some uncertainty about the actual value of the measured quantity. Traditionally, in science and engineering, this uncertainty is characterized by a probability distribution; however, often, we do not know this probability distribution exactly. So, to get a more adequate description of this uncertainty, we must consider *classes* of possible probability distributions. A natural question is: Are all possible classes needed for this description? In this paper, we show that even for simple situations, we indeed need arbitrary closed convex classes of probability distributions.

Traditional Description of Uncertainty in Science and Engineering and its Drawbacks

Uncertainty is typical in science and engineering. A large portion of knowledge in science and engineering comes from measurements. Some of this knowledge comes not from measurements but from the expertise of scientists and engineers; however, to make sure that this additional knowledge is indeed correct, we must check it against the results of the measurements. In short, measurements are the basis of modern science and engineering.

By definition, a measurement means measuring the value of a physical quantity. Ideally, we would like to get the exact value of this quantity, but in real life, measurements are never 100% accurate. Due to inevitable noise, inaccuracies, etc., the results of the measurement are never absolutely accurate. After the measurement, we do not get the exact value of the measured quantity, because several close values are consistent with the same measurement result.

Therefore, when we process measurement results in science and engineering, we must take the measurement uncertainty into consideration.

Traditional approach to describing measurement uncertainty. To characterize the measurement uncertainty, we must know:

- which exactly values of the measured quantity are possible (provided the given measurement results), and
- which of these possible values are more probable.

In accordance with this idea, traditionally in science and engineering, uncertainty of the measurement result is characterized by describing the set of all possible values of the measured quantity, and in describing, for each possible value of the measured quantity, the probability that this value is the actual one. This probability can be viewed as a frequency of this particular value among all experiments in which we get the given measurement results.

Let us make this intuitive description more mathematically accurate. For continuous measured quantities, the probability of each exact value is usually 0, we can only talk about the probability of this value to be in a certain interval; so, instead of a probability of each value, we get a probability distribution on the set of all possible values. So, the traditional approach to describing uncertainty means that we have:

- a set of possible values of measured quantity, and
- a probability measure (distribution) on this set of possible values.

The traditional engineering approach enables us to process this statistical uncertainty.

The main problem with the traditional approach. In traditional probabilistic approach, we describe measurement uncertainty by describing probabilities of different values of the measured quantity. The main problem with this traditional approach is that in real life, we often do now know the exact values of these probabilities; sometimes, we do not have any information about these probabilities at all.

Towards A General Description of Uncertainty

Main idea. In traditional approach, we characterize measurement uncertainty by describing:

- the set X of possible values of the measured quantity and
- a probability measure P on the set X.

The main problem with this approach is that we do not know the exact probability measure. Therefore, it is natural to characterize the uncertainty by describing:

- the set X of possible values of the measured quantity and
- the *class* $\mathcal{P}$ of probability measure P on the set X which may describe the probability of different actual values.

We can consider two extreme particular cases of this general definition:

- The traditional statistical case, when the probability distribution is known, corresponds to the class $\mathcal{P} = \{P\}$ which consists of a single probability distribution P.
- In the other extreme, if we do not have any information about the probability measure, we must take, as $\mathcal{P}$, the class of all possible probability measures on the set X.

Towards formalization of the main idea: the class $\mathcal{P}$ must be convex. In the above description, we simply mentioned, informally, that the class $\mathcal{P}$ must describe all possible probability distributions which are consistent with our knowledge (and with our measurement results). How can we describe this condition in precise mathematical terms?

We will show that this informal description leads to two mathematically precise requirements on the class $\mathcal{P}$. The first of these requirements – convexity of the class $\mathcal{P}$ – can be deduced from the following argument.

