SAND2008-XXXX
Unlimited release
Printed XXXX 2008

DRAFT

Projecting uncertainty through black boxes

Scott Ferson and Vladik Kreinovich

Applied Biomathematics
100 North Country Road
Setauket, New York 11733 USA

Abstract

Computational models whose internal details are not accessible to the analyst are called black boxes. They arise because of security restrictions or because of the loss of the source code for legacy software programs. Computational models whose internal details are extremely complex are also sometimes treated as black boxes. It is often important to assess the uncertainty that should be ascribed to the output from a black box owing to uncertainty about its input quantities, their statistical distributions, or interdependencies. Sensitivity or 'what-if' studies are commonly used for this purpose. In such studies, the space of possible inputs is sampled as a vector of real values which is then provided to the black box to compute the output(s) that corresponds to those inputs. Such studies are often cumbersome to implement and understand, and they generally require many samples, depending on the complexity of the model and the dimensionality of the inputs. This report reviews methods that can be used to propagate about inputs through black boxes, especially 'hard' black boxes whose computational complexity restricts the total number of samples that can be evaluated. The focus is on methods that estimate the uncertainty of the outputs from the outside inward. That is, we are interested in methods that produce conservative characterizations of uncertainty that become tighter and tighter as the total computational effort increases.

Acknowledgments

We thank William Oberkampf, Jon Helton, Martin Pilch and Laura Swiler of Sandia National Laboratories, David S. Myers, Janos Hajagos, W. Troy Tucker and Lev Ginzburg of Applied Biomathematics, R. Ross Barmish of University of Wisconsin, William Huber of Quantitative Decisions, and Jason Augenbaugh of the University of Texas at Austin. The opinions expressed herein are solely those of the authors. This report was developed for Sandia under contract 19094, under the project direction of William Oberkampf.

This report will live electronically at http://www.ramas.com/outside_in.pdf. Please direct any questions, corrections, or suggestions for improvements to scott@ramas.com.

Table of contents

1. Introduction

A 'black box' is a model whose internal computational details are unknown to the analyst. Black box models are common in some engineering settings when, for example, the internal details kept hidden for the sake of security, confidentiality, or intellectual property concerns, or when the internal details are available, but are so complex that projecting analyses through them is impractical. Some businesses and government agencies actually use legacy computer models for which the original source code has been lost. Sampling, in which the model is applied to a given set of input values and returns one or more output values, is often the only effective means to study black box models. This makes the study of computational black boxes much like empirical scientific inquiry of the nature world in that we can see the outcomes generated under particular input conditions, but cannot directly see into the inner workings that produced those results.

A *hard* black box model is one for which the number of samples is tightly constrained because of computational difficulty or other limits. The larger and more computer-intensive codes become, the harder the black box models are. Although the raw computational power available to analysts is still exponentially expanding, computer simulation codes are often developed with similarly increasing scientific and engineering complexity at the limits of practical computability. For practical purposes, we are interested in methods that can be applied to black box models in general, and especially in any methods that can be applied to hard black box models for which relatively few samples will be available.

<<introduce Monte Carlo>>

Monte Carlo simulation and its kin (Latin hypercube sampling, importance sampling, etc.) are considered the methods of choice for propagating aleatory uncertainty of input variables through

<<mention RSM, defer to section below>>

<<introduce kinds of uncertainty and uncertain numbers>>

Uncertain numbers characterize quantities with aleatory uncertainty (variability) or epistemic uncertainty (partial ignorance) or both kinds of uncertainty. They generalize real numbers, intervals, probability distributions, interval bounds on probability distribution (probability boxes), and finite Dempster-Shafer structures whose elements are closed intervals of the real line.

<<purpose of report>>

This report reviews several methods that can be used to compute a surely or approximately conservative projection of uncertainty through a black box, i.e., a model whose internal details are not known to the analyst except through sampling. Because such strategies trade-off optimality of the results for computational convenience in obtaining them, they could be useful in screening assessments for problems where sampling is severely limited. The idea is to make calculations in a way that is sure to be conservative about uncertainty (i.e., sure not to underestimate uncertainty) but which can be completed without exhausting computational effort. In last year's research, three approximation methods were studied: the Cauchy deviate method, the Kolmogorov-Smirnov method, and the Saw et al. inequality method. In the Cauchy deviate method, complicated

Dempster-Shafer structures representing inputs were replaced with coarse intervals that enclose the structure completely and an approximate sampling strategy based on Cauchy deviates is used to propagate the intervals through a black box function. The deviates straddle and extend beyond the ranges of the intervals, but corrections to the summary results allow good estimates of the output interval when the function is roughly linear or the uncertainties are relatively small. In the Kolmogorov-Smirnov method, a black box function is treated as an oracle that produces a "sample" output for any set of inputs. Because these outputs are independent (if the sets of inputs are) and identically distributed, Kolmogorov-Smirnov confidence limits for a distribution can be used to obtain bounds on the distribution that account for uncertainty arising from the small sample size.

We also study some other methods, including the uniformity principle which computes conservative risk estimates in a distributionally robust way and the Saw et al. method which appeals to a generalization of the Chebyshev inequality, which yields bounds on the tail risks of a quantity given the mean and variance of the quantity.

The sampling of the black box function is thus reduced to a problem of estimating the mean and variance of the output. Follow-on research this year will explore other possible approximation strategies. The effort will include comparing the performance characteristics of the approximation strategies in realistic example problems and contrasting the strategies in terms of their usefulness in different situations.

We assume that the computation inside the black box is deterministic, although some of the methods described herein may also be useful if it is stochastic.

In particular situations, there may also be extra information available about the model, such as that it is an approximately linear or monotone function or that it is a smooth or slowing changing function of its inputs. Such knowledge may arise from the programmer or scientific understanding of the underlying physics.

This report reviews several methods that can be used to compute a surely or approximately conservative projection of uncertainty about inputs through a black box, i.e., a model whose internal details are not known to the analyst except through sampling.

Being conservative about uncertainty is being sure not to underestimate uncertainty.

We know how to do it intrusively (see section <<RSM>>).

The advantages outlined above of the new methods are limited to models involving explicitly known calculations. Probability bounds analysis and Dempster-Shafer theory cannot be rigorously applied to black box models, at least with algorithms currently available. Various sampling schemes have been proposed to extend these methods to black boxes (e.g., Helton et al. 2004a,b; 2006c; Bruns et al. 2006), but, because they are necessarily approximation methods, they abandon the guarantee that the results will enclose the true distributions. This means that the sampling-based methods do not provide "automatic result verification" (Adams and Kulisch 1993), although numerical simulations suggest that overall error rates can be made reasonably small if the black box permits many samples to be computed.

Furthermore, although it is often easy to apply the new methods to explicitly known models in a way that produces results that rigorously bound the possible output distributions (given the inputs),

it can often be difficult to make the results *best possible* bounds whenever uncertain variables appear multiple times in a calculation as often occurs in complex models involving dependencies among subcomponents. In general, computing best possible bounds is computationally intensive, and optimality of the bounds becomes harder and harder to maintain as the size of the problem and the interconnectedness of its elements increase. In practical assessments, however, optimality may not be essential. For instance, if an assessment can show that the risks are surely below some threshold of concern, because the upper distributional bound of risk is below it, then no further study is necessary. Likewise, if the result indicates that the risks are surely intolerably high, because the lower distributional bound on risk is above the threshold, then the implied decision is also clear. In practice we find analyses producing results with clear implications for decisions and management to be surprisingly common, even in cases with large input uncertainties that might have been expected to cloud the results. This frees available analytical resources to be concentrated where they are needed most: on problems where the attendant uncertainty makes the decision unclear.

<<Roadmap>>

2. Response surface modeling

The general strategy of response surface modeling (Myers 1971; Morton 1983; Downing et al. 1985; Kleijnen 1992; Myers 1999; Myers et al. 2004) allows an indirect application of Dempster-Shafer theory and probability bounds analysis to black boxes that might often be useful and effective. Response surface modeling is widely employed in engineering to replace a black box that is too hard to study directly with a statistical model of the black box that is more amenable to detailed analysis. In principle, the response surface models can have any form, but usually a linear or low-order polynomial model is employed, which is often characterized as a "model of the model" in that it is a nakedly phenomenological model of a much richer, physics-based model.

Many analysts (e.g., Frey and Patil 2002) suggest that it will often be advantageous to limit the number of inputs that are included in the response surface model to those that are identified as the most important using some screening sensitivity analysis. Apart from the chicken-and-egg problem of having to decide what is important to a sensitivity analysis before one conducts a sensitivity analysis, there is a more fundamental objection to this suggestion: several individually unimportant variables may, in aggregate, be important. Although it may be reasonable to drop small-impact terms when trying to make an approximation, this is not a good idea when trying to *bound uncertainty*. In the context of epistemic uncertainty, it may be far more reasonable to simplify the problem in other ways, such as replacing a complex Dempster-Shafer or p-box representation of an input variable with its interval support. Such replacements cannot lead to underestimates of uncertainty, even if there are many of them. This strategy would therefore be preferable in many situations to simply omitting variables. On the other hand, the strategy cannot by itself overcome the problem of having very few samples.

Another important consideration also argues against omitting any of the inputs before computing the regression. In a reduced regression analysis, the regression coefficients cannot be directly interpreted as sensitivities associated with the terms of the regression. This is because regression coefficients can change, sometimes dramatically or even in sign, when the regression model is altered. When developing a response surface model, at least analysts know the correct inputs to use for their black-box model. (In this way they are better off than regression analysts in general

who do not have such information.) Omitting some of the inputs to simplify the response surface is problematic because it creates the same disadvantage of varying regression coefficients for the response surface modeler.

The selection of the inputs to be used in sampling is an important consideration in response surface modeling, and this problem is treated in the broad statistical literature on sampling and experimental design. Typically, the inputs can be chosen by the analyst, although the design of inputs for sampling may occasionally not be under the analyst's control. This can happen when, for example, samples were obtained for other purposes (such as calibration) and additional sampling would be costly. When an analyst can specify the inputs for the samples, randomness of sample design is often a good strategy in many situations and usually simplifies statistical inferences, but various stratified sampling strategies such as Latin hypercube sampling or importance sampling may more commonly be preferable (Helton and Davis 2000a; 2002; 2003).

Perhaps even more critical than where the points will be is the issue of how many points there will be. If there are D input dimensions, one needs a minimum of D+1 sample points in general position to specify a linear model. Many more points would be necessary to specify a full quadratic or higher-order model. If there are more input dimensions than there are sample points, then the regression is underdetermined and cannot be performed by objective statistical methods. There are infinitely many planes that pass through two points. It may still be possible, nevertheless, to use response surfaces even in these extreme cases if the analyst can interject mechanistic knowledge of the physics of the underlying process to specify the model.

The difference between the original sample output and the output that would be predicted from the response surface model applied to the corresponding sample input is called the residual. The statistical fit of a response surface model and the normality of residuals can be studied using various standard well-known techniques such as the Kolmogorov-Smirnov, Anderson-Darling or chi-squared tests. For many hard black boxes, however, the goodness of fit of the response surface model is rarely an issue because the number of available sample points is so few relative to the dimensionality of the model. If a response surface can be selected to pass through all the available points, the residuals are zero.

In applications where there are non-zero residuals, the uncertainty that they embody should not be neglected in the subsequent analysis. A response surface model fitted by least-squares regression, for example, will have a form like

$$y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

where the response variable y is explained as a sum of inputs x_i multiplied respectively by associated regression coefficients β_i, and an error term ε which is represented by a normal distribution having mean zero and standard deviation σ (which is essentially the standard deviation of the residuals). Typically, ε is assumed to be independent of the other terms in the regression model. This error term should be incorporated into any subsequent sensitivity or uncertainty analyses based on this response surface model. Failure to do so could clearly understate the true uncertainty.

Once a black box model is represented by a response surface model, the methods of uncertainty and sensitivity analysis outlined in the previous sections of this report can be directly applied, including ordinary Dempster-Shafer theory and probability bounds analysis and pinching analyses

of various kinds. The computational difficulty associated with these applications may be fairly low if the response surface is a first-order linear model. If it includes repeated variables such as squares or higher powers or cross products representing interactions in addition to linear terms, then more careful strategies that account for the repeated variables will be needed to obtain best possible results. Strategies that may be useful in such cases are reviewed by Kreinovich et al. (2006).

There is a loss of guaranteed rigor in the use of a response surface model rather than the original model. This means that, even if the uncertainty of the input variables is surely captured by their uncertain number representations (Dempster-Shafer structures or p-boxes) and original model is an exact representation of the underlying process, the fitting of a response surface model is a statistical exercise and it may be imperfect. Indeed, it would be expected to be imperfect when there are few sample points available to inform the regression. We know of no method that would allow an analyst to rigorously propagate uncertainty through a black box model without assumptions that make the results contingent on the presumption that the response surface model is correct. The absence of the guarantee means that the uncertainty and sensitivity analyses of hard black box models will be approximate. Nevertheless, these approximations can often be good enough for use throughout engineering.

3. Using uncertain numbers in black box calculations

An algorithm for computing a function $f(x_1, \ldots, x_n)$ of uncertain quantities $x_1, \ldots, x_n$ may sometimes be implemented in a program available only as an executable computer code, with no source code at hand. In such situations, when we have no easy way to analyze the code and decompose it into a sequence of arithmetic operations, the only thing we can do is take this program as a *black box*, i.e., to apply it to different real-valued inputs and use the results of this sampling to compute the desired p-box or Dempster-Shafer structure.

We are often interested in the value of a quantity y that is difficult or impossible to measure directly. In this case, a natural idea is to measure easier-to-measure quantities $x_1,\ldots, x_n$ that are related to the desired quantity y, and try to estimate y based on the results $\tilde{x}_1,...,\tilde{x}_n$ of these measurements. To be able to produce such an estimate, we need to have an algorithm $f(x_1,\ldots, x_n)$ that, based on the values $x_1,\ldots, x_n$ of the directly measured quantities, reconstructs the value y of the desired quantity as $y = f(x_1,\ldots, x_n)$. Once we have such an algorithm, we plug in the measured values of x_i into this algorithm f, and get an estimate for y as $\tilde{y} = f(\tilde{x}_1,\ldots,\tilde{x}_n)$.

Measurements are never 100% accurate; as a result, the actual values x_i of the measured quantities may somewhat differ from the measured values. In other words, we know the inputs to the algorithm f only with some (measurement-related) uncertainty. Because of this input uncertainty $\tilde{x}_i \neq \tilde{x}_i$, our estimate $\tilde{y} = f(\tilde{x}_1,\ldots,\tilde{x}_n)$ is generally different from the actual value $y = f(x_1,\ldots, x_n)$ of the desired quantity. In other words, uncertainty in the inputs leads to the uncertainty in the output as well. It is therefore desirable to estimate this output uncertainty. So, we arrive at the following problem:

We know

- the algorithm $f(x_1,...,x_n)$;

- the measured values $\tilde{x}_1, \ldots, \tilde{x}_n$; and
- the information about the uncertainty $\Delta x_i \stackrel{\text{def}}{=} \tilde{x}_i - x_i$ of each direct measurement.

We must estimate the uncertainty $\Delta y = \tilde{y} - y$ of the algorithm's output.

In order to solve this problem, we must know what are the possible types of information that we can have about the uncertainty of each measurement error Δx_i. We do not know the exact values of the measurement errors Δx_i; as a result, in real life, we may have (and often we do have) several situations in which we get exactly exactly the same measurement results $\tilde{x}_1, \ldots, \tilde{x}_n$, but the actual values $x_1, \ldots, x_n$ of the measured quantity are different. Thus, to describe the uncertainty, we need to know what the possible values of Δx_i are, and how often different possible values can occur.

In the ideal case, when we have a complete description of uncertainty, we know the exact frequency (probability) of all possible error combinations $(\Delta x_1, \ldots, \Delta x_n)$. In other words, we know the exact probability distribution of the set of all n-dimensional vectors $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$. Often, the measurement errors corresponding to different measurements are independent, so it is sufficient to know the distribution of each variable x_i. This distribution can be described, e.g., by a cumulative density function (cdf) $F_i(t) \stackrel{\text{def}}{=} \text{Prob}(x_i \le t)$.

Most traditional methods of processing uncertainty in science and engineering (see, e.g., Wadsworth 1990) are based on the assumption that we have a *probabilistic uncertainty*, i.e., that the error distributions are independent, and that we know the probability distribution $F_i(t)$ for each of the variables x_i. However, in real life, we often do not have all this information.

In some real-life situations, we do not have any information about the frequency of different measurement error Δx_i; all we know is the range $[\Delta_i^-, \Delta_i^+]$ of possible values of this error. In this case, the only information that we have about the actual measured value $x_i = \tilde{x}_i - \Delta x_i$ of i-th quantity is that x_i must be in the interval $[\underline{x}_i, \overline{x}_i]$, where we denoted $\underline{x}_i \stackrel{\text{def}}{=} \tilde{x}_i - \Delta_i^+$ and $\overline{x}_i \stackrel{\text{def}}{=} \tilde{x}_i - \Delta_i^-$. The corresponding uncertainty is called *interval uncertainty*.

So far, we have describe two extreme situations:

in the case of probabilistic uncertainty, we have a complete information about which values Δx_i are possible, and what are the frequencies of different possible values;

in the case of interval uncertainty, we only know the range of possible values of Δx_i, we do not have any information about the frequencies at all.

Sometimes we have an intermediate situation: we have some (partial) information about the frequencies (probabilities) of different values of Δx_i, but we do not have the complete information about these frequencies.

How can we describe such situations? To describe the complete information about the probabilities of different values of Δx_i, we must describe, for every real number t, the value $F_i(t)$ of the corresponding cdf. Thus, when we have a partial information about these probabilities, it means that, instead of the exact value $F_i(t)$, we only have the range $[\underline{F}_i(t), \overline{F}_i(t)]$ of possible values of $F_i(t)$. Thus, to describe such an intermediate situation, we must describe the *bounds* $\underline{F}_i(t)$ and $\overline{F}_i(t)$ for the cdf. These bounds are called *probability boxes* (or *p-boxes*, for short, Ferson 2002).

Both probability distributions and intervals can be described as particular cases of p-boxes:

a probability distribution $F_i(t)$ can be described as a degenerate p-box $[F_i(t), F_i(t)]$; and

an interval $[a^-, a^+]$ can be described as a p-box $[\underline{F}_i(t), \overline{F}_i(t)]$ in which:

$\underline{F}_i(t) = 0$ for $t < a^+$ and $\underline{F}_i(t) = 1$ for $t \geq a^+$;

$\overline{F}_i(t) = 0$ for $t < a^-$ and $\overline{F}_i(t) = 1$ for $t \geq a^-$.

So, p-boxes are the most general way of representing these types of uncertainty.

Another way to describe partial information about the uncertainty is by using the Dempster-Shafer approach. In this approach, for each variable x_i, instead of a single interval $[\underline{x}_i, \overline{x}_i]$, we have several intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$ with probabilities $p_i^{(k)}$ attached to each such interval (so that for every i, $p_1^{(k)} + p_2^{(k)} + ... = 1$). For example, we may have several experts who provide us with different intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$, and $p_i^{(k)}$ is the probability that k-th expert is right. The collection of intervals with probabilities attached to different intervals constitutes a Dempster-Shafer knowledge base.

