How to Measure a Degree of Mismatch Between Probability Models, p-Boxes, etc.: A Decision-Theory-Motivated Utility-Based Approach

Luc Longpré
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
email longpre@utep.edu

Scott Ferson and W. Troy Tucker
Applied Biomathematics
100 North Country Road
Setauket, New York 11733
scott@ramas.com, troy@tucker.com

***Abstract*—Different models can be used to describe real-life phenomena: deterministic, probabilistic, fuzzy, models in which we have interval-valued or fuzzy-valued probabilities, etc. Models are usually not absolutely accurate. It is therefore important to know how accurate is a given model. In other words, it is important to be able to measure a mismatch between the model and the empirical data. In this paper, we describe an approach of measuring this mismatch which is based on the notion of utility, the central notion of utility theory.**

We also show that a similar approach can be used to measure the loss of privacy.

I. Formulation of the Problem

A. Models are usually approximate

In most areas of science and engineering, we only have *approximate* models for the real-world phenomena, i.e., models which are not 100% accurate. Since the models are approximate, their predictions are also only approximate.

B. It is desirable to gauge the accuracy of a model

In order to understand how accurate are the models' predictions, we need to know how accurate are the models themselves.

An ideal way to gauge the quality of a model is to compare it with the empirical data, i.e., to *validate* this model.

C. Simplest case: deterministic phenomena

Let us start with the simplest situation, when we have a deterministic phenomenon and we have a deterministic model which describes this phenomenon. In this situation, we can simply compare the measured value of the desired quantity with the values predicted by the model.

In such a situation, the difference between the actual and predicted values is a reasonable measure of a mismatch between the real-life phenomenon and the model.

D. Real-life situation: non-deterministic phenomena

In real life, many phenomena are non-deterministic. For such phenomena, we cannot predict the exact values of the corresponding quantities; at best, we can predict the *probabilities* of different values of these quantities.

To validate such models, we must therefore compare the predicted probability distribution with the empirical probability distribution. In such situations, it is not completely clear how we can measure the mismatch between the corresponding probability distributions, i.e., how we can gauge the validity of the probabilistic models.

E. Additional complexity and relation to fuzzy techniques

In practice, the situation is even more complex. Based on a finite sample of real-life events, we cannot uniquely determine the corresponding empirical distribution: we can only provide, with different degrees of confidence, bounds on the corresponding probabilities.

In other words, for each event, instead of a single value of its probability, we get a nested family of confidence intervals corresponding to different levels of uncertainty. Nested families are, in effect, equivalent to fuzzy numbers; see, e.g., [2], [7], [8], [9], so hopefully, techniques for processing fuzzy numbers will be helpful in this comparison.

F. What we do in this paper

In this paper, we mainly consider the case of probability distributions. The last section discusses the possibility of extending these results to a more general case of interval-valued probability distributions (p-boxes) and nested (= fuzzy) families of such interval-valued objects.

We also show that a similar approach can be used to measure the degree of privacy loss.

Comment. Some of our privacy-related results appeared in [1].

II. Our Main Idea: Utility Approach

A. Utility approach: a reminder

In decision making theory, it is proven that under certain reasonable assumptions, a person's preferences are defined by his or her *utility function* $U(x)$ which assigns to each possible outcome x a real number $U(x)$ called *utility*; see, e.g., [6], [10].

In many real-life situations, a person's choice s does not determine the outcome uniquely, we may have different outcomes $x_1, \ldots, x_n$ with probabilities, correspondingly, $p_1, \ldots, p_n$. For such a choice, we can describe the utility $U(s)$ associated with this choice as the expected value of the utility of outcomes:

$$U(s) = E[U(x)] = p_1 \cdot U(x_1) + \ldots + p_n \cdot U(x_n).$$

Among several possible choices, a user selects the one for which the utility is the largest: a possible choice s is preferred to a possible choice s' (denoted $s > s'$) if and only if

$$U(s) > U(s').$$

In the general case, when we have a (1-dimensional) probability distribution with the cumulative distribution function (cdf) $F(x)$, the utility is described by a similar formula

$$U(s) = E[U(x)] = \int U(x)\, dF(x).$$

In particular, in the continuous case, when we have a probability distribution with the probability density (pdf) $\rho(x)$, the utility is described by a formula

$$U(s) = E[U(x)] = \int U(x) \cdot \rho(x)\, dx.$$

B. Application of utility approach to the problem of measuring a mismatch between probability distributions

Since our preferences are characterized by the utility values, it is reasonable to measure mismatch by the possible decrease in utility. Specifically, let $F_1(x)$ denote the cdf of the model, and $F_2(x)$ denote the (usually unknown) cdf of the actual distribution. Similarly, in the continuous case, we will denote the pdf corresponding to the model by $\rho_1(x)$ and the actual pdf by $\rho_2(x)$.