We have already mentioned that if we know the probabilities, then the class $\mathcal{P}$ consists of a single probability measure; in all other cases, the class $\mathcal{P}$ contains at least two different probability measures. Let P_1 and P_2 be such measures. By definition of a class $\mathcal{P}$, the fact that both these measures belong to the class $\mathcal{P}$ means that both these measures can describe the frequency of different measured values in different situations. In other words, there is a situation type s_1 in which the frequencies are described by the measure P_1, and a situation type s_2 in which the frequencies are described by the measure P_2. Now, we can fix some integer N, pick N situations of type s_1 and N situations of type s_2, and consider these $2N$ situations as a new situation type s. A randomly chosen situation of this type is, with probability 1/2, of type 1, and, with probability 1/2, of type 2. We already know that for every event E, its frequency in the situations of the first type is equal to $P_1(E)$, and its frequency in the situations of the second type is equal to $P_2(E)$. Therefore, the frequency $P(E)$ of the event E in situations of the mixed type s is equal to $P(E) = (1/2) \cdot P_1(E) + (1/2) \cdot P_2(E)$. In other words, the new situation type s is characterized by a probability measure $P = (1/2) \cdot P_1 + (1/2) \cdot P_2$.

Let us summarize the above argument. We started with the assumption that P_1 and P_2 are two probability measures from the class $\mathcal{P}$, i.e., two probability measures which can represent frequencies in different situation types. We ended up by concluding that under this assumption, the convex combination $P = (1/2)\cdot P_1 + (1/2)\cdot P_2$ of these given probability measures can also describe frequencies in some situations, and therefore, this convex combination also belongs to the class $\mathcal{P}$.

Similarly, if instead of the equal numbers of situations of two types s_1 and s_2, we consider unequal numbers of situations, we can conclude that for every two probability measures $P_1, P_2 \in \mathcal{P}$, and for every real number $\alpha \in (0, 1)$, the convex combination $P = \alpha \cdot P_1 + (1 - \alpha) \cdot P_2$ of the probability measures P_i also describes frequencies and therefore, also belongs to the class $\mathcal{P}$. Thus, the class $\mathcal{P}$ must be *convex*.

The second requirement on the class $\mathcal{P}$: this class must be closed. By definition, the class $\mathcal{P}$ consists of all probability measures which correctly describe the frequencies in different situations. How can we experimentally confirm that a given probability measure P correctly describes the frequencies? In real life, we can only observe finitely many measurement results, so the resulting frequencies can only approximately describe probabilities. Thus, a more accurate way of saying that a probability measure P is consistent with the experiments is to say that for every accuracy $\varepsilon > 0$, we can find a situation type for which, within this accuracy, the probability measure P is consistent with the observed frequencies.

Let us show that from this "more accurate" informal definition, we can deduce the requirement that the class $\mathcal{P}$ be closed. To be more precise, we want to show that if the class $\mathcal{P}$ contains a sequence $\{P_n\}$ of probability measures, and this sequence tends to a limit P, then this limit also belongs to this class $\mathcal{P}$. To prove that the limit P belongs to the class $\mathcal{P}$, let us pick $\varepsilon > 0$ and show that there exists a situation type in which frequencies are ε-close to P. Indeed, by definition of a convergence, there exists an n for which the probability measure P_n is $(\varepsilon/2)$-close to P. Since P_n belongs to the class $\mathcal{P}$, it means that for every $\delta > 0$, in particular for $\delta = \varepsilon/2$, there exists a situation type for which the frequencies are δ-close to P_n. So, the frequencies are $\delta = (\varepsilon/2)$-close to P_n, and P_n is $(\varepsilon/2)$-close to P.

Hence, the frequencies corresponding to this situation s are $(\varepsilon/2)+(\varepsilon/2)=\varepsilon$-close to P. Since such a situation exists for every ε, we can conclude that the limit probability measure P belongs to the class $\mathcal{P}$. In other words, the class $\mathcal{P}$ is indeed closed.

Final step towards formalization: when are two probability measures ε-close? In the above argument, we did not specify what exactly we meant by saying that the two probability measures are ε-close; our argument is valid for an arbitrary choice of a metric on the set of all probability measures.

