THE VALUE OF REPRESENTING EPISTEMIC UNCERTAINTY IN ENGINEERING DESIGN

Jason Matthew Aughenbaugh
gtg224k@mail.gatech.edu

Christiaan J. J. Paredis
chris.paredis@me.gatech.edu

Systems Realization Laboratory
G. W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0405
www.srl.gatech.edu

KEYWORDS
Epistemic uncertainty, aleatory uncertainty, engineering design, probability bounds analysis, PBA

ABSTRACT
Engineering design decisions inherently are made under uncertainty. In this paper, we recognize a difference between aleatory uncertainty (due to inherent randomness) and epistemic uncertainty (due to lack of knowledge). Our hypothesis is that, in engineering design decisions, it is valuable to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty. In this paper, we support this hypothesis with a computational experiment in which a pressure vessel is designed using two approaches, both variations of utility-based decision making. In the first approach, designers use a purely probabilistic, best-fit normal distribution to represent uncertainty. In the second approach, designers explicitly express epistemic uncertainty distinctly from aleatory uncertainty using a probability box, or p-box. When the epistemic uncertainty is large, this latter approach results on average in designs with expected utilities that are greater than those for designs created with the purely probabilistic approach. In the context of decision theory, this suggests that there are design problems for which it is valuable to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty.

INTRODUCTION
Of the many challenges in engineering design, one of the greatest is uncertainty. During the design process, engineers must make decisions without being certain of the outcomes of these decisions. Prior to making a decision, engineers can remove some of this uncertainty by expending resources to acquire more knowledge, for example, by modeling the product or by performing experiments. Engineers also reduce uncertainty by making decisions that reduce design freedom and make the product specification more certain. However, even once a complete description of the product has been developed, uncertainty about its performance may remain due to variations in the manufacturing process and the product’s environment of use.

Engineers have developed or adopted methods to support design decisions under uncertainty, such as safety factors [1], utility theory [2-5], probabilistic risk assessments [6], reliability based design optimization [7], and robust design [8, 9]. We believe that current methods are still limited in their ability to clearly and quantitatively reflect the true nature of the uncertainty encountered in engineering design.

In this paper, we recognize two types of uncertainty—aleatory uncertainty due to inherent randomness, and epistemic uncertainty due to lack of knowledge. We discuss how existing probabilistic methods for engineering design often confound these two types of uncertainty. Our hypothesis is that it is valuable in engineering design to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty. In the context of decision theory and utility theory, one design method is more valuable than another if on average it yields designs with higher expected utilities than the other approach. In this paper we consider an example pressure vessel design problem and present a computational experiment that demonstrates the value, in comparison to a purely probabilistic approach, of explicitly representing epistemic uncertainty distinctly from aleatory uncertainty.
UNCERTAINTY IN ENGINEERING DESIGN

Definitions of uncertainty

In this paper, we view uncertainty in the context of decision theory [10, 11]. In this context, Nikolaidis defines uncertainty indirectly from the definition of certainty [12]. He defines certainty as the condition of knowing everything necessary to choose the course of action whose outcome is most preferred. As shown in Figure 1(a), a decision-maker’s uncertainty is the gap between certainty and what he or she actually knows, which is defined as the present state of knowledge.

Researchers in the risk assessment community recently have emphasized a fundamental difference between uncertainty due to inherent randomness and uncertainty due to a lack of knowledge [13-16], respectively named aleatory and epistemic uncertainty. Although some authors have argued that it is philosophically incorrect to distinguish aleatory and epistemic uncertainty [17], most authors recognize that in practice it is often useful to make the distinction [13-17].

Aleatory uncertainty (from the Latin aleator, for dice thrower or gambler) is due to natural random behavior in a physical process or property and is also known as variability, stochastic uncertainty, objective uncertainty, and irreducible uncertainty. The term irreducible is often used because there is no feasible way for a decision-maker to know a particular outcome of a random process with certainty. The future outcome is random and at best can be described probabilistically.

Epistemic uncertainty (from the Greek episteme for knowledge), is due to a lack of knowledge or information and sometimes is called imprecision, ignorance, or reducible uncertainty. The term reducible indicates that epistemic uncertainty can be reduced by expending resources to acquire more knowledge. Epistemic uncertainty has been represented in terms of intervals [18-20] and more elaborate methods, including rough sets, fuzzy sets, possibility theory, copulas, and evidence theory [21]. In this paper we focus on the value of interval representations of epistemic uncertainty.

With the distinction between aleatory and epistemic uncertainty made, a perfect state of knowledge, shown in Figure 1(b), is defined as the state in which the decision-maker cannot possibly gain more certainty about a future outcome. In a state of perfect knowledge, aleatory uncertainty will remain if there is inherent randomness. In most engineering problems, the present state of knowledge falls short of this perfect state of knowledge; therefore, engineers face epistemic uncertainty in addition to aleatory uncertainty.

To make the distinction between aleatory and epistemic clearer, consider the following example. Assume a decision-maker knows that the manufacture of a steel plate is a random process, and the thickness of the plate is described by a gamma distribution with known parameters. This is a perfect state of knowledge because the gap between the present state of knowledge and certainty is due entirely to aleatory uncertainty. The decision-maker cannot get closer to certainty because it is impossible to know the thickness of a particular plate until it is built.

