Mechanical Systems and Signal Processing 37 (2013) 30–42

Contents lists available at SciVerse ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Analysing uncertainties: Towards comparing Bayesian and interval probabilities'

David Blockley

Emeritus Professor and Senior Research Fellow, University of Bristol, Bristol, UK

ARTICLE INFO

Article history:
Received 8 August 2011
Received in revised form
14 May 2012
Accepted 18 May 2012
Available online 15 June 2012

Keywords:
Uncertainty
Probability
Incompleteness
Imprecision
Bayes theorem

ABSTRACT

Two assumptions, commonly made in risk and reliability studies, have a long history. The first is that uncertainty is either aleatoric or epistemic. The second is that standard probability theory is sufficient to express uncertainty. The purposes of this paper are to provide a conceptual analysis of uncertainty and to compare Bayesian approaches with interval approaches with an example relevant to research on climate change. The analysis reveals that the categorisation of uncertainty as either aleatoric or epistemic is unsatisfactory for practical decision making. It is argued that uncertainty emerges from three conceptually distinctive and orthogonal attributes FIR i.e., fuzziness, incompleteness (epistemic) and randomness (aleatory). Characterisations of uncertainty, such as ambiguity, dubiety and conflict, are complex mixes of interactions in an FIR space. To manage future risks in complex systems it will be important to recognise the extent to which we 'don't know' about possible unintended and unwanted consequences or unknown–unknowns. In this way we may be more alert to unexpected hazards. The Bayesian approach is compared with an interval probability approach to show one way in which conflict due to incomplete information can be managed.

1. Introduction

Two assumptions, commonly made in risk and reliability studies, have a long history. The first is that uncertainty is either aleatoric or epistemic ([14]). The second is that standard probability theory is sufficient to express uncertainty (for recent discussions see Der Kiureghian and Ditlevsen [12], Aven [1], Dubois [13]). However as Corotis [10] states 'The world is less precise than our models, even our probability ones … advances … require bridging the gap … between theory and uncertainty measures'.

Hacking [14] traces the history of chance right back to the talus, the heel bone of a running animal such as a deer or horse used as an ancient predecessor of the die but then wonders why theories of frequency, betting, randomness and probability did not surface until the 15th century and were not really developed until the 17th with the work of Pascal. Probability emerging from that time is essentially dual—aleatory (gamble or stable frequency) and epistemic (belief). Aleatory is effectively another name for randomness and epistemic is 'what we know'.

Der Kiureghian and Ditlevsen [12] recognise that a characterisation of uncertainty is a pragmatic choice according to the purpose of a model. They accept that there is a 'degree of subjectivity in the selection of models …. It constitutes the "art" of engineering … that most distinguishes the quality of its practice'. They admit that from a linguistic point of view all uncertainties 'are the same as lack of knowledge'. They go on to state that the choice of model depends on context,

E-mail address: D.Blockley@bris.ac.uk

0888-3270/$ - see front matter
http://dx.doi.org/10.1016/j.ymssp.2012.05.007

application and convenience. They remark that aleatoric uncertainty is intrinsic and irreducible whilst epistemic uncertainty can be reduced by using auxiliary non-physical variables defining statistical dependencies which arise among different components that have common uncertainties. By asserting that human errors tend to be random in nature and hence aleatoric they seem to fail to take into account the whole literature of the social scientists ([17], [18], [20]) and their theories of human and social explanations of failure.

The mathematics of probability has to be interpreted by engineering practitioners when used to inform decision making. I will argue that classical probability theory (with a completeness axiom that requires probabilities on the sample space to add up to unity) is inadequate for the task of managing risks in complex systems. The paper is aimed at practitioners and its purpose is (1) to discuss a deeper conceptual analysis of uncertainty by contrasting its use in natural languages with formal languages; (2) to compare Bayesian approaches and contrast them with interval approaches and (3) to provide an example relevant to research on climate change.

2. Risk

Risk is in the future. Managing it requires us to form judgements concerning evidence about the past and current states ([7]) together with projected possible future states so that we can make decisions that will lead to future success or failure. The new process ([8]) of getting to that future state may be thought of in terms of activities e.g., travelling by car, climbing some scaffold; or in terms of hazards e.g., a live electric wire, or flooding; or in terms of events e.g., an accident, industrial action; or in terms of consequences such as financial loss or death. For numerous examples of this approach in civil engineering and construction see [2,5].