Which of the known metrics is the most appropriate for our case? In this paper, we will mainly consider measurements in which we measure a single physical quantity. In such measurements, possible values of the measured quantity form either the real line R, or a subset of the real line. Therefore, the corresponding probability measures are probability measures on a real line. It is reasonable to require that for every real number x, the event $\xi \le x$ has a probability. For such probability measures, we can define the probability distribution function $F(x)=P(\xi \le x)$. It is known that the probability distribution function uniquely defines such a probability measure. Therefore, we can describe the probability distribution by describing its probability measure. Correspondingly, the class $\mathcal{P}$ of probability measures can be described as a class of probability distribution functions, and closeness between probability measures can be described a closeness between distribution functions.

We would like to define this closeness in such a way that if some results of measuring the distribution function are consistent with some distribution F, and G is sufficiently close to F, then these same measurement results should be consistent with G as well. Intuitively, in order to experimentally determine the value $F(x)$ of a distribution function, we must describe a pair $(\widetilde{x},\widetilde{p})$, where $\widetilde{x}$ is a measured value of the quantity x, and $\widetilde{p}$ is the (approximate) measured value of the probability $P(\xi \le x)$. For example, we can take the ratio $N(x)/N$ as an approximate value of this probability, where N is the total number of outputs that we analyzed, and $N(x)$ is the total number of outputs for which the measured value was $\le x$.

For each measurement, both x and p are known only approximately. Let us denote the accuracy with which we know both numbers, by ε. Since the value x is known only approximately, the measured value $\widetilde{p}$ of $F(x)$ may represent not only $F(x)$, but also $F(x')$ for some value $x' \approx x$. If the density function $F(x)$ is continuous, then from the fact that x' is close to x, we can conclude that $F(x')$ is also close to $F(x)$. However, the function $F(x)$ can have discontinuities, i.e., the left limit $F(x-0)=\lim F(x-\delta)$ (for $\delta \downarrow 0$) may be different from the value $F(x)$. In this case, when x' is close to x, we cannot guarantee that the measured value of the probability is close to $F(x)$; we can only require that it should be close to some value in between $F(x-0)$ and $F(x)$.

We can then express the closeness as follows: for each distribution function, we can form the graph $F(x)$ of the distribution function. If this distribution function is continuous, this graph is all we need. If it is not, we complement it by drawing the graph of the "inverse" function $\varphi(p)=\sup\{x \mid F(x)\le p\}$. In effect, this addition makes the graph connected.

In this case, it is reasonable to say that F and G are ε-close if for an arbitrary x, the pair (x,p) on thus extended graph can (with accuracy ε) be interpreted as corresponding to G. In other words, we require that for every pair (x,p) on the extended graph of F, there exists a point (y,q) on the extended graph of G for which $|x-y|\le\varepsilon$ and $|p-q|\le\varepsilon$.

This seemingly reasonable definition has to be somewhat modified if we take into consideration the fact that the imperfection of the measuring instrument is reflected not only in its inaccuracy, but also in the fact that each measuring instrument has only a limited range. The larger the range, the close this instrument to the idea one. Therefore, instead of requiring that the above property hold for every x, we only require it for values x within the range, e.g., for the values x for which $|x|\le 1/\varepsilon$.

Now, we are ready for precise definitions:

A General Definition of Uncertainty: Precise Definitions

Definition 1. *Let $\varepsilon>0$. We say that the probability distributions $F(x)$ and $G(x)$ are ε-close if the following two properties hold:*

- *for every pair (x,p) on the extended graph of F for which $|x|\le 1/\varepsilon$, there exists a point (y,q) on the extended graph of G for which $|x-y|\le\varepsilon$ and $|p-q|\le\varepsilon$, and*
- *for every pair (y,q) on the extended graph of G for which $|y|\le 1/\varepsilon$, there exists a point (x,p) on the extended graph of F for which $|x-y|\le\varepsilon$ and $|p-q|\le\varepsilon$.*

Definition 2. *For every two probability distributions F and G, the distance $\rho(F,G)$ is defined as the smallest $\varepsilon>0$ for which F and G are ε-close.*

Comments.