We now assume that the decision-maker has ten samples of plate thicknesses instead of perfect knowledge of the randomness. Due to random sampling, these ten samples do not necessarily reflect the true nature of the process. In this case, the decision-maker faces a gap—epistemic uncertainty—between the present state of knowledge and a perfect state of knowledge, because in a perfect state of knowledge the decision-maker would be able to characterize the aleatory uncertainty with certainty. In this example, the decision-maker can reduce this gap, and hence reduce the epistemic uncertainty, by acquiring more knowledge in the form of additional data samples.

The exact way in which decision-makers represent the uncertainty inherent in a limited data sample depends somewhat on their interpretation of probability, as discussed in the next section. However, our hypothesis is that, regardless of how probability is interpreted, it is valuable to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty.

Interpretations of probability

There are at least 11 different significant interpretations of probability, but the most commonly adopted in engineering design are the frequentist and subjective [22]. It is important to understand and distinguish these two interpretations because a single probability distribution should be interpretable as either representing aleatory uncertainty under a frequentist interpretation or as an expression of uncertainty under a subjective interpretation [23].

The frequentist interpretation, also known as objective or relativist, is based on the notion of relative frequencies of outcomes. Under a frequentist interpretation, a probability represents the ratio of times that one outcome occurs compared to the total number of outcomes in a series of identical and possibly random trials. Therefore, a frequentist probability distribution can only express aleatory uncertainty. Construction of a frequentist probability distribution requires knowledge of the relative frequencies. This presents a challenge in engineering design because these frequencies can
only be known after observation [22], but engineers often lack large data histories, especially for novel designs. Even when data is available, there is no guarantee that a particular sample is representative of the true relative frequency. Although in theory the relative sample frequency approaches the true relative frequency as the sample size goes to infinity, an infinite sample size is impossible in practice. Consequently, engineers can never know the true probabilities with certainty. This epistemic uncertainty cannot be expressed by a frequentist probability distribution.

Proponents of a subjective or Bayesian interpretation of probability claim that there is no such thing as a true or objective probability, but rather probabilities are an expression of belief based on an individual’s willingness to bet [17, 22, 24]. However, one’s willingness to enter a bet depends not only on probabilities but also on preferences, which some argue should be kept separate from risk assessments [25]. In addition, people are inherently bad at assessing probabilities [26]. Consequently, the construction of subjective probabilities involves its own challenges, especially when it comes to aggregating subjective probabilities from multiple sources [27].

By proposing that there is no underlying true frequency, subjectivists blur the distinction between epistemic uncertainty and aleatory uncertainty. Subjectivists capture their overall uncertainty—epistemic and aleatory—in a single probability distribution. However, once an individual’s belief about both epistemic and aleatory uncertainty is captured in a single probability distribution, there is no way to distinguish the two types (Section 1.2.4 of [28]), and therefore it is difficult to draw useful insights from the distribution [29]. This confounding of uncertainty is a limitation of a subjective interpretation of probability, especially when it comes to propagating uncertainty. It is, in general, incorrect to combine a probability distribution encoding epistemic uncertainty with a distribution that encodes aleatory uncertainty (Section 3.4.1 of [28]). The two distributions instead should be treated independently at different levels, with the two types of uncertainty expressed separately.

In the preceding paragraphs, we have explained the apparently limited ability of probability distributions to represent epistemic uncertainty clearly under either a frequentist or subjective interpretation of probability. These limitations suggest that additional means are necessary to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty. The only concrete way to determine the value of explicitly and distinctly representing epistemic uncertainty in engineering design is in the context of decision theory. Informally, one method of representing uncertainty is better than another if it allows designers to make better decisions. With this in mind, we discuss decision theory in the following section.

**DECISION-MAKING UNDER UNCERTAINTY**

Engineering design decisions are hampered by the uncertainty inherent in predicting the outcomes and payoffs of different actions. In this section, we discuss a safety-factor approach and traditional statistical decision theory.

**Design with safety factors**

One way engineers deal with uncertainty is by using safety factors. For example, if engineers are building a pressure vessel with material yield strength \( \sigma_y \), the requirement to avoid failure is that \( \sigma_y \) exceeds the maximum stress in the pressure vessel, \( \sigma_{\text{max}} \). That is: \( \sigma_y > \sigma_{\text{max}} \). Engineers rarely, if ever, know these parameters exactly. In a safety factor approach, engineers employ a safety factor SF > 1 with their best point estimates \( \hat{\sigma}_y \) and \( \hat{\sigma}_{\text{max}} \) respectively, and design the pressure vessel such that

\[
\hat{\sigma}_y > \text{SF} \times \hat{\sigma}_{\text{max}}.
\]

The safety factor is an attempt to wrap all uncertainty—epistemic and aleatory—into one number. The engineers hope that by designing the pressure vessel with a safety margin around their estimates, the true yield strength \( \sigma_y \) (which may be much less than \( \hat{\sigma}_y \)) will exceed the true \( \sigma_{\text{max}} \) (which may be much larger than \( \hat{\sigma}_{\text{max}} \)).

