Introduction to Uncertainty Quantification
Module 1.3: Mathematical Tools for Aleatory and Epistemic Uncertainty

1 Interpretation of Probability

One of the most important mathematical theories for the treatment of uncertainty is probability theory. We will introduce concepts in probability theory in the next module. But, before we dig into the theory, it's important that we understand the concept of probability. To understand the concept, we need to answer the question "What does it mean to say that an event will occur with a certain probability?" It turns out, the answer is not necessarily simple. Throughout the history and evolution of probability, its meaning has been interpreted in different ways.

These interpretations began with the *classical interpretation* of probability, which emerged from the original works of Bernoulli and Laplace. The classical interpretation of probability deals strictly by reducing all possible outcomes to a set of events that are equally likely. The probability is then the ratio of the number of possible cases that are produce a specified outcome to the total number of possible outcomes. This is a consequence of the *principle of indifference*, which states that unless evidence shows otherwise all possible outcomes should be treated as equally possible.

The classical interpretation has come under a great deal of criticism since its inception and it's not our intention to discuss this in detail. The important result of this criticism is that, from the classical interpretation, two common philosophies have emerged that provide different interpretations of probability: the *frequentist interpretation* and the *Bayesian interpretation*. To this day, the differences in interpretation cause contention among mathematicians and philosophers of science that remains unresolved. Our intention, at this juncture, is to introduce these two interpretations as motivation for our chosen means of quantifying uncertainty.

1.1 Frequentist Probability

The *frequentist interpretation* states that the probability of an event is the limit of its relative frequency over a very large number of trials (i.e. the long-run frequency of occurrence). Stated mathematically, the probability of an event $\mathcal{A}$ is given by

$$P(\mathcal{A}) = \lim_{n\to\infty} \frac{n_a}{n} \tag{1}$$

where n_a is the number of observations of event $\mathcal{A}$ and n is the total number of experiments.

The frequentist interpretation states that probability is objectively defined only through a well-defined experiment that can be repeated a large (theoretically infinite) number of times. In other words, subjectivity, belief, and opinion play no role in the determination of probabilities. Probabilities are therefore inherently deterministic; having true underlying values that can only be discovered through a set of repeated objective experiments. This, of course, gives rise to some criticism about how probabilities can be treated in the absence of many repeated experiments and more generally, how uncertainty in the identification of deterministic probabilities can be treated. This will be discussed more in the following sections.

1.2 Bayesian Probability

The *Bayesian interpretation*, on the other hand, states that probability expresses our current belief, given the present state of knowledge, in the chances of an event occurring. That is, Bayesian probabilities are *subjective* and therefore two individuals can reasonably assign different probabilities to the same event, even if they are presented with the same evidence. This is because different individuals may start with different *prior* beliefs. We will formalize this later, but for now we can suffice it to say that they start with different beliefs before the evidence is presented and therefore they arrive at different probabilities after the evidence is presented.

A consequence of this interpretation is that probabilities are not necessarily deterministic. The probability of an event is uncertain, and can be treated with probabilities. Although the Bayesian mathematical formalism provides a means to estimate this uncertainty (addressing a critical shortcoming of the frequentist interpretation), it also gives rise to certain criticisms as we will see in the following sections.

2 Treatment of Uncertainties

Given the different interpretations of probability, it's natural to ask how we should mathematically treat the many uncertainties we encounter – whether in engineering analysis, data science, financial modeling, social sciences, or otherwise. Should we rely on the frequentist interpretation and treat probabilities as absolute truth that can only be discovered through many observations? Should be treat uncertainties using the Bayesian interpretation and allow our beliefs, prior knowledge, and existing data to influence our assessment of uncertainties? Or should we use different tools altogether?

The answer to these questions is not trivial and their answers may be governed by the circumstances, the objectives of our investigation, and the type of uncertainty that we're concerned with. Indeed, it is very often suggested that aleatory and epistemic uncertainties require different mathematical treatments (e.g. [1–3]). However, as we'll see, no scientific consensus on the "correct" approach has emerged. In fact, there remains a vigorous scientific debate over the appropriate mathematical framework in which to treat epistemic uncertainty in particular. In the following sections, we'll explore how aleatory and epistemic uncertainty are treated in turn.

