Sandia Technical Report, SAND2007-2670C
Submitted to: Joint Statistical Meetings, July 29-Aug.2, 2007

Aleatory and Epistemic Uncertainty Quantification for Engineering Applications

L. P. Swiler*, A. A. Giunta

* Corresponding Author: US Postal Address: Sandia National Laboratories, P.O. Box 5800, Mail Stop 1318 Albuquerque, NM 87185-1318, USA; Email: lpswile@sandia.gov; Telephone: 505-844-8093.

Sandia National Laboratories, Albuquerque, NM 87185 USA

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Abstract

Most computer models for engineering applications are developed to help assess a design or regulatory requirement. As part of this task, the capability to quantify the impact of variability and uncertainty in the decision context is critical. The requirement is often stated as: the probability that some system response quantity exceeds a threshold value is less than some required probability. This presentation will provide an outline and comparison of methods that are used for analyzing and propagating aleatory and epistemic uncertainties. The methods are all available in a software tool called DAKOTA. We will specifically focus on five classes of methods: Latin Hypercube sampling, analytic reliability methods, polynomial chaos expansions, Dempster-Shafer theory of evidence, and "second-order" probability analysis. Examples of each of the methods as applied to a simple engineering model will be provided.

1. Introduction

Most computer models for engineering applications are developed to help assess a design or regulatory requirement. As part of this task, the capability to quantify the impact of variability and uncertainty in the decision context is critical. The requirement is often stated as: the probability that some system response quantity exceeds a threshold value is less than some required probability. This paper will provide an outline and comparison of methods that are used for analyzing and propagating aleatory and epistemic uncertainties in engineering models. We will specifically focus on five classes of methods: Latin Hypercube sampling, analytic reliability methods, polynomial chaos expansions, Dempster-Shafer theory of evidence, and "second-order" probability analysis. Examples of each of the methods as applied to a simple engineering model will be provided. The computational cost of engineering simulation models is often quite expensive: high fidelity finite-element simulations may require hours or days to run on dozens of processors. Thus, understanding how the uncertainty propagation methods work and their relative advantages and costs is very important. Additionally, response surface approximations (RSA) are often used as inexpensive replacements for computationally expensive computer simulations. Once a RSA has been computed, it is cheap to evaluate this "meta-model" or surrogate, and thus the RSA is often used in a variety of contexts, including uncertainty quantification. This paper will not focus on RSA. However, some of the uncertainty quantification methods rely on such approximations and they will be mentioned where appropriate.

The outline of the paper is as follows: Section 2 provides background on uncertainty quantification (UQ). Section 3 describes an engineering example which will be used as a basis for examples and comparison of methods. Sections 4-8 describe particular UQ methods, and Section 9 summarizes the methods and presents some current research areas of interest.

2. Uncertainty Quantification

Incertitude (commonly referred to as "uncertainty") can be formally classified as *aleatory uncertainty* and *epistemic uncertainty*. Guidance from a Department of Energy document which provides guidelines for quantifying margins and uncertainties using modeling and simulation [Diegert et al.] states: "Where it is practical, calculation input characterizations should separate aleatory and epistemic uncertainties."

Aleatory uncertainty characterizes the inherent randomness in the behavior of the system under study. Alternative terminologies include: variability, stochastic uncertainty, irreducible uncertainty, and Type A uncertainty. Aleatory uncertainty is irreducible except

through design modifications. Examples of aleatory uncertainty are component failures or material properties derived from statistically significant testing under conditions relevant to the application. Aleatory uncertainties are characterized by frequency distributions; and aleatory uncertainties propagated through a model will result in distributions for key system response quantities that should also carry a frequensic interpretation.

Epistemic uncertainty characterizes the lack of knowledge about the appropriate value to use for a quantity that is assumed to have a fixed value in the context of a specific application. Alternative terminologies include: state of knowledge uncertainty, subjective uncertainty, reducible uncertainty, and Type B uncertainty. Epistemic uncertainties are reducible through increased understanding (research), or increased data, or through more relevant data. Epistemic uncertainties are characterized degrees of "belief" and should not be given a frequensic interpretation.

