Introduction to Uncertainty Quantification
Module 1.1: Types of Uncertainty

1 What is Uncertainty?

Uncertainty is something that makes us all uncomfortable. In our personal lives, as in our professional lives, we aim to avoid uncertainty wherever possible. But, we cannot escape it. Indeed it permeates every aspect of our lives and plays a huge role in the decisions that we make. Whether conscious or unconscious, uncertainty influences how we live our lives. Reflecting on your life, how many times have you made a decision based only on the fact that one outcome was predictable and another was not – even if that outcome was less than ideal or even undesirable? How many times have you taken the path with higher uncertainty? Perhaps that path has higher potential rewards, but it may also have greater potential downside or consequences.

These decision occur frequently; maybe even daily. But we seldom consciously make these decisions by formally analyzing uncertainties. Maybe you make a pros and cons list, or use a similar mechanism to aid us in your decision making. But, do you ever quantitatively evaluate the risk associated with a decision? Unless you're trained in such methods, and do so for a living, then the answer is most likely no. If you work in an industry where decisions are made under uncertainty (and this is true in almost all industries), then your decision-making processes can most likely benefit from a quantitative evaluation of uncertainty, or uncertainty quantification (UQ).

This is especially true in engineering analysis where uncertainties can have significant consequences in terms of product/system performance, cost (of development, maintenance, and operations), and safety. As we'll see, uncertainties are ubiquitous in engineering systems of all types. But the benefits of UQ are not limited to engineering. Although this course will focus largely on engineering applications, uncertainties play a major role in decision making for other industries including, but not limited to: insurance risk assessment; both macro- and micro-economic supply and demand projections; economic, environmental, social or other policy impacts; meteorology and weather forecasting; corporate financial / stock market projections; and countless others. The principles we discuss here are universally applicable across all of these domains, and many more.

Uncertainty quantification is therefore an important component of many, if not most, decision making processes. The aim of this course is to introduce you to the essential tools necessary to perform UQ and integrate the results into your decision making. Before we do so, we must first develop an understanding of uncertainty. Much the way we handle uncertainties subconsciously in our daily lives, we likewise do not often contemplate the meaning of uncertainty. But, in order to gain a quantitative grasp of uncertainties, and how to treat them, we must first learn to define and identify them. So we ask "What is uncertainty?"

2 Types of Uncertainty

If we simply look in the dictionary, it's difficult to find a definition of uncertainty that provides insight into the nature of uncertainty. For example, www.dictionary.com provides the following six definitions for uncertain:

1. not definitely ascertainable or fixed, as in time of occurrence, number, dimensions, or quality;
2. not confident, assured, or free from hesitancy;
3. not clearly or precisely determined; indefinite, unknown;
4. vague; indistinct; not perfectly apprehended;
5. subject to change; variable; capricious; unstable;
6. ambiguous; unreliable; undependable;

Certainly these definitions tell us what uncertainty means, but to develop quantitative tools to deal with uncertainty we need to develop a deeper understanding of uncertainty. We need to know where uncertainty comes from, why it arises, and ultimately how we can treat it. We begin by introducing two types of uncertainty: aleatory uncertainty and epistemic uncertainty.

2.1 Aleatory Uncertainty

Aleatory uncertainty is uncertainty that arises from inherent randomness in the data generation and/or collection process. The word *aleatory* derives from the Latin *alea*, which means the rolling of dice, gambling, or a game of chance. As such, aleatory uncertainty deals with outcomes that cannot be known because they are random in nature. Because aleatory uncertainties are random, they cannot be reduced or eliminated. For this reason, they are often referred to as *irreducible uncertainties*.

Although we cannot reduce or eliminate aleatory uncertainty, we can often statistically characterize or quantify it. Consequently, aleatory uncertainty may also be referred to as *statistical uncertainty*. The statistical characterization of aleatory uncertainty is typically unbiased. That is, statistics describing aleatory uncertainties tend toward certain values and repeated (fair) observations will not tend to systematically overestimate or underestimate those values. For example, if the mean value of an aleatory quantity is 5, then the statistical average of many observations of this quantity will tend toward 5 and will not tend systematically to another value. We will discuss this more in Section 3 below.