In these terms, if we make a decision based on the model distribution $F_1(x)$, then the expect value of utility is

$$U_1 = \int U(x)\, dF_1(x).$$

Since the actual distribution is different, the actual value of the expected utility is equal to

$$U_2 = \int U(x)\, dF_2(x).$$

If the actual expected utility is smaller than the what we planned, i.e., if $U_2 < U_1$, then we have a loss caused by the mismatch. It is therefore reasonable to characterize the mismatch by this loss $U_1 - U_2$.

This loss describes the effect of the mismatch on a specific problem characterized by a specific utility function $U(x)$. Usually, a model is used for *many* different applications, with many different utility functions. In some applications, the difference between the two probability distribution may be irrelevant for our objectives; in this case, there is no loss. In other situations, this different may lead to a significant loss.

It is reasonable to gauge the mismatch by the worst possible loss caused by this mismatch.

C. Which functions $U(x)$ should we consider

In different situations, we may have different utility functions $U(x)$ that describe the dependence of a (predicted) gain on the (unknown) actual value of the corresponding parameter x.

This prediction only makes sense only if we can predict $U(x)$ for each situation with a reasonable accuracy, e.g., with an accuracy $\varepsilon > 0$. Measurements are never 100% accurate, and measurement of x are not exception. Let us denote by δ the accuracy with which we measure x, i.e., the upper bound on the (absolute value of) the difference $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ between the measured value $\widetilde{x}$ and the (unknown) actual value x. Due to this difference, the estimated value $U(\widetilde{x})$ is different from the ideal prediction $U(x)$. Usually, measurement errors Δx are small, so we can expand the prediction inaccuracy

$$\Delta U \stackrel{\text{def}}{=} U(\widetilde{x}) - U(x) = U(x + \Delta x) - U(x)$$

in Taylor series in Δx and ignore quadratic and higher order terms in this expansion, leading to $\Delta U \approx U'(x) \cdot \Delta x$, where $U'(x)$ denotes the derivative of the utility function $U(x)$.

Since the largest possible value of Δx is δ, the largest possible value for ΔU is thus $|U'(x)| \cdot \delta$. Since this value should not exceed ε, we thus conclude that $|U'(x)| \cdot \delta \le \varepsilon$, i.e., that $|U'(x)| \le M \stackrel{\text{def}}{=} \varepsilon/\delta$.

Thus, we arrive at the following definition.

III. Definition and the Main Result

Definition 1. *Let $F_1(x)$ and $F_2(x)$ be two probability distributions on a real line, and let $M > 0$ be a real number. By the* degree of mismatch $d_M(F_1, F_2)$ *between the distributions, we mean the largest possible value of the difference*

$$\int U(x)\, dF_1(x) - \int U(x)\, dF_2(x)$$

over all possible functions $U(x)$ for which $|U'(x)| \le M$ for all x.

Proposition 1. *For every two distributions,*

$$d_M(F_1, F_2) = M \cdot \int |F_1(x) - F_2(x)|\, dx.$$

Comment. In view of this result, it is reasonable to measure the mismatch between two probability distributions by the following L_1-*metric*:

$$d(F_1, F_2) \stackrel{\text{def}}{=} \int |F_1(x) - F_2(x)|\, dx.$$

This metric was indeed proposed and successfully used in model validation [4], [5]. The above result shows that this metric is not only reasonable, it follows from the general decision theory-motivated utility-based approach.

IV. Proof of the Main Result

The desired difference $\Delta U = U_1 - U_2$ can be reformulated as the integral

$$\Delta U = \int U(x)\,(dF_1(x) - dF_2(x)).$$

Integrating this expression by parts, we conclude that

$$\Delta U = \int (F_1(x) - F_2(x)) \cdot U'(x)\,dx.$$

Since $|U'(x)| \le M$, we conclude that

$$(F_1(x) - F_2(x)) \cdot U'(x) \le |(F_1(x) - F_2(x)) \cdot U'(x)| =$$

$$|F_1(x) - F_2(x)| \cdot |U'(x)| \le M \cdot |F_1(x) - F_2(x)|$$

and thus,

$$\Delta U = \int (F_1(x) - F_2(x)) \cdot U'(x)\,dx \le$$

$$\int M \cdot |F_1(x) - F_2(x)|\,dx = M \cdot \int |F_1(x) - F_2(x)|\,dx.$$

So, we have $\Delta U \le d_M(F_1, F_2)$ for all possible utility functions $U(x)$. Thus, the largest possible value of ΔU cannot exceed $d_M(F_1, F_2)$.