Currently, the following methods are commonly used to characterize and propagate aleatory uncertainty in input parameters through computer models. These are all based on probability theory:

1. Sampling
2. Analytic reliability methods
3. Polynomial Chaos Expansions

Currently, we have the following methods to characterize and propagate epistemic uncertainty:

1. Dempster-Shafer Theory of Evidence

Finally, we have one way which is generally used to propagate uncertainty in combined analysis, where we are propagating both aleatory and epistemic uncertainty. This is called "second-order" probability. The second-order refers to the two types of uncertainty and does not refer to a second-order approximation of the underlying function.

1. Second-order probability

Sections 4-8 describe these UQ methods in more detail. Note that all of the methods presented in this paper are available for use in the DAKOTA software tool [Eldred et al., available at http://www.cs.sandia.gov/DAKOTA].

3. Cantilever Beam Example

This example is taken from the reliability optimization community. The problem involves a simple cantilever beam as shown in Figure 1. The goal is to understand how the deflection of the beam varies with respect to the length, width, and height of the beam as well as to applied load and elastic modulus of the beam.

Figure 1. Cantilever Beam Schematic

The variables characterizing the beam problem are shown in Table 1, along with nominal values.

Variable	Description	Nominal Value
L	Length	1 m
W	Width	1 cm
H	Height	2 cm
I	Area Moment of Inertia	$1/12\ WH^3$
P	Load	100 N
E	Elastic Modulus of Aluminum 6061-T6	69 GPa

Table 1. Variables characterizing the Cantilever Beam Problem

The formula for beam deflection (assuming simplifying assumptions such as isotropic material properties of the beam, elastic displacement, and negligible beam mass) is: Deflection = $PL^3/3EI$. For this problem with 100N load, the deflection is approximately 7.2 cm.

4. Sampling

The most common method of incorporating uncertainty into simulations is to assume certain distributions on the uncertain input values, sample from those distributions, run the model with the sampled values, and do this repeatedly to build up a distribution of the outputs. This is classical propagation of uncertainty. The sampling techniques can be a variety of Monte Carlo methods, including stratified sampling (Latin hypercube sampling) which spread the samples over the space, or quasi-Monte Carlo sampling, which is a way of generating sequences that approximate a uniform distribution. Figure 2 shows the concept of sampling, where N samples are taken from distributions on inputs (in this case, 3 input variables), each one is run through the simulation model, resulting in N realizations on the outputs.

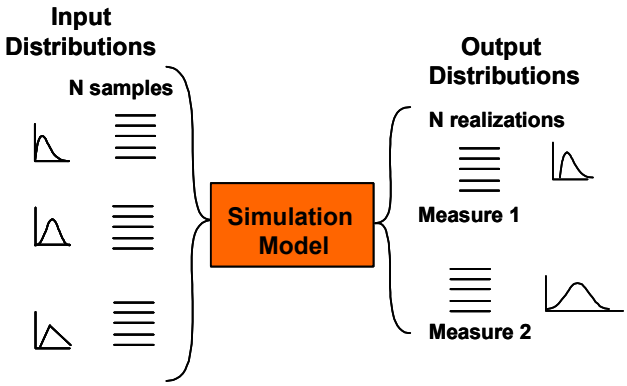

Figure 2. Sampling Approach to Propagate Uncertainties

At Sandia, we often use Latin Hypercube Sampling (LHS) [Swiler and Wyss]. LHS is a stratified sampling method where one sample is chosen from each of N equally probable strata over the support of the distribution. LHS is more efficient than pure Monte Carlo in the sense that it requires fewer samples to achieve the same accuracy in statistics (variance of the mean, for example).

A rule of thumb is to have the number of samples be at least 10 times the number of uncertain variables. For example, if you have 15 uncertain variables, you should specify at least 150 samples.

For the cantilever beam example, we treat multiple inputs as uncertain, distributed with the probability distributions specified in Table 2. Note that 3 variables are uncertain in this problem, L, P, and E.