Aleatory uncertainties are ubiquitous and examples are all around us because they reflect results that are either truly random or so complex that the outcome is unpredictable. Some generic examples include:

- A coin flip:

  Why is it aleatory? In theory, a coin flip is predictable. If we know the exact motion of the coin (including initial position, velocity, etc.), the exact properties of the material (density, stiffness, etc.), the exact geometry of the coin, the exact fluid properties of the air (density, temperature, humidity, etc.), and the exact material properties of the surface on which the coin will land then Newtonian mechanics will give us an accurate prediction of the outcome. But in reality many of these things are unknowable. For example, even with careful attention to exact repetition, a human throwing a coin will make subtle changes that drastically alter the initial trajectory of the coin and change the outcome.

- The roll of a die:
  Why is it aleatory? Same reasons as the coin flip.

- Radioactive decay:

  Why is it aleatory? At the atomic scale, radioactive decay is a stochastic process. According to quantum mechanics, we cannot predict when a particular atom will decay, regardless of how long the atom has existed.

- Measurement Noise:

  Why is it aleatory? Sometimes, the measurements we make are polluted with noise. When this noise is random and we cannot reduce it, we would categorize it as aleatory.

Since our particular interest is in engineering applications, we also provide the following examples in engineering:

- Properties of Engineering Materials:

  Why is it aleatory? Every engineering material has small variations in its structure, even when produced under identical conditions. These small variations give rise to deviations in material properties that cannot be tested. We cannot test the strength of every material specimen that is used in the construction of a building, bridge, automobile, or aircraft. Therefore, we must recognize that there are random variations from one component to the next.

- Loads on a structure:

  Why is it aleatory? It is impossible to know what loads a structure will see in the future. Whether we're considering loads due to seismic ground motions on a building, turbulent wind pressures on a long-span bridge, or ocean wave loads on an offshore platform – these loads are inherently random and cannot be known in advance.

- Demands on an engineering system:

  Why is it aleatory? Although we can ascertain general trends in the demand on engineering systems such as power grids, transportation networks, and water distribution systems, we cannot predict the exact demand at any given point in time. The demand follows a random process driven by the behavior of all users of the system simultaneously.

2.2 Epistemic uncertainty

The word epistemic derives from the Latin *episteme*, which means knowledge. *Epistemic Uncertainty* is therefore caused by a lack of knowledge or data. It may also be referred to as *systematic uncertainty* Because epistemic uncertainty is associated with a lack of knowledge, we have the opportunity to improve our state of knowledge and therefore reduce this form of uncertainty. For this reason, it is also commonly referred to as *reducible uncertainty*. Moreover, it is (at least in principle) possible to eliminate epistemic uncertainty in the presence of perfect knowledge. That is, a variable with epistemic uncertainty is actually deterministic. In other words, it possesses a true value. We simply don't know that true value.

In practice, reducing epistemic uncertainty can prove very difficult and impractical for a number of reasons. Perhaps the most common reason is that the cost associated with collecting the data, performing the work, or acquiring the equipment necessary to reduce epistemic uncertainty can be prohibitive. Moreover, we may be limited in the data we can collect due to technological factors and/or sociopolitical factors. For example, if we consider the epistemic uncertainty associated with the precision of a measurement instrument. The precision of the instrument may be limited by existing technology and hence we cannot, in practice, reduce the uncertainty because we cannot build a better instrument. Additionally, within an organization or society, there may simply not be the will to invest in reducing a given uncertainty – even if that uncertainty is reducible.

Epistemic uncertainty may also be biased. That is, uncertain observations of an epistemic variable may not accurately reflect the underlying truth. There are many possible reasons that epistemic uncertainties may be biased. For example, a common approach to handling epistemic uncertainty is to consult subject matter experts. Regardless of their level of expertise, these expert opinions may be biased (even from the most esteemed experts). Returning to the issue of epistemic uncertainty in measurements, we may find that the instrument we use to make a measurement introduces some bias; perhaps systematically underestimating or overestimating a value. Finally, epistemic uncertainty may be biased if a mathematical model used for the system makes biased predictions of the system output or response.