- One can easily show that if F is ε-close to G, and G is δ-close to H, then F is $(\varepsilon+\delta)$-close to H. Thus, ρ is indeed a metric on the set of all possible probability measures.
- When we consider only probability measures which are located on a given interval, then the condition $|x|\le 1/\varepsilon$ is automatically satisfied for sufficiently small ε. For such ε, the above metric $\rho(F,G)$ becomes a *Hausdorff distance* between the (extended) graphs of the functions F and G with respect to sup-norm on the plane

$$\|(x_1,p_1)-(x_2,p_2)\|=\max(|x_1-x_2|,|p_1-p_2|).$$

In probability theory, this distance between probability measures is also known as *Levý distance* (see, e.g., [Shiryaev 1984], p. 314, [Levy 1937], or [Ito 1993], Vol. 2, pp. 1262–1263). Convergence in this metric is equivalent to so-called *weak convergence* of the distributions.

- In general, the above metric is a generalization of a Hausdorff metric. The topology corresponding to this generalized Hausdorff metric is described, e.g., in [Busemann 1955].

Definition 3. *By an uncertainty class $\mathcal{F}$, we mean a convex closed class of probability distributions (closed w.r.t. the metric ρ).*

Comment. As we have mentioned, the following are the basic examples of such classes:

- For every probability distribution F, a one-element class $\mathcal{F} = \{F\}$ consisting of this distribution F is an uncertainty class.
- For every interval $[a, b]$, the class of all distributions located on $[a, b]$, i.e., the class of all distributions F for which $F(a-0) = 0$ and $F(b) = 1$, also forms an uncertainty class. This class will be denoted by $\mathcal{F}_{[a,b]}$.

Data Processing Under General Uncertainty

Uncertainty of measurements: distributions or intervals. In some situations, measurement is a goal in itself: we are interested in the values of some physical quantity x, and we measure the value of this quantity. In such situations, all we need to know is the corresponding uncertainty class.

For such situations, the most widely used description of uncertainty is the uncertainty corresponding to a single probability distribution. In some real-life situations, we only know the upper bound Δ on the measurement error Δx defined as the difference $\Delta x = \widetilde{x} - x$ between the measured value $\widetilde{x}$ and the actual value x. In this situations, after we get the measurement result $\widetilde{x}$, the only thing we know about the actual value x of the measured quantity is that it belongs to the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. Therefore, it is reasonable to consider interval situations as well (for examples, see, e.g., [Kearfott et al. 1996], or http://cs.utep.edu/interval-comp).

So, on the basic level, to describe uncertainty of direct measurements, we only need two types of uncertainty classes described above. The natural question is: do we need to use more complicated uncertainty classes? If yes, which ones? To answer this question, we must take into consideration that direct measurements are not always possible and that often, we need indirect measurements and data processing.

Why data processing. In some situations, measurement is a goal in itself: we are interested in the values of some physical quantity x, and we measure the value of this quantity. In such situations, all we need to know is the corresponding uncertainty class.

In many real-life situations, however, it is difficult or even impossible to directly measure the value of the quantity y in which we are interested. For example, we cannot directly measure the distance to a quasar or the amount of oil in a well. In such situations, we can often determine the value of this quantity *indirectly*: namely,

- we measure the quantities $x_1, \ldots, x_n$ which are related to y and which can be measured directly, and then,
- we use the known relationship between x_i and y to estimate the value of y based on the measured values of x_i.

In many cases, this relationship has a functional form $y = f(x_1, \ldots, x_n)$ for some known function f. In such cases, as the result of indirect measurement, we take the result $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ of applying the function f to the results $\widetilde{x}_i$ of measuring the auxiliary quantities x_i. This application of an (algorithmically implemented) function f is called *data processing*.