Variable	Distribution	Distribution Parameters
L	Normal	Mean = 1m Std. Dev. = 0.01 m
W	Fixed	1 cm
H	Fixed	2 cm
P	Normal	Mean = 100 N Std. Dev. = 5 N
E	Normal	Mean = 69 GPa Std. Dev. = 13.8 GPa

Table 2. Uncertain Variable Specification for Sampling

When generating 1000 samples from these three normal distributions and then using the sample values to calculate displacement of the beam, we obtain the histogram of displacement shown in Figure 3:

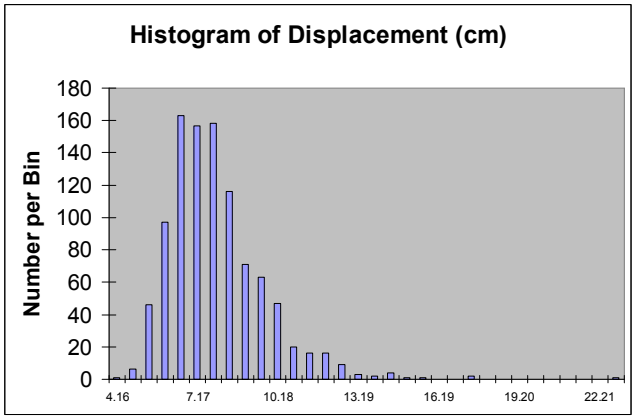

Figure 3. Histogram of 1000 Input Samples of Displacement

If we plot the output samples cumulatively, we get the cumulative distribution function (CDF) shown in Figure 4:

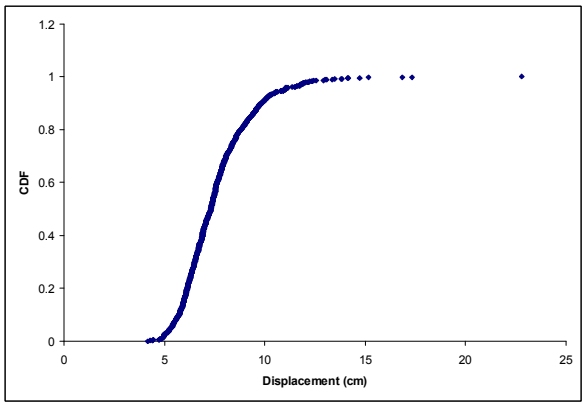

Figure 4. CDF of 1000 Input Samples of Displacement

We can also visually look at the correlations of the output with respect to the inputs. For example, the following two scatter plots in Figures 5 and 6 show that the displacement is not strongly correlated to the change in load but very strongly negatively correlated to the modulus of elasticity. The graphs also allow one to quickly identify outliers.