The simplicity of employing safety factors is its biggest advantage, but the unanswered question in this approach is how large of a safety factor is necessary to meet reliability and performance requirements given the existing uncertainty. In certain industries, safety factors are well established and have been used with great success. In the presence of large data histories, safety factors actually can be linked to probabilistic characterizations of uncertainty and reliability [1]. However, this link requires the random process’s true distribution to be known with certainty, thus assuming no epistemic uncertainty.

In practice, particularly for novel design tasks, engineers do not have perfect knowledge and often must resort to an ad hoc choice of a safety factor. Because engineers do not want structures to fail, engineers usually choose safety factors that are much larger than necessary. This frequently results in costly over-design of the product. At the same time, the ad hoc method fails to provide any guarantee of reliability. Consequently, engineers have pursued more formal methods for dealing with uncertainty, such as traditional statistical decision theory.

**Traditional Statistical Decision Theory**

In traditional statistical decision theory [30], utility analysis, as originally proposed by von Neumann and Morgenstern [31], is used for making decisions under uncertainty. *Utility* is a measure of preference — more preferred decision outcomes are assigned higher utility values. However, rather than merely an ordinal ranking, utilities provide cardinal ranking — the difference in utility is a measure for how much one alternative is preferred over another. If chosen correctly, such a cardinal ranking reflects the decision-maker’s preferences even under uncertainty. By applying the expected value operator, the decision-maker weights all possible outcomes according to their likelihood of occurring. Von Neumann and Morgenstern developed an approach based on simulated lotteries to help decision-makers define utility functions for which the expected utility accurately reflects their preferences. Utility theory has been studied extensively by economists and decision theorists, and, in the last two decades, there has been an increasing
interest in applying utility methods to engineering problems, as in [2-5].

In statistical decision theory, one characterizes the uncertainty with a probability density function. The implications of this probability density function depend on the adopted interpretation of probability. Under a frequentist interpretation of probability, a probability distribution can only represent aleatory uncertainty. By adopting this interpretation in statistical decision theory, a designer is forced to ignore or eliminate epistemic uncertainty. In order to eliminate the epistemic uncertainty, the designer must expend resources to acquire the missing knowledge, thus increasing the costs of the design process. If instead, the designer ignores epistemic uncertainty, he or she is overstating the true knowledge by making assumptions that may be wrong or misleading. Since both of these approaches have significant limitations, a frequentist interpretation of probability requires an extension to existing statistical decision theory.

A subjective interpretation of probability has its own limitations. Under the subjective interpretation, the designer expresses both epistemic and aleatory uncertainty in a probability distribution that reflects his or her belief. However, the designer then faces the limitations of confounding the two types of uncertainty that we described in the previous section. In order to avoid these limitations, a subjective interpretation of probability also requires a more general statistical formalism.

In this paper, we investigate the hypothesis that by generalizing the statistical formalism to allow for the explicit representation of epistemic uncertainty distinct from aleatory uncertainty, better decisions can be obtained. The following example design problem and computational experiment support this hypothesis.

PROBABILITY BOUNDS ANALYSIS
In order to consider epistemic uncertainty explicitly and distinctly from aleatory uncertainty, we use a recently developed formalism that extends traditional probability theory to include both aleatory and epistemic uncertainty—probability bounds analysis (PBA) [32, 33]. There are other methods for dealing with both epistemic and aleatory uncertainty, such as double-loop or joint Monte Carlo sampling [34]. However, PBA has additional appealing properties, such as its ability to propagate uncertainty in a computationally efficient fashion, as described in the Discussion and Future Work section.

PBA basics
PBA expresses uncertainty in a structure called a probability-box, or p-box. The p-box incorporates both epistemic and aleatory uncertainty by expressing interval bounds on the cumulative probability distribution function (CDF) for a random variable. This marriage of probability theory and interval analysis is captured in the analogy that a p-box is a “stretched-out” distribution function in the same way that an interval is a stretched out scalar [35]. More formally, the bounds on a p-box, such as shown in Figure 2(a), are given by two CDFs (F₁ and F₂) that enclose a set of CDFs that are, under some interpretation, consistent with the current state of knowledge.

![Figure 2: Example P-box and distributions](image)

Interpreting a p-box
While there are several ways to construct p-boxes [36, 37], we choose a practical method based on traditional confidence intervals, as explained in Appendix A. The method by which a p-box is constructed influences its exact interpretation, but for illustration of meaning, it is easiest to consider a p-box as constructed at the 100% confidence level. In this case, the p-box expresses the range of all CDFs that are still deemed possible based on existing knowledge. For example, assume there is aleatory uncertainty about a particular outcome, but engineers have strong theoretical knowledge that a parameter X is normally distributed with known variance of 1. However, the engineers are uncertain about the distribution’s true mean, which introduces epistemic uncertainty into the design problem. In the notation of probability,

\[ X \sim N(\mu, \sigma^2) \]

means that parameter X is distributed normally with mean \( \mu \) and variance \( \sigma^2 \). In this hypothetical example, the variance is known with certainty (\( \sigma^2=1 \)), while the engineers only know with certainty (100% confidence) that the true mean is between zero and one, which can be expressed as the interval \( \mu=[0,1] \).