What is the uncertainty of the result of data processing? For a direct measurement, we want to know not only the measured value of the measured quantity x, but also which values of x are actually possible, and what are the possible probability distributions on the set of all possible values of x. In other words, we want to know the *uncertainty class*.

Similarly, for an indirect measurement, we want to know not only the result $\widetilde{y}$ of the indirect measurement, we also want to know which values of $y = f(x_1, \ldots, x_n)$ are actually possible, and what are the possible probability distributions on the set of all possible values of y. In other words, we want to know the *uncertainty class* corresponding to y.

Two basic cases: independent distributions and no information about correlation. Even when we know exactly the probability distributions corresponding to the auxiliary quantities x_i, the uncertainty in the result $y = f(x_1, \ldots, x_n)$ of data processing is not determined uniquely: it depends on the correlation between the distributions x_i.

In traditional applications to science and engineering, it is usually assumed that the measurement errors are independent and therefore, that the probability distributions corresponding to different quantities x_i are *independent*.

- In some real-life measurement situations, we know that errors of different measurements come from completely independent sources and it is therefore reasonable to assume that the corresponding probability distributions are independent.
- In many other real-life situations, however, errors of different measurements may come from the same source, so independence is no longer justified; moreover, in many situations, the independence assumption turns out to be experimentally false.

Ideally, we would like to know the exact correlation between different measurement errors, but this information is rarely available. So, typically, we just know the probability distributions corresponding to different values x_i, and we have no information at all about the correlation between them. In terms of probability theory, the only information we have about the joint distribution of the variables $x_1, \ldots, x_n$ is the marginal distributions, i.e., the distributions corresponding to each variable x_i.

Thus, on the basic level, it makes sense to consider two types of situations:

- the situations when we know that the distributions corresponding to different x_i are independent, and
- the situations in which we have no knowledge about the dependence between the variables x_i.

Summary: basic situations of indirect measurement. We already know the basic types of uncertainty classes corresponding to each direct measurement: these classes correspond either to a probability distribution or to an interval. We have also described the basic ways of combining these uncertainty classes: independence or no information about dependence. Thus, to describe the basic uncertainty classes corresponding to indirect measurement, we must apply both basic combination ways to the basic uncertainty classes describing the variables x_i.

For simplicity, we will consider the simplest case when we only have two variables x_i. In some sense, these results will be the most general:

- first, some of these results will be easily generalizable to the case of several inputs x_i;
- second, even for two inputs, we will get, in some situations, the most general uncertainty classes, and so, adding more inputs will not change the results;
- third, sometimes, data processing of three or more variables can be described by first combining two of them, and then doing some additional data processing with the resulting combination and the third one; e.g., the arithmetic average $a = (x_1 + x_2 + x_3)/3$ can be described as first combining x_1 and x_2 into the preliminary average $a_p = (x_1 + x_2)/2$, and then combining this average with x_3 into $a = (2/3) \cdot a_p + (1/3) \cdot x_3$; in such situations, it is sufficient to be able to predict what will happen with uncertainty when we combine two variables.

Let us describe different basic cases of uncertainty in indirect measurements one by one:

- we will start with the cases when both (or all) x_i are characterized by probability distributions;
- then, we will describe the situations in which both (or all) x_i are characterized by interval uncertainty; and
- finally, we will consider *mixed* situations in which one of the variables is characterized by a probability distribution and another is characterized by an interval.

General description. Before we describe the results of this analysis, let us give a general formulation of this problem in precise terms.