2.1 Aleatory Uncertainty

Recall that aleatory uncertainty is associated with inherent randomness. For this reason, aleatory uncertainty is naturally treated with probability theory – which deals with the occurrence of random events. Regardless of the interpretation (Bayesian or frequentist), probability theory is universally used for the treatment of aleatory uncertainty. Aleatory uncertainties are treated by assigning probability distributions to the possible outcomes of a random experiment (where any aleatory outcome is, by its nature, a realization of a random experiment). When provided with data from a set of aleatory experiments, statistical methods (again Bayesian or frequentist) can be used to quantify the associated distributions and/or properties of their distributions (e.g. statistical moments, see Module 2). In the frequentist interpretation, the probability of an event is identified as the long-run frequency of occurrence of that event from many experiments and moments are estimated using standard statistical methods. In the Bayesian interpretation, prior beliefs / knowledge / assumptions are used to infer probabilities from data. This will be covered in detail in Module 5.

2.2 Epistemic Uncertainty

Epistemic uncertainty is more difficult to treat mathematically because epistemic uncertainties are associated with quantities / outcomes that are inherently *deterministic*, but nonetheless unknown. In other words, the true value exists. We simply don't know it. Since probability theory deals specifically with outcomes that are *random*, it has been argued that probability theory is not necessarily appropriate for treating ignorance [1]. Others argue that probability theory can be used when interpreted from a Bayesian perspective thanks to its subjectivity and specifically its ability to quantitatively update beliefs probabilistically from data. Yet, the matter is far from resolved. There remains a strong philosophical debate over the manner in which epistemic uncertainties should be appropriately treated. We *will not* attempt to resolve this debate. Instead, the following presents a brief summary of some of the prevailing approaches. In this course, we will specifically apply a Bayesian probabilistic approach, with full recognition that criticisms and shortcomings of this approach remain.

Note that the following review is not intended to be comprehensive, nor is it intended to provide theoretical details for any of the approaches. It is intended as a *high-level, conceptual* presentation intended to introduce some of the mathematical approaches that have been used for quantifying epistemic uncertainty.

2.2.1 Bayesian Probability Theory

From the Bayesian perspective, an epistemic variable is treated mathematically as a random variable (see Module 2). But here, the randomness in the outcome is not interpreted as true randomness. Instead, randomness is a means of modeling ignorance. Inherently, we understand that a true *deterministic* value exists for an epistemic variable, but we assume that different possible true values have associated probabilities. There are several advantages of this approach, as we'll see in more detail in Module 5. First, there is a direct means of incorporating our current (or *prior*) knowledge. This prior knowledge is then easily updated to account for new data. That is, as new data are collected Bayesian probability allows us to probabilistically update our beliefs and, as more and more data are collected, reduce the uncertainty and converge toward the underlying truth (again, assuming that a true value exists).

Philosophically, some critics argue that there is no basis for modeling ignorance as randomness. Moreover, these critics argue that assigning probability distributions to epistemic variables can, under certain circumstances, equate to asserting knowledge or information that one cannot justify. The fundamental example of this is derives from the *principle of indifference*, discussed above, which states that in the lack of evidence all possible outcomes should be treated as equally probable. Critics of this viewpoint would point out that there is, in fact, no justification for treating all outcomes as equally probable in this case and that doing so imposes an arbitrary belief.

Indeed, treating uncertainty as randomness is a fundamental assumption of using probability theory, whether Bayesian or otherwise, to model epistemic uncertainty. Some probabilists, however, have provided justifications for this assumption. Critics of the Bayesian approach specifically argue that it places too much emphasis on and is highly influenced by (subjective) prior beliefs as discussed above. This has been addressed, to some extent, through the introduction of hierarchical Bayesian probabilities, in which we consider *probabilities of probabilities* (or meta-probabilities). In hierarchical Bayesian approaches, prior knowledge is treated probabilistically, again using Bayesian probability and introducing a so-called *pre-prior* belief that is the "prior belief of the prior belief." This can continue hierarchically for as many levels as one desires, but quickly descends into abstractions whose consequences are beyond our scope.