Extending the notation of probability distributions, one can write:

\[ X \sim N([0,1], 1) \]

The corresponding p-box is shown in Figure 2(a). In this case, the bounds on the p-box are defined by the two distributions, \( F_1 \sim N(0,1) \) and \( F_2 \sim N(1,1) \). The true CDF is unknown, and any of the infinite number of CDFs inside the p-box could be the true one, such as those shown in Figure 2(b). However, any distribution that falls partially or entirely outside of the p-box is definitely inconsistent with the present state of knowledge.

Vertical slices of the p-box yield intervals on the CDF for a particular realization. For example, a vertical slice at zero yields the interval for the CDF of [0.1587, 0.5]. This means that the probability that X is less than zero is between 0.1587 and 0.5, but one does not have enough knowledge to specify the relative probabilities of points within that interval. Horizontal slices of a p-box result in intervals on the quantiles of the CDF. For example, a slice at the median (CDF=0.5) gives the interval [0,1] for the median.
In summary, a p-box is a more expressive generalization of both traditional probability distributions and interval representations, as is illustrated in Figure 3. A general p-box expresses both aleatory uncertainty (represented by the shapes of the boundary CDFs) and epistemic uncertainty (represented by the separation between the upper and lower bounds) in separate dimensions.

**Deterministic**

<table>
<thead>
<tr>
<th>Perfect state of knowledge</th>
<th>Epistemic Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Precise knowledge of the deterministic value (a crisp number)</td>
<td>(b) Epistemic uncertainty about the true deterministic value (a pure interval)</td>
</tr>
<tr>
<td>(c) Precise knowledge of the aleatory uncertainty (a single CDF)</td>
<td>(d) Epistemic uncertainty and aleatory uncertainty (a general p-box)</td>
</tr>
</tbody>
</table>

**Figure 3: Dimensions of uncertainty**

**MEASURING THE VALUE OF REPRESENTING EPISTEMIC UNCERTAINTY**

In this section, a pressure vessel is designed under incomplete information and knowledge according to several different design approaches. Descriptions of the design scenario and the computational experiment follow.

**Design scenario**

One needs to design a pressure vessel which is to contain 0.15 m$^3$ of gas under 7 MPa of pressure. Due to space limitations, certain maximum dimensions are imposed. The goal is to determine the dimensions of the vessel, shown in Figure 4, for which the overall utility, defined in the next section, is maximized. Since the vessel will be used in a human-occupied location, the cost of failure is weighted heavily. The vessel will be made of a new type of steel for which the yield strength is not well characterized. The material production process produces random variations in the material properties, so there is aleatory uncertainty about the true material yield strength of any particular sample. Because the material is new and testing is relatively expensive, variations in yield strength have only been measured 30 times in independent tension tests. These tests can at best give an estimate of the true distribution, so in addition to inherent randomness (aleatory uncertainty), engineers also face epistemic uncertainty—they do not know the parameters that describe the aleatory uncertainty with certainty. A detailed description of the design assumptions, requirements, and parameters is given in Appendix B.

The general layout of the experiment is shown in Figure 5. The experiment consists of two designers: one using a single best-fit normal distribution (approach A), and the other (approach B) using a p-box to represent the uncertainty about the yield strength. The details of these approaches are explained in the sections following this general overview of the experiment.

Both approaches start with the same information about the uncertain yield strength. This information is a set $\Sigma$ of $n$ random samples from the true normal distribution:

$$\Sigma = \{\sigma_y, \ldots\}.$$

Both designers assume that the true distribution is normal, but they use their own approach to represent their uncertainty in their knowledge of the distribution’s parameters. According to decision policies explained in the following sections, each designer then selects an optimal design, denoted as $a_x^c = \{R_x^c, t_x^c, L_x^c\}$ and $b_x^c = \{R_y^c, t_y^c, L_y^c\}$ respectively for approach A and approach B.

The supervisor compares each of the design solutions to determine the expected utility evaluated under perfect knowledge for each solution. For approach A this is written as

$$E_{a_x^c, \Sigma}[U(\sigma_y, a_x^c)],$$

and for approach B as

$$E_{a_x^c, \Sigma}[U(\sigma_y, b_x^c)].$$

**Figure 4: Schematic and design variables for pressure vessel**

**The computational experiment**

The goal of the experiment is to compare the utility of the design solutions that result from different approaches for representing uncertainty applied to the same design problem. The comparison is made possible in this experiment because we assume that overseeing the experiment is a supervisor who is in a state of perfect knowledge about the steel’s material properties. From the supervisor’s perspective, only aleatory uncertainty about the yield strength of the material exists — expressed as a normal distribution with a mean of 180 MPa and a standard deviation of 15 MPa. No epistemic uncertainty is present because the supervisor has perfect knowledge of this random distribution. The supervisor can therefore determine with certainty the dimensions of the pressure vessel that result in the maximum expected utility. This optimal design under the perfect state of knowledge is the benchmark for comparison for the other design approaches.

The computational experiment follow.
In order to compare the value of the two approaches, the supervisor computes the difference in expected utility under perfect knowledge. Using the fact that the two approaches start with the same knowledge (sample $\Sigma$), the value of approach B over approach A can then be expressed as

$$V(B) = E_{\Sigma}[U(\sigma_y, b_k^*) - U(\sigma_y, a_k^*)].$$

It is necessary to note that this value was for only one sample $\Sigma$—the yield strength measurements with which each designer starts. Due to the randomness in $\Sigma$, one trial is not sufficient to judge the relative value of each approach; the supervisor needs to repeat the above experiment many times in order to determine which design approach performs best on average, over $m$ different sample sets $\Sigma$. Mathematically, the expectation must be taken with respect to $\Sigma$ in order to calculate the average expected value of approach $B$ over $A$, written

$$E_{\Sigma}[V(B)] = E_{\Sigma}[E_{\Sigma}[U(\sigma_y, b_k^*)] - U(\sigma_y, a_k^*)]].$$

While not required, the addition of the word average emphasizes that this quantity is the expectation over the samples of the expected utility of particular design solutions.