Low chance risks with extremely high consequences are particularly difficult to deal with because data and experience is almost always sparse and often ill-defined. The more we understand of the complex interconnectedness and interdependence of natural and man-made systems the more likely the often made assumptions of independence by some expert analysts may be shown to be misleading. In an age of complexity we need specifically to acknowledge vagueness and imprecision in situations where we genuinely 'don't know' or are unsure.

3. The nature of evidence

Evidence is information that proves or disproves, supports or contradicts, makes plain or clear, an indication or a sign of something. It is crucial in managing risk. It may take the form of a measurement, a judgement, opinion or even a gamble and it may be a written document or a verbal statement or testimony. Inevitably evidence will be uncertain.

In the 17th century the word probability meant the extent to which it is right to do something i.e., worthy of approbation, approval, commendation and sanction by an authority such as the church or a wise person ([14]). Knowledge was not simply justified true belief but rather was necessary (true in all contexts) demonstrable and universal. Opinion was not demonstrable but referred to belief resulting from reflection or argument. Increasing the probability of opinion might bring certain belief but it would not bring knowledge. At that time testimony as the evidence of witnesses and authority was ample but evidence provided by things as they demonstrably appeared in reality was lacking. The latter idea was required in order to state the problem of induction which David Hume did in 1739. Hume doubted that evidence from the past could be used as evidence for the future. Such scepticism however does not prevent people from distinguishing good inductive reasons from bad ones and a new logic of induction was born. Frequency and credibility were linked and a dual concept of probability became possible—men could start to order the different degrees to which hypotheses are supported. Knowledge was not now considered to be universal truth but rather the use of first principles, demonstrations and comparisons of ideas. Cause was not necessary (i.e., in all possible circumstances) for effect but rather constant 'habitual' conjunctions. Evidence that provides good reason for believing a proposition could be represented in a formal language and that language was probability. A new logic of induction was born where objective evidence that provides good reason for believing a proposition can be represented in a formal language. That language was probability and $p(A/E)$ becomes a measure of the degree to which the evidence (E) is a reason for a proposition (A).

As a consequence of these historical developments aleatory uncertainty has come to be seen by many theorists as 'objective' in the sense it does not depend on an observer. Epistemic, on the other hand is 'subjective' and personal. The word aleatoric derives from gambling and other combinatorial problems—it means 'dependant on chance or accidental events or other contingencies'. It is unclear how this is different from randomness in a stochastic process and here so we will consider aleatory, random and stochastic processes as equivalent. The modern use of the word epistemic means 'pertaining to knowledge or the conditions for acquiring it'. This is insufficient for capturing the nuances in the meaning of uncertainty since actually all uncertainties are due to limitations in what we know or the conditions for acquiring or understanding. In this sense all uncertainty is knowledge-based. For example random uncertainty is epistemic since randomness is the lack of pattern or specific order in data. It must therefore depend on our understanding and consequent models of what we mean by order. Some (but not all) of the epistemic uncertainties seem to arise from assumptions made in the modelling of a system, the methods of analysis and limitations of data. They may be reduced by using more resource to perform more detailed analysis and obtaining more data but there will always be a residual uncertainty because we can never have certainty—there is always a gap between what we think we know and what we do resulting in the possibility of unintended and unwanted consequences. It is this residual uncertainty which is the real challenge to risk managers since it can sometimes lead to unintended, unforeseen and unwanted disastrous consequences such as 9/11, Chernobyl, the near collapse of the world's banking system in 2008 and the damage to nuclear reactors by the tsunami after the 2011 earthquake in Japan. Social scientists such as Turner and Pidgeon [20] have argued that the preconditions to major disasters can incubate or develop in a way that it may be possible in some instances to identify before a final disastrous event. They argue that we need to develop methods for identifying those preconditions with sufficient dependability to enable decision makers to make such politically difficult and potentially expensive decisions to avoid the even greater costs and consequences of a disaster.

In this paper we will define judgements about evidence as sensible opinions or beliefs expressed in a context by experienced people such as practicing engineers and medics. We consider only those people who are particularly qualified to make considered decisions based on those judgement and for which they are held to account by society through a legal duty of care. The engineer and the medic are not necessarily concerned with the absolute truth of a theory or model—rather they are concerned with his responsibility to act on the basis of a theory or model. This we call the dependability of a proposition ([3,4]). This taking of responsibility implies not that he has earned the right to be right or even sufficiently right but that he has taken precautions that can be reasonably expected to take against being wrong.