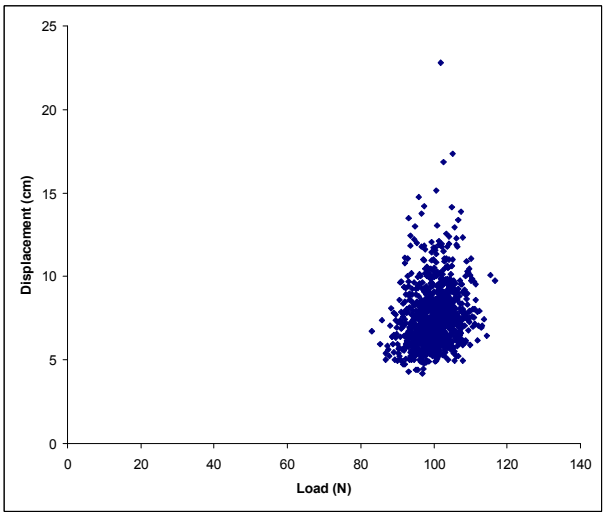

Figure 5. Scatter plot of displacement vs. load

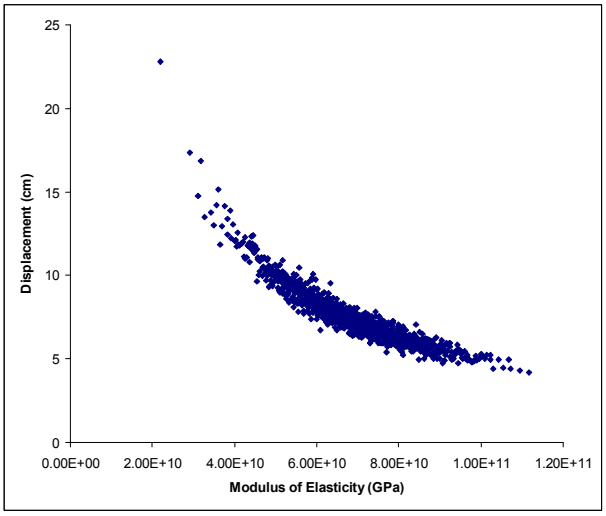

Figure 6. Scatter plot of displacement vs. modulus

The example presented in Figures 3-6 above, with 3 uncertain variables (L, P, and E), will be used as the basis for the subsequent uncertainty analysis methods.

5. Analytic Reliability

A full discussion of the theory behind analytic reliability methods will not be presented here. There is a good discussion in Section 6.3 of the DAKOTA User's Manual [Eldred et al.] and in [Haldar and Mahadevan]. Reliability methods provide an alternative approach to uncertainty quantification which can be less computationally demanding than sampling techniques. Reliability methods for uncertainty quantification are based on probabilistic approaches that compute approximate response function statistics based on specified uncertain variable distributions. These response statistics include response mean, response standard deviation, and cumulative or complementary cumulative distribution functions (CDF/CCDF). These methods are often more efficient at computing statistics in the tails of the response distributions (events with low probability) than sampling based approaches since the number of samples required to resolve a low probability can be prohibitive.

The methods all answer the fundamental question: Given a set of uncertain input variables, X, and a scalar response function, g, what is the probability that the response function is below or above a certain level, $\bar{z}$? This probability calculation involves a multi-dimensional integral over an irregularly shaped domain of interest. The reliability methods all involve the transformation of the user-specified uncertain variables in x-space to a space of independent standard Gaussian random variables, possessing a mean value of zero and unit variance called u-space. The region of interest is also mapped to the transformed space. In the transformed space, probability contours are circular in nature and the multi-dimensional integrals can be approximated by simple functions of a single parameter, β, called the reliability index. β is the minimum Euclidean distance from the origin in the transformed space to the response surface. This point is also known as the most probable point (MPP). The distance of the MPP from the origin has the meaning of the number of input standard deviations separating the mean response from a particular response threshold. There are many reliability methods: most of them involve different ways to approximate the limit state, solve for the MPP, or integrate to calculate the probabilities.

Note that although analytic reliability methods may require fewer function evaluations than sampling methods, they do require that finite difference or analytic gradients of the output variable with respect to the uncertain inputs be provided. Also, Hessian information, if available, tends to help the performance of these methods. Finally, note that each point in a cumulative distribution function requires performing an optimization run.

Figure 7 shows an example CDF from an analytic reliability calculation, using the inputs specified in Table 2. Note that 11 points were calculated along the CDF, for

a total of 267 function evaluations. This is a relatively "easy" function for reliability analysis, since the inputs are normal and the function is not noisy or ill-behaved with respect to the inputs. The statistics from the CDF in Figure 7 compare well with the statistics from sampling.