Thus, depending on the information that we have about the uncertainty in x_i, we can have five different formulations of the above problem:

we know the probability distribution $F_i(t)$ for each variable x_i, we know that these distributions are independent, and we must find the distribution $F(t)$ for $y = f(x_1, ..., x_n)$;

we know the interval $[\underline{x}_i, \overline{x}_i]$ of possible values of each variable x_i, and we must find the interval $[\underline{y}, \overline{y}]$ of possible values of y;

we know the p-boxes $[\underline{F}_i(t), \overline{F}_i(t)]$ that characterize the distribution of each variable x_i, we know that the corresponding distributions are independent, and we must find the p-box $[\underline{F}(t), \overline{F}(t)]$ that describe the variable y;

we know the Dempster-Shafer knowledge bases

$$\{[\underline{x}_i^{(1)}(t), \overline{x}_i^{(1)}(t)], p_i^{(1)}\}, \{[\underline{x}_i^{(2)}(t), \overline{x}_i^{(2)}(t)], p_i^{(2)}\}, \ldots$$

that characterize the distribution of each variable x_i, we know that the corresponding distributions are independent, and we must find the Dempster-Shafer knowledge base that describe the variable y;

we may also have different types of uncertainty for different variables x_i: e.g., we may have probabilistic uncertainty or x_1 and interval uncertainty for x_2.

It is also reasonable to consider the formulations in which the corresponding distributions may be dependent.

There exist efficient methods for solving these problems (see Ferson 2002 and references therein). Almost all of these methods are based on the fact that we know the algorithm f; so, instead of applying this algorithm step-by-step to the measured values $\tilde{x}_1, \ldots, \tilde{x}_n$, we apply this same algorithm step-by-step to the corresponding "uncertain numbers": probability distributions, intervals, and/or p-boxes.

In several practical situations, however, the algorithm is given as a *black box*: we do not know the sequence of steps forming this algorithm; we can only plug in different values into this algorithm and see the results. This situation is reasonably frequent, both with commercial software, where the software's owners try to prevent competitors from using their algorithms, and with classified or security-related software, where efficient security-related algorithms are kept classified to prevent the adversary from using them. In some practical cases, the situation is made even more difficult by the fact that the software $f(x_1, \ldots, x_n)$ is so complex and requires so much time to run that it is only possible to run it a few times. This complex black-box situation is what we will analyze in this text. Even for a black-box function, it may be possible to run more simulations if we use the actual black-box function $f(x_1, \ldots, x_n)$ to provide an approximating easier-to-compute model $f_{\text{appxo}}(x_1, \ldots, x_n) \approx f(x_1, \ldots, x_n)$, and then, we use this approximate models to estimate the uncertainty of the results (see Kreinovich et al. 2008). So, if our preliminary computations show that we need more simulations that the black-box function can give us, it does not necessarily mean that the corresponding uncertainty estimation method cannot be applied to our case: we may still be able to apply it to the approximate function f_{approx}.

4. Known methods for solving the problem

Monte Carlo techniques

Let us first consider the case of the probabilistic uncertainty, when we know that the values Δx_i are distributed according to the cdf $F_i(t)$, and that the corresponding random variables Δx_i are independent. In this case, we are interested to know the distribution $F(t)$ of Δy.

In the probabilistic case, a natural idea is to use Monte Carlo simulations. Specifically, on each iteration k:

for each input variable x_i, we simulate the values $x_i^{(k)}$ distributed according to the known distribution $F_i(t)$;

then, we plug the simulated values $x_i^{(k)}$ the algorithm f, and thus get the value $y^{(k)} = f(x_1^{(1)},...,x_n^{(k)})$.

After N iterations, we get N values $y^{(k)}$.

Since the inputs $x_i^{(k)}$ are independently distributed according to the corresponding input distributions $F_i(t)$, the outputs $y^{(k)}$ are distributed according to the desired distribution $F(t)$. Thus, the N values $y^{(k)}$ are a sample from the unknown distribution $F(t)$. It is therefore necessary to extract information about $F(t)$ from this sample.

Kolmogorov-Smirnov approach

One way to extract this information is to use the Kolmogorov-Smirnov confidence limits (see, e.g., Dixon and Massey 1969; Gibbon 1990; Sheskin 2004). These limits are based on considering the value

$$M \stackrel{\text{def}}{=} \max_t |F(t) - F_N(t)|,$$

where $F_N(t)$ is an empirical distribution related to the sample:

$$F_N(t) \stackrel{\text{def}}{=} \frac{\#\{k : y^{(k)} \le t\}}{N}.$$

It is known that for every given confidence level α and for every sample size N, there exists a value $M_\alpha(N)$ such that $M \le M_\alpha(N)$ with certainty $\le 1-\alpha$. For each α, as N grows, the value $M_\alpha(N)$ tends to 0.

Generalization of Chebyshev inequality

An alternative method has also been proposed for extracting $F(t)$ from the empirical data. This method is based on the generalization of Chebyshev inequality proposed by Saw et al. (1984; 1988; see also Konijn 1987; Stuart and Ord 1987; Woo 1991; Young et al. 1988). Specifically, Chebyshev inequality states that for a random variable X with mean μ and standard deviation

σ, for every real number $\lambda > 1$, we have $|X - \mu| \le \lambda \cdot \sigma$ with probability $\ge 1 - 1/\lambda^2$. Saw et al. produce a similar inequality based on the sample mean

$$\overline{\mu} = \frac{X^{(1)} + ... + X^{(N)}}{N}$$

and a sample standard deviation $\overline{\sigma} = \sqrt{\overline{V}}$, where

$$\overline{V} = \frac{(X^{(1)} - \overline{\mu})^2 + ... + (X^{(N)} - \overline{\mu})^2}{N - 1}.$$

According to Saw et al. (1984), we have $\text{Prob}(|X - \overline{\mu}| \ge \lambda \cdot \overline{\sigma}) \le f(N, \lambda)$ for some function $f(N, \lambda)$. Thus, if we have know N samples $y^{(1)}, ..., y^{(N)}$ from the unknown distribution $F(t)$, then we can conclude that with probability $\ge 1 - f(N, \lambda)$, the random variables y is located within the interval $[\overline{y} - \lambda \cdot \overline{\sigma}_y, \overline{y} + \overline{\sigma}_y]$, where

$$\overline{y} = \frac{y^{(1)} + ... + y^{(N)}}{N}$$

and $\overline{\sigma}_y = \sqrt{\overline{V}_y}$, with

$$\overline{V}_y = \frac{(y^{(1)} - \overline{y})^2 + ... + (y^{(N)} - \overline{y})^2}{N - 1}.$$

Cauchy deviates method

Yet another method for extracting information about $F(t)$ comes from considering the interval case.

In the interval case, we have intervals $[\underline{x}_i, \overline{x}_i]$ of possible values of each input x_i, and we are interested in finding the corresponding interval $[\underline{y}, \overline{y}]$ of possible values of y.

It is convenient to represent each interval $[\underline{x}_i, \overline{x}_i]$ by its midpoint

$$x_i^{\text{mid}} \stackrel{\text{def}}{=} \frac{\underline{x}_i + \overline{x}_i}{2}$$

and by its half-width

$$\Delta_i \stackrel{\text{def}}{=} \frac{\overline{x}_i - \underline{x}_i}{2},$$

so that each such interval takes the form $[x_i^{\text{mid}} - \Delta_i, x_i^{\text{mid}} + \Delta_i]$. In this representation, instead of the original variables x_i that take values from $\underline{x}_i$ to $\overline{x}_i$, it is often convenient to consider auxiliary variables $\delta x_i \stackrel{\text{def}}{=} x_i - x_i^{\text{mid}}$ that take values from $-\Delta_i$ to Δ_i.

For the case when the dependence $y = f(x_1, ..., x_n)$ is given as a black box, there is an algorithm for estimating the interval $[\underline{y}, \overline{y}]$ that is based on using Cauchy deviates (Kreinovich and Ferson 2004; Trejo and Kreinovich 2001). This method works when the function $f(x_1, ..., x_n)$ is reasonable smooth and the box $[\underline{x}_1, \overline{x}_1] \times ... \times [\underline{x}_n, \overline{x}_n]$ is reasonably small, so that on this box, we can reasonably approximate the function f by its linear terms:

$$f(x_1^{\text{mid}} + \delta x_1, ..., x_n^{\text{mid}} + \delta x_n) \approx y^{\text{mid}} + \delta y,$$

where

$$\delta y \stackrel{\text{def}}{=} c_1 \cdot \delta x_1 + ... + c_n \cdot \delta x_n,$$

$y^{\text{mid}} \stackrel{\text{def}}{=} f(x_1^{\text{mid}}, ..., x_n^{\text{mid}})$, and

$$c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}.$$

One can easily show that when each of the variables δx_i takes possible values from the interval $[-\Delta_i, \Delta_i]$, then the largest possible value of the linear combination δy is

$$\Delta = |c_1| \cdot \Delta_1 + ... + |c_n| \cdot \Delta_n, \qquad (1)$$

and the smallest possible value of δy is $-\Delta$. Thus, in this approximation, the interval of possible values of δy is $[-\Delta, \Delta]$, and the desired interval of possible values of y is $[y^{\text{mid}} - \Delta, y^{\text{mid}} + \Delta]$.

Cauchy distribution with a parameter Δ is a distribution described by the following density function:

$$\rho(x) = \frac{\Delta}{\pi \cdot (x^2 + \Delta^2)}.$$

It is known that if $\xi_1, ..., \xi_n$ are independent variables distributed according to Cauchy distributions with parameters Δ_i, then, for every n real numbers $c_1, ..., c_n$, the corresponding linear combination $c_1 \cdot \xi_1 + ... + c_n \cdot \xi_n$ is also Cauchy distributed, with the parameter Δ described by the formula (1).

Thus, if we for some number of iterations N, we simulate $\delta x_i^{(k)}$ $(1 \le k \le N)$ as Cauchy distributed with parameter Δ_i, then, in the linear approximation, the corresponding differences

$$\delta y^{(k)} \stackrel{\text{def}}{=} f(x_1^{\text{mid}} + \delta x_1^{(k)}, \ldots, x_n^{\text{mid}} + \delta x_n^{(k)}) - y^{\text{mid}}$$

are distributed according to the Cauchy distribution with the parameter Δ. The resulting values $\delta y^{(1)}, \ldots, \delta y^{(N)}$ are therefore a sample from the Cauchy distribution with the unknown parameter Δ. Based on this sample, we can estimate the value Δ.

Simulation can be based on the functional transformation of uniformly distributed sample values:

$$\delta x_i^{(k)} = \Delta_i \cdot \tan(\pi \cdot (r_i - 0.5)),$$

where r_i is uniformly distributed on the interval $[0,1]$.

In order to estimate Δ, we can apply the Maximum Likelihood Method which leads to the following equation:

$$\frac{1}{1+\left(\frac{\delta y^{(1)}}{\Delta}\right)^2} + \ldots + \frac{1}{1+\left(\frac{\delta y^{(N)}}{\Delta}\right)^2} = \frac{N}{2}.$$

The left-hand side of this equation is an increasing function that is equal to $0 (< N/2)$ for $\Delta = 0$ and $> N/2$ for $\Delta = \max\left|\delta y^{(k)}\right|$; therefore the solution to this equation can be found by applying a bisection method to the interval $\left[0, \max\left|\delta y^{(k)}\right|\right]$.

<<extra copy>>

Trejo and Kreinovich (2000) suggested a Monte Carlo algorithm for computing the interval bound on black box output $f(x_1, \ldots, x_n)$ for the case when the inputs $x_1, \ldots, x_n$ are intervals. Because Dempster-Shafer structures and p-boxes are, in effect, just collections of intervals, this method can in principle be naturally generalized to these structures.

We first outline the interval algorithm. Suppose we know that $x_i \in [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$, and we want to compute the upper bound Δ on the error $\tilde{y} - y$, where $\tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$ and $y = f(x_1, \ldots, x_n)$. If we get this upper bound, we will then compute the interval Y for y as $[\tilde{y} - \Delta, \tilde{y} + \Delta]$. The following algorithm makes use of deviates from the Cauchy distribution: For $k = 1, 2, \ldots, N$, repeat the following:

- Use a standard random number generator to compute n real numbers $r_i^{(k)}$, $i = 1, 2, \ldots, n$, that are uniformly distributed on the interval [0,1].
- Compute $\delta_i^{(k)} = \Delta_i \tan(\pi(r_i^{(k)} - 0.5))$.

- Compute the (Euclidean) length $\delta^{(k)}=\| \underline{\delta}^{(k)} \|$ of the vector $\underline{\delta}^{(k)}=(\delta_1^{(k)}, \ldots, \delta_n^{(k)})$.
- Compute the normalized coefficient $K^{(k)}_{\text{norm}} = \delta^{(k)}/\delta_0$ (for an appropriate small constant δ_0).
- Compute the auxiliary vector $\underline{\beta}^{(k)}= \underline{\delta}^{(k)}/K^{(k)}_{\text{norm}}$ with components $\beta_i^{(k)} = \delta_i^{(k)}/K^{(k)}_{\text{norm}}$.
- Substitute $\tilde{x}_i+\beta_i^{(k)}$ into the program f and compute $c^{(k)}=(\delta^{(k)}/\delta_0)(f((\tilde{x}_1+\beta_1^{(k)}, \ldots, \tilde{x}_n+\beta_n^{(k)})-\tilde{y})$.
- Compute Δ by applying the bisection method to solve the equation

$$\frac{1}{1+\left(\frac{c^{(1)}}{\Delta}\right)^2}+\ldots+\frac{1}{1+\left(\frac{c^{(N)}}{\Delta}\right)^2}=\frac{N}{2}$$

  on the interval $[0, \max |c^{(k)}|]$.

This method works well when the intervals are sufficiently narrow relative to the curvature of the function evaluated that its nonlinearity in each local interval is small.

It is worth acknowledging that the use of Cauchy distribution in the above algorithm may seem somewhat counterintuitive. Indeed, in the interval setting, we do not know the exact probability distribution of each error $\Delta x_i = \tilde{x}_i - x_i$, but we do know that each error Δx_i belongs to the corresponding interval $[-\Delta_i, \Delta_i]$, so the actual (unknown) probability distribution for Δx_i must be located on this interval with probability 1. At first glance, then, if we want to design a simulation-type technique for computing Δ, we should use one of these possible distributions in our simulations. Instead, we use a Cauchy distribution for which the probability to be outside the interval $[-\Delta_i, \Delta_i]$ is non-zero. In other words, in order to make the simulations work, we use the distributions which are inconsistent with our knowledge. The reason why such impossible distributions are useful here is that it can be shown that if we select, for simulations, a distribution within the corresponding p-box, we end up with a wrong estimate because any random sampling strategy for such a distribution will underestimate its dispersion.

As mentioned above, because any p-box or Dempster-Shafer structure is, in effect, a collection of intervals with probability masses, it is possible to generalize this black-box algorithm to handle them as inputs, at least under the assumption that these structures are stochastically independent of one another. The first step is to translate any inputs that are given as p-boxes to their associated Dempster-Shafer structures. This should be done in such a way as to minimize the number of focal elements consistent with a good approximation. This is important because the execution time of the calculation is a function of the product of the number of focal elements that must be considered.

Suppose there are D Dempster-Shafer structures $S_1, S_2, \ldots, S_D$, each having interval focal elements. Let card(S_i) denote the number of focal elements within the i^{th} Dempster-Shafer structure. To propagate these structures through the black box, create a Cartesian product of the elements of all D Dempster-Shafer structures (cf. sections <<>> and <<>>). This will be an D-dimensional array consisting of M elements, where $M = \Pi$ card(S_i). Each element of this array has

two parts, the (interval) location and the (point) probability mass. The probability mass of each element is computed, under independence among all the inputs, by multiplying the respective probability masses from the D marginal Dempster-Shafer structures. Because we are assuming that Dempster-Shafer structures are finite collections, these calculations will always be possible. Moreover, because the masses of each Dempster-Shafer structure individually summed to one, the overall sum of the M masses will also sum to one. The second part of each of the M elements in the n-dimensional array is the solution to an interval-blackbox calculation $f(A_1, A_2, \ldots, A_n)$, where A_i is the respective element of the i^{th} Dempster-Safer structure in the Cartesian product. Because we are assuming that the focal elements of the Dempster-Shafer structures are intervals, these calculations will always be possible.

Thus the basic probability assignment of the result of the blackbox propagation of Dempster-Shafer structures is therefore

$$m^*(C) = \sum_{C=f(A_1,A_2,\ldots,A_n)} \prod_{i}^{D} m_i(A_i)$$

where m_i is the basic probability assignment for the i^{th} Dempster-Shafer structure. In this formulation, the summation is of all results C of the black-box calculation that might arise from different combinations of input focal elements. Note that only the ability to *evaluate* $f(A_1, A_2, \ldots, A_D)$ is required; we do not need to know what mathematical operations it entails. In computational practice, we would not actually bother to perform the summation, but would instead be content with computing the cumulative plausibility and belief functions, which can be done simply by remembering all of the left and right endpoints of C's and cumulating them separately. The cumulative histogram of the left endpoints forms the estimated cumulative plausibility function, and the cumulative histogram of the right endpoints forms the estimated cumulative belief function. This pair of functions is the probability box for the output of the blackbox calculation.

The calculation strategy outlined here is, of course, extremely computationally intensive. Unless the most advanced computers can be used for the task, it would be unlikely that this approach could be used if the function evaluation is itself expensive or when the number of input variables exceeds more than a handful. There are, of course, many strategies that could be employed to accelerate the calculation. Finding the best of these is the subject of concurrent research.

It is possible to apply this algorithm without necessarily assuming stochastic independence among all the input variables. Appealing to the Fréchet (1935) inequality, we could replace the expression $\Pi m_i(A_i)$ used to compute the probability mass of each element of the Cartesian product with the interval expression

$$[\ \max(0,\ 1 - D + \Sigma m_i(A_i)),\ \ \min\ m_i(A_i)\ ].$$

This would discharge all dependence assumptions used in the calculation. It is also possible to combine independence assumptions for some variables with the general case for others by an obvious combination of multiplication and Fréchet intervals. The result of this approach would be to produce a Cartesian product whose masses are not point values but intervals. Although this appears to constitute a fundamentally generalized kind of Dempster-Shafer structure, it is associated with an ordinary Dempster-Shafer structure by canonical discretization of the

cumulative plausibility and belief functions. However, when no precise dependency structure is assumed, the result is not guaranteed to be best possible and would be expected to overestimate somewhat the uncertainty in the output Dempster-Shafer structure. This may not be too much of a problem if, despite this uncertainty inflation, it is still clear that a particular result is acceptable in its risk management context.

<<extra copy>>

Comparison with sensitivity analysis

It is worth mentioning that the Cauchy deviates method is beneficial only if we have a large number of inputs D (thus, it is not beneficial to apply this method to the simplified functions $(a+b)^{\alpha}$ or to the spring function).