Let us now show that the largest possible value of ΔU is actually equal to $d_M(F_1, F_2)$, i.e., the value $\Delta U = d_M(F_1, F_2)$ is attained for some utility function $U(x)$. Indeed, as such a utility function, we can take

$$U(x) \stackrel{\text{def}}{=} \int_{-\infty}^{x} M \cdot \text{sign}(F_1(t) - F_2(t))\,dt,$$

where $\text{sign}(u)$ is defined as usual:

- when $u > 0$, we define $\text{sign}(u) = +1$;
- when $u < 0$, we define $\text{sign}(u) = -1$;
- when $u = 0$, we define $\text{sign}(u) = 0$.

For this utility function, $U'(x) = M \cdot \text{sign}(F_1(x) - F_2(x))$ and thus, $|U'(x)| \le M$ for all x. On the other hand, for this function,

$$\Delta U = \int (F_1(x) - F_2(x)) \cdot U'(x)\,dx =$$

$$\int (F_1(x) - F_2(x)) \cdot M \cdot \text{sign}(F_1(x) - F_2(x))\,dx.$$

For each value u, we have $u \cdot \text{sign}(u) = |u|$. Thus,

$$\Delta U = \int M \cdot |F_1(x) - F_2(x)|\,dx,$$

i.e., indeed, $\Delta U = d_M(F_1, F_2)$. The proposition is proven.

V. How to Extend the Measure of Mismatch to p-Boxes and to Fuzzy-Valued p-Boxes

A. Extension to p-boxes

In practice, based on the empirical data, we cannot uniquely determine the corresponding probabilities $F(x)$. Instead, we can have confidence intervals $[\underline{F}(x), \overline{F}(x)]$ that contain the (unknown) values of these probabilities; see, e.g., [11], [12]. Such an interval-valued function that assigns, to every real number x, the corresponding interval $[\underline{F}(x), \overline{F}(x)]$ is called a *p-box*; see, e.g., [3]. Once we fix the confidence level, we thus have a *p-box* that contains all probability distributions which are consistent with the given empirical data.

In this situation, when the empirical data is describe by a p-box, how can we describe to what extent a given probability model is consistent with the empirical data? If the model $F(x)$ fits within the p-box $\mathbf{F}$ ($F \in \mathbf{F}$), i.e., if $F(x) \in [\underline{F}(x), \overline{F}(x)]$ for all x, this means that we have a perfect match.

In general, it is reasonable to define the degree of mismatch as the smallest possible mismatch between a model F_1 and distributions from a given p-box:

$$d(F_1, \mathbf{F}) = \min_{F_2 \in \mathbf{F}} d(F_1, F_2).$$

A model itself is not necessarily formulated in precise probabilistic terms. For example, we can say that according to our model, we have a normal distribution with the mean between -0.1 and 0.1. In this case, a model is also naturally described by a p-box. In such situations, it is reasonable to define the mismatch between the p-box $\mathbf{F}_1$ describing the model and the p-box $\mathbf{F}_2$ describing the empirical distribution as the smallest possible mismatch between the probability distributions $F_1 \in \mathbf{F}_1$ and $F_2 \in \mathbf{F}_2$:

$$d(\mathbf{F}_1, \mathbf{F}_2) = \min_{F_1 \in \mathbf{F}_1, F_2 \in \mathbf{F}_2} d(F_1, F_2).$$

B. Case of fuzzy uncertainty

Instead of fixing a single confidence level, it is reasonable to consider confidence intervals $\mathbf{F}^{(\alpha)}(x)$ corresponding to different confidence levels α. The resulting nested family of intervals can be naturally viewed as a fuzzy number for which these intervals are α-cuts; see, e.g., [2], [7], [8], [9]. Alternatively, the probabilities $F(x)$ may be given by experts and thus, can be naturally represented as fuzzy numbers. In both case, it is reasonable to characterize the mismatch between the corresponding "fuzzy-valued" probability distributions $\mathbf{F}_1$ and $\mathbf{F}_2$ as a function that assigns, to every level α, the degree of mismatch between the corresponding α-cuts:

$$d^{(\alpha)}(\mathbf{F}_1, \mathbf{F}_2) = \min_{F_1 \in \mathbf{F}_1^{(\alpha)}, F_2 \in \mathbf{F}_2^{(\alpha)}} d(F_1, F_2).$$

VI. How to Measure Loss of Privacy: An Auxiliary Problem

A. Measuring loss of privacy is important

Privacy means, in particular, that we do not disclose all the information about ourselves. If some of the originally undisclosed information is disclosed, some privacy is lost. To

compare different privacy protection schemes, we must be able to gauge the resulting loss of privacy.