4. Three independent parameters of uncertainty FIR

In previous work ([5]) uncertainty was classified using three conceptually distinctive and orthogonal separate characteristics FIR—fuzziness, incompleteness and randomness. The motivation was that the aleatoric/epistemic classification is not rich enough for practical decision making. Fig. 1 illustrates previous assertions that other characteristics of uncertainty such as ambiguity, confusion, contingency, indeterminacy and conflict emerge from mixes and interactions between these basic three attributes as they are interpreted in any real world context. However before discussing Fig. 1 we need first to state clearly what we mean by FIR and its relationship with aleatoric and epistemic uncertainty.

Fuzziness is imprecision or vagueness of definition and is implicit in such statements as 'this is a long span bridge' or 'the residual stresses in this welded steel structure are considerable'. It is implicit in assumptions about levels of performance in earthquake engineering such as immediate occupancy, life safety and prevention of collapse. It is an entire subject in the form of theories of approximate reasoning based on fuzzy sets and fuzzy logic ([2,7]). It is epistemic in the sense that it is reducible by expending resource to obtain more precise information but as Zadeh [22] asserted in one of his first papers on fuzzy sets 'as the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics'. He went on to state that 'we need approaches which do not make a fetish of precision, rigour, and mathematical formalism, and which employ instead a methodological framework which is tolerant of imprecision and partial truths.' He later wrote ([23]) that he began to feel that complex systems cannot be dealt with effectively by the use of conventional approaches largely because the description languages based on classical mathematics are not sufficiently expressive to serve as a means of characterisation on input-output relations in

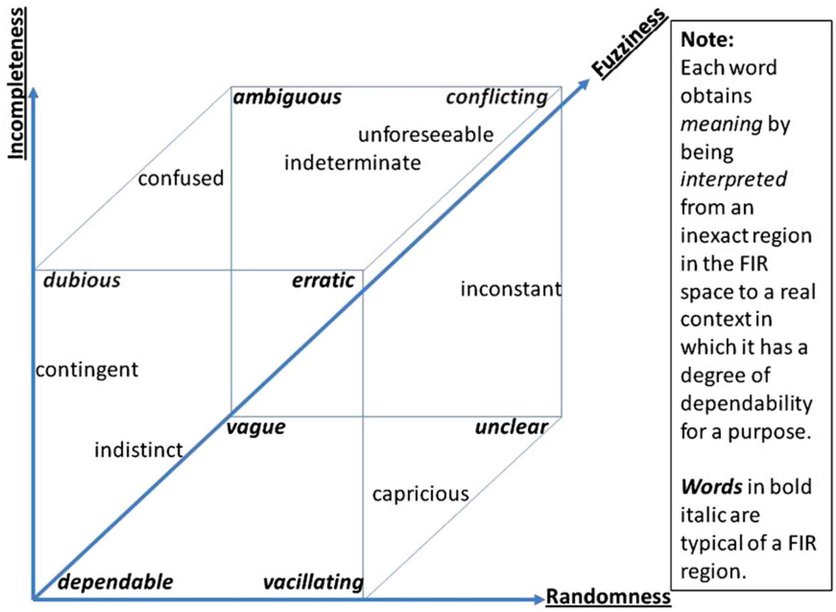

Fig. 1. Some interpretations in the FIR space of uncertainty.

an environment of imprecision, uncertainty, and incompleteness of information. In what follows the precision (or fuzziness) with which propositions are defined will be captured within a level of definition as demonstrated by Blockley and Godfrey [5] according to the nature of the practical problem being addressed. Fuzziness will not be modelled mathematically as in the theory of fuzzy sets.

Incompleteness refers to that which we do not know—it is perhaps the most neglected characteristic of epistemic uncertainty. Some theorists even deny it exists. It is not recognised in the axioms of classical probability theory because the probabilities of events or statement in the sample space must sum to unity i.e., everything in the sample space is totally known and identified. This requirement of totality is dropped in fuzzy sets and in interval probability theory. There is a need, in any methodology, explicitly to decouple measures for and against the truth or dependability of an event or logical statement and hence to allow a state of 'don't know'. As we shall see later interval probability theory, interpreted as Italian Flags as described by Blockley and Godfrey [5] and by Blockley [6], was designed to facilitate a direct handling of incompleteness.

Randomness is interpreted here as the lack of a specific pattern or purpose in some data and is explicit in probability theory and its derivatives such as reliability theory. These theories address only this one aspect of uncertainty, randomness, making it difficult to include epistemic uncertainty in models of the physical phenomena. Indeed it forces researchers such as Der Kiureghian and Ditlevsen [12] either to straightjacket uncertainties into an ill fitting framework by asserting a need for auxiliary non-physical variables that ignore unknown unknowns ([7]) or to neglect whole areas of parallel research in other disciplines such as the social sciences. Randomness is aleatoric but, as remarked earlier, aleatoric uncertainty is also epistemic.