Indeed, for small D, <<in sequel, n is D>> we can use the following *sensitivity analysis* method – a method that is applicable not only for approximately linear functions $f(x_1,\ldots,x_n)$, but also for all functions that are monotonic (increasing or decreasing) with respect of each of its variables. Specifically, in the sensitivity analysis method:

First, we apply f to the results $\tilde{x}_1,\ldots,\tilde{x}_n$ of direct measurements, resulting in the value $\tilde{y}=f(\tilde{x}_1,\ldots,\tilde{x}_n)$. Then, for each of n inputs x_i, we modify this input to $x'_i \neq \tilde{x}_i$ and, leaving other inputs, apply f again. By comparing the values $f(\tilde{x}_1,\ldots,\tilde{x}_i,x'_i,\tilde{x}_{i+1},\ldots,\tilde{x}_n)$ and $\tilde{y}=f(\tilde{x}_1,\ldots,\tilde{x}_n)$, we decide whether f in increasing or decreasing in x_i. Finally, we apply f two more times to get the desired bounds for y as follows: $\underline{y}=f(x_1^-,\ldots,x_n^-)$ and $\overline{y}=f(x_1^+,\ldots,x_n^+)$, where, for the variables x_i for which f increases with x_i, we take $x_i^-=\underline{x}_i$ and $x_i^+=\overline{x}_i$, and, for the variables x_i for which f decreases with x_i, we take $x_i^-=\overline{x}_i$ and $x_i^+=\underline{x}_i$.

Use with nonlinear functions

The Cauchy deviate method can be generalized in a couple of ways for functions that are clearly nonlinear relative to the breadth of uncertainty in the inputs. Beck (2004) described a second-order Taylor version of the method that would be appropriate for functions that are approximately quadratic. However, the generalization needs samples sizes that are roughly the square of those needed by the linear method as it is essentially a nested approach. This approach is therefore unlikely to be useful for propagating uncertainty through hard black boxes where the number of samples is limited.

Another approach when we cannot reasonably approximate f by a linear expression on the entire box is to divide the box into a few sub-boxes on each of which f is approximately linear. For example, if the dependence of f on one of the variables x_i is strongly nonlinear, then we can divide the interval $[\underline{x}_i,\overline{x}_i]$ of possible values of this variable into two (or more) subintervals, e.g., $[\underline{x}_i,x_i^{\mathrm{mid}}]$ and $[x_i^{\mathrm{mid}},\overline{x}]$, and consider the corresponding sub-boxes

$$[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_{i-1}, \overline{x}_{i-1}] \times [\underline{x}_i, x_i^{\text{mid}}] \times [\underline{x}_{i+1}, \overline{x}_{i+1}] \times \cdots \times [\underline{x}_n, \overline{x}_n]$$

and

$$[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_{i-1}, \overline{x}_{i-1}] \times [x_i^{\text{mid}}{}_i, \overline{x}_i] \times [\underline{x}_{i+1}, \overline{x}_{i+1}] \times \cdots \times [\underline{x}_n, \overline{x}_n]$$

By using the Cauchy deviates methods, we compute the range of *f* over each of these sub-boxes, and then take the union of the resulting range intervals.

Use with other forms of uncertainty

As we have mentioned, the Cauchy deviates method can be used not only to handle the case of interval uncertainty, it can also be used to handle the case of probabilistic uncertainty. Specifically, if for each input variable x_i, its distribution $F_i(t)$ is located on some interval $[\underline{x}_i, \overline{x}_i]$ with probability 1, then we can, in principle, ignore the probability values and simply use the fact that x_i is located on the interval $[\underline{x}_i, \overline{x}_i]$. In this case, the Cauchy deviates method produces an interval $[\underline{y}, \overline{y}]$ that is guaranteed to contain y.

Similarly, we can treat the case when for each x_i, we know the p-box $[\underline{F}_i(t), \overline{F}_i(t)]$ that describes its distribution. Let us assume that the probability distribution $\underline{F}_i(t)$ is located, with probability 1, on an interval $[\underline{x}_i^-, \overline{x}_i^-]$, i.e.:

$\underline{F}_i(\underline{x}_i^-) = 0$, and $\underline{F}_i(\underline{x}_i^- + \varepsilon) > 0$ for arbitrarily small $\varepsilon > 0$,

$\underline{F}_i(\overline{x}_i^-) = 1$, and $\underline{F}_i(\overline{x}_i^- - \varepsilon) < 1$ for arbitrarily small $\varepsilon > 0$.

We also assume that the probability distribution $\overline{F}_i(t)$ is located, with probability 1, on an interval $[\underline{x}_i^+, \overline{x}_i^+]$, i.e.:

$\overline{F}_i(\underline{x}_i^+) = 0$, and $\overline{F}_i(\underline{x}_i^+ + \varepsilon) > 0$ for arbitrarily small $\varepsilon > 0$,

$\overline{F}_i(\overline{x}_i^+) = 1$, and $\overline{F}_i(\overline{x}_i^+ - \varepsilon) < 1$ for arbitrarily small $\varepsilon > 0$.

The p-box means that the actual distribution $F(t)$ is in between $\underline{F}_i(t)$ and $\overline{F}_i(t)$, i.e., that $\underline{F}_i(t) \le F_i(t) \le F^+(t)$. So:

From $\overline{F}_i(\underline{x}_i^+) = 0$, we conclude that $F_i(\underline{x}_i^+) \le \overline{F}_i(\underline{x}_i^+) = 0$, i.e., that $F_i(\underline{x}_i^+) = 0$. Thus, with probability 1, we have $x_i \ge \underline{x}_i^+$. Similarly, from $\underline{F}_i(\overline{x}_i^-) = 1$, we conclude that $F_i(\overline{x}_i^-) \ge \underline{F}_i(\overline{x}_i^-) = 1$, i.e., that $F_i(\overline{x}_i^-) = 1$. Thus, with probability 1, we have $x_i \le \overline{x}_i^-$. Thus, for every i, with probability 1, we have $x_i \in [\underline{x}_i^+, \overline{x}_i^-]$. Similarly to the probabilistic case, we can

ignore the probability values and simply use the fact that x_i is located on the interval $[\underline{x}_i^+, \overline{x}_i^-]$. In this case, the Cauchy deviates method produces an interval $[\underline{y}, \overline{y}]$ that is guaranteed to contain y.

Finally, for the Dempster-Shafer case, if we know that x_i belongs to the one of the intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$ with certain probabilities, then we can ignore these probability and conclude that x_i belongs to the narrowest interval that contains all of these. i.e., to the interval $[\underline{x}_i, \overline{x}_i]$, where $\underline{x}_i = \min_k \underline{x}_i^{(k)}$ and $\overline{x}_i = \max_k \overline{x}_i^{(k)}$. In this case, the Cauchy deviates method produces an interval $[\underline{y}, \overline{y}]$ that is guaranteed to contain y.

In this report, we analyze and compare the existing methods and, if necessary, suggest some possibly better methods of estimating uncertainty of *y* from hard black boxes.

5. Optimal statistics under probabilistic uncertainty

Let us formulate the problem in precise terms. We start our analysis with the case of probabilistic uncertainty. In this case, we have a sample $X_1, \ldots, X_N$ taken from an unknown probability distribution, and we want to estimate this probability distribution $F(t)$. To be more precise, we would like to provide bounds $[\underline{F}(t), \overline{F}(t)]$ on $F(t)$ that hold with some degree of certainty (ideally, as large degree of certainty as possible).

There may be different formulations of this problem. A very important case is when there is a critical threshold t_0 for y, then all we are interested in is what is the probability that for our system, the value y exceeds this critical threshold. In this case, we are not really interested in the entire probability distribution; all we want to know is the probability P_{crit} that $y > t_0$ – i.e., equivalently, the probability $F(t_0) = 1 - P_{\text{crit}}$ that $y \le t_0$.

In this case, to estimate the desired value $F(t_0)$ of the unknown distribution function $F(t)$ from below, we must find a statistic $L_{t_0}(X_1, \ldots, X_N)$ (i.e., in mathematical terms, a real-valued function of N real variables) for which, for $\underline{F}(t_0) \stackrel{\text{def}}{=} L_{t_0}(X_1, \ldots, X_n)$, we get $\underline{F}(t_0) \le F(t_0)$ with a certainty $\ge 1 - \alpha$. In other words, we want a function $L_{t_0}(X_1, \ldots, X_N)$ for which

$$\text{Prob}(L_{t_0}(X_1, \ldots, X_N) \le F(t_0)) \ge 1 - \alpha. \quad (2)$$

Of course, since $X_1, \ldots, X_N$ form a random sample, the statistic $L_{t_0}(X_1, \ldots, X_N)$ cannot depend on the order of the values $X_1, \ldots, X_N$, i.e., it must be permutation-invariant.

Similarly, to estimate the desired value $F(t_0)$ of the unknown distribution function $F(t)$ from above, we must find a statistic $U_{t_0}(X_1, \ldots, X_N)$ (i.e., in mathematical terms, a real-valued function

of N real variables) for which, for $\overline{F}(t_0) \stackrel{\text{def}}{=} U_{t_0}(X_1,...,X_n)$, we get $F(t_0) \le \overline{F}(t_0)$ with a certainty $\ge 1-\alpha$. In other words, we want a function $U_{t_0}(X_1,...,X_N)$ for which

$$\text{Prob}(F(t_0) \le U_{t_0}(X_1,...,X_N)) \ge 1-\alpha. \qquad (3)$$

The statistic $U_{t_0}(X_1,...,X_N)$ cannot on the order of the values $X_1,...,X_N$ either, i.e., it must also be permutation-invariant.

Among all such statistics, we need to find a pair that is, in some reasonable sense, optimal.

Scale invariance

For each physical quantity, the numerical values X_k depend on the choice of scale for measuring this quantity. For example, from the physical viewpoint, it is quite reasonable to describe the strength of an *earthquake* by its energy E, but when we talk about its consequences, it is much more convenient to use the logarithmic Richter scale which is some constant times the log of E. Nonlinear scales are also used in many other application areas, such as in the use of decibels in electrical engineering to describe noise and amplifier performance, or the in hardness scales of minerals in the geology. A general overview of different scales and rescalings is given by Suppes et al. (1971; 1989; 1989).

The physical problem remains the same no matter what scale we use: in the new scale $X' = f(X)$, the same sample takes the form $X'_1 = f(X_1),..., X'_n = f(X_n)$, and the threshold takes the form $t'_0 = f(t_0)$. Since the physical problem remains the same, it is reasonable to require that the estimates for the probability of exceeding the threshold should not depend on what scale we use. In other words, it is reasonable to require that $L_{t'_0}(X'_1,...,X'_N) = L_{t_0}(X_1,...,X_N)$, i.e., that

$$L_{f(t_0)}(f(X_1),...,f(X_N)) = L_{t_0}(X_1,...,X_N) \qquad (4)$$

for all strictly increasing 1-1 functions f. Similarly, it is reasonable to require that

$$U_{f(t_0)}(f(X_1),...,f(X_N)) = U_{t_0}(X_1,...,X_N) \qquad (5)$$

for all strictly increasing 1-1 functions f.

In Appendix 1, we show that this requirement is not only reasonable, it also follows from our desire to have *optimal* statistics L_t and U_t.

Non-degeneracy

Another reasonable requirement is non-degeneracy. The distribution of X is usually continuous, so the probability that two random numbers are exactly equal is 0. In practice, a measurement result has only finite many digits; as a result, we may get identical values $X_i = X_j$ for some $i \ne j$, but this identity is clearly a measurement artifact. So, we can safely assume that in reality,

all the values X_i are different. Similarly, we can safely assume that none of the values X_i is exactly equal to the threshold, i.e., $X_i \neq t_0$ for all i.

The conclusion from this argument is that both bounds on $F(t)$ depend only on the empirical distribution function $F_N(t)$. Let us show that due to scale-invariance (4)-(5), the estimates L_{t_0} and U_{t_0} can only depend on $F_N(t_0)$, i.e., on the ratio k/N, where k is the number of the values F_i that are $\le t_0$. Indeed, let us assume that we have two different samples $X_1,\ldots,X_N$ and $X'_1,\ldots,X'_N$, with different thresholds t_0 and t'_0, for which $F_N(t_0) = F'_N(t'_0)$. In other words, we assume that in each of the two samples, we have exactly

$$k \stackrel{\text{def}}{=} N \cdot F_N(t_0) = N \cdot F'_N(t'_0)$$

values that are smaller than the threshold. Because of this requirement, if we sort each sample, the corresponding threshold falls exactly in between k-th and $(k+1)$-st values:

$$X_{(1)} < \ldots < X_{(k)} < t_0 < X_{(k+1)} < \ldots < X_{(n)};$$

$$X'_{(1)} < \ldots < X'_{(k)} < t'_0 < X'_{(k+1)} < \ldots < X'_{(n)}.$$

We can easily design a strictly increasing function $f(x)$ for which

$$f(X_{(1)}) = X'_{(1)},\ldots, f(X_{(k)}) = X'_{(k)}, f(t_0) = t'_0,$$

$$f(X_{(k+1)}) = X'_{(k+1)},\ldots, f(X_{(n)}) = X'_{(n)};$$

for example, we can construct this function by linear interpolation. For this function f, invariance (4) leads to $L_{t'_0}(X'_1,\ldots,X'_N) = L_{t_0}(X_1,\ldots,X_N)$, and invariance (5) leads to $U_{t'_0}(X'_1,\ldots,X'_N) = U_{t_0}(X_1,\ldots,X_N)$. Thus, indeed, both statistics are functions only of $k/N = F_N(t)$.

In short, for every t, the optimal interval $[\underline{F}(t), \overline{F}(t)]$ that contains $F(t)$ with a given certainty can only depend on $F_N(t)$. Since it is reasonable to include the frequency $F_N(t)$ within this interval, we thus conclude that the optimal estimate must be Kolmogorov-Smirnov-type, with a interval around $F_N(t)$ whose width only depends on N and on $F_N(t)$.

Note that these bounds do not have to be identical to the Kolmogorov-Smirnov ones. In the original Kolmogorov-Smirnov bounds, all intervals $[\underline{F}(t), \overline{F}(t)]$ have exactly the same width. In general, we may have different widths depending on the actual value of $F_N(t)$. Moreover, if we use, e.g., smaller widths for smaller values $F_N(t)$, we get better estimates for the tails of the desired distribution.

6. Analysis of the generalized Chebyshev inequality

It may be somewhat surprising that our analysis of possible methods resulted in techniques similar to the Kolmogorov-Smirnov-type statistics but not in the use of the generalization of Chebyshev inequality proposed by Saw et al. (1984). The surprise comes from the fact that, at first glance, the Saw et al. method has a clear advantage over the Kolmogorov-Smirnov approach: the Kolmogorov-Smirnov confidence limits provide us with bounds on $F(t)$ that are only true with a certain confidence, while the Saw's method seems to provide us with the bounds that are 100% guaranteed.

However, as we will now show, this seeming advantage of Saw's method is based on a misunderstanding of the corresponding inequality

$$\text{Prob}(|X-\overline{\mu}| \geq \lambda \cdot \overline{\sigma}) \leq f(N,\lambda).$$

Indeed, what we want is to be able, based on the observed sample $X^{(1)},\ldots,X^{(N)}$ from a random variable X with an unknown distribution, to make conclusions about the value $F(t)=\text{Prob}(X \leq t)$. At first glance, it may seem that this is exactly what the above inequality is doing: for example, from the above inequality, we can conclude that

$$\text{Prob}(X \geq \overline{\mu}+\lambda \cdot \overline{\sigma}) \leq \text{Prob}(|X-\overline{\mu}| \geq \lambda \cdot \overline{\sigma}) \leq f(N,\lambda),$$

and therefore, that

$$\text{Prob}(X \leq \overline{\mu}+\lambda \cdot \overline{\sigma}) \geq 1-f(N,\lambda),$$

i.e., that $F(t^+) \geq 1-f(N,\lambda)$ for $t^+ \stackrel{\text{def}}{=} \overline{\mu}+\lambda \cdot \overline{\sigma}$.

However, as we will show, in our interpretation of the probability, this inequality cannot be 100% true. Indeed, consider, e.g., the case when the actual distribution $F(t)$ is the standard normal distribution, with 0 mean and standard deviation 1. Let us fix a value $\lambda > 1$, and let us select N to be sufficiently large so that $f(N,\lambda) < 1$ – i.e., that this inequality does not degenerate into a meaningless inequality $F(t) \geq 1$.

Since the function $F(t)$ is strictly increasing, the inequality $F(t^+) \geq 1-f(N,\lambda)$ is equivalent to $t^+ \leq t_0$, where $t_0 \stackrel{\text{def}}{=} F^{-1}(1-f(N,\lambda))$. Thus, in our interpretation, we would conclude that with absolute certainty, for N samples from the normal distribution, we have $t^+ = \overline{\mu}+\lambda \cdot \overline{\sigma} \leq t_0$, where $\overline{\mu}$ is the sample average and $\overline{\sigma}$ is the sample standard deviation. Since $\lambda > 1 \geq 0$ and $\overline{\sigma} \geq 0$, we can thus conclude that $\overline{\mu} \leq t_0$ with certainty (i.e., with probability 1).

On the other hand, we know that for the standard normal distribution X (with 0 mean and standard deviation 1), the sample mean

$$\overline{\mu} = \frac{X_1 + \ldots + X_N}{N}$$

is also distributed according to the normal distribution, with 0 mean and standard deviation $1/\sqrt{N}$. A normal distribution is not localized on any interval. In other words, for a normal distribution, no matter how large real number t_0 we select, there is always a positive probability that the random value is $\geq t_0$ – and thus, the probability that $\overline{\mu} \leq t_0$ is always less than 1.

We thus get a contradiction between the above-described interpretation of Saw's inequality and the known facts about the normal distribution. According to the above interpretation of Saw's inequality, for normal distribution X, the probability that $\overline{\mu} \leq t_0$ is equal to 1; on the other hand, this same probability must be smaller than 1.

Does this contradiction mean that the paper Saw et al. (1984) is based on a mathematical mistake? No, the proof is correct; the contradiction simply means that this paper interprets probability differently from what we want. What we want is, based on a fixed sample $X_1, \ldots, X_N$, to predict the probability that X is within the certain interval. In other words, we want to bound a probability over a random variable X, and we want the corresponding bound to hold for all samples. What the paper is doing is considering the probability over all possible *pairs* $(X_1, \ldots, X_n), X$ consisting of a sample $X_1, \ldots, X_N$ and a value X. In other words, we want the probability $F_{X_1,\ldots,X_N}(t)$ corresponding to an individual sample $X_1, \ldots, X_n$, while Saw's inequality provides us only with the *average* $E_{X_1,\ldots,X_n} F_{X_1,\ldots,X_n}(t)$ of the desired probability $F_{X_1,\ldots,X_N}(t)$ over all possible samples.

Thus, based on their inequality, when we have a single sample $X_1, \ldots, X_n$, we cannot make any conclusion about $F(t)$. We can, however, make a conclusion that is true "on average" if we repeat this same experiment for multiple samples. For some individual samples, the inequality is true, for some individual samples, the inequality is false. In other words, contrary to the appearance, this generalized Chebyshev inequality does not provide us with any guaranteed bounds for X – all we can hope for is certainty bounds, and this inequality does not even provide us with these certainty bounds.

The difference between our interpretation of probability and the interpretation implicitly assumed by Saw et al. (1984) becomes even clearer if we consider the follow-up paper by Konijn (1987). This paper starts by comparing Saw's result with a similar previously known result, a result over which Saw et al. improve: that for every sample $X_1, \ldots, X_N$, with probability $\leq (N-1)/(N+1)$, we have $X \in [X_{(1)}, X_{(N)}]$, where, as usual in statistics, $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}$ denotes the ordering of the sample in the increasing order – so that $X_{(1)} = \min(X_1, \ldots, X_N)$ and $X_{(N)} = \max(X_1, \ldots, X_N)$.