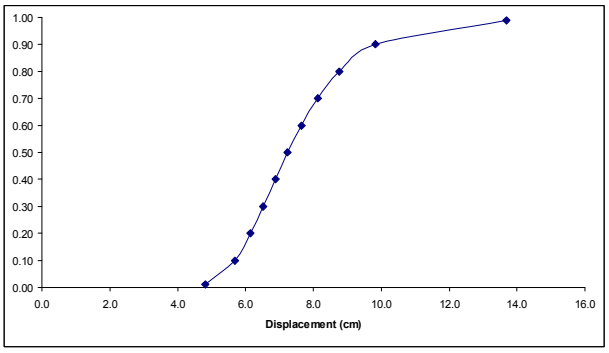

Figure 7. CDF based on Reliability Method

6. Polynomial Chaos Expansions

The polynomial chaos expansion (PCE) method is a general framework for the approximate representation of random response functions in terms of finite-dimensional series expansions. PCE is a type of stochastic response surface method [Ghanem and Red-Horse]. These methods approximate both the inputs and the outputs of the uncertain system through series expansions of standard random variables. In PCE, the series expansion is based on unit Gaussian random variables. The expansion is based on Hermite polynomials which are functions of the Gaussian random variables. The goal of a PCE analysis is to determine the unknown coefficients of the Hermite polynomials in the series expansion. Usually, these coefficients can be calculated from a limited number of model simulations.

Conceptually, the propagation of input uncertainty through a model using PCE consists of the following steps: (1) input uncertainties are expressed in terms of a set of unit Gaussian random variables, (2) a functional form such as Hermite polynomials is assumed for selected outputs, and (3) the parameters of the functional approximation are determined.

An important distinguishing feature of the PCE methodology is that the solution series expansions are expressed as random processes, not merely as statistics as in the case of many nondeterministic methodologies. PCE can be most efficient when there is some additional information from the physics of the problem about what type of approximating polynomials are most appropriate. In DAKOTA, we have implemented a "black-box" or "non-intrusive" version of PCE. In this approach, a number of samples are taken initially of the real function, then those are used to help solve for the coefficients of the expansion. The samples can be based on Latin Hypercube sampling, quadrature points, imported from a file, or stochastic collocation points. The DAKOTA results can give you percentile values if you request them. The DAKOTA output also lists the coefficients of the expansion so that you could, if desired, sample from many unit random normal variables and construct a distribution of the output based on plugging the unit normal variables, the Hermite polynomials, and the expansion coefficients into the correct expansion formula.

Figure 8 shows two cumulative distribution function from PCE calculations, again based on the inputs shown in Table 2. One CDF is based an initial set of 250 LHS samples, with finite differencing to calculate derivatives, for a total of 1750 function evaluations. The second CDF is based on quadrature points. The order of the quadrature is five in each uncertain dimension, for a total of 125 integration points. The expansion order (the highest order of the orthogonal polynomials appearing in the expansion) is two for both CDFs. The percentiles for both CDFs were based on using the chaos expansion formulas and evaluating them at 10,000 LHS sample points.

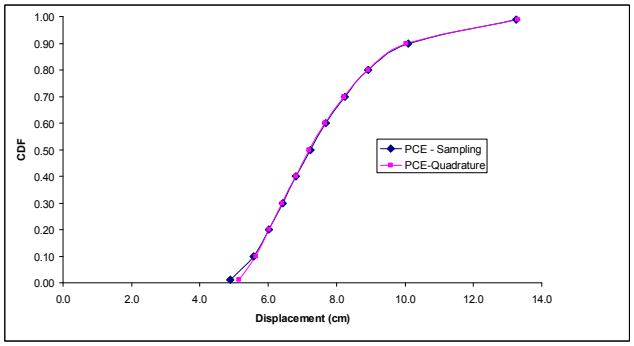

Figure 8. CDF based on Polynomial Chaos Expansion Method

7. Dempster-Shafer Evidence Theory (Epistemic)

The previous uncertainty propagation methods (sampling, analytic reliability, and polynomial chaos expansions) dealt with uncertain input variables that can be specified by a probability distribution. These approaches apply primarily to propagating uncertain variables characterized by aleatory uncertainty. This section discusses one approach to propagating epistemic uncertain variables.

There are many ways of representing epistemic uncertainty, including fuzzy sets, possibility theory, and imprecise probability. At Sandia we have chosen to use Dempster-Shafer, in part because it is a generalization of classical probability theory which allows the simulation

code to remain black-box (it is non-intrusive to the code) and because the Dempster-Shafer calculations use much of the probabilistic framework we already have in place. [Oberkampf and Helton] Dempster-Shafer has two measures of uncertainty, belief and plausibility. Together, these define an interval-valued probability distribution, not a single probability distribution.