2.2.2 Set Theory & Intervals

Perhaps the most commonly employed non-probabilistic approaches for treating epistemic uncertainty use *set theory* and *intervals*. These approaches define sets, or intervals, of possible outcomes but do not attempt to associate any measure (e.g. a probability) to the members of the set. Instead, these models deal strictly with outcomes that are either possible (in the set) or not possible (not in the set) and make no attempt to judge the likelihood of these outcomes. These methods have a rigorous mathematical foundation, but have the drawback of also being restrictive, or rigid, in regard to the integration of existing knowledge. These approaches do not lend themselves naturally, for example, to cases where some evidence does in fact exist that a certain outcome (or outcomes) is (are) more likely than others.

2.2.3 Imprecise Probabilities

A variety of methods have been made to generalize probability theory that we will generally refer to as *imprecise probabilities*. These generalizations – which include the imprecise probability formulation of Walley [4], Dempster-Shafer evidence theory [5, 6], possibility thoery [7], and fuzzy sets [8], among others – aim to strike a compromise between the rigidly non-informative approach of set theory and the (perhaps overly or incorrectly) informative approach of Bayesian probability theory. They do so in different ways, but they all work notionally by introduce some generalized measure of imprecision by working with sets of probabilities. This is, of course, an over-simplification and does not attempt to express the nuances of these methods in any theoretical detail. The interested reader is encouraged to explore these different mathematical theories, perhaps by starting as simply as exploring the Wikipedia pages for these theories or through a more rigorous investigation using the references provided. Excellent overviews of these methods are provided in [9] for machine learning applications and [10, 11] for engineering analysis.

3 Further Inquiry

As we can see from this brief introduction, the mathematical treatment of epistemic uncertainty is not straightforward. Indeed, the scientific community has not yet reached a consensus on its appropriate treatment. We have attempted to provide a brief, **conceptual** introduction to the prevailing approaches and have deliberately avoided any formal mathematical formulation of these theories. However, students are highly encouraged to explore these various theories on their own. Some select foundational references are provided below, while less formal introductions can be found by exploring the following Wikipedia pages:

- https://en.wikipedia.org/wiki/Uncertainty_quantification
- https://en.wikipedia.org/wiki/Bayesian_probability
- https://en.wikipedia.org/wiki/Frequentist_probability
- https://en.wikipedia.org/wiki/Imprecise_probability
- https://en.wikipedia.org/wiki/Dempster-Shafer_theory
- https://en.wikipedia.org/wiki/Fuzzy_set

As a cautionary note, these sites can provide some rudimentary mathematical introduction to the concepts discussed in this lesson but should not be used as primary references. Indeed, the descriptions on these pages are incomplete, imprecise, and potentially erroneous. In fact, one might say that these sites have high epistemic uncertainty. As primary references, students are encouraged to review the reference list below.

References

[1] S. Ferson, L. R. Ginzburg, Different methods are needed to propagate ignorance and variability, Reliability Engineering & System Safety 54 (2-3) (1996) 133–144.

[2] M. H. Faber, On the treatment of uncertainties and probabilities in engineering decision analysis (2005).

[3] A. Der Kiureghian, O. Ditlevsen, Aleatory or epistemic? does it matter?, Structural safety 31 (2) (2009) 105–112.

[4] P. Walley, Statistical reasoning with imprecise probabilities (1991).

[5] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, in: Classic works of the Dempster-Shafer theory of belief functions, Springer, 2008, pp. 57–72.

[6] G. Shafer, A mathematical theory of evidence, Vol. 42, Princeton university press, 1976.

[7] D. Dubois, H. Prade, Possibility theory: an approach to computerized processing of uncertainty, Springer Science & Business Media, 2012.

[8] L. A. Zadeh, Fuzzy sets, Information and control 8 (3) (1965) 338–353.

[9] E. Hüllermeier, W. Waegeman, Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods, Machine Learning 110 (3) (2021) 457–506.

[10] S. Ferson, C. A. Joslyn, J. C. Helton, W. L. Oberkampf, K. Sentz, Summary from the epistemic uncertainty workshop: consensus amid diversity, Reliability Engineering & System Safety 85 (1-3) (2004) 355–369.

[11] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses, Mechanical systems and signal processing 37 (1-2) (2013) 4–29.

Nomenclature

Functions

$P(\cdot)$ Probability measure

Variables

$\mathcal{A}$ An event

n The total number of experiments

n_a The number of observations of event $\mathcal{A}$