Let us show that this statement cannot be true in our interpretation of probability. Indeed, let us assume that it is true for every distribution. In particular, it is then true for the uniform distribution

on the interval $[0,1]$. On this uniform distribution, the probability that a random value X belongs to an interval $[a,b]$ is equal to the width $b-a$ of this interval. Thus, the above statement would mean that for every sample $X_1,...,X_N$, we have

$$\max(X_1,...,X_N)-\min(X_1,...,X_N)\le\frac{N-1}{N+1}.$$

On the other hand, if $X_1\le 1/(N+1)$ and $X_2\ge 1-1/(N+1)=N/(N+1)$ (the probability of both inequalities is $1/(N+1)\cdot 1/(N+1)>0$), then we have

$$\max(X_1,...,X_N)-\min(X_1,...,X_N)\ge X_2-X_1>\frac{N-1}{N+1}.$$

The contradiction shows that both in Saw's and in Konijn's papers, the probability is interpreted as an *average* probability over all possible samples, not the desired probability based on a single sample.

The inequality of Saw et al. is probably not useful for propagating aleatory and epistemic uncertainty through black box models. A more detailed analysis of Konijn (1987) reveals that, in effect, the approach described in Saw's and Konijn's papers leads to a (partially) exact characterization of the average value $F^{\text{av}}(t)\stackrel{\text{def}}{=}E_{X_1,...,X_N}F(t)$: namely, $F^{\text{av}}(X_{(k)})=k/(N+1)$. In other words, this average is almost identical to the empirical cdf that serves as a base for the Kolmogorov-Smirnov confidence limits. From this viewpoint, Saw's inequality does *not* help us understand the difference between the actual distribution $F(t)$ and the empirical distribution $F^{\text{av}}(t)$; in contrast, this inequality is, in effect, a property of the empirical distribution.

So, from the viewpoint of our problem, Saw's inequality is not helpful at all: instead of using this inequality that binds the empirical distribution $F^{\text{av}}(t)$, we can as well take the actual empirical distribution (as described above) and thus, get the exact expression for $F^{\text{av}}(t)$, at least for $t=X_1,...,X_N$.

The difference in interpretations also helps to understand a seemingly paradoxical situation mentioned in Saw et al. (1984) that the generalized Chebyshev inequality sometimes leads to better bounds than the original Chebyshev inequality. The Chebyshev inequality $\text{Prob}(\mu-\lambda\cdot\sigma\le X\le\mu+\lambda\cdot\sigma)\ge 1-1/\lambda^2$ enables us, once we know the exact values of mean μ and standard deviation σ, to get guaranteed bounds for $F(t)$:

$$\text{Prob}(\mu-\lambda\cdot\sigma\le X\le\mu+\lambda\cdot\sigma)=F(t^+)-F(t^-)\ge 1-\frac{1}{\lambda^2},$$

where $t^-\stackrel{\text{def}}{=}\mu-\lambda\cdot\sigma$ and $t^+\stackrel{\text{def}}{=}\mu+\lambda\cdot\sigma$.

If we do not know the exact values μ and σ, then we cannot use the Chebyshev inequality, but we can use Saw's inequality based on the sample mean and sample standard deviation.

In the second case, we have less information than in the first case, so, intuitively, we expect to get wider bounds that for the original Chebyshev inequality. However, sometimes the bounds produced by the generalized Chebyshev inequality are *narrower*. In view of our analysis, however, this is not a paradox, because the original Chebyshev inequality provides us with the actual guaranteed bound for $F(t)$, while the generalized Chebyshev inequality only bounds the "average" value of $F(t)$. Since the generalized inequality is less informative than the original one, it is no wonder that it can correspond to a narrower interval: the interval is narrower, but we no longer provide a guarantee that X is within interval with a certain probability (we only guarantee the *average* probability of X being there).

7. Kolmogorov-Smirnov-Type Statistics: Analysis

In this section, we consider which of the Kolmogorov-Smirnov-type statistics we should choose. In the previous text, we have shown that the optimal way to estimate $F(t)$ is to use Kolmogorov-Smirnov-type estimates, i.e., estimates for which, for every t, we generate a confidence interval $[\underline{F}(t), \overline{F}(t)]$ around the empirical frequency $F_N(t)$.

In the original Kolmogorov-Smirnov statistic, we simply fix a half-width $\varepsilon > 0$, and we enclose each value $F_N(t)$ into an interval $[F_n(t) - \varepsilon, F_N(t) + \varepsilon]$ of this half-width. However, within our result, this is not the only possible choice: we could as well select different widths for different values $F_N(t)$. A natural question is: which of the possible Kolmogorov-Smirnov-type statistics should we choose?

Analysis of the problem

To answer this question, let us analyze the accuracy with which the estimate $F_N(t)$ approximates the actual value $F(t)$. This estimation can be, in effect, reformulated as one of the basic statistical problems. Indeed, for each t, $f = F_N(t)$ is the frequency, of having $X \le t$; based on this frequency, we want to estimate the probability $p = F(t)$ that $X \le t$.

The problem of estimating the probability p based on the frequency f is a well-known problem in probability theory, a problem that can be also described as the problem of estimating the parameter p of the binomial distribution (see, e.g., Dixon and Massey 1969).

It is known that for large N, the difference $f - p$ is normally distributed with the mean $\mu = 0$ and the standard deviation

$$\sigma = \sqrt{\frac{p \cdot (1-p)}{N}}.$$

If we know the certainty $1-\alpha$ that we want to achieve, then we can, based on the normal distribution, find the corresponding coefficient k_α for which the probability of a normal distribution to be outside the interval $[\mu - k_\alpha \cdot \sigma, \mu + k_\alpha \cdot \sigma]$ is exactly α. For example, for $\alpha = 0.05$, we get $k_\alpha = 2$, and for $\alpha = 0.1\%$, we get $k_\alpha = 3$.

Thus, for each p, with certainty $\geq 1-\alpha$, we can conclude that the frequency f belongs to the interval $[p - k_\alpha \cdot \sigma, p + k_\alpha \cdot \sigma]$, i.e., to the interval

$$\left[p - k_\alpha \cdot \sqrt{\frac{p\cdot(1-p)}{N}}, p - k_\alpha \cdot \sqrt{\frac{p\cdot(1-p)}{N}}\right].$$

When p is not too close to 0 or 1, then $p \approx f$, so $p\cdot(1-p) \approx f\cdot(1-f)$. Hence, from

$$|f-p| \leq k_\alpha \cdot \sqrt{\frac{p\cdot(1-p)}{N}},$$

we can conclude that

$$|f-p| \leq k_\alpha \cdot \sqrt{\frac{f\cdot(1-f)}{N}},$$

hence

$$p \in \left[f - k_\alpha \cdot \sqrt{\frac{f\cdot(1-f)}{N}}, f - k_\alpha \cdot \sqrt{\frac{f\cdot(1-f)}{N}}\right].$$

When p is small, e.g., when $f = 0$ is possible, then this approximate formula would lead to $p = 0$, which is not necessarily true. For $f = 0$:

the inequality $f \leq p + k_\alpha \cdot \sigma$ is always true, so

the only inequality that we need to check is $f = 0 \geq p - k_\alpha \cdot \sigma$, i.e., $p \leq \sqrt{p(1-p)/N}$.

Squaring this inequality and taking into consideration that $p \approx 0$ hence $1-p \approx 1$, we conclude that $p \leq k_\alpha^2/N$.

A similar inequality holds for $p \approx 1$: $p \geq 1 - k_\alpha^2/N$. Summarizing:

for small p and for $p \approx 1$, the bound on $f-p$ is k_α^2/N;

for larger p, this bound is $f + k_\alpha \cdot \sqrt{f\cdot(1-f)/N}$.

The first bound is larger until these bounds are equal, i.e., until $f + k_\alpha \cdot \sqrt{f \cdot (1-f)/N} \approx f + k_\alpha \cdot \sqrt{f/N}$. We know that this happens around $f \approx k_\alpha^2/N$, so we look for the solution of the form $f = z \cdot k_\alpha^2/N$ for some z. Substituting this expression into the above equation, we conclude that $z + \sqrt{z} = 1$, i.e., that $\sqrt{z} = (\sqrt{5}-1)/2$ hence $z = (3-\sqrt{5})/2 \approx 0.35$. Thus:

for $f \le 0.35 \cdot k_\alpha^2/N$, we conclude that $p \le k_\alpha^2/N$;

for $0.35 \cdot k_\alpha^2/N \le f \le 1 - 0.35 \cdot k_\alpha^2/N$, we have

$$p \in \left[f - k_\alpha \cdot \sqrt{\frac{f \cdot (1-f)}{N}}, f - k_\alpha \cdot \sqrt{\frac{f \cdot (1-f)}{N}} \right];$$

for $f \ge 1 - 0.35 \cdot k_\alpha^2$, we have $p \ge 1 - k_\alpha^2/N$.

Proposed approach

In our case, we want similar estimates for all N points $y^{(1)}, \ldots, y^{(N)}$. In other words, we want to guarantee that with the certainty $\ge 1-\alpha$, we do not exceed the bounds for each of these N values. To get the overall probability bounded by α, it is reasonable to make each probability bounded by α/N, i.e., select the following values:

for $F_N(t) \le 0.35 \cdot \frac{k_{\alpha/N}^2}{N}$, we conclude that $F(t) \le \frac{k_{\alpha/N}^2}{N}$;

for

$$0.35 \cdot \frac{k_{\alpha/N}^2}{N} \le F_N(t) \le 1 - 0.35 \cdot \frac{k_{\alpha/N}^2}{N},$$

we have

$$F(t) \in \left[F_n(t) - k_{\alpha/N} \cdot \sqrt{\frac{F_N(t) \cdot (1-F_N(t))}{N}}, F_N(t) - k_{\alpha/N} \cdot \sqrt{\frac{F_N(t) \cdot (1-F_N(t))}{N}} \right];$$

for $F_N(t) \ge 1 - 0.35 \cdot \frac{k_{\alpha/N}^2}{N}$, we have $F(t) \ge 1 - \frac{k_{\alpha/N}^2}{N}$.

In particular, for a realistic case $\alpha = 0.05$ and $N = 50$, we have $\alpha/N = 0.1\%$, hence $k_{\alpha/N} = 3$.

When N is small, then, instead of the above normal distribution formulas, we should use the more accurate formulas of estimating the parameters of the binomial distribution (see Dixon and Massey 1969).

Convergence

For large N, just like for estimating probability from frequency, the estimating intervals converge to the actual values of $F(t)$, and the widths decrease, in effect, as $1/\sqrt{N}$.

Actually, the decrease is a little bit slower because, as N increases, we need to take $k_{\alpha/N}$ instead of k_α. However, since the Gaussian distribution exponentially decreases, the value $k_{\alpha/N}$ grows logarithmically with N, so this growth is much slower than the $1/\sqrt{N}$ decrease.

Comparison with the traditional Kolmogorov-Smirnov statistic

Is what we are proposing better than the traditional Kolmogorov-Smirnov statistic? Yes, and here is why: in Kolmogorov-Smirnov statistic, in effect, we pick the width in such a way that even for the case of the largest σ (i.e., for $F_N(t) \approx 0.5$), we still fit within the tolerance bound (e.g., 2 sigma or 3 sigma).

For much smaller $F_N(t)$, the traditional Kolmogorov-Smirnov statistic uses the exact same width. However, e.g., for $f = F_N(t) = 0.05$, the standard deviation of the difference $p - f$ is proportional to $\sqrt{0.05 \cdot 0.95} \approx 0.22$ and is, thus, more than twice smaller than the standard deviation at $f \approx 0.5$ which is proportional to $\sqrt{0.5 \cdot 0.5} = 0.5$. Thus, the certainty bound that is a 3 sigma bound for $F_N(t) \approx 0.5$ becomes a 6 sigma bound for $F_N(t) \approx 0.05$. The probability of exceeding 6 sigma is so miniscule ($\approx 10^{-6}\%$) that we can safely decrease this bound by, say, 50%, and still get the exceeding probability much smaller than at $F_N(t) \approx 0.5$.

In other words:

In the traditional Kolmogorov-Smirnov formula, we can drastically decrease the widths of the certainty intervals corresponding to small values $F_n(t)$ without changing the overall certainty level.

In the propose statistic, the width of a certainty interval at each value t is already adjusted to make sure that it is not too wide.

8. Analysis of the Cauchy deviate method

It is important to note first that the Cauchy deviate method does not propagate probabilistic uncertainty. Our analysis of the case of probabilistic uncertainty showed that for this uncertainty, Kolmogorov-Smirnov-type methods are better. Thus, the Cauchy deviates methods are not the recommended choice here. This conclusion makes perfect sense, because the proposed use of the Cauchy deviates technique ignores all the information about the probability distributions and thus,

can only lead to inferior estimates. However, in contrast to the case of probabilistic uncertainty, where Cauchy deviates methods are not recommended, these methods are highly recommended for the case of interval uncertainty. As we have mentioned, in the case of interval uncertainty, if the function can be reasonably approximated by a linear expression, we can use the Cauchy deviates techniques to estimate the desired parameter Δ: the width of the interval of possible values of y.

In the previous descriptions of the Cauchy deviates method (Kreinovich and Ferson 2004; Trejo and Kreinovich 2001), we concentrated on finding the number of iterations N that would provide the desired accuracy (usually, 20% accuracy in estimating Δ). The difference between the actual value Δ and its estimate $\tilde{\Delta}$ is distributed, for large N, according to normal distribution, with 0 mean and standard deviation $\sigma_e = \Delta \cdot \sqrt{2/N}$. Thus, e.g., to get a 20% accuracy $0.2 \cdot \Delta$ with 95% certainty (corresponding to $2\sigma_e$), we need $N = 200$ runs.

After 200 runs, we can conclude that $\Delta \le 1.2 \cdot \tilde{\Delta}$ with certainty 95%.

What if we cannot perform this many sample evaluations of the black box? It may not be possible to run the program f 200 times. In this case, we can still use the Cauchy deviates estimates, but we need to come up with new formulas that translate the numerical estimate into the enclosure for Δ. If N is large enough so that the difference $\tilde{\Delta} - \Delta$ is still Gaussian, we can conclude that

$$\Delta \le \tilde{\Delta} \cdot \left(1 + k_0 \cdot \sqrt{\frac{2}{N}}\right)$$

(where $k_0 = 2$), with certainty 95%. (If we want, e.g., 99.9% certainty, which corresponds to 3 sigma, then we should take $k_0 = 3$.) Thus, e.g., for $N = 50$, we conclude that $\Delta \le 1.4 \cdot \tilde{\Delta}$. This may yield a workable estimate.

Very few evaluations

When the number of iterations is even smaller, then we can no longer assume that the distribution of the error $\tilde{\Delta} - \Delta$ is Gaussian. In this case, to find the bounds on Δ with, e.g., 95% certainty, we must perform numerical experiments. The possibility of such experiments is caused by the fact that, as we have mentioned in the above description of the Cauchy deviates method, the distribution of the results $\delta y^{(k)}$ always follows the Cauchy distribution, no matter how small N is.

So, to find out the confidence bounds on the Cauchy deviate estimates, it is sufficient to make experiments with the Cauchy distribution. The Cauchy distribution with a parameter Δ can be obtained by multiplying the Cauchy-distributed random variable with parameter $\Delta_0 = 1$ by the number Δ. Thus, it is sufficient to test the method on Cauchy deviates with parameter 1.

For each N and α, we want to find $k(N,\alpha)$ for which $\Delta \le k(N,\alpha)\cdot \overline{\Delta}$ with certainty $1-\alpha$, i.e., for which $\overline{\Delta} \ge (1/k(N,\alpha))\cdot \Delta$ with probability $1-\alpha$. Since we will be using Cauchy distribution with $\Delta = 1$, we must thus find $k(N,\alpha)$ for which $\overline{\Delta} \ge 1/k(N,\alpha)$ with probability $1-\alpha$.

To find such value, we do the following. We pick a large number of iterations M (the relative accuracy of our estimate of $k(N,\alpha)$ will be $\approx 1/\sqrt{M}$). Then:

For each m from 1 to M:

we simulate Cauchy distribution (with parameter $\Delta_0 = 1$) N times, producing N numbers

$$\delta y_1^{(m)} = \tan(\pi\cdot(r_1^{(m)} - 0.5)), \ldots, \delta y_N^{(m)} = \tan(\pi\cdot(r_N^{(m)} - 0.5));$$

we then apply the above Maximum Likelihood Method to find $\overline{\Delta}_m$ as the solution to the following equation:

$$\frac{1}{1+\left(\frac{\delta y_1^{(m)}}{\overline{\Delta}_m}\right)^2} + \ldots + \frac{1}{1+\left(\frac{\delta y_N^{(m)}}{\overline{\Delta}_m}\right)^2} = \frac{N}{2};$$

we solve this equation by applying a bisection method to the interval

$$\left[0, \max_i \left|\delta y_i^{(m)}\right|\right].$$

After that, we sort the values $\overline{\Delta}_m$ into an increasing sequence

$$\overline{\Delta}_{(1)} \le \ldots \le \overline{\Delta}_{(M)}.$$

We take the value $\overline{\Delta}_{(\alpha\cdot M)}$ for which the probability to be greater than this number is exactly $1-\alpha$, and estimate $k(N,\alpha)$ as $1/\overline{\Delta}_{(\alpha\cdot M)}$.

Simulation results

We wrote a C program that implements this algorithm; this program is listed in Appendix 2. For $\alpha = 0.05$, the results of applying this program are:

For $N = 20$, we get $k \approx 1.7$, which fits very well with the above Gaussian-based formula $k_{\text{norm}} \approx 1 + 2 \cdot \sqrt{2/20} \approx 1.7$.

For $N = 10$, we get $k \approx 2.1$, which is slightly higher than the Gaussian-based formula $k_{\text{norm}} \approx 1 + 2 \cdot \sqrt{2/10} \approx 1.9$.

For $N = 5$, we get $k \approx 5$, which is already much higher than the Gaussian-based value $k_{\text{norm}} \approx 1 + 2 \cdot \sqrt{2/5} \approx 2.3$.

9. Sampling entire distributions

Various sampling-based schemes have been proposed that might be used to extend the methods of uncertainty projection based on Dempster-Shafer evidence theory and probability bounds analysis to black boxes. For example, Helton et al. (2004a,b) suggested decomposing the problem into a Cartesian product (à la Yager 1986) and solving the resulting matrix of interval problems by using black box sampling to estimate the ranges of these intervals with the observed ranges of the sample outputs. The approach is designed for problems involving Dempster-Shafer inputs, but it can be immediately applied to problems with p-boxes using the basic conversions described in Ferson et al. (2003). In principle, various ancillary strategies might be used to accelerate the convergence of this approach, such as methods that take account of overlap among the input intervals or the likely association of extreme values of the output variables with extreme values of the input variables, or employ strategic simplifications to reduce the dimensionality of the problem (Helton et al. 2006c). The approach was illustrated for a problem involving an algebraic expression in Helton et al. (2004b), and for a much more complex problem involving competing failure risks of strong and weak switches in Helton et al. (2004a).