B. Seemingly natural idea: measuring loss of privacy by the acquired amount of information

Since privacy means that we do not have complete information about a person, a seemingly natural idea is to gauge the loss of privacy by the amount of new information that we gained about this person.

C. Why information is not always a perfect measure of loss of privacy

In our opinion, the amount of new information is not always a good measure of the loss of privacy because it does not distinguish between:

- crucial information that may seriously affect a person, and
- irrelevant information – that may not affect a person at all.

To make a distinction between these two types of information, let us estimate potential financial losses caused by the loss of privacy.

D. Example when loss of privacy can lead to a financial loss

As an example, let us consider how a person's blood pressure x affects the premium that this person pays for his or her health insurance.

From the previous experience, insurance companies can deduce, for each value of blood pressure x, the expected (average) value of the medical expenses $f(x)$ of all individuals with this particular value of blood pressure. So, when the insurance company knows the exact value x of a person's blood pressure, it can offer this person an insurance rate $F(x) \stackrel{\text{def}}{=} f(x) \cdot (1+\alpha)$, where α is the general investment profit. Indeed:

- If an insurance company offers higher rates, then its competitor will be able to offer lower rates and still make a profit.
- On the other hand, if the insurance company is selling insurance at a lower rate, then it will not earn enough profit, and investors will pull their money out and invest somewhere else.

To preserve privacy, we only keep the information that the blood pressure of all individuals from a certain group is between two bounds L and U, and we do not know have any additional information about the blood pressure of different individuals. Under this information, how much will the insurance company charge to insure people from this group?

Based on the past experience, the insurance company is able to deduce the relative frequency of different values $x \in [L, U]$ – e.g., in the form of the corresponding probability density $\rho(x)$. In this case, the expected medical expenses of an average person from this group are equal to

$$E[f(x)] \stackrel{\text{def}}{=} \int \rho(x) \cdot f(x)\, dx.$$

Thus, the insurance company will insure the person for a cost of

$$E[F(x)] = \int \rho(x) \cdot F(x)\, dx.$$

Let us now assume that for some individual, the privacy is lost, and for this individual, we know the exact value x_0 of his or her blood pressure. For this individual, the company can now better predict its medical expenses as $f(x_0)$ and thus, offer a new rate $F(x_0) = f(x_0) \cdot (1+\alpha)$. When

$$F(x_0) > E[F(x)],$$

the person whose privacy is lost also experiences a financial loss $F(x_0) - E[F(x)]$. We will use this financial loss to gauge the loss of privacy.

E. Need for a worst-case comparison

In the above example, there is a financial loss only if the person's blood pressure x_0 is worse than average. A person whose blood pressure is lower than average will only benefit from reduced insurance rates.

However, in a somewhat different situation, if the person's blood pressure is smaller (better) than average, this person's loss or privacy can also lead to a financial loss. For example, an insurance company may, in general, pay for a preventive medication that lowers the risk of heart attacks – and of the resulting huge medical expenses. The higher the blood pressure, the larger the risk of a heart attack. So, if the insurance company learns that a certain individual has a lower-than-average blood pressure and thus, a lower-than-average risk of a heart attack, this risk may not justify the expenses on the preventive medication. Thus, due to a privacy loss, the individual will have to pay for this potentially beneficial medication from his/her own pocket – and thus, also experience a financial loss.

So, to gauge a privacy loss, we must consider not just a single situation, but several different situations, and gauge the loss of privacy by the worst-case financial loss caused by this loss of privacy.

F. Which functions $F(x)$ should we consider

Similarly to the main part of the text, we should consider functions $F(x)$ for which $|F'(x)| \le M$ for some given number $M > 0$.