For discussion, it is also useful to define the average expected utility evaluated under perfect knowledge of approach A and approach B, respectively

$$E_{\Sigma}[E_{\Sigma}[U(\sigma_y, a_k^*)]]$$

This distinguishes these quantities from what was earlier defined as the expected utility evaluated under perfect knowledge of a particular design solution.

All of the design approaches are based on utility theory maximization and use the same preference structure and utility function given by:

$$U(\sigma_y, DV) = P_{\text{selling}} - C_{\text{material}}(DV) - C_{\text{failure}}(\sigma_y, DV),$$

where:

- $P_{\text{selling}} \equiv$ selling price = $200$
- $C_{\text{material}} \equiv$ material cost per volume = $8500/m^3$
- $\sigma_y \equiv$ true yield strength of pressure vessel
- $DV \equiv$ design variables (radius, thickness, length)
- $C_{\text{failure}}(\sigma_y, DV) \equiv \begin{cases} 0 & \text{if } \sigma_y \geq \sigma_{\text{max}}(DV) \\
1,000,000 & \text{otherwise} \end{cases}$

As defined in the previous line, for given yield strength and design, the failure cost will either be zero (no failure) or a constant (the cost of the damage, lost productivity, etc. when the pressure vessel fails). Also note that in this experiment, only the yield strength is uncertain, and in the utility function, only the last term depends on the random variable $\sigma_y$.

There are actually three design approaches of interest. In addition to approach A and approach B, the supervisor is able to create a design using perfect knowledge. The three approaches are described in the following sections.

**Supervisor’s design under perfect knowledge**

The supervisor can create a design using the true distribution, since he or she is in a perfect state of knowledge. The supervisor therefore knows with certainty that $\sigma_y \sim \text{Normal}(180MPa, (15MPa)^2)$. If we define this as approach $K$ (for perfect knowledge, not shown in the figure), the supervisor then chooses the design variables

$$DV = k = \{R_K, t_K, L_K\}$$

such that the expected utility $E_{\Sigma}[U(\sigma_y, k)]$ is maximized. This leads to the optimal design under perfect knowledge:

$$k^* = \{R^*_K, t^*_K, L^*_K\}.$$

This optimal design, with expected utility denoted $E[U(k^*)]$ for brevity, serves as the baseline for comparison because no other approach can yield a higher expected utility on average.
**Design using approach A**

Designer A does not have access to perfect knowledge, but instead only has access to the limited data samples $\Sigma$. Because the designer does not know the true distribution of $\sigma$, he or she must make an approximation, denoted $\tilde{\sigma}$. In this approximation, the representation of $\tilde{\sigma}$ depends on both the approach, in this case A, and the random sample $\Sigma$. Designer A represents the uncertainty as a normal distribution, using the sample mean and sample variance as unbiased estimates of the true mean and true variance, respectively. This yields the probabilistic model

$$\tilde{\sigma}(B, \Sigma) \sim \text{Normal}(\frac{1}{n} \sum \sigma \theta, \frac{1}{n} \sum (\sigma \theta - \frac{1}{n} \sum \sigma \theta)^2).$$

Designer A therefore chooses design variables $a^*_\Sigma = \{R_\theta, l_\theta, L_\theta\}$ that maximize the estimated expected utility

$$E_{\tilde{\sigma}(\cdot)(\cdot)|\Sigma}[U(\tilde{\sigma}(A, \Sigma), a^*_\Sigma)],$$

given his or her knowledge about the randomness. The expected utility is only estimated because Designer A does not have access to perfect knowledge about the true aleatory uncertainty in $\sigma$. The expected utility maximization results in the optimal design using approach A given samples $\Sigma$, denoted:

$$a^*_{\Sigma} = \{R_\theta, l_\theta, L_\theta\}.$$

**Design using approach B**

Designer B takes a different approach for capturing the uncertainty in the yield strength. Specifically, designer B represents the uncertainty in $\sigma$ by $\tilde{\sigma}$, where the nature of $\tilde{\sigma}$ is expressed as a p-box rather than a pure normal distribution. The construction of this p-box is addressed in Appendix A. The estimated expected utility under design approach B is defined as

$$E_{\tilde{\sigma}(\cdot)(\cdot)|\Sigma}[U(\tilde{\sigma}(B, \Sigma), b^*_\Sigma)]$$

where $b^*_\Sigma$ is the designer’s chosen design action given samples $\Sigma$. Because the p-box expresses a range of possible distributions, the expected utility is no longer a crisp number but rather an interval defined by lower-bound $E^L$ and upper-bound $E^U$, such that

$$E_{\tilde{\sigma}(\cdot)(\cdot)|\Sigma}[U(\tilde{\sigma}(B, \Sigma), b^*_\Sigma)] = [E^L, E^U].$$