Bruns et al. (2006a) described an alternative direct sampling approach called "optimized parameter sampling". This approach can be applied in situations where the inputs are "parameterized" p-boxes, which are essentially collections of distributions of a given shape (such as normal) specified by one or two parameters from within given intervals. In an outer sampling loop, distributions are selected from all the *k* input p-boxes by randomly picking the scalar parameter values from their respective intervals. For each collection of *k* (precise) probability distributions, sampling-based techniques are employed in an inner loop to solve the twin optimization problems of finding the upper and lower bounds on the expectation or any percentile of the result. Bruns et al. (2006a) illustrated the sampling strategy and evaluated its efficiency on a problem estimating the first passage time for a thermocouple temperature.

Bruns et al. (2006b) also described yet another direct sampling approach called p-box convolution sampling, although it can be applied immediately to Dempster-Shafer structures too. It involves taking random samples from each of the *k* inputs. A random sample from an uncertain number is the interval corresponding to the $(r\ 100)^{\text{th}}$ percentile where *r* is a random number uniformly distributed on the unit interval [0,1]. This generalizes the selection of a random value from a precise probability distribution (Ferson and Ginzburg 1995; Cooper et al. 1996). These *k* intervals

are then projected through the black box model using sampling-based optimization techniques to find the largest and smallest output values given the input intervals, or the Cauchy-deviate sampling strategy described by Trejo and Kreinovich and (2001).

It does not seem possible to account for uncertainty about dependence among the inputs by generalizing the approach of Helton et al. (2004a,b) or either the optimized parameter sampling* or p-box convolution sampling approaches of Bruns et al. (2006a,b). Even if it is possible to relax this restriction, all of these entire distribution sampling approaches will remain computationally intensive and would likely produce reliable results only when fairly many sample evaluations of the black box can be made. Despite their limitations and computational costs, these sampling approaches could make the new methods workable for a variety of problems in engineering. These methods of sampling entire distributions convert the problem of propagating uncertainty characterized by Dempster-Shafer structures or p-boxes into several easier problems of propagating precise probability distributions with pure aleatory uncertainty.

Bernardini and Tonon (2008) describe a way to convert any p-box or Dempster-Shafer structure to a list of extreme distributions associated with the credal set, that is, the space of all distribution functions consistent with a specified p-box or Dempster-Shafer structure. The theory of imprecise probabilities (Walley 1990) holds that it is these extremal distributions that determine the possible ranges of the mean, tail risks and other features of the combination of epistemic and aleatory uncertainty. It is not yet clear how this ability to identify the extreme distributions of the credal set can be used to inform a practical strategy for calculation, although

10. P-boxes and Dempster-Shafer knowledge bases

In the previous sections, we described and analyzed different methods for estimating uncertainty in the cases when we have probabilistic or interval uncertainty in the inputs. What if the uncertainty in each input x_i is characterized, e.g., by the Dempster-Shafer knowledge bases?

One reason why this problem is difficult is that it is not even clear how we can represent the Dempster-Shafer knowledge base corresponding to the output. Indeed, a Dempster-Shafer knowledge base for each input variable x_i means that we may have different intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$, with different probabilities $p_i^{(k)}$. For each combination of intervals,

$$[\underline{x}_1^{(k_1)}, \overline{x}_1^{(k_1)}], \ldots, [\underline{x}_n^{(k_n)}, \overline{x}_n^{(k_n)}],$$

*Bruns et al. (2006a) suggested that the optimized parameter sampling approach does not generalize to account for uncertainty about distribution family or to general Dempster-Shafer structures. However, this generality might be possible by devising a scheme for "sampling" discrete probability distributions from a non-parametric uncertain number. The theory of Chebyshev systems (Karlin and Studden 1966) suggests that the bounds on uncertain numbers are set by degenerate discrete probability distributions having mass only on a minimal number of points. It should be possible to generate example distributions that are consistent with given p-boxes or Dempster-Shafer structures.

we can use the known techniques to find the corresponding interval $[\underline{y}^{(k_1,...,k_n)}, \overline{y}^{(k_1,...,k_n)}]$ for the output. Since we know the probability $p_i^{(k_i)}$ of each interval $[\underline{x}_i^{(k_i)}, \overline{x}_i^{(k_i)}]$, and we assume that these probabilities are independent, we can compute the probability $p^{(k_1,...,k_n)}$ of the corresponding output interval as the product $p^{(k_1,...,k_n)} = p_1^{(k_1)} \cdot ... \cdot p_n^{(k_n)}$.

At first glance, this may sound like a reasonable solution to our problem, but in reality, this solution is not practical at all: even in the simplest case, when for each variable, we have two possible intervals, for $n = 50$ inputs, we will have an astronomical number of $2^{50} \approx 10^{15}$ output intervals $[\underline{y}^{(k_1,...,k_n)}, \overline{y}^{(k_1,...,k_n)}]$.

Thus, although the resulting uncertainty is still a Dempster-Shafer uncertainty, we can no longer represent it as we represented the uncertainty for each input: by listing all the intervals and the corresponding probabilities.

Thus, not only it is not clear how to compute the resulting uncertainty, it is not even clear what exactly we want to compute.

Our idea comes from the fact that the Dempster-Shafer uncertainty is a generalization of interval uncertainty, a generalization in which, for each inputs x_i, instead of a single interval $[\underline{x}_i, \overline{x}_i]$, we have several possible intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$, with different probabilities $p_i^{(k)}$. For the interval uncertainty, in a realistic case when the black-box function is linearizable, we can use the Cauchy deviates method to estimate the interval uncertainty of the output. Let us see whether it is possible – at least, under some reasonable assumptions – to extend the Cauchy deviates method to the more general Dempster-Shafer case.

The fact that the black-box function is linearizable means that we have

$$f(x_1,...,x_n) = \tilde{y} + \sum_{i=1}^{n} c_i \cdot (x_i - \tilde{x}_i),$$

where $\tilde{y} \stackrel{\text{def}}{=} f(\tilde{x}_n,...,\tilde{x}_n)$ and for every i, c_i denotes the (unknown) value of the partial derivative $\partial f / \partial x_i$ of the black-box function $f(x_1,...,x_n)$ with respect to i-th input x_i.

If we know the exact values $x_1,...,x_n$ of all the inputs, then we can simply plug in the values x_i and get the desired value.

If for each i, we know the interval $[x_i^{\text{mid}} - \Delta_i, x_i^{\text{mid}} + \Delta x_i]$, then, in the linearized case described above, the corresponding range of y can be described by the interval $[y^{\text{mid}} - \Delta, y^{\text{mid}} + \Delta]$, where:

$$y^{\text{mid}} = \tilde{y} + \sum_{i=1}^{n} c_i \cdot (y_i^{\text{mid}} - \tilde{y}_i); \quad (6)$$

$$\Delta = \sum_{i=1}^{n} |c_i| \cdot \Delta_i. \qquad (7)$$

In the Dempster-Shafer case, for each i, instead of a single pair $(y_i^{\text{mid}}, \Delta_i)$, we have different pairs with different probabilities. Due to the formulas (6) and (7), the vector $(y^{\square}, \Delta)$ is a linear combination of the vectors $(y_i^{\text{mid}}, \Delta_i)$ corresponding to different inputs x_i.

If one of these vectors was prevailing, then we would have a single input (or a few dominating inputs), and there would be no need to consider the uncertainty in all n inputs. Thus, the only case when this problem makes sense is when the contributions of all n vectors is approximately of the same size (or at least the same order of magnitude). In this case, the vector (y^{mid}, Δ) is a linear combination of n independent vectors of approximately the same size.

This situation is exactly the case covered by the Central Limit Theorem, the case when in the limit $n \to \infty$, we have a normal 2-D distribution and hence, for sufficient large n, with a good approximation, we can assume that the pair (y^{mid}, Δ) is normally distributed.

Note that, strictly speaking, the distribution is almost normal, but it is not exactly normal. From the purely *theoretical* viewpoint, the distribution of the pairs (y^{mid}, Δ) cannot be exactly normal, because the interval half-width Δ is always non-negative, while for every normally distributed random variable, there is a non-zero probability that this value attains negative values. However, *in practice*, every normal distribution with mean μ and standard deviation σ is located within the interval $[\mu - k\,\sigma, \mu + k\,\sigma]$ with practically a certainty, i.e., with probability ≈ 1. This means that, for $k = 3$, the probability to be outside the 3 sigma interval is $\approx 0.1\%$, and for $k = 6$, the probability to be outside the 3-sigma interval is $\approx 10^{-6}\%$, etc. Thus, if $\mu \geq k\,\sigma$, then, for all practical purposes, the half-width Δ is indeed always non-negative.

It is therefore reasonable to conclude that for large n, the uncertainty in y can be characterized as follows: we have different intervals $[y^{\text{mid}} - \Delta, y^{\text{mid}} + \Delta]$, and the probability of an interval is described by a 2-D normal distribution on the (y^{mid}, Δ) plane. To describe a 2-D normal distribution, it is sufficient to know 5 parameters: the means and standard deviations of both variables and the covariance (that describes their dependence). At first glance, it may seem like we are abandoning our approach: we started with the idea of having non-parametric estimates, and we ended up with a 5-parametric family. However, realistically, to exactly describe a generic distribution, we must use infinitely many parameters. In reality, we only have finitely many runs of the black-box function f with reasonable accuracy, and based on their results, we can only estimate finitely many parameters anyway.

Even in the ideal case of Monte Carlo tests, we need N experiments to get a value of each parameter with an accuracy of $1/\sqrt{N}$. Thus, to get a reasonably low accuracy of 30% (everything worse makes it order-of-magnitude qualitative estimate), we need ≈ 10 runs. With 50 runs, we can therefore determine the values of no more than 5 parameters anyway. The above 5-parametric

family is reasonable, its justification is very similar to the common justification of the use of Gaussian distributions.

Determining the parameters

How can we determine the parameters of this model. If we simply take the midpoints $x_i^{(k)\text{mid}}$ of the corresponding intervals in our simulations, then the resulting value $y^{(k)}$ are normally distributed, with the distribution corresponding to y^{mid}. We can therefore estimate the mean and standard deviation of y^{mid} as simply the sample mean and the sample variance of the values $y^{(1)}, y^{(2)}, \ldots$

For Δ, from the formula (7), we conclude that

$$E[\Delta] = \sum_{i=1}^{n} |c_i| \cdot E[\Delta_i] \quad (8)$$

and

$$\sigma[\Delta] = \sqrt{\sum_{i=1}^{n} |c_i|^2 \cdot \sigma^2[\Delta_i]}. \qquad (9)$$

Due to the formula (8), we can use the Cauchy deviates technique to estimate $E[\Delta]$ if for each input x_i, we use the average half-width

$$E[\Delta_i] = p_i^{(1)} \cdot \Delta_i^{(1)} + p_i^{(1)} \cdot \Delta_i^{(1)} + \ldots$$

of the corresponding interval.

Due to the fact that $|c_i|^2 = c_i^2$, the formula (9) means that we can compute $\sigma[\Delta]$ by using the standard Monte Carlo simulation technique: namely, we simulate δx_i to be normally distributed with 0 mean and standard deviation $\sigma[\Delta_i]$, then the resulting value of $\delta y = \sum c_i \cdot \delta x_i$ is also normally distributed, with the standard deviation equal to (9). We can thus estimate (9) as a sample variance of the corresponding simulated values $\delta y^{(k)}$.

We thus know how to estimate 4 of 5 parameters that describe the desired uncertainty. The only remaining problem is how to estimate the covariance between y^{mid} and Δ. For this, we propose the following idea.

The non-zero covariance means, in particular, that the conditional average $E[\Delta \,|\, y^{\text{mid}} \le E[y^{\text{mid}}]]$ of Δ over the cases when y^{mid} is smaller than its average $E[y^{\text{mid}}]$ is different from the conditional average $E[\Delta \,|\, y^{\text{mid}} \ge E[y^{\text{mid}}]]$ of Δ over the cases when y^{mid} is larger than its

average $E[y^{\text{mid}}]$. From the difference between these two conditional averages, we can determine the desired value of the covariance.

To compute the conditional averages, we can use the Cauchy deviates idea. Namely, at each simulation, for each variable x_i, we select one of the intervals $[\underline{x}_i^{(k)}, \overline{x}_i^{(k)}]$ with the corresponding probability $p_i^{(k)}$, and we apply the black box function f to the centers of the corresponding intervals, to get the result y^{mid}. We then apply the Cauchy techniques with the corresponding intervals and get the value distributed according to the Cauchy distribution with the width corresponding to selected intervals for x_i.

The main difference between what we propose to do here and the previously described Cauchy deviates methods is that, in the previously described Cauchy deviates method, we combine all the results of Cauchy simulation into a single sample, and we then compute the parameter Δ based on this sample, but, in the proposed methods, we separate the results of Cauchy simulation into two different samples, that is, a sample containing all the cases in which $y^{\text{mid}} \le E[y^{\text{mid}}]$, and a sample containing all the cases in which $y^{\text{mid}} \ge E[y^{\text{mid}}]$.

In the previous described approach, in all simulations, we had *the same* interval width, so the results of the simulation belong to the same Cauchy distribution. In the new method, we have *different* widths with different probabilities, so the resulting distribution is a combination of different Cauchy distributions, with different probabilities.

For each sample, we can safely assume that the distribution of the width Δ is a Gaussian distribution, with mean μ and standard deviation σ. Thus, our sample corresponds to the combination in which the Cauchy distribution with parameter Δ occurs with the Gaussian probability density

$$\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma}\cdot\exp\left(-\frac{(\Delta-\mu)^2}{2\sigma^2}\right).$$

Cauchy-distributed random variable ξ with the parameter Δ can be described by its characteristic function $E[\exp(\mathrm{i}\cdot\omega\xi)]=\exp(-|\omega|\cdot\Delta)$. Thus, the above-described probabilistic combination of Cauchy distributions can be described by the corresponding probabilistic combination of these characteristic functions:

$$E[\exp(\mathrm{i}\cdot\omega\cdot\xi)]=\int\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma}\cdot\exp\left(-\frac{\Delta-\mu}{2\sigma^2}\right)\cdot\exp(-|\omega|\cdot\Delta)\,\mathrm{d}\Delta. \qquad (10)$$

As we show in Appendix 3, this integral is equal to:

$$\exp\left(\frac{1}{2}\cdot\sigma^2\cdot\omega^2-\mu\cdot|\omega|\right). \qquad (11)$$

We can estimate the characteristic function by its sample value

$$E[\exp(\mathrm{i}\cdot\omega\cdot\xi)] \approx \frac{1}{N}\cdot\sum_{k=1}^{N}\cos(\omega\cdot y^{(k)})$$

(Since the expression (11) is real, it makes sense to only consider the real part of $\exp(\mathrm{i}\cdot\omega\cdot\xi)$, i.e., $\cos(\omega\cdot\xi)$.)

So, we arrive at the following algorithm for computing μ and σ from the sample values $y^{(1)},\ldots,y^{(N)}$:

for different real values $\omega_1,\ldots,\omega_k > 0$, compute $l(\omega_k) \stackrel{\text{def}}{=} -\ln(c(\omega_k))$, where

$$c(\omega_k) \stackrel{\text{def}}{=} \frac{1}{N}\cdot\sum_{k=1}^{N}\cos(\omega\cdot y^{(k)});$$

use the Least Squares Method to find the values μ and σ for which

$$\mu\cdot\omega_k - \frac{1}{2}\sigma^2\cdot\omega_k^2 \approx l(\omega_k).$$

The resulting value μ is the average Δ.

Thus, when we repeat this algorithm for both samples, we get the two desired conditional averages of Δ – from which we can then compute the covariance.

Application to p-boxes

It is known that a p-box can be described as a Dempster-Shafer knowledge base (Ferson et al. 2003). A p-box $[\underline{F}(t),\overline{F}(t)]$ is a generalization of a cumulative distribution function $F(t)$. A cumulative distribution function can be represented by an explicit formula, or it can be represented by listing the quantiles corresponding to different probability levels. For instance, one could list all the percentiles for probability levels p = 0, 0.01, 0.02, …, 0.99, 1. A sufficiently finely discretized list would be useful for almost any practical application. In mathematical terms, quantiles are the values of the inverse function $f(p) = F^{-1}(p)$. Using equally spaced probability values, each Δp far apart, a variable with a probability distribution $F(t)$ can be approximately described by the values $f(0), f(\Delta p), f(2\Delta p)$, etc. The discretization suggests that the variable takes on each of these quantiles with equal probability Δp. Similarly, a p-box can be alternatively represented by listing, for each p, the interval $[\underline{f}(p),\overline{f}(p)]$ of the possible quantile values where the function $\underline{f}(p)$ is an inverse function to $\overline{F}(t)$, and the function $\overline{f}(p)$ is an inverse function to $\underline{F}(t)$. Because of this description, we can interpret the p-box as the Dempster-Shafer knowledge base, in which, with equal probability Δp, we can have intervals $[\underline{f}(0),\overline{f}(0)]$, $[\underline{f}(\Delta p),\overline{f}(\Delta p)]$, etc. Thus, whatever method we have for Dempster-Shafer knowledge bases, we can apply it to p-boxes as well.

How can we describe the resulting p-boxes? We have just mentioned that, in principle, we can interpret each p-box as a Dempster-Shafer knowledge base, and apply the above Dempster-Shafer method to describe the uncertainty of the output. The result, however, is a Dempster-Shafer knowledge base. How can we describe the corresponding "Gaussian" Dempster-Shafer knowledge base as a p-box?

For a Dempster-Shafer knowledge base, i.e., for a probabilistic distribution on the set of intervals $[\underline{x}, \overline{x}]$, the probability $F(t) = \text{Prob}(X \le t)$ attains its largest possible value $\overline{F}(t)$ if for each interval, we take the smallest possible value $\underline{x}$. Similarly, the probability $F(t) = \text{Prob}(X \le t)$ attains its smallest possible value $\underline{F}(t)$ if for each interval, we take the largest possible value $\overline{x}$. Thus, $\overline{F}(t)$ is a probability distribution for the lower endpoints $y^{\min} - \Delta$, and $\underline{F}(t)$ is a probability distribution for the upper endpoints $y^{\min} + \Delta$ of the corresponding intervals. Since the 2-D distribution of the pairs (y^{mid}, Δ) is Gaussian, the distributions of both linear combinations $y^{\min} - \Delta$ and $y^{\min} + \Delta$ are Gaussian as well. Therefore, as a result of this procedure, we get a p-box $[\underline{F}(t), \overline{F}(t)]$ for which both bounds $\underline{F}(t)$ and $\overline{F}(t)$ correspond to Gaussian distributions.

Strictly speaking, the distributions are almost normal but not exactly normal. Let us denote the cdf of the standard Gaussian distribution, with 0 mean and standard deviation 1 by $F_0(t)$. Then, an arbitrary Gaussian distribution, with mean μ and standard deviation σ, can be described as $F(t) = F_0((t-\mu)/\sigma)$. In particular, if we denote the mean and the standard deviations of the Gaussian distribution $\underline{F}(t)$ by $\underline{\mu}$ and $\underline{\sigma}$, and the mean and the standard deviations of the Gaussian distribution $\overline{F}(t)$ by $\overline{\mu}$ and $\overline{\sigma}$, then we conclude that $\underline{F}(t) = F_0((t-\underline{\mu})/\underline{\sigma})$ and $\overline{F}(t) = F_0((t-\overline{\mu})/\overline{\sigma})$. From the theoretical viewpoint, for thus defined functions $\underline{F}(t)$ and $\overline{F}(t)$, we cannot always have $\underline{F}(t) \le \overline{F}(t)$, because, due to monotonicity of $F_0(t)$, this would be equivalent to

$$\frac{t-\underline{\mu}}{\underline{\sigma}} \le \frac{t-\overline{\mu}}{\overline{\sigma}}$$

for all t, i.e., to one straight line being always below the other – but this is only possible when they are parallel.