Dempster-Shafer Theory of Evidence may be used to perform epistemic analysis. In Dempster-Shafer evidence theory, the epistemic uncertain input variables are modeled as sets of intervals. Note that each variable may be defined by one or more intervals. The user assigns a basic probability assignment (BPA) to each interval, indicating how likely it is that the uncertain input falls within the interval. The BPAs for a particular uncertain input variable must sum to one. The intervals may be overlapping, contiguous, or have gaps. Dempster-Shafer has two measures of uncertainty, belief and plausibility. The intervals are propagated to calculate belief (a lower bound on a probability value that is consistent with the evidence) and plausibility (an upper bound on a probability value that is consistent with the evidence). Together, belief and plausibility define an interval-valued probability distribution, not a single probability distribution.

The method for calculating Dempster-Shafer intervals is computationally very expensive. Many hundreds of thousands of samples are taken over the space. Each combination of input variable intervals defines an input "cell." By interval combination, we mean the first interval of the first variable paired with the first interval for the second variable, etc. Within each interval calculation, it is necessary to find the minimum and maximum function value for that interval "cell." These min and max values are aggregated to create the belief and plausibility curves. The Dempster-Shafer method may use a surrogate model and/or optimization methods. The accuracy of the Dempster-Shafer results is highly dependent on the number of samples and the number of interval combinations. If one has a lot of interval cells and few samples, the estimates for the minimum and maximum function evaluations is likely to be poor.

We demonstrate two examples of the propagation of uncertainty with Dempster-Shafer structures. The Dempster-Shafer structures are shown in Table 3. Note that for both of these examples, width and height of the beam are fixed at 1cm and 2cm, respectively. The three uncertain variables are L, E, and P. The first example is very simple: each uncertain variable is represented by one interval, which means that the uncertain variable can take any number of possible values between the upper and lower bound on that interval. For example, the load P can take a value anywhere between 85 and 115 Newtons. Note that this does NOT imply that any value within that interval is equally likely (as would be the case with a uniform distribution). Rather, the interpretation is that any value within that interval is a possibility or possible realization. Table 3b shows the second example, where each variable is subdivided into 3 possible intervals. The middle interval is the most likely interval, in that it has a basic probability assignment of 0.5: there is a 50% chance that P will fall in the interval [90,110], for example. The outer intervals each have a BPA of 0.25. The intervals in this example were constructed for the purpose of demonstration. Ideally, one would elicit expert judgment to construct the intervals and their associated BPAs. Also, the intervals defined in Table 3b are all contiguous, but there is no requirement that they be: the intervals can be overlapping or disjoint.

Variable	Intervals	BPA
L	[0.97, 1.03] m	1.0
P	[85,115] N	1.0
E	[27.6,110.4]GPa	1.0

Table 3a. Epistemic Variables for the Cantilever Beam Problem, Example 1

Variable	Intervals	BPA
L	[0.97, 0.98] [0.98, 1.02] [1.02,1.03] m	0.25, 0.5, 0.25
P	[85,90] [90,110] [110,115] N	0.25, 0.5, 0.25
E	[27.6,41.4] [41.4, 96.6] [96.6,110.4]GPa	0.25, 0.5, 0.25

Table 3b. Epistemic Variables for the Cantilever Beam Problem, Example 2

Based on the intervals defined for L, P, and E, we performed the Dempster-Shafer calculations. The resulting belief and plausibility bounds are shown in Figure 9. For example 1, since each uncertain variable is defined by one interval, all we can say is that the output falls within one interval. In example 1, displacement falls between 3.75 cm and 20.53 cm. Note that the Dempster Shafer calculations for this example were based on 1000 samples; the results will change slightly as the number of samples changes.

The cumulative belief and plausibility functions for Example 2 are much tighter than those in Example 1. This is due to the fact that there was more structure to the intervals defining the three uncertain variables in Example 2. Together, belief and plausibility provide bounds as to where the true, unknown cumulative distribution function may fall. For a given displacement value, the lower bound of that interval is the belief and the upper bound is the plausibility The bounds on the unknown "true" CDF given by belief and plausibility are

often quite large, reflecting the uncertainty in the input variables as defined by the intervals.