Definition 4. *Let $f(x_1, \ldots, x_n)$ be a continuous function of n real variables, and let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be uncertainty classes (i.e., closed convex classes of probability distributions on the real line).*

- *We say that a distribution F_i on a real line is possible for x_i if $F_i \in \mathcal{F}_i$.*
- *We say that a joint distribution $P^{(n)}$ on R^n is possible if for every i, the marginal distribution is possible.*
- *We say that a joint distribution $P^{(n)}$ on R^n is possible under independence if $P^{(n)}$ is possible and $P^{(n)} = F_1 \times \ldots \times F_n$ (i.e., all x_i are independent random variables).*
- *We say that a distribution $F^{(1)}$ on R is possible if it corresponds to the distribution of the variable $y = f(x_1, \ldots, x_n)$ for some possible distribution $P^{(n)}$ on R^n.*
- *We say that a distribution $F^{(1)}$ on R is possible under independence if it corresponds to the distribution of the variable $y = f(x_1, \ldots, x_n)$ for some distribution $P^{(n)}$ on R^n which is possible under independence.*
- *By an uncertainty class $f(\mathcal{F}_1, \ldots, \mathcal{F}_n)$ corresponding to $y = f(x_1, \ldots, x_n)$, we mean the closed convex hull of the class of all possible distributions $F^{(1)}$.*
- *By an uncertainty class $f_{\text{indep}}(\mathcal{F}_1, \ldots, \mathcal{F}_n)$ corresponding to $y = f(x_1, \ldots, x_n)$ under independence, we mean the closed convex hull of the class of all distributions $F^{(1)}$ which are possible under independence.*

Comment. In the definition of the class $f(\mathcal{F}_1, \ldots, \mathcal{F}_n)$, we do not need to require that we take a convex hull.

Case 1 (traditional): independent probability distributions. If we know the probability distributions for $x_1, \ldots, x_n$, and we know that these distributions are independent, then we have the complete description of the join distribution for these n variables $x_1, \ldots, x_n$. Based on this joint distribution, we can determine the probability distribution for the function $f(x_1, \ldots, x_n)$. In this case, we start with probability distributions, and we end up with a probability distribution.

Formally, $f_{\text{indep}}(\{F_1\}, \ldots, \{F_n\})$ is a one-element uncertainty class, i.e., an uncertainty class consisting of a single probability distribution.

Case 2: possible dependent probability distributions. This case is analyzed by copula theory; for latest developments, see, e.g., [Nelsen 1999].

Cases 3 and 4: independent and possibly dependent intervals. In these cases, we have a continuous function $f(x_1, \ldots, x_n)$, and each variable x_i, $1 \le i \le n$, is described by an interval $[x_i^-, x_i^+]$, i.e., by a class $\mathcal{F}_{[x_i^-, x_i^+]}$ of all probability measures which are located on this interval. In this case, as we will see from the following result, the uncertainty class corresponding to $y = f(x_1, \ldots, x_n)$ is also an interval, namely, the image interval

$$[y^-, y^+] = f([x_1^-, x_1^+], \ldots, [x_n^-, x_n^+]) =$$

$$\{y = f(x_1, \ldots, x_n) \,|\, x_1 \in [x_1^-, x_1^+], \ldots, x_n \in [x_n^-, x_n^+]\}.$$

Proposition 1. *Let $f(x_1, \ldots, x_n)$ be a continuous function of n real variables, and let $[x_1^-, x_1^+]$, $\ldots$, $[x_n^-, x_n^+]$ be n intervals. Then,*

$$f_{\rm indep}\left(\mathcal{F}_{[x_1^-, x_1^+]}, \ldots, \mathcal{F}_{[x_n^-, x_n^+]}\right) = \mathcal{F}_{[y^-, y^+]};$$

$$f\left(\mathcal{F}_{[x_1^-, x_1^+]}, \ldots, \mathcal{F}_{[x_n^-, x_n^+]}\right) = \mathcal{F}_{[y^-, y^+]}.$$

In other words, interval uncertainty for x_i leads to interval uncertainty for $y = f(x_1, \ldots, x_n)$ irrespective on whether the inputs are independent or not.

Comment. It is worth mentioning that, in contrast to the comment after Definition 4, this result will not hold if we do not define the desired class $f_{\rm indep}(\mathcal{F}_{[x_1^-, x_1^+]}, \ldots, \mathcal{F}_{[x_n^-, x_n^+]})$ as a convex hull.