However, as we have mentioned while describing the similar situation with the Dempster-Shafer knowledge bases, in practice, we can have this inequality if we ignore the values t for which $F_0(t)$ is very small, and therefore not practically possible. Alternatively, we can assume that the inequality $\underline{F}(t) \le \overline{F}(t)$ holds for all t – but the distributions $\underline{F}(t)$ and $\overline{F}(t)$ are only approximately—although not exactly—normal.

This approach can also be applied if we have different types of uncertainty for different inputs. If we have different types of uncertainty for different inputs, we can transform them to p-boxes, and hence to Dempster-Shafer knowledge bases, and use a similar approach.

11.Conclusions

When the uncertainty about model inputs is *purely aleatory* and characterized by probability distributions, traditional Monte Carlo methods can be used to estimate the distributions of the output variables from a black box. Assuming the samples were selected at random (i.e., independently of each other), confidence bands can be computed for the empirical distributions of the output values that conservatively characterize the epistemic uncertainty about that output given the limited number of Monte Carlo replications that could be computed. When the number of samples grows large, these bands converge to the true distribution of the output variable which is some convolution of the input distributions. The confidence bands can be selected to have any prescribed coverage probability, i.e., the probability that the true falls entirely within the computed confidence bands. The confidence bands may incorporate assumptions about the distribution shape of the output quantity or they may make no a priori assumption about that shape. In principle, these bands may be computed from any number of sample evaluations of the black box, although at least five samples are necessary for the bands to be non-trivial.

When the uncertainty about the model inputs is *purely epistemic* in nature and characterized by intervals, several related methods can be to estimate the interval range of an output variable from the black box model so long as the function in the black box is mathematically well behaved. For instance, if the model is linear over the uncertain variables, then as few as $D + 1$ evaluations of the black box may be needed to compute the range of the output. If the model is approximately linear, then this strategy will yield approximate results. If the model is not approximately linear, but is monotone in each uncertain variable, then 2^D evaluations would be needed to compute the output range. If the black box function is decidedly nonlinear, but can be approximated by a piecewise linear function with P segments, then these strategies can be used on each segment with the final result taken as the union over the segment results. In this case, $P(D+1)$ evaluations of the black box would be needed. Such directed sampling is useful when the number of uncertain inputs is small relative to the cost of each evaluation of the black box. When the number of uncertain variables D is very large, however, this approach may not be practical. In such cases, if the black box function is approximately linear over the range of interval uncertainty, then the Cauchy deviate method (section <<>>) can be used. The coverage characteristics of this approach are approximate if the function is only approximately linear. Its level of conservativism can be prescribed by the analyst.

When the model inputs have both aleatory and epistemic uncertainty, the problem becomes much more complex, and no single method seems to be preferable for use in the black box problem. The table below summarizes the performances of several possible schemes for projecting mixed uncertainty through black boxes in terms of each method's conservativism, statistical coverage properties, whether it eventually convergence to the true answer, and cost in sample evaluations.

Method	*Conservative*	*Coverage*	*Convergence*	*Cost*
condense & interval	Yes	Approximate	No	*K*
supp-core interval	Yes	Approximate	No	2*K*
entire distribution sampling	Maybe	Approximate	Yes	*MD*
Saw-Yang-Mo inequality	No	No	No	any
uniformity principle*	Yes		No	
u,*v*-sampling	No	No	Yes	*MDL*
u,end-sampling	No	No		
u-sampling & interval			Yes	*DK*
confidence & *u*,*v*-sampling	Maybe	No	Yes	*MDL*
confidence & *u*,end-sampling	Maybe	No		
confidence & *u*-sampling & interval	Yes	Yes	Yes	*DK*

The first column in the table names the various methods considered in this report. When "interval" is mentioned as part of the method name, it refers to any of the related interval propagation schemes discussed in section <<>> such as those for linear or monotone functions, the Cauchy deviate method or its variants. When "confidence" is mentioned, it refers to any of the procedures for confidence bands about distribution functions such as the Kolmogorov-Smirnov method or related distributional methods discussed in section <<>>. The second column indicates whether the method yields a conservative result that would be appropriate for an outside-in approximation strategy. The coverage column indicates whether the result has a guaranteed statistical performance, i.e., whether an analyst can prescribe the probability with which the true distribution function will lie entirely inside the computed result. The notation in the convergence column tells whether the result of applying the method when there are asymptotically many samples converges to the correct answer that would be obtained with a best-possible analysis using intrusive methods. Finally, the cost column gives an indication of the relative cost of applying the method, where *K* denotes the cost in sample evaluations of propagating intervals (such as by the direct sampling or Cauchy deviate methods), *M* is the number of sample evaluations needed to project a probability distribution, *D* is the number of uncertain inputs.

Inequality of Saw et al.

The generalization of the Chebyshev inequality proposed by Saw et al. (1984) at first seemed very promising as a way for an analyst to escape specifying the particular shape of the underlying distribution. It seemed to promise to give a conservative answer over all possible distribution shapes even when the number of samples is small. Such a method would be ideal for propagating uncertainty through a black-box model. Even substantial conservativism might be a cost worth paying for a distribution-free method to rigorously assess tail risks from small samples. However, analysis of the inequality reveals that this seeming advantage is based on a misunderstanding, whose origin lies in different interpretations of probability that are in common circulation.

In fact, the difference in interpretations also explains why the Saw inequality sometimes leads to tighter bounds than the original Chebyshev inequality (a real anomoly since the original bounds are best possible). Thus, the inequality proposed by Saw et al. (1984) is arguably not really a generalization of the classical Chebyshev inequality. Instead, it might better be thought of as merely being inspired by it, but answering a completely different question. Although the question it answers might be interesting for some applications under a subjectivist interpretation of probability, it does not provide the answers we are seeking when we propagate aleatory and epistemic uncertainties through black boxes. It is evident that the inequality of Saw et al. is not useful for the black box problem.

Appendix 1: Optimal estimates of $F(t)$ are scale-invariant

We analyze the following question: which statistic for estimating $F(t)$ is the best? When we say "the best", we mean that on the set of all such statistics, there must be a relation $\geq$ describing which statistic is better or equal in quality. This relation must be transitive (if A is better than B, and B is better than C, then A is better than C). This relation is not necessarily asymmetric, because we can have two statistics of the same quality. However, we would like to require that this relation be *final* in the sense that it should define a unique *best* statistic L_{opt}, i.e., the unique statistic for which $\forall L\,(L_{\text{opt}} \geq L)$. Indeed, if none of the statistics is the best, then this criterion is of no use, so there should be *at least one* optimal family.

If *several* different statistics are equally best, then we can use this ambiguity to optimize something else: e.g., if we have two statistics with the same approximating quality, then we choose the one which is easier to compute. As a result, the original criterion was not final: we get a new criterion ($A \geq_{\text{new}} B$ if either A gives a better approximation, or if $A \sqsupset_{\text{old}} B$ and A is easier to compute), for which the class of optimal statistics is narrower. We can repeat this procedure until we get a final criterion for which there is only one optimal statistic.

It is reasonable to require that the relation $A \geq B$ should not change after rescaling.

Definition 1. *Let $N > 0$ be an integer. By a N-sample statistic, we mean a function $L : (R^{N+1})_{\neq} \to R$ that maps $(N+1)$-tuples of pairwise different real numbers $t_0, X_1, \ldots, X_N$ into a real number, and that is invariant under all possible permutations of $X_1, \ldots, X_N$.*

Definition 2. By an optimality criterion, we mean a transitive relation $\geq$ on the set of all N-sample statistics. We say that a criterion is final if there exists one and only one optimal statistic, i.e., a statistic L_{opt} for which $\forall L\,(L_{\text{opt}} \geq L)$.

Definition 3. By a rescaling, we mean a strictly increasing 1-1 function $f : R \to R$. For every statistic L and for every rescaling f, the result $f(L)$ of applying f to L is defined as follows:

$$(f(L))_t(X_1, \ldots, X_n) \stackrel{\text{def}}{=} L_{f(t)}(f(X_1, \ldots, f(X_N))).$$

Definition 4. *We say that a criterion $\geq$ is scale-invariant if for every rescaling f and for every two statistics L and L', $L \geq L'$ implies $f(L) \geq f(L')$.*

Proposition 1. Let $\geq$ be a final scale-invariant optimality criterion. Then, the corresponding optimal statistic L_{opt} is scale-invariant.

Proof. This proof is similar to the ones given by Nguyen and Kreinovich (1997). Indeed, let f be an arbitrary rescaling. Since L_{opt} is optimal, i.e., better than every other statistic, we conclude that for every other statistic L, we have $L_{\text{opt}} \geq f^{-1}(L)$ (where f^{-1} means the inverse rescaling

transformation). Since the optimality criterion $\geq$ is invariant, we conclude that $f(L_{\text{opt}}) \geq f(f^{-1}(L)) = L$, i.e., $f(L_{\text{opt}}) \geq L$.

Since this is true for every statistic L, the statistic $f(L_{\text{opt}})$ is also optimal. But since our criterion is final, there is only one optimal statistic and therefore, $f(L_{\text{opt}}) = L_{\text{opt}}$. In other words, the optimal statistic is indeed scale-invariant. The proposition is proven.

Appendix 2: C program for the Cauchy deviates method for small *N*

```
// This program estimates the bounds for Cauchy estimates for given values N and alpha

include <stdio.h>
include <stdlib.h>
include <time.h>
include <math.h>

define N 20             //number of iterations N to estimate the accuracy
define alpha 0.05       //confidence level
define M 10000          //number of iterations
define PI 3.14159265358979323846264338327795

double simulate(void)
double deltay[N+1];
double r;               // uniform random variable on [0,1]
double deltaminus;      // lower bound for the desired delta
double deltaplus;       // upper bound for the desired delta
double delta;           // desired value of delta
double sum;             // auxiliary sum (should equal to N/2 in the equation for delta)
int i;                  // auxiliary variable used for loops

// simulating Cauchy distributed values
for (i = 1; i <= N; i++)

// simulated uniform distribution on [0,1]
r = (double) rand() / (double) RANDMAX

// transform uniform deviate into Cauchy deviate
detay[i] = tan(PI * (r – 0.5));

// upper bound for delta is set at max of deltay[i]
deltaplus = 0.0;
for (i=1; i <= N; i++)
if (deltaplus < fabs(deltay[i])) deltaplus = fabs(deltay[i]);

// lower bound for delta is set at 0
deltaminus = 0.0;

// finding delta by bisection; we stop when the lower bound
// for delta is within 5while (deltaminus < 0.95 * deltaplus)
// compute the midpoint of the interval [deltaminus,deltaplus]
delta = (deltaminus + deltaplus) / 2;

// compute the value of the sum at this midpoint
sum = 0.0;
```

```
for (i = 1; i <= N; i++) sum = sum + (delta*delta)/(delta*delta + deltay[i]*deltay[i]);

// depending on whether this sum is > N/2 or < N/2 conclude
// that delta belongs to the corresponding half-interval
if (sum > (N / 2)) deltaplus = delta;
else deltaminus = delta;
return delta;

int main(void)
double delta[M + 1];   // simulation results
int i,j;               // auxiliary variables for loops
double smallest;       // auxiliary variables for sorting
int location; //
double temp;           // temporary variable for a swap
int index = (int) (alpha * (double) M); //computing alpha * M

// seeding the random number generator
srand( (unsigned)time( NULL ) );

//generating the simulated values delta
for (i = 1; i <= M; i++) delta[i] = simulate();

//sorting the simulated values delta
for (i = 1; i <= index; i++)

//find smallest values among i, i + 1, ...
smallest = delta[i];
location = i;
for (j = i; j <= M; j++)
if (delta[j] < smallest)  smallest = delta[j];
location = j;

// placing the smallest value in ith place by swapping the current
// value delta[i] with the place where this smallest value is located
temp = delta[i];
delta[i] = smallest;
delta[location] = temp;

printf("index is:
printf("k is:
return 0;
```

Appendix 3: Characteristic function for combination of Cauchy variates

The integrated expression from (10) has the form

$$\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma}\cdot\exp(-Q),$$

where

$$Q\stackrel{\text{def}}{=}\frac{(\Delta-\mu)^2}{2\sigma^2}+|\,\omega\,|\cdot\Delta=\frac{1}{2\sigma^2}\cdot Q',$$

and

$$Q'\stackrel{\text{def}}{=}(\Delta-\mu)^2+2\cdot\sigma^2\cdot|\,\omega\,|\cdot\Delta=\Delta^2-2\Delta(\mu-\sigma^2\cdot|\,\omega\,|)+\mu^2.$$

Thus, by separating the full square, we conclude that

$$Q'=(\Delta-\mu+\sigma^2\cdot|\,\omega\,|)^2-(\mu-\sigma^2\cdot|\,\omega\,|)^2+\mu^2=(\Delta-\mu+\sigma^2\cdot|\,\omega\,|)^2-\sigma^4\cdot\omega^2+2\mu\cdot\sigma^2\cdot|\,\omega\,|.$$

Therefore,

$$Q=\frac{1}{2\sigma^2}\cdot Q'=\frac{1}{2\sigma^2}\cdot(\Delta-\mu+\sigma^2\cdot|\,\omega\,|)^2-\frac{1}{2}\cdot\sigma^2\cdot\omega^2+\mu\cdot|\,\omega\,|.$$

Hence, the integrated expression is equal to

$$\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma}\cdot\exp(-Q)=\exp\left(\frac{1}{2}\cdot\sigma^2\cdot\omega^2-\mu\cdot|\,\omega\,|\right)\cdot\left[\frac{1}{2\sigma^2}\cdot\exp\left(-\frac{1}{2\sigma^2}\cdot(\Delta-\mu+\sigma^2\cdot|\,\omega\,|)^2\right)\right].$$

The expression before the large parenthesis does not depend on Δ at all. The expression inside the large square brackets is the probability density function for a Gaussian distribution with mean $\mu-\sigma^2\cdot|\,\omega\,|$ and standard deviation σ, so the integral of this expression is exactly 1. Thus, the integral (1) takes the desired form

$$\exp\left(\frac{1}{2}\cdot\sigma^2\cdot\omega^2-\mu\cdot|\,\omega\,|\right).$$

References

Adams, E., and U. Kulisch. 1993. *Scientific Computing with Automatic Result Verification.* Academic Press, Boston.

Alefeld, G., and J. Herzberger. 1983. *Introduction to Interval Computations.* Academic Press, New York.

Barmish, B.R., and C.M. Lagoa 1996. The uniformity principle: a new tool for probabilistic robustness control. *Proceedings of the 30th Annual Conference on Information Science and Systems*, Princeton.

Beck, J. 2004. <<>>. Dissertation, University of Texas at El Paso.

Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Berleant, D. 1993. Automatically verified reasoning with both intervals and probability density functions. *Interval Computations* 1993(no. 2): 48-70.

Berleant, D. 1996. Automatically verified arithmetic on probability distributions and intervals. Pages 227-244 in *Applications of Interval Computations*, edited by B. Kearfott and V. Kreinovich, Kluwer Academic Publishers.

Bernardini, A., and F. Tonon 2008a. Extreme probability distributions of random/fuzzy sets and p-boxes. Pages 61-81 in *Proceedings of the 3rd International Workshop on Reliable Engineering Computing, NSF Workshop on Imprecise Probability in Engineering Analysis and Design*, February 20-22, 2008, Savannah, Georgia, USA, R.L. Muhanna and R.L. Mullen (eds.), http://www.gtsav.gatech.edu/workshop/rec08/documents/REC08_Paper_Bernardini.pdf.

Bernardini, A., and F. Tonon 2008b. "Extreme probability distributions of random/fuzzy sets and p-boxes". PowerPoint presentation, available on-line at http://www.gtsav.gatech.edu/workshop/rec08/documents/REC2008Bernardini-Tonon.ppt.

Breth, M. 1978. Bayesian confidence bands for a distribution function. *Annals of Statistics* 6: 649-657.

Bruns, M., C.J.J. Paredis, and S. Ferson. 2006a. Computational methods for decision making based on imprecise information. Pages 341-368 in *Proceedings of the NSF Workshop on Reliable Engineering Computing: Modeling Errors and Uncertainty in Engineering Computations*, February 22-24, 2006, Savannah, Georgia, USA, R.L. Muhanna and R.L. Mullen (eds.), http://www.gtrep.gatech.edu/workshop/rec06/papers/Burns_paper.pdf [sic].

Bruns, M., C.J.J. Paredis, and S. Ferson. 2006b. "Computational methods for decision making based on imprecise information". PowerPoint presentation, available on-line at http://www.gtrep.gatech.edu/workshop/rec06/presentations/REC2006_Bruns.ppt.

Cheng, R.C.H., and T.C. Iles. 1983. Confidence bands for cumulative distribution functions of continuous random variables. *Technometrics* 25: 77-86.

Cheng, R.C.H., and T.C. Iles. 1988. One-sided confidence bands for cumulative distribution functions. *Technometrics* 30: 155-159.

Cheng, R.C.H., B.E. Evans and J.E. Williams. 1988. Confidence band estimations for distributions used in probabilistic design. *International Journal of Mechanical Sciences* 30: 835-845.

Cooper, J.A., S. Ferson and L.R. Ginzburg. 1996. Hybrid processing of stochastic and subjective uncertainty data. *Risk Analysis* 16: 785–791.

Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment*. Plenum Press.

Dixon, W.J., and F. J. Massey, Jr. 1969. *Introduction to Statistical Analysis*, McGraw-Hill, New York.

Downing, D.J., R.H. Gardner and F.O. Hoffman. 1985. Response surface methodologies for uncertainty analysis in assessment models. *Technometrics* 27: 151-163.

Dwyer, P. 1951. *Linear Computations*. John Wiley, New York.

EPA [U. S. Environmental Protection Agency] 2001. Risk Assessment Guidance for Superfund (RAGS), Volume III - Part A: Process for Conducting Probabilistic Risk Assessment. EPA 540-R-02-002, Office of Emergency and Remedial Response, U.S.EPA, Washington, D.C. http://www.epa.gov/superfund/programs/risk/rags3a/index.htm.

EPA [U.S. Environmental Protection Agency]. 2002-2005. "EPA NE: GE/Housatonic River Site - Rest of River – Reports". http://www.epa.gov/region1/ge/thesite/restofriver-reports.html.

Ferson, S. 1994. Naive Monte Carlo methods yield dangerous underestimates of tail probabilities. *Proceedings of the High Consequence Operations Safety Symposium*, Sandia National Laboratories, SAND94–2364, J.A. Cooper (ed.), pp. 507–514.