Because the expected utility is now an interval, the designer cannot choose design variables $b^*_\Sigma$ that maximize the expected utility in the traditional sense. Instead, a new decision rule is required. While other approaches exist, in this experiment, a conservative best-worst case, or maxi-min, rule is used. Designer B therefore chooses the design action $b^*_\Sigma = \{R_\theta, l_\theta, L_\theta\}$ that has the highest lower bound $E^L$ on the expected utility. This results in an optimal design decision using approach B given the observed samples $\Sigma$, denoted:

$$b^*_\Sigma = \{R_\theta, l_\theta, L_\theta\}.$$

**EXPERIMENTAL RESULTS**

**Value of representing epistemic uncertainty for 30 samples of the true yield strength**

We first conduct the experiment with a sample size of $n=30$, meaning the designers are given the results of 30 independent yield stress tests. As measured by the supervisor under perfect knowledge and averaged over $m=100,000$ samples, approach B yields designs with greater expected utility than approach A. Specifically, we find the 95% confidence interval (CI) on the value of approach B over A to be:

95% CI on $V(B)$ is [$16, 22$].

To put this result in perspective, the expected utility of the supervisor’s design, which is the best possible because it is designed under a perfect state of knowledge, is $E[U(k^*)] = 104$. Thus the CI on the expected value of approach B over A can also be expressed as [16%, 21%] of the optimal utility $E[U(k^*)]$. This is a substantial deviation and therefore suggests that there is value in using the p-box approach for this design problem. We also note that the average expected utilities realized under approach A and B are relatively close to the optimal $E[U(k^*)]$

- Approach A: $E_{\Sigma}[E_{\tilde{\sigma}(\cdot)(\cdot)|\Sigma}(\sigma, a^*_{\Sigma})] = 70$
- Approach B: $E_{\Sigma}[E_{\tilde{\sigma}(\cdot)(\cdot)|\Sigma}(\sigma, b^*_{\Sigma})] = 89$

The total deviations from the optimal ($34$ for approach A and $15$ for approach B), coupled with the relative value of approach B over A, indicate that B is a better approach at this level of epistemic uncertainty. In the next section, we explore how these results differ at varying levels of epistemic uncertainty.

**Variation of value with level of epistemic uncertainty**

The previous discussion dealt with a fixed sample size of 30 material strength tests. While those results demonstrated that it was valuable to use the p-box approach in that case, a more general result is desirable. By varying the number of material strength tests, we can vary the amount of epistemic uncertainty. The supervisor’s design yields the best possible expected utility $E[U(k^*)]$, and hence the designs of designers A and B can at best equal it. In Figure 6, we plot (in log-log scale) the percent deviation from this best possible expected utility under perfect knowledge for approach A and approach B.

For large amounts of epistemic uncertainty, approach B performs significantly better than approach A. For example, at a sample size of 10, a 95% CI on the value of approach B over A is [540%, 580%] of $E[U(k^*)]$. There is no doubt that this value is significant. Around sample size 50, the two approaches yield similar results. Past this point, approach A performs better, but not by much. For example, for a sample size of 100, the 95% CI on the value of approach B is [-0.7%, -0.2%] of $E[U(k^*)]$. This difference appears insignificant, but there may be cases in which such small percentages matter. The difficulty for designers is knowing if they are in this situation, without the perfect knowledge baseline of this
In this section, we provide some insight into the results of the computational experiment. The preceding results are for two representative cases. The overall results of the experiment, discussed previously, indicate that, on average, the latter case dominates. Approach A is more likely to overestimate the true material strength, and when it does, the consequences are disastrous — a high probability of failure. Approach B is more conservative, resulting in higher material costs, but, on average, these material costs are offset by the reduced failure costs. As the sample size increases, the best-fit normal distribution of approach A becomes, on average, closer to the true distribution, such as the example sample $\Sigma_3$ of size $n=200$ shown in Figure 10. For this sample, the optimal designs of the three approaches converge, and therefore yield similar utilities. For low levels of epistemic uncertainty, this results in value of approach B over A near zero.

Designers A and B chose a thickness ($t_A^*$ and $t_B^*$ respectively) that is optimal according to their approach and estimated expected utilities. However, they will not, on average, realize the expected utility estimates that they used to guide their selections of thickness. Instead, they will realize an expected utility based on the true curve, $E_{\sigma_t}[U(\sigma_t, t)]$, which the designers do not know.

In Figure 8 (based on sample $\Sigma_1$), the true curve is between the curve from approach A and the upper bound from approach B. By noting $t_A^*$ and $t_B^*$, we can read off the true expected utility evaluated under perfect knowledge for each approach from $E_{\sigma_t}[U(\sigma_t, t)]$. For this particular sample $\Sigma_1$, the expected utility realized from approach A is about $20$ higher than the expected utility realized from approach B. This means for sample $\Sigma_1$ the value of approach B is negative: $V(B) = -20$, indicating approach A performs better.

In Figure 9 (based on sample $\Sigma_2$), the true curve is near the lower bound of approach B. For this particular sample $\Sigma_2$, the expected utility from approach B is about $60$ more than that from approach A, so the value of approach B for sample $\Sigma_2$ is $V(B)=60$.