Mixed case: independent probability distribution and interval. In this case, as we will see, we get, in effect, an arbitrary uncertainty class.

To be more precise, we will show that for every uncertainty class $\mathcal{F}$ and for every $\varepsilon > 0$, we can get, in this way, an uncertainty class which is ε-close to $\mathcal{F}$. To formulate this result precisely, we need to define the notion of "closeness" (i.e., metric) on the set of all uncertainty classes. This can be done as follows:

- We have already defined a metric $\rho(F, F')$ on the set of all probability distributions. This metric is a minor modification of a Hausdorff distance between the (extended) graphs of the corresponding distribution functions, where the Hausdorff distance between the two sets A and B is defined as the smallest ε for which:
  - for every $a \in A$, there exists a $b \in B$ for which $\|a - b\| \le \varepsilon$, and
  - for every $b \in B$, there exists an $a \in A$ for which $\|a - b\| \le \varepsilon$.
- We can now go one step further and define a distance $\rho_H(\mathcal{F}, \mathcal{G})$ between the two uncertainty classes $\mathcal{F}$ and $\mathcal{G}$ as the Hausdorff distance between these classes, i.e., as the smallest ε for which:
  - for every $F \in \mathcal{F}$, there exists a $G \in \mathcal{G}$ for which $\rho(F, G) \le \varepsilon$, and
  - for every $G \in \mathcal{G}$, there exists an $F \in \mathcal{F}$ for which $\rho(F, G) \le \varepsilon$.

Theorem 1. *For every uncertainty class $\mathcal{F}$, and for every $\varepsilon > 0$, there exists a probability distribution F, an interval $[a, b]$, and a continuous function $f(x_1, x_2)$ for which the uncertainty class $f_{\rm indep}(\{F\}, \mathcal{F}_{[a,b]})$ is ε-close to $\mathcal{F}$ (in the sense of the Hausdorff metric ρ_H).*

Comments.

- In other words, even in the simplest case when we only have two independent variables, one of which has a precisely known distribution and the other is characterized by an interval of possible values, we can get an arbitrary uncertainty classes. Thus, arbitrarily complicated uncertainty classes are not simply the result of a mathematical definition, they are practically possible.
- Of course, this result does not necessarily means that we have to use arbitrarily complex uncertainty classes; we can restrict ourselves to a collection of simpler uncertainty classes such as p-bounds (see, e.g., [Ferson et al. 1995]). In this case, our result says that in some real-life situations, the description by these simple uncertainty classes will not be precise, it will only be an *approximation*.
- We are currently working on the applications of this general formalism to problems ranging from ecology to Space Shuttle control to optimal organization of web-based kiosks.

Acknowledgments. This work was supported in part by NASA under cooperative agreement NCC5-209, by NSF grants No. DUE-9750858 and CDA-9522207, by United Space Alliance, grant No. NAS 9-20000 (PWO C0C67713A6), by the Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518, and by the National Security Agency under Grant No. MDA904-98-1-0561 and MDA904-98-1-0564.

References

[Busemann 1955] H. Busemann, *The geometry of geodesics*, Academic Press, N.Y., 1955.

[Ferson et al. 1995] S. Ferson and L. Ginzburg, "Hybrid arithmetic", In: B. M. Ayyub (ed.), *Proceedings of the ISUMA-NAFIPS'95*, IEEE Computer Society Press, Los Alamitos, CA, 1995, pp. 619–623.

[Ito 1993] K. Ito (ed.), *Encyclopedic dictionary of mathematics*, MIT Press, Cambridge, MA, 1993.

[Kearfott et al. 1996] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

[Levy 1937] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.

[Nelsen 1999] R. B. Nelsen, *An introduction to copulas*, Springer-Verlag, New York, 1999.

[Shiryaev 1984] A. N. Shiryaev, *Probability*, Springer-Verlag, N.Y., 1984.