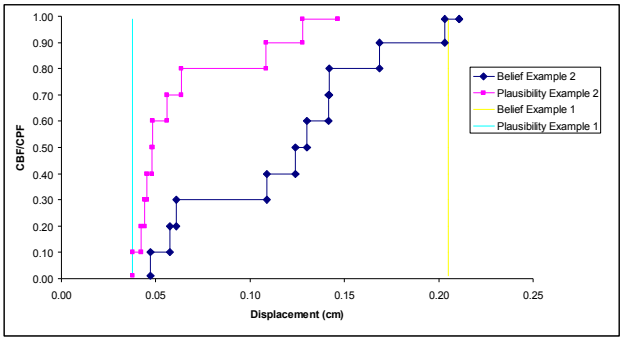

Figure 9. CBF/CPF based on Dempster-Shafer Evidence Theory

8. Second-Order Probability

This section discusses the case where we are trying to propagate both aleatory and epistemic uncertainty. A common situation is where one may know the form of the probability distribution for an uncertain variable (for example, that it is distributed normally or lognormally), but one is not sure of the parameters governing the distribution. In this case, the analysis is done with an outer loop and an inner loop. In the outer loop, the epistemic variables are specified. In this example, the epistemic variables are specified as intervals on parameter values such as means or standard deviations of uncertain variables. A particular value is selected from within the specified intervals. Then, this value is sent to the inner loop. In the inner loop, the values of the distribution parameters are set by particular realizations of the epistemic variables, and the inner loop performs sampling on the aleatory variables in the usual way (e.g., a LHS sample is taken). Figure 10 shows the sampling structure of second-order probability.

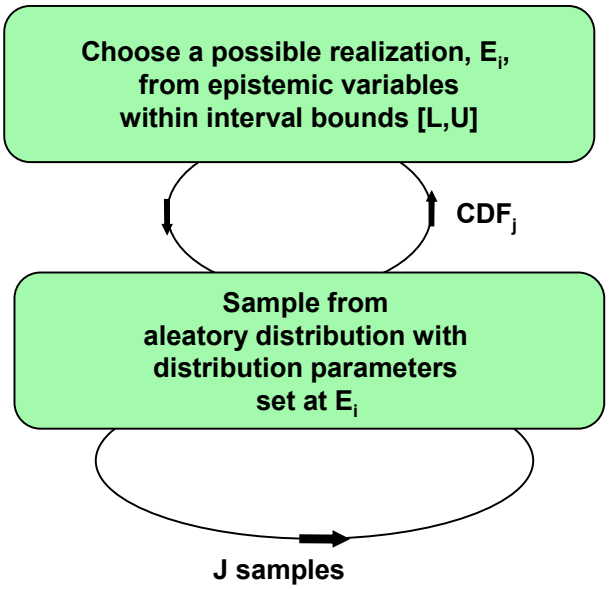

Figure 10. Second-order Probability

Second-order probability may be expensive since we have two sampling loops. However, it has the advantage that it is easy to separate and identify the aleatory vs. epistemic uncertainty. Each particular set of epistemic variable values generates an entire CDF for the response quantities based on the aleatory uncertainty. So, for example, if one had 50 values or samples taken of the epistemic variables, one would have 50 CDFs. When you plot the 50 CDFs, you get the upper and lower bound on the family. Thus, the results look very similar to a Dempster-Shafer analysis.

Table 4 outlines the uncertain variables for second-order probability sampling for the cantilever beam problem.

Variable	Epistemic Mean	Distribution
L	[0.98, 1.02] m	Normal(epistemic mean, 0.01) m
P	[90,110] N	Normal(epistemic mean, 5) N
E	[41.4,96,6] GPa	Normal(epistemic mean, 13.8) GPa

Table 4. Epistemic Variables for the Cantilever Beam Problem, Example 1

For the results shown in Figure 11, we took 20 outer loop samples and 100 inner loop samples. We get inner and outer bounds on the family of CDFs generated by the 20 outer loop samples on the epistemic variables. Two particular realizations of CDFs (based on 100 inner loop samples) are also plotted.