G. Resulting definitions

Thus, we arrive at the following definition:

Definition 2. *Let $\mathcal{P}$ be a class of probability distributions on a real line, and let $M > 0$ be a real number. By the* amount of privacy *$A(\mathcal{P})$ related to $\mathcal{P}$, we mean the largest possible value of the difference $F(x_0) - \int \rho(x) \cdot F(x)\, dx$ over:*

- *all possible values x_0,*
- *all possible probability distributions $\rho \in \mathcal{P}$, and*
- *all possible functions $F(x)$ for which $|F'(x)| \le M$ for all x.*

The above definition involves taking a maximum over all distributions $\rho \in \mathcal{P}$ which are consistent with the known information about the group to which a given individual belongs. In some cases, we know the exact probability distribution, so the family $\mathcal{P}$ consists of only one distribution. In other situations, we may not know this distribution. For example, we may only know that the value of x is within the interval $[L, U]$, and we do not know the probabilities of different values within this interval. In this case, the class $\mathcal{P}$ consists of all distributions which are located on this interval (with probability 1).

When we learn new information about this individual, we thus reduce the group and hence, change from the original class $\mathcal{P}$ to a new class $\mathcal{Q}$. This change, in general, decreases the amount of privacy.

In particular, when we learn the exact value x_0 of the parameter, then the resulting class of distribution reduces to a single distribution concentrated on this x_0 with probability 1 – for which $F(x_0) - \int \rho(x) \cdot F(x)\,dx = 0$ and thus, the privacy is 0. In this case, we have a 100% loss of privacy – from the original value $A(\mathcal{P})$ to 0. In other cases, we may have a partial loss of privacy.

In general, it is reasonable to define the *relative loss of privacy* as a ratio

$$\frac{A(\mathcal{P}) - A(\mathcal{Q})}{A(\mathcal{P})}. \tag{1}$$

In other words, it is reasonable to use the following definition:

Definition 3.

- *By a* privacy loss*, we mean a pair $\langle \mathcal{P}, \mathcal{Q}\rangle$ of classes of probability distributions.*
- *For each privacy loss $\langle \mathcal{P}, \mathcal{Q}\rangle$, by the* measure of a privacy loss*, we mean the ratio (1).*

Comment. At first glance, it may sound as if these definitions depend on an (unknown) value of the parameter M. However, it is easy to see that the actual measure of the privacy loss does not depend on M:

Proposition 2. *For each pair $\langle \mathcal{P}, \mathcal{Q}\rangle$, the measure of the privacy loss is the same for all $M > 0$.*

Proof. To prove this proposition, it is sufficient to show that for each $M > 0$, the measure of privacy loss is the same for this M and for $M_0 = 1$. Indeed, for each function $F(x)$ for which $|F'(x)| \le M$ for all x, for the re-scaled function $F_0(x) \stackrel{\text{def}}{=} F(x)/M$, we have $|F_0'(x)| \le 1$ for all x, and

$$F(x_0) - \int \rho(x) \cdot F(x)\,dx =$$

$$M \cdot \left(F_0(x_0) - \int \rho(x) \cdot F_0(x)\,dx\right). \tag{2}$$

Vice versa, if $|F_0'(x)| \le 1$ for all x, for the re-scaled function $F(x) \stackrel{\text{def}}{=} M \cdot F_0(x)$, we have $|F'(x)| \le M$ for all x, and (2). Thus, the maximized values corresponding to M and $M_0 = 1$ different by a factor M. Hence, the resulting amounts of privacy $A(\mathcal{P})$ and $A_0(\mathcal{P})$ corresponding to M and M_0 also differ by a factor M: $A(\mathcal{P}) = M \cdot A_0(\mathcal{P})$. Substituting this expression for $A(\mathcal{P})$ (and a similar expression for $A(\mathcal{Q})$) into the definition (1), we can therefore conclude that $\dfrac{A(\mathcal{P}) - A(\mathcal{Q})}{A(\mathcal{P})} = \dfrac{A_0(\mathcal{P}) - A_0(\mathcal{Q})}{A_0(\mathcal{P})}$, i.e., that the measure of privacy is indeed the same for M and $M_0 = 1$. The proposition is proven.

H. The new definition of privacy loss is in good agreement with intuition

Let us show that the new definition adequately describes the difference between learning that the parameter is in the lower half of the original interval and that the parameter if even.

Proposition 3. *Let $[l, u] \subseteq [L, U]$ be intervals, let $\mathcal{P}$ be the class of all probability distributions located on the interval $[L, U]$, and let $\mathcal{Q}$ be the class of all probability distributions located on the interval $[l, u]$. For this pair $\langle \mathcal{P}, \mathcal{Q}\rangle$, the measure of the privacy loss if equal to $1 - \dfrac{u - l}{U - L}$.*

Proof. Due to Proposition 2, for computing the measure of the privacy loss, it is sufficient consider the case $M = 1$. Let us show that for this M, we have $A(\mathcal{P}) = U - L$.