The results can also be understood in terms of the expected utility curves used in the experiment. In Figure 8 and Figure 9 we illustrate the expected utility as a function of thickness for two different samples, $\Sigma_1$ and $\Sigma_2$ respectively, both of size 30. The four curves shown in each figure are:

1. Estimated expected utility under approach A: $E_{\sigma_t,A}[U(\hat{\sigma}_t(A, \Sigma), t)]$
2. Estimated lower bound on expected utility under approach B: $E_{\sigma_t,B}[U(\hat{\sigma}_t(B, \Sigma), t)]$
3. Estimated upper bound on expected utility under approach B: $E_{\sigma_t,B}[U(\hat{\sigma}_t(B, \Sigma), t)]$
4. The true expected utility in a perfect state of knowledge: $E_{\sigma_t}[U(\sigma_t, t)]$

In many cases, approach B yields a design with a lower utility than approach A. However, in some cases, approach B yields a design with a much higher expected utility. The tails of the distribution for $n=10$ and $n=30$ extend much farther than shown in the figure. For example, the maximum expected value of approach B seen in any trial of $n=10$ was $95,000$. Results such as these skew the overall distribution such that on average, approach B yields a design with a higher expected utility than approach A yields. As the amount of epistemic uncertainty decreases, both the skewness and value of approach B over A also decrease.

Figure 6: Variation of value with amount of epistemic uncertainty

Figure 7: Histogram of value of p-box approach


**Summary of results**

For this design problem, the experimental results support the hypothesis that it is valuable to use a design approach that explicitly represents epistemic uncertainty distinctly from aleatory uncertainty. When the amount of epistemic uncertainty is large, such an approach performs significantly better. When the amount of epistemic uncertainty is small, the difference between the two approaches is insignificant. This computational experiment has therefore demonstrated that there are scenarios in which it is valuable to represent epistemic uncertainty explicitly and distinctly from aleatory uncertainty.

**DISCUSSION AND FUTURE WORK**

In this paper, we present a specific and simplified design problem and computational experiment that demonstrates the value of representing epistemic uncertainty in engineering design. The demonstration of the value in one design problem leads to the natural question: what are the characteristics of a design problem that make it valuable to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty? The specific use of PBA in this example leads to the question: are there more effective approaches making decisions under epistemic uncertainty? These questions form the basis for future work.

**Net value of representing epistemic uncertainty**

In this paper, we have demonstrated the potential value of a method that represents epistemic uncertainty explicitly and distinctly. However, designers are more interested in the gain, or net value, which is defined as the difference between value and cost [38].

One measure of cost is computational cost. The probabilistic approach often requires Monte Carlo analysis or related methods in order to calculate expected values [39]. These methods have an inherently high computational cost, even when used in combination with surrogate modeling [40, 41] to reduce the computational burden. In this example, the design problem was simplified to the point that Monte Carlo analysis could be avoided. First, all of the distributions considered were normal, and second, the utility function was simple enough that the expected utility could be calculated analytically using the cumulative normal distribution, a commonly available statistic.

This computational simplicity extended to the PBA approach in this example. The boundaries of the p-box were piece-wise normal CDFs. Thus, the expected utility interval could be calculated from the p-box using the cumulative normal distribution, too. Consequently, the computational cost of the two design methods was comparable. Since the PBA approach yielded a greater expected value under large epistemic uncertainty at a similar cost, in this example, the net value (gross value less cost) of moving to an approach based on PBA is positive, and is therefore worth further research. For example, the ability of PBA to propagate uncertainty in more complicated design problems should be examined.
Propagation of uncertainty using p-boxes

In this experiment, only one parameter is uncertain, so there is no need to combine uncertainties from different sources. However, most engineering design problems contain multiple sources of uncertainty and thus require a means to propagate uncertainty. PBA includes algorithms with foundations in interval analysis for propagating p-boxes through calculations [35, 42]. Ferson and Ginzburg [13] claim that these methods are on average less computationally expensive for computing with epistemic and aleatory uncertainty than double-loop Monte Carlo methods. However, it does not appear that these methods have been explored in detail for engineering design applications.

Another consideration is the flexibility of the algorithms. In Monte Carlo analysis, engineers often assume specific correlations between variables as a matter of convenience that enables the use of standard statistics. This understates the true epistemic uncertainty and can have serious consequences. PBA provides methods [35] that can propagate uncertainty under various conditions of dependence, including the extremes of fully known and completely unknown dependencies. Despite this promise, it is not yet clear how well PBA can propagate uncertainty from many sources, or how well PBA can be integrated with more complex design tools and models, such as discrete-event simulations and optimization methods. These issues need to be resolved before PBA can be put to use in general engineering design problems.

Variation in the true aleatory uncertainty

In this experiment, the true aleatory uncertainty was normally distributed. However, the nature of the aleatory uncertainty, such as its distribution type and parameters, may affect the value of various design approaches. For example, if the true distribution is not normal, then both of these approaches may perform much worse compared to the reference case of perfect information. It is therefore important to explore different cases of aleatory uncertainty.

Decision policies and preferences

In this experiment, the relatively high cost of failure, in comparison to material costs, means it is very important to avoid failure. Because the PBA approach—including the maximum decision policy—used in this experiment is conservative compared to a normal fit approach, it may be that PBA provides more value only when there is a strong incentive to avoid failure. Therefore the performance of PBA should be explored for different preference structures. Similarly, given an interval on expected utility, there are alternative (and less conservative) methods in interval theory [19] for making decisions that should be explored, including midpoint and maximum entropy policies.