As we will see in the upcoming sections, epistemic uncertainty can be more difficult to quantify and the mathematical tools used to quantify it are not so clear to identify. But first let's look at some examples of epistemic uncertainty.

- Determination of the fairness of a coin or die:

  Why is it epistemic? Although the flipping of a coin or rolling of a die has aleatory uncertainty (the outcome is random), this assumes a fair coin or die where the probabilities of a given outcome (i.e. heads or tails) are known. Consider the case where we don't know that the coin (or die) is a fair coin having equal probability of all outcomes. Perhaps the coin is slightly weighted such that heads is modestly more likely than tails. This uncertainty in the probability of each outcome is epistemic because we can, in principle, know the probabilities of each outcome even if they are not equal. It would require us to flip the coin (roll the die) a very large number of times and observe the frequency of occurrence of each outcome.

- Simplifying assumptions or inadequacies in a model:

  Why is it epistemic? When we simplify a model or reduce it's accuracy, this produces uncertainty in the outcome of the system that the model aims to predict. This uncertainty is epistemic because, in principle, we can improve the model and thereby reduce uncertainty in predictions. As stated previously, this particular form of epistemic uncertainty is likely to be biased.

- Imprecision in a measurement:

  Why is it epistemic? Here, we want to distinguish between imprecision in a measurement (which is epistemic) and noise in a measurement (which may be random and therefore aleatory). In principle, we can reduce imprecision by developing a better (i.e. higher resolution) instrument. Therefore, the uncertainty is epistemic. In the case of noise, it may be possible to reduce the noise – in which case the uncertainty is partly epistemic. But the uncertainty associated with the random noise itself is aleatory because it is random in nature.

- Data that is missing, withheld, or hidden:

  Why is it epistemic? Missing data is a common source of uncertainty and is epistemic in nature because, if those data are recovered or provided, the uncertainty can be reduced.

Again, some examples in engineering include:

- A lack of data characterizing the properties of an engineering material:

  Why is it epistemic? Simply put, we can collect more data to characterize the material and reduce uncertainty.

- A physics-based model that has missing components, such as a model for the trajectory of a ball that neglects wind resistance.

  Why is it epistemic? In this case, we can improve the model by introducing the missing physics and therefore reduce the uncertainty.

- Inability to resolve small features in a medical image

  Why is it epistemic? If medical images lack resolution, they provide an imprecise measurement. Imprecision can, in principle, be reduced by developing a better imaging device, which would reduce uncertainty.

- Instrumentation failure or loss of signal causing intermittent data collection

  Why is it epistemic? In this case, we can reduce uncertainty by fixing the instrument and collecting data that does not have these faults.

2.3 Mixed Uncertainty

It is very common for a given problem to have sources of both aleatory and epistemic uncertainty. In these cases, it's not always trivial to determine which uncertainties are aleatory and which are epistemic. This was the case in the previous example related to measurement noise. When noise is present in an observation, the presence of the noise is likely epistemic because we can perhaps reduce the noise. But the noise itself is aleatory because of the inherent randomness of its variations. More generally, we might consider that we have a variable in our problem that is inherently random and can be characterized by probabilities. But, we may not know enough to determine those probabilities. In such general cases, we have an aleatory variable that has associated epistemic uncertainty. This case arises quite frequently and can be particularly difficult to handle.

References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis: the primer, John Wiley & Sons, 2008.

[2] I. Sobol´, Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp. 1 (1993) 407.

[3] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering & System Safety 52 (1) (1996) 1–17.

[4] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Computer physics communications 145 (2) (2002) 280–297.

[5] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability engineering & system safety 93 (7) (2008) 964–979.

[6] T. Crestaux, O. Le Maıtre, J.-M. Martinez, Polynomial chaos expansion for sensitivity analysis, Reliability Engineering & System Safety 94 (7) (2009) 1161–1172.