Let us first show that for every $x_0 \in [L, U]$, for every probability distribution $\rho(x)$ on the interval $[L, U]$, and for every function $F(x)$ for which $|F'(x)| \le 1$, the privacy loss $F(x_0) - \int \rho(x) \cdot F(x)\,dx$ does not exceed $U - L$.

Indeed, since $\int \rho(x)\,dx = 1$, we have

$$F(x_0) = \int \rho(x) \cdot F(x_0)\,dx$$

and hence,

$$F(x_0) - \int \rho(x) \cdot F(x)\,dx = \int \rho(x)\,(F(x_0) - F(x))\,dx.$$

Since $|F'(x)| \le 1$, we conclude that

$$|F(x_0) - F(x)| \le |x_0 - x|.$$

Both x_0 and x are within the interval $[L, U]$, hence $|x_0 - x| \le U - L$, and $|F(x_0) - F(x)| \le U - L$. Thus, the average value $\int \rho(x) \cdot (F(x_0) - F(x))\,dx$ of this difference also cannot exceed $U - L$.

Let us now show that there exists a value $x_0 \in [L, U]$, a probability distribution $\rho(x)$ on the interval $[L, U]$, and a function $F(x)$ for which $|F'(x)| \le 1$, for which the privacy loss $F(x_0) - \int \rho(x) \cdot F(x)\,dx$ is exactly $U - L$. As such an example, we take $F(x) = x$, $x_0 = U$, and $\rho(x)$ located at a point $x = L$ with probability 1. In this case, the privacy loss is equal to $F(U) - F(L) = U - L$.

Similarly, we can prove that $A(\mathcal{Q}) = u - l$, so we get the desired measure of the privacy loss. The proposition is proven.

Comment. In particular, if we start with an interval $[L, U]$, and then we learn that the actual value x is in the lower half $[L, (L + U)/2]$ of this interval, then we get a 50% privacy loss.

What about the case when we assume that x is even? Similarly to the proof of the above proposition, one can prove that if both L and U are even, and $\mathcal{Q}$ is the class of all distributions $\rho(x)$ which are located, with probability 1, on even values x, we get $A(\mathcal{Q}) = A(\mathcal{P})$. Thus, the even-values restriction lead to a 0% privacy loss.

Thus, the new definition of the privacy loss is indeed in good agreement with our intuition.

ACKNOWLEDGMENTS

The authors are thankful to Ivo Babuska, Vladik Kreinovich, and to all the participants of the NSF Workshop on Reliable Engineering Computing at Georgia Tech (February 2008) for valuable discussions, and to anonymous referees for valuable suggestions.

REFERENCES

[1] M. Ceberio, G. Xiang, L. Longpré, V. Kreinovich, H. T. Nguyen, and D. Berleant, "Two Etudes on Combining Probabilistic and Interval Uncertainty: Processing Correlations and Measuring Loss of Privacy", *Proceedings of the 7th International Conference on Intelligent Technologies InTech'06*, Taipei, Taiwan, December 13–15, 2006, pp. 8–17.

[2] D. Dubois and H. Prade, Operations on fuzzy numbers, *International Journal of Systems Science*, 1978, Vol. 9, pp. 613–626.

[3] S. Ferson. *RAMAS Risk Calc 4.0*. CRC Press, Boca Raton, Florida, 2002.

[4] S. Ferson, W. L. Oberkampf, and L. Ginzburg, "Validation of imprecise probability models", In: R. L. Muhanna and R. L. Mullen (eds.), *Proceedings of the International Workshop on Reliable Engineering Computing REC'08*, Savannah, Georgia, February 20–22, 2008, pp. 23–44.

[5] S. Ferson, W. L. Oberkampf, and L. Ginzburg, "Model validation and predictive capability for the thermal challenge problem", *Computer Methods in Applied Mechanics and Engineering*, to appear.

[6] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives*, John Wiley and Sons, New York, 1976.

[7] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, Prentice Hall, New Jersey, 1995.

[8] H. T. Nguyen and V. Kreinovich, Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications, In: R. B. Kearfott and V. Kreinovich, eds., *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

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

[10] H. Raiffa, *Decision Analysis*, Addison-Wesley, Reading, Massachusetts, 1970.

[11] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[12] H. M. Wadswort, (ed.), *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York, 1990.