SUMMARY

Engineers inherently lack knowledge during the engineering design process. Existing design methods do not explicitly represent this epistemic uncertainty distinctly from aleatory uncertainty. The pressure vessel design example in this paper demonstrates that when the designers only have access to a small set of sample data, a PBA approach to handling uncertainty in design decisions can lead to better designs than a purely probabilistic approach that emphasizes only aleatory uncertainty. It is therefore valuable to explicitly represent epistemic uncertainty distinctly from aleatory uncertainty in at least some engineering design problems.

ACKNOWLEDGMENTS

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. This research is further sponsored in part by NASA Ames Research Center under cooperative agreement NNA04CK40A. Additional support is provided by the G.W. Woodruff School of Mechanical Engineering at Georgia Tech. The authors would like to thank Rich Malak, Steve Rekuc, Jay Ling, Morgan Bruns, and Scott Duncan for their valuable insights and feedback.

APPENDIX A: CONSTRUCTING A P-BOX FROM LIMITED DATA

PBA is a recently developed method, and most publications have focused on introducing the concept of PBA, explaining its theoretical value, and introducing methods for computing with p-boxes [33, 35, 43]. Only more recently have researchers addressed the construction of p-boxes from observed sample data [36].

While several approaches exist [36], we choose a pragmatic approach based on standard confidence intervals. In this example, the designers know the true yield strength is normally distributed, but with unknown mean and variance. For clarity here, we denote the random variable for yield strength as $X$ to distinguish it from standard statistical notations that use $\sigma$ for the standard deviation. Thus we have:

$$X \sim \text{Normal}(\mu, \sigma^2).$$

Because the true $\mu$ and $\sigma^2$ are unknown, the designer uses the unbiased point estimates

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad \hat{\sigma}^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

where the $X_i$’s are the sample observations, and $n$ is the sample size. These quantities are called respectively the sample mean and sample variance. So far this is exactly what we have defined as the pure probabilistic approach. In order to construct a p-box, we broaden these point estimates to confidence intervals. For the mean we find a 100(1- $\alpha$)% confidence interval as

$$[\mu_{\alpha/2}, \mu_{1-\alpha/2}] = [\hat{\mu} - z_{\alpha/2}/\hat{\sigma}_X, \hat{\mu} - z_{1-\alpha/2}/\hat{\sigma}_X]$$

where $\hat{\sigma}_X = \sqrt{\hat{s}^2}$.

$$z_{\alpha/2} = \Phi^{-1}(\alpha/2).$$

A table of these values is found in most probability and statistic books, such as Devore [45]. In this experiment, a 99% confidence interval is used, and thus $z_{0.05/2} = 2.58$. 

\[10\] Copyright © #### by ASME
The 100(1 - α)% confidence interval on the variance estimate is a bit more complicated, but it can be found [46] to be

\[ \left[ \sigma^2, \bar{\sigma}^2 \right] = \left[ s^2 - z_{\alpha/2} \sqrt{\text{var}(s)}, s^2 - z_{\alpha/2} \sqrt{\text{var}(s)} \right] \]

where: \( \text{var}(s^2) = \frac{(n-1)\bar{s}^2}{n} \).

APPENDIX B: DETAILS OF EXPERIMENT

As discussed in the body of the paper, the experiment consists of the design of a pressure vessel by different design approaches. In this appendix, we present the details of this experiment, including assumptions and numerical values.

Goal:
The goal of this problem is to design a pressure vessel with volume of 0.15 m³ that can sustain an internal gas pressure of 7 MPa such that the expected utility of the design is maximized subject to the following constraints. As noted in the body of the paper, the utility function is given by

\[ \text{Utility} = P_{\text{selling}} - C_{\text{material}} \cdot \text{volume} - C_{\text{failure}}. \]

In this utility function, material cost is directly proportional to the volume of material used. This assumption is a simple aggregation of total cost of manufacture. Because the pressure vessel will be used in a location occupied by humans, the cost of failure is relatively high when the vessel fails.

Constraints and assumptions

- Total length of pressure vessel less than 1.5 meters
- Total height of the pressure vessel less than 0.6 meters
- Vessel internal pressure P=7MPa
- Manufacture of the pressure vessel results in a uniform thickness \( t \), and a constant material strength throughout
- It is assumed the pressure vessel will fail due to the stress in the walls. Therefore, the critical point for failure is when the maximum stress in the material of the walls exceeds the yield strength, or \( \sigma_{\text{max}} > \sigma_y \).
- The maximum stress \( \sigma_{\text{max}} = \max(\sigma_c, \sigma_s) \), where \( \sigma_c \) is the stress in the cylindrical section of the pressure vessel and \( \sigma_s \) is the stress in the spherical section. These quantities are given by the following approximations [47], which are meaningful for \( R \geq 5t \):
  \[ \sigma_c = P \cdot \left[ (R + 0.5t)/2t \right] \]
  \[ \sigma_s = P \cdot \left[ (R + t)^2 + R^2 \right]/(2Rt + t^2) \]

REFERENCES