Ferson, S. 1996. Automated quality assurance checks on model structure in ecological risk assessments. *Human and Environmental Risk Assessment* 2:558-569.
Ferson, S. 2002. RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers. Lewis Publishers, Boca Raton, Florida.
Ferson, S. and M.A. Burgman. 1995. Correlations, dependency bounds and extinction risks. *Biological Conservation* 73:101-105.
Ferson, S. and L. Ginzburg. 1995. Hybrid arithmetic. *Proceedings of the 1995 Joint ISUMA/NAFIPS Symposium on Uncertainty Modeling and Analysis*, IEEE Computer Society Press, Los Alamitos, California, pp. 619–623.
Ferson, S., and J.G. Hajagos. 2004. Arithmetic with uncertain numbers: rigorous and (often) best-possible answers. *Reliability and Engineering and System Safety* 85: 135-152.
Ferson, S., and J.G. Hajagos. 2006. Varying correlation coefficients can underestimate uncertainty in probabilistic models. *Reliability Engineering and System Safety* 91: 1461-1467.
Ferson, S.. and W.T. Tucker. 2006. Sensitivity analysis using probability bounding. *Reliability Engineering and System Safety* 91: 1435-1442.
Ferson, S., V. Kreinovich, L. Ginzburg, K. Sentz and D.S. Myers. 2003. *Constructing Probability Boxes and Dempster-Shafer Structures*. Sandia National Laboratories, SAND2002-4015, Albuquerque, New Mexico. http://www.ramas.com/unabridged.zip
Ferson, S., R.B. Nelsen, J. Hajagos, D.J. Berleant, J. Zhang, W.T. Tucker, L.R. Ginzburg and W.L. Oberkampf. 2004. *Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis*. Sandia National Laboratories, SAND2004-3072, Albuquerque, New Mexico. http://www.ramas.com/depend.zip
Ferson, S., V. Kreinovich, J. Hajagos, W. Oberkampf and L. Ginzburg. 2007. *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty.* SAND2007-0939, Sandia National Laboratories, Albuquerque, NM. Available on-line at http://www.ramas.com/intstats.pdf.
Frey, H.C. and S.R. Patil. 2002. Identification and review of sensitivity analysis methods. *Risk Analysis* 22: 553-578. Available on-line at http://www.ce.ncsu.ed/risk/pdf/frey.pdf.
Fu, M.C. 1994. Optimization via simulation: a review. *Annals of Operations Research* 53: 199-247.
Gibbons, J.D. 1990. Nonparametric statistics, pages 11.1–11.26 in *Handbook of Statistical Methods for Engineers and Scientists*, H. M. Wadsworth (ed.), McGraw-Hill, New York.
Helton, J.C. 1993. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering and System Safety* 42: 327-367.
Helton, J.C. 1994. Treatment of uncertainty in performance assessments for complex systems. *Risk Analysis* 14(4): 483-511.
Helton, J.C. 1997. Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. *Journal of Statistical Computation and Simulation* 57: 3-76.
Helton, J.C., and F.J. Davis 2000a. *Sampling-based Methods for Uncertainty and Sensitivity Analysis.* Sandia National Laboratories, SAND99-2240, Albuquerque, New Mexico.
Helton, J.C., and F.J. Davis 2000b. Sampling-based methods. Pages 101-153 in *Sensitivity Analysis*, edited by A. Saltelli, K. Chan, and E.M. Scott, John Wiley, New York.
Helton, J.C., and F.J. Davis 2002. *Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems*. SAND2001-0417, Sandia National Laboratories, Albuquerque, New Mexico,.
Helton, J.C., and F.J. Davis 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety* 81(1): 23-69.
Helton, J.C., J.D. Johnson, M.D. McKay, A.W. Shiver and J.L. Sprung. 1995. Robustness of an uncertainty and sensitivity analysis of early exposure results with the MACCS reactor accident consequence model. *Reliability Engineering and System Safety* 48: 129-148.
Helton, J.C., W. Beyeler and S.C. Hora. 1997a. Conceptual basis of a systems prioritization methodology for the Waste Isolation Pilot Plant. *Reliability Engineering and System Safety* 57: 203-222.
Helton, J.C., D.R. Anderson, B.L. Baker, J.E. Bean, J.W Berglund, W. Beyeler, R. Baline, K. Economy, J.W. Garner, S.C. Hora, R.C. Lincoln, M.G. Marietta, F.T. Mendenhall, N.H. Prindle, D.K. Rudeen, J.D. Schreiber, A.W. Shiver, L.N. Smith, P.N. Swift and P. Vaughn. 1997b. Computational implementation of a systems prioritization methodology for the Waste Isolation Pilot Plant: a preliminary example. *Reliability Engineering and System Safety* 57: 223-266.

Helton, J.C., J.D. Johnson, W.L. Oberkampf. 2004a. Probability of Loss of Assured Safety in Temperature Dependent Systems with Multiple Weak and Strong Links. SAND2004-5216, Sandia National Laboratories.

Helton, J.C., J.D. Johnson, W.L. Oberkampf. 2004b. An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering and System Safety* 85: 39-72.

Helton, J.C., F.J. Davis, J.D. Johnson. 2005a. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability Engineering and System Safety* 89: 305-330.

Helton, J.C., W.L. Oberkampf, J.D. Johnson. 2005b. Competing failure risk analysis using evidence theory. *Risk Analysis* 25: 973-995.

Helton, J.C., J.D. Johnson, W.L. Oberkampf. 2006a. Probability of loss of assured safety in temperature dependent systems with multiple weak and strong links. *Reliability Engineering and System Safety* 91:320-348.

Helton, J.C., J.D. Johnson, C.J. Sallaberry and C.B. Storlie. 2006b. *Survey of Sampling-based Methods for Uncertainty and Sensitivity Analysis*. Sandia National Laboratories, SAND2006-2901, Albuquerque, New Mexico. Also in press in *Reliability Engineering and System Safety*.

Helton, J.C., J.D. Johnson, W.L. Oberkampf, and C.B. Storlie. 2006c. A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory. Submitted to *Computer Methods in Applied Mechanics and Engineering*.

Hettmansperger, T., and M. Keenan. 1980. Tailweight, statistical inference, and families of distributions–a brief survey. *Statistical Distributions in Scientific Work*, Volume 1, G.P. Patil et al. (eds.), Kluwer, Boston.

Hoffman, F. O., and J.S. Hammonds. 1994. Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis* 14(5): 707-712.

Hosmer, D.W., and S. Lemeshow. 1989. *Applied Logistic Regression*. John Wiley, New York.

Iman, R. L. 1987. A matrix-based approach to uncertainty and sensitivity analysis for fault trees. *Risk Analysis* 7: 21-33.

Iman, R.L., and W.J. Conover. 1980. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. *Communications in Statistics* A9:1749-1842.

Iman, R.L., and W.J. Conover. 1982. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics B11*:311-334.

Iman, R.L., and J.C. Helton. 1985. A Comparison of Uncertainty and Sensitivity Analysis Techniques for Computer Models. NUREG/CR-3904. National Technical Information Service, Springfield, Virginia.

Iman, R.L., and J.C. Helton. 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* 8: 71-90.

Iman, R.L., and J.C. Helton. 1991. The repeatability of uncertainty and sensitivity analyses for complex probabilistic risk assessments. *Risk Analysis* 11: 591-606.

Iman, R.L., and S.C. Hora. 1990. A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Analysis* 10: 401-406.

Iman, R.L., J.C. Helton and J.E. Campbell. 1981a. An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. *Journal of Quality Technology* 13:174-183.

Iman, R.L., J.C. Helton and J.E. Campbell. 1981b. An approach to sensitivity analysis of computer models, Part 2. Ranking of input variables, response surface validation, distribution effect and technique synopsis. *Journal of Quality Technology* 13:232-240.

Insua, D.R., and Ruggeri, F. (eds.) 2000. *Robust Bayesian Analysis*. Lecture Notes in Statistics, Volume 152. Springer-Verlag, New York.

Jaulin, L., M. Kieffer, O. Didrit and E. Walter. 2001. *Applied Interval Analysis*. Springer, London.

Jaynes, E.T. 2003. *Probability Theory*, edited by L. Bretthorst. Cambridge University Press.

Jeng, S.-L., and W.Q. Meeker 2001. Parametric simultaneous confidence bands for cumulative distributions from censored data. *Technometrics* 43: 450-461.

Kaplan, S. 1981. On the method of discrete probability distributions in risk and reliability calculations—applications to seismic risk assessment. *Risk Analysis* 1: 189-196.

Kanofsky, P., and R. Srinivasan. 1972. An approach to the construction of parametric confidence bands on cumulative distribution functions. *Biometrika* 59: 623-631.
Karlin, S. and W.J. Studden. 1966. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Interscience Publishers, John Wiley and Sons, New York.
Klir, G. and B. Yuan. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, Upper Saddle River, NJ.
Klir, G.J. 2006. *Uncertainty and Information: Foundations of Generalized Information Theory*. Wiley-Interscience, Hoboken, New Jersey.
Konijn, H.S. 1987. Distribution-free and other prediction intervals, *The American Statistician* 41: 11–15.
Kriegler, E. and H. Held. 2004. Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning* 39: 185-209.
Kreinovich, V., and S.A. Ferson [sic] 2004. A new Cauchy-based black-box technique for uncertainty in risk analysis. *Reliability Engineering and Systems Safety* 85: 267–279.
Kreinovich, V., A. Lakeyev, J. Rohn, and P. Kahl. 1997. *Computation Complexity and Feasibility of Data processing and Interval Computations*. Kluwer, Dordrecht.
Kreinovich, V., J. Hajagos, W. Troy Tucker, L. Ginzburg and S. Ferson. 2008. *Propagating uncertainty through a quadratic response surface model*.
Lagoa, C.M., P.S. Shcherbakov and B.R. Barmish 1998. Probabilistic ehhancement of classical robustness margins: the unirectangularity concept. *Systems and Control Letters* 35: 32-43.
Lagoa, C.M. and B.R. Barmish <<year>>. Monte Carlo simulation without probability distributions: a tutorial. [<<unpublished manuscript>>].
Lavine, M. 1991. Sensitivity in Bayesian statistics, the prior and the likelihood. *Journal of the American Statistical Association* 86 (414): 396-399.
Leamer, E.E. 1990. Sensitivity analysis would help. *Modelling Economic Series*, edited by C.W.J. Granger, Clarendon Press, Oxford.
Li, G., C. Rosenthal and H. Rabitz. 2003. High-dimensional model representations. *Journal of Physical Chemistry* 105(33):7765-7777.
Manes, E.G. 1982. A class of fuzzy theories. Journal of Mathematical Analysis and Its Applications 85.
MacDonald, D.D. et al. 2002. *Calcasieu Estuary Remedial Investigation/Feasibility Study (RI/FS): Baseline Ecological Risk Assessment (BERA)*. Report to U.S. Environmental Protection Agency, Region 6, Dallas, Texas. Document control number 3282-941-RTZ-RISKZ-14858. Available on-line at http://www.epa.gov/earth1r6/6sf/sfsites/calcinit.htm.
McCarthy, M.A., M.A. Burgman and S. Ferson. 1995. The use of logistic regression in sensitivity analysis of population viability models. *Biological Conservation* 73:93-100.
Montgomery, V., F. Coolen and A. Hart 2008. Bayesian probability boxes in risk assessment. *Journal of Statistical Theory and Practice* [to appear].
Mosleh, A., and V.M. Bier. 1996. Uncertainty and Probability: A Reconciliation with the Subjectivist Viewpoint, *IEEE Trans. on Systems, Man and Cybernetics, Part A; Systems and Humans* 26:303-310.
Moore, R.E. 1966. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.
Morgan, M.G., and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, England.
Morton, R.H. 1983. Response surface methodology. *Mathematical Scientist* 8: 31-52.
Myers, R.H. 1971. *Response Surface Methodology*. Allyn and Bacon, Boston.
Myers, R.H. 1999. Response surface methodology—current status and future directions. *Journal of Quality Technology* 31: 30-44.
Myers, R.H., D.C. Montgomery, G.G. Vining, C.M. Borror and S.M. Kowalski. 2004. Response surface methodology: a retrospective and literature review. *Journal of Quality Technology* 36: 53-77.
Nelsen, R.B. 1999. *An Introduction to Copulas*. Lecture Notes in Statistics 139, Springer-Verlag, New York.
Neumaier, A. 1990. Interval Methods for Systems of Equations. Cambridge University Press, Cambridge.
Nguyen, H.T., and V. Kreinovich 1997. *Applications of Continuous Mathematics to Computer Science*, Kluwer, Dordrecht.
Oakley, J. and A. O'Hagan. 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society Series* B 66: 751-769.
Oberkampf, W.L., J.C. Helton. 2005. Evidence theory for engineering applications. In *Engineering Design Reliability Handbook*, E. Nikolaidis, D.M. Ghiocel, S. Singhal (eds.), CRC Press, New York.

Owen, A.B. 1995. Nonparametric likelihood confidence bands for a distribution function. *Journal of the American Statistical Association* 90: 516-521.
Paté-Cornell, M.E. 1996. Uncertainties in risk analysis: six levels of treatment. *Reliability Engineering and System Safety* 54: 99-111.
R Development Core Team. (2004). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.
Rabitz, H. an O.F. Alis. 1999. General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry* 25(2-3): 197-233.
Rekuc, S.J., J.M. Aughenbaugh, M. Bruns, and C.J.J. Paredis. 2006. Eliminating design alternative based on imprecise information. *Proceedings of the Society of Automotive Engineering World Congress* 2006-01-0272.
Rosenkrantz, W.A. 2000. Confidence bands for quantile functions: a parametric and graphic alternative for testing goodness of fit. *The American Statistician* 54: 185-190.
Ross, T., J. Lucero, K. Hench, and K. Kline. 2005. "Expressions of total uncertainty in engineering problems involving sparse and uncertain information". Unpublished slide presentation.
Rubinstein, R., and A. Shapiro. 1993. Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization via the Score Function Method, Wiley, Chichester.
Rubio, E., J.W. Hall and M.G. Anderson. 2004. Uncertainty analysis in a slope hydrology and stability model using probabilistic and imprecise information. *Computers and Geotechnics* 31: 529-536.
Saltelli, A. 2002. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications* 154: 280-297.
Saltelli, A. 2003. Global sensitivity analysis: novel settings and methods. Pages 71ff in Proceedings of the International Workshop on Uncertainty, Sensitivity, and Parameter Estimation for Multimedia Environmental Modeling, T.J. Nicholson, J.E. Babendreier, P.D. Meyer, S. Mohanty, B.B. Hicks and G.H. Leavesley (eds.), Federal Interagency Steering Committee on Multimedia Environmental Models (ISCMEM), U.S. Nuclear Regulatory Commission, NUREG/CP-0187, Washington, DC, http://iscmem.org/ Documents/Proceedi_02.pdf.
Saltelli, A., K. Chan and E.M. Scott (eds.). 2000. *Sensitivity Analysis*. John Wiley, New York.
Saltelli, A., S. Tarantola, F. Compolongo and M. Ratto. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley, New York.
Saw, J.G., M.C.K. Yang, and T.C. Mo 1984. "Chebyshev inequality with estimated mean and variance", *The American Statistician* 38: 130–132. [correction in 1988 *The American Statistician* 42: 166]
Sheskin, D.J. 2004. *Handbook of parametric and nonparametric statistical procedures*, Chapman & Hall/CRC, Boca Raton, Florida.
Sobol', I.M. 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment* 1(4): 407-414.
Springer, M.D. 1979. *The Algebra of Random Variables*. Wiley, New York.
Stuart, A., and J.K. Ord 1987. *Kendall's Advanced Theory of Statistics*, Oxford University Press, New York.
Suppes, P., D. H. Krantz, R.D. Luce, and A. Tversky 1971. *Foundations of measurement,* Volume I. Academic Press, San Diego.
Suppes, P., D. H. Krantz, R.D. Luce, and A. Tversky 1989. *Foundations of measurement,* Volume II. Academic Press, San Diego.
Suppes, P., D. H. Krantz, R.D. Luce, and A. Tversky 1989. *Foundations of measurement,* Volume III. Academic Press, San Diego.
Swartzman, G.L., and S.P. Kaluzny. 1987. *Ecological Simulation Primer.* Macmillan, New York.
Tonon, F., A. Bernadini and I. Elishakoff. 1999. Concept of random sets as applied to the design of structures and analysis of expert opinions for aircraft crash. *Chaos, Solitons and Fractals* 10: 1855-1868.
Tonon, F., A. Bernadini and A. Mammino. 2000a. Determination of parameters range in rock engineering by means of Random Set Theory. *Reliability Engineering and System Safety* 70: 241-261.
Tonon, F., A. Bernadini and A. Mammino. 2000b. Reliability analysis of rock mass respsonse by means of Random Set Theory. *Reliability Engineering and System Safety* 70: 263-282.
Trejo, R., and V. Kreinovich. 2001. Error estimations for indirect measurements: randomized vs. deterministic algorithms for 'black-box' programs. Pages 673-729 in *Handbook on Randomized Computing*, S. Rajasekaran, P. Pardalos, J. Reif, and J. Rolim (eds.), Kluwer, Dordrecht.

Wadsworth, H.M. (ed.) 1990. *Handbook of Statistical Methods for Engineers and Scientists*, McGraw-Hill, New York.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

Williamson, R.C., and T. Downs. 1990. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* 4: 89-158.

Woo, G. 1991. A quitting rule for Monte Carlo simulation of extreme risks. *Reliability Engineering and System Safety* 31: 179–189.

Yager, R.R. 1986. Arithmetic and other operations on Dempster-Shafer structures. *International Journal of Man-machine Studies* 25: 357-366.

Young, R.C. 1931. The algebra of multi-valued quantities. *Mathematishe Annalen* 104: 260-290. Available on-line at http://www.cs.utep.edu/interval-comp/young.pdf.

Young, D.M., J. W. Seaman Jr., D. W. Turner, and V. R. Marco 1988. Probability inequalities for continuous unimodal random variables with finite support. *Communications in Statistics: Theory and Methods* 17: 3505–3519.