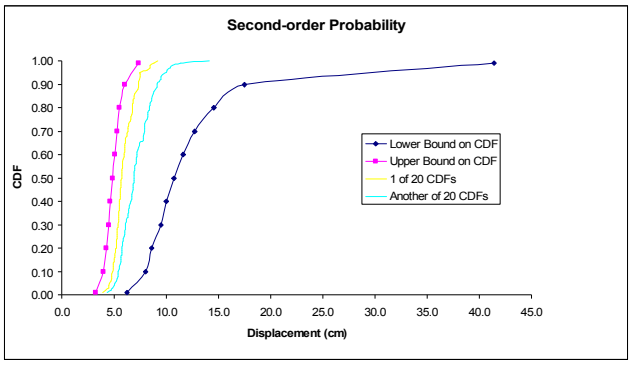

Figure 11. Second-order Probability Results for the Cantilever Beam Problem

9. Summary of UQ Methods

This paper has outlined five methods that are used by the engineering community to propagate uncertainty in computational models. These five methods are all available in the DAKOTA software tool. Three of these methods focus on propagation of aleatory uncertainties: sampling, analytic reliability, and polynomial chaos expansions. Dempster-Shafer evidence theory is used to propagate epistemic uncertainty, and second-order probability can handle both. When choosing a method, one should consider a variety of factors, including the computational cost of the function evaluations, the transparency of the method, the type of outputs generated, and the question being addressed (e.g., is the goal to get an accurate assessment of the response level at a particular percentile, a reliability/failure probability, the mean or variance of the output distribution, is an entire output CDF necessary, etc.). The transparency of the method is very important in regulatory applications and environments where results may have to be repeated under a variety of conditions. Sampling methods are the most "transparent" in terms of being able to trace a particular set of inputs to a particular output value on the CDF. One class of methods this paper did not address is Bayesian methods. Bayesian methods require the specification of a likelihood function (i.e., the likelihood of obtaining a particular set of computational results based on data.) In general, this is difficult to do in a general-purpose framework without a significant set of additional statistical assumptions. For this reason, Bayesian methods have not been included in this presentation.

References

Diegert, K., Klenke, S., Novotny, G., Paulsen, R., Pilch, M. and T. Trucano. "Toward a More Rigorous Application of Margins and Uncertainties within the Nuclear Weapons Life Cycle – A Sandia Perspective." Sandia Technical Report SAND2007-6219, October 2007.

Eldred, M.S., Brown, S.L., Adams, B.M., Dunlavy, D.M., Gay, D.M., Swiler, L.P., Giunta, A.A., Hart, W.E., Watson, J.-P., Eddy, J.P., Griffin, J.D., Hough, P.D., Kolda, T.G., Martinez-Canales, M.L. and Williams, P.J., "DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.0 Users Manual," Sandia Technical Report SAND2006-6337, October 2006.

Eldred, M.S., Brown, S.L., Adams, B.M., Dunlavy, D.M., Gay, D.M., Swiler, L.P., Giunta, A.A., Hart, W.E., Watson, J.-P., Eddy, J.P., Griffin, J.D., Hough, P.D., Kolda, T.G., Martinez-Canales, M.L. and Williams, P.J., "DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.0 Reference Manual," Sandia Technical Report SAND2006-4055, October 2006.

Ghanem, R. and J. R. Red-Horse. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element technique. Physica D 133: 137-144, 1999.

Haldar, A., and S. Mahadevan, Probability, Reliability, and Statistical Methods in Engineering Design, John Wiley and Sons, 2000.

Oberkampf, W.L. and J.C. Helton. Evidence Theory for Engineering Applications. Sandia National Laboratories Technical Report, SAND2003-3559P.

Swiler, L.P. and Wyss, G.D., "A User's Guide to Sandia's Latin Hypercube Sampling Software: LHS UNIX Library Standalone Version," Sandia Technical Report SAND2004-2439, July 2004.