External Distribution

M. A. Adams
Jet Propulsion Laboratory
4800 Oak Grove Drive, MS 97
Pasadena, CA 91109

Bilal Ayyub
Department of Civil Engineering
University of Maryland
College Park, MD 20742-3021

Ivo Babuska
TICAM
Mail Code C0200
University of Texas at Austin
Austin, TX 78712-1085

S. Balachandar
Mechanical and Aerospace Engineering
University of Florida
231 MAE-A, PO Box 116250
Gainesville, FL 32611-6205

Osman Balci
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061

Timothy Barry
Science Policy, Planning and Evaluation
USEPA Headquarters, MC2174
401 M Street, S.W.
Washington, DC 20460

John Benek
AFRL/VAAC
2210 Eighth St.
Wright-Patterson AFB, OH 45433

James Berger
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Vicki M. Bier
Department of Industrial Engineering
University of Wisconsin
Madison, WI 53706

Lynne Billard
Department of Statistics
University of Georgia
Athens, GA 30602-1952

Mark Brandyberry
Computational Science and Engineering
2264 Digital Computer Lab, MC-278
1304 West Springfield Ave.
University of Illinois
Urbana, IL 61801

John A. Cafeo
General Motors R&D Center
Mail Code 480-106-256
30500 Mound Road
Box 9055
Warren, MI 48090-9055

Andrew Cary
The Boeing Company
MC S106-7126
P.O. Box 516
St. Louis, MO 63166-0516

James C. Cavendish
General Motors R&D Center
Mail Code 480-106-359
30500 Mound Road
Box 9055
Warren, MI 48090-9055

Chun-Hung Chen
Department of Systems Engineering &
Operations Research
George Mason University
4400 University Drive, MS 4A6
Fairfax, VA 22030

Wei Chen
Department of Mechanical Engineering
Northwestern University
2145 Sheridan Road, Tech B224
Evanston, IL 60208-3111

Kyeongjae Cho
Dept. of Mechanical Engineering
MC 4040
Stanford University
Stanford, CA 94305-4040

Hugh Coleman
Department of Mechanical &
Aero. Engineering
University of Alabama/Huntsville
Huntsville, AL 35899

Raymond Cosner (2)
Boeing-Phantom Works
MC S106-7126
P. O. Box 516
St. Louis, MO 63166-0516

Thomas A. Cruse
AFRL Chief Technologist
1981 Monahan Way
Bldg. 12, Room 107
Wright-Patterson AFB, OH 45433-7132

Department of Energy (6)
Attn: Christopher Deeney, NA-113
Dimitri Kusnezov, NA-114
Njema Frazier, NA-114
Karen Pao, NA-114
Kevin Greenaugh, NA-115
Jamlieh Soudah, NA-116
Forrestal Building
1000 Independence Ave., SW
Washington, DC 20585

David Dolling
Department of Aerospace Engineering
& Engineering Mechanics
University of Texas at Austin
Austin, TX 78712-1085

Isaac Elishakoff
Dept. of Mechanical Engineering
Florida Atlantic University
777 Glades Road
Boca Raton, FL 33431-0991

Ashley Emery
Dept. of Mechanical Engineering
Box 352600
University of Washingtion
Seattle, WA 98195-2600

Scott Ferson (20)
Applied Biomathematics
100 North Country Road
Setauket, New York 11733-1345

Marc Garbey
Dept. of Computer Science
Univ. of Houston
501 Philipp G. Hoffman Hall
Houston, Texas 77204-3010

Roger Ghanem
254C Kaprielian Hall
Dept. of Civil Engineering
3620 S. Vermont Ave.
University of Southern California
Los Angles, CA 90089-2531

Michael Ginevan
307 Hamilton Avenue
Silver Spring, MD 20901

James Glimm
Dept. of Applied Math & Statistics
P138A
State University of New York
Stony Brook, NY 11794-3600

Melody S. Goodman
Preventive Medicine
Health Sciences Center, 3-071
Stony Brook, NY 11794-8338

James Gran
SRI International
Poulter Laboratory AH253
333 Ravenswood Avenue
Menlo Park, CA 94025

Ramana Grandhi
Dept. of Mechanical and Materials Engineering
3640 Colonel Glenn Hwy.
Dayton, OH 45435-0001

John W. Green
DuPont Stine-Haskell Research Center
1090 Elkton Road
P.O. Box 50
BLDG H-1 / 423B
Newark, DE 19714-0050

Bernard Grossman
The National Institute of Aerospace
100 Exploration Way
Hampton, VA 23666

Raphael Haftka
Dept. of Aerospace and Mechanical
Engineering and Engr. Science
P. O. Box 116250
University of Florida
Gainesville, FL 32611-6250

Yacov Haimes
Systems and Information Engineering
University of Virginia
P.O. Box 400747
151 Engineer's Way
Charlottesville, VA 22904

Achintya Haldar
Dept. of Civil Engineering
& Engineering Mechanics
University of Arizona
Tucson, AZ 85721

Tim Hasselman
ACTA
2790 Skypark Dr., Suite 310
Torrance, CA 90505-5345

Dale Hattis
CENTED
Clark University
950 Main Street
Worcester, MA 01610

George Hazelrigg
Design, Manufacturing & Innovation
Room 508N
4201 Wilson Blvd.
Arlington, VA 22230

Richard Hills
Mechanical Engineering Dept.
New Mexico State University
P. O. Box 30001/Dept. 3450
Las Cruces, NM 88003-8001

F. Owen Hoffman
SENES
102 Donner Drive
Oak Ridge, TN 37830

Luc Huyse
Reliability & Materials Integrity, B-128
Southwest Research Institute
P.O. Drawer 28510
San Antonio, TX 78228-0510

Gianluca Iaccarino
Center for Turbulence Research
Building 500, Room 500A
488 Escondido Mall,
Stanford University
Stanford, CA 94305-3035

Mark Inlow
Department of Mathematics
Rose-Hulman Institute of Technology
5500 Wabash Avenue
Terre Haute, IN 47803

George Ivy
Northrop Grumman Information Technology
222 West Sixth St.
P.O. Box 471
San Pedro, CA 90733-0471

Mark Jablonowski
Post Office Box 521
Tariffville, CT 06081

Leo Kadanoff
Research Institutes Building
University of Chicago
5640 South Ellis Ave.
Chicago, IL 60637

George Karniadakis
Division of Applied Mathematics
Brown University
192 George St., Box F
Providence, RI 02912

Alan Karr
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Vladik Kreinovich (5)
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968

Bryce Landenberger
Health & Environmental Research
The Dow Company
1803 Building
Midland, MI 48674

John Lehner
Energy Sciences and Technology
Building 475, Mailstop 475C
Brookhaven National Laboratory
Upton, NY 11973-5000

W. K. Liu
Northwestern University
Dept. of Mechanical Engineering
2145 Sheridan Road
Evanston, IL 60108-3111

Robert Lust
General Motors, R&D and Planning
MC 480-106-256
30500 Mound Road
Warren, MI 48090-9055

Sankaran Mahadevan
Dept. of Civil &
Environmental Engineering
Vanderbilt University
Box 6077, Station B
Nashville, TN 37235

Hans Mair
Institute for Defense Analysis
Operational Evaluation Division
4850 Mark Center Drive
Alexandria VA 22311-1882

Richard McGrath
Weston Solutions
10 Lyman Street
Pittsfield, MA 01201

Gregory McRae
Dept. of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

John G. Michopoulos
Naval Research Laboratory,
Special Projects Group, Code 6303
Computational Multiphysics Systems Lab
Washington DC 20375, USA

Sue Minkoff
Dept. of Mathematics and Statistics
University of Maryland
1000 Hilltop Circle
Baltimore, MD 21250

Granger Morgan
Carnegie-Mellon University
Dept. of Engineering & Public Policy
Pittsburgh, PA 15213

Max Morris
Department of Statistics
Iowa State University
304A Snedecor-Hall
Ames, IW 50011-1210

Zissimos P. Mourelatos
Dept. of Mechanical Engineering
School of Engr. and Computer Science
Rochester, MI 48309-4478

Rafi Muhanna
Regional Engineering Program
Georgia Tech
210 Technology Circle
Savannah, GA 31407-3039

NASA/Ames Research Center (2)
Attn: Unmeel Mehta, MS 229-3
David Thompson, MS 269-1
Moffett Field, CA 94035-1000

NASA/Glenn Research Center (2)
Attn: John Slater, MS 86-7
Chris Steffen, MS 5-11
21000 Brookpark Road
Cleveland, OH 44135

NASA/Langley Research Center (9)
Attn: Dick DeLoach, MS 236
Lawrence Green, MS 462
Michael Hemsch, MS 499
Jim Luckring, MS 286
Joe Morrison, MS 128
Ahmed Noor, MS 369
Sharon Padula, MS 159
Thomas Zang, MS 449
Hampton, VA 23681-0001

Efstratios Nikolaidis
MIME Dept.
4035 Nitschke Hall
University of Toledo
Toledo, OH 43606-3390

William Oberkampf
Aerospace and Ocean Engineering Dept.
215 Randolph Hall
Virginia Tech
Blacksburg, VA 24061-0203

Tinsley Oden
TICAM
Mail Code C0200
University of Texas at Austin
Austin, TX 78712-1085

Michael Ortiz
Graduate Aeronautical Laboratories
California Institute of Technology
1200 E. California Blvd./MS 105-50
Pasadena, CA 91125

Chris Paredis
School of Mechanical Engineering
Georgia Institute of Technology
813 Ferst Drive, MARC Rm. 256
Atlanta, GA 30332-0405

Chris L. Pettit
Aerospace Engineering Dept.
MS-11B
590 Holloway Rd.
Annapolis, MD 21402

Joseph Powers
Aerospace and Mechanical Engineering
University of Notre Dame
Notre Dame, IN 46556-5637

P. Radovitzky
Graduate Aeronautical Laboratories
California Institute of Technology
1200 E. California Blvd./MS 105-50
Pasadena, CA 91125
Sheldon Reaven
343-A Harriman Hall
State University of New York
Stony Brook, NY 11794-3760

J. N. Reddy
Dept. of Mechanical Engineering
Texas A&M University
ENPH Building, Room 210
College Station, TX 77843-3123

John Renaud
Dept. of Aerospace & Mechanical Engr.
University of Notre Dame
Notre Dame, IN 46556

Patrick J. Roache
1215 Apache Drive
Socorro, NM 87801

F. James Rohlf
Dept. Ecology and Evolution
State University of New York
Stony Brook, NY 11794

Tim Ross
Dept. of Civil Engineering
University of New Mexico
Albuquerque, NM 87131

Chris Roy
Aerospace and Ocean Engineering Dept.
215 Randolph Hall
Virginia Tech
Blacksburg, VA 24061-0203

Randall Ryti
Neptune and Company
1505 15th Street
Los Alamos, NM 87544

J. Sacks
Inst. of Statistics and Decision Science
Duke University
Box 90251
Durham, NC 27708-0251

Larry Sanders
Defense Threat Reduction Agency
Attn: DTRA/ASC
8725 John J. Kingman Rd
Mail Stop 6201
Fort Belvoir, VA 22060-6201

Len Schwer
Schwer Engineering & Consulting
6122 Aaron Court
Windsor, CA 95492

Teddy Seidenfeld
Department of Philosophy, BH 135J
Carnegie Mellon University
Pittsburgh, PA 15213

Mark Shephard
Rensselaer Polytechnic Institute
Scientific Computation Research Center
Troy, NY 12180-3950

Tom I-P. Shih
Dept. of Aerospace Engineering
2271 Howe Hall, Room 1200A
Iowa State University
Ames, IA 50011-2271

Don Simons
Northrop Grumman Information Tech.
222 W. Sixth St.
P.O. Box 471
San Pedro, CA 90733-0471

Munir M. Sindir
Boeing - Rocketdyne Propulsion & Power
MS GB-11
P. O. Box 7922
6633 Canoga Avenue
Canoga Park, CA 91309-7922

Bill Spencer
Civil Engineering & Geological Sciences
University of Notre Dame
Notre Dame, IN 46556-0767

G. R. Srinivasan
Org. L2-70, Bldg. 157
Lockheed Martin Space & Strategic Missiles
1111 Lockheed Martin Way
Sunnyvale, CA 94089

Eugene Stakhiv
U.S. Army Corps of Engineers
Institute for Water Resources
Casey Building
7701 Telegraph Road
Alexandria, VA 22315

Fred Stern
Professor Mechanical Engineering
Iowa Institute of Hydraulic Research
The University of Iowa
Iowa City Iowa 52242

D. E. Stevenson
Computer Science Department
Clemson University
442 Edwards Hall, Box 341906
Clemson, SC 29631-1906

Glenn Suter
US Environmental Protection Agency
Nat. Center for Environmental Assessmt
26 W Martin Luther King Dr, MC-A-130
Cincinnatti, OH 45267

Raul Tempone
School of Computational Science
400 Dirac Science Library
Florida State University
Tallahassee, FL 32306-4120

Ben Thacker
B-128
Southwest Research Institute
6220 Culebra Road
P. O. Drawer 28510
San Antonio, TX 78284-0510

Kimberly Thompson
Harvard School of Public Health
617 Huntington Avenue
Boston, MA 02115

Fulvio Tonon
Dept. of Civil Engineering
University of Texas at Austin
1 University Station C1792
Austin, TX 78712-0280

Robert W. Walters
Aerospace and Ocean Engineering
Virginia Tech
215 Randolph Hall, MS 203
Blacksburg, VA 24061-0203

Justin Y-T Wu
8540 Colonnade Center Drive, Ste 301
Raleigh, NC 27615

Ronald R. Yager
Machine Intelligence Institute
Iona College
715 North Avenue
New Rochelle, NY 10801

Ren-Jye Yang
Ford Research Laboratory
MD2115-SRL
P.O.Box 2053
Dearborn, MI 4812

Simone Youngblood
The Johns Hopkins University
Applied Physics Laboratory
11100 Johns Hopkins Road
Mail Stop 13N409
Laurel, MD 20723

Foreign Distribution

Tom Aldenberg
RIVM
PO Box 1
NL-3720 BA Bilthoven
NETHERLANDS

Michael Beer
Dept. of Civil Engineering
National University of Singapore
BLK E1A, #05-04, 1 Engineering Drive 2
SINGAPORE 117576

Yakov Ben-Haim
Department of Mechanical Engineering
Technion-Israel Institute of Technology
Haifa 32000
ISRAEL

Mark Burgman
School of Botany
University of Melbourne
Victoria, 3010
AUSTRALIA

Roger M. Cooke
Delft Institute of Applied Mathematics
Delft University of Technology
PO Box 5031, 2600 GA, Delft
Mekelweg 4, 2628 CD, Delft
THE NETHERLANDS

Frank Coolen
Department of Mathematical Sciences
Durham University
Science Laboratories
South Road
Durham DH1 3LE
UNITED KINGDOM

Gert de Cooman
Universiteit Gent
Onderzoeksgroep, SYSTeMS
Technologiepark - Zwijnaarde 9
9052 Zwijnaarde
BELGIUM

Peter Craig
Department of Mathematical Sciences
University of Durham
South Road
Durham DH1 3LE
UNITED KINGDOM

Luis Eca
Instituto Superior Tecnico
Department of Mechanical Engineering
Av. Rovisco Pais
1096 Lisboa CODEX
PORTUGAL

Dominique Guyonnet
BRGM-Environnemental Department
3 avenue Claude Guillemin
BP 6009
45060 Orléans Cedex 2
FRANCE

Andy Hart
Central Science Laboratory
Sand Hatton
York Y041 1LZ
UNITED KINGDOM

Charles Hirsch
NUMECA International
Avenue Franklin Roosevelt, No. 5
B-1050 Brussels
BELGIUM

Bulent Korkem
P.O. Box 18 Kavaklidere
06692 Ankara
TURKEY

Igor Kozin
Systems Analysis Department
Riso National Laboratory
P. O. Box 49
DK-4000 Roskilde
DENMARK

David Moens
K. U. Leuven
Mechanical Engineering, Div. PMA
Kasteelpark Arenberg 41
B – 3001 Heverlee
BELGIUM

Bernd Möller
Institut für Statik und Dynamik der
Tragwerke
Technische Universität Dresden
01062 Dresden
GERMANY

Dwayne Moore
Cantox Environmental
1550A Laperriere Avenue, Suite 103
Ottawa, Ontario
CANADA K1Z 7T2

Kung Ng
Computational Dynamics Ltd.
200 Shepherds Bush Road
London W6 7NY
UNITED KINGDOM

Nina Nikolova - Jeliazkova
Institute of Parallel Processing
Bulgarian Academy of Sciences
25a "acad. G. Bonchev" str.
Sofia 1113
BULGARIA

K. Papoulia
Inst. Seismology & Earthquake Engineering
P.O. Box 53, Finikas GR-55105
Thessaloniki
GREECE

Dominique Pelletier
Genie Mecanique
Ecole Polytechnique de Montreal
C.P. 6079, Succursale Centre-ville
Montreal, H3C 3A7
CANADA

Willem Roelofs
Central Science Laboratory
Sand Hatton
York Y041 1LZ
UNITED KINGDOM

Graham de Vahl Davis
CFD Research Laboratory
University of NSW
Sydney, NSW 2052
AUSTRALIA

Reza Vaziri
Civil Engineering/Materials Engineering
The University of British Columbia
6250 Applied Science Lane
Vancouver, B.C. V6T 1Z4
CANADA

Marco Zaffalon
IDSIA
Galleria 2, CH-6928 Manno (Lugano)
SWITZERLAND

Department of Energy Laboratories

Los Alamos National Laboratory (10)
Mail Station 5000
P.O. Box 1663
Los Alamos, NM 87545
Attn: Mark C. Anderson, MS T080
Jerry S. Brock, MS F663
Scott Doebling, MS T080
S. Eisenhawer, MS K557
Francois Hemez, MS F699
Karen Hench, MS P946
David Higdon, MS F600
James Kamm, MS D413
Mandy Rutherford, MS T080
Alyson G. Wilson, MS F600

Lawrence Livermore National Laboratory (7)
7000 East Ave.
P.O. Box 808
Livermore, CA 94551
Attn: Frank Graziani, MS L-095
Henry Hsieh, MS L-229
Richard Klein, MS L-023
Roger Logan, MS L-125
Cynthia Nitta, MS L-096
Joe Sefcik, MS L-160
Carol Woodward, MS L-561

Idaho National Laboratory (3)
2525 Fremont Ave.
P.O. Box 1625
Idaho Falls, ID 83415
Attn: Richard W. Johnson, MS 5226
Dana A. Knoll, MS 3855
Richard R. Schultz, MS 3890

Sandia Internal Distribution

1	MS 0139	1220	P. Yarrington
1	MS 0139	1221	M. Pilch
1	MS 1318	1410	D. E. Womble
1	MS 1318	1411	J. R. Stewart
1	MS 1318	1411	B. M. Adams
1	MS 1318	1411	M. S. Eldred
1	MS 1318	1411	D. M. Gay
1	MS 1318	1411	E. T. Phipps
1	MS 1318	1411	L. P. Swiler
1	MS 0370	1411	T. G. Trucano
1	MS 1318	1414	K. F. Alvin
1	MS 1318	1414	B. G. van Bloemen Waanders
1	MS 1320	1414	R. B. Lehoucq
1	MS 1320	1416	S. S. Collis
1	MS 1320	1416	M. Heroux
1	MS 0378	1431	R. M. Summers
1	MS 1322	1435	J. B. Aidun
1	MS 0825	1510	M. R. Prairie
1	MS 0847	1520	P. J. Wilson
1	MS 0821	1530	A.L. Thornton
1	MS 0380	1540	H. S. Morgan
1	MS 0828	1544	A. R. Black
1	MS 0828	1544	K. J. Dowding
1	MS 0828	1544	A. A. Giunta
1	MS 0776	1544	J. C. Helton
1	MS 0557	1544	T. L. Paez
1	MS 0828	1544	J. R. Red-Horse
1	MS 0828	1544	V. J. Romero
1	MS 0847	1544	W. R. Witkowski
1	MS 0751	6315	L. S. Costin
1	MS 0751	6315	R. M. Brannon
1	MS 0751	6315	A. F. Fossum
1	MS 0757	6442	G. D. Wyss
1	MS 0776	6783	R. J. MacKinnon
1	MS 0778	6783	P. N. Swift
1	MS 0776	6783	R. P. Rechard
1	MS 9154	8244	R. E. Oetken
1	MS 9042	8776	E. P. Chen
1	MS 9159	8962	H. R. Ammerlahn
1	MS 9159	8962	P. D. Hough
1	MS 0830	12335	K. V. Diegert
1	MS 0829	12337	J. M. Sjulin
1	MS 0829	12337	S. V. Crowder
1	MS 0829	12337	B. M. Rutherford
1	MS 0829	12337	F. W. Spencer
1	MS 0899 (electronic)	9536	Technical Library