For practitioners the term 'ignorance' as used by many theoreticians such as Hacking [14] and Huber [15] to express the idea of unknown–unknowns is unfortunate. In a practical context it can imply negligence as a lack of learning, being uneducated or uninformed or not properly qualified and hence not exercising a proper duty of care. Of course ignorance of the law is no excuse but here we are referring to effects that no-one properly qualified can reasonably foresee. Norton [16] attempts to capture ignorance by adopting the 'Principle of Indifference' and a 'Principle of the invariance of ignorance'. In this state we have no grounds for preferring one proposition over any other and so we assign equal belief to an outcome and its negation symmetrically.

Fig. 1 is indicative of one set of possible 'mixes' for these expressions of uncertainty in natural language. In natural language the subtlety of meaning of each word will depend on an interpretation of the inexact region of an FIR mix in a context (Fig. 2). Each usage may have a different mix of FIR depending on specific situations. So for example ambiguity emerges from interacting fuzziness and incompleteness that gives rise to a potential for more than one interpretation of the meaning of a statement. If someone ambiguously asserts that a proposition is 'somewhat true and somewhat false' they may mean that there is some rather vague evidence for and some evidence against. It is associated with dubiety where one hesitates to believe through mistrust due to incompleteness even when there is no explicit fuzziness as when a politician answers a question by leaving out important aspects of an issue. Erratic uncertainty emerges from interacting incompleteness and randomness so that interpretations are deviating, wandering and not fixed. The worst kind of uncertainty is where there is conflicting interpretations that are either not comparable or are incompatible or simply disagree. This can happen when all three parameters of uncertainty FIR occur simultaneously. The traditional way of handling these kinds of uncertainties is to demand clearer and more precise statements. Of course this is a rational and wholly correct reaction and should be used wherever possible—but in complex problems that process if carried through insensitively may result in the loss of valuable evidence concerning poorly and incompletely understood phenomena.

In summary an expression $p(A)$ will here be interpreted as an interval probability measure of a region in an FIR space (Fig. 2) in contrast to traditional probability theory where $p(A)$ is a measure on the R space only. In other words the conceptual interpretation of the term A is that the uncertainty associated with it may be variously fuzzy, incomplete or random. For example A might be the precise proposition 'The deflection at the centre of a given bridge is 31 mm' or the fuzzy proposition that 'The deflection at the centre of a given bridge is acceptably small' or the fuzzy and incomplete proposition, that practitioners often have to grapple with, that 'There is evidence that the deflection at the centre of a given

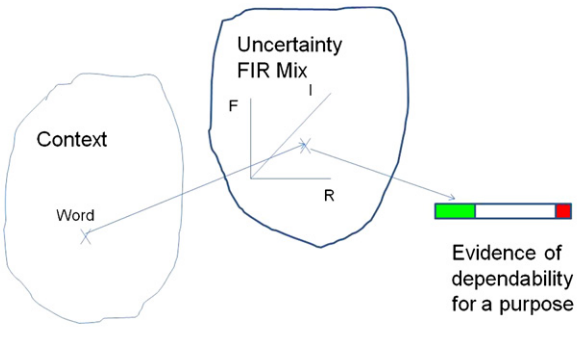

Fig. 2. A mapping from context to FIR space to an Italian Flag.

bridge is small but there is also evidence that it is unacceptable'. Clearly the measure $p(A)$ will be different for each of these versions of A but in each case the FIR components of A can be varied independently.

4.1. Probability: Statistics, Bayes and evidence

Probability theory is normally interpreted in four main ways (though there are many others)—as statistics of frequencies or propensities (tendencies), as degrees of belief or as measures of evidence. The first two are usually assumed as aleatory and the second two as epistemic. No matter how it is interpreted, in essence, probability is a mathematical measure on the interval [0,1] which behaves according to axioms ([14], [21]). The three most important axioms, for the purposes of this paper, are that the sum of all probabilities on the sample space or reference set must sum to 1, that probabilities of mutually exclusive sets are countably additive and that a conditional probability $p(H/E)=p(E\cap H)/p(E)$. Walley [21] gives a detailed mathematical account of developments in uncertainty measures including coherent previsions, fuzzy sets, maximum entropy and the Dempster–Shafer theory [19]. We can illustrate the impact of these developments with some simplified examples.

Table 1 shows the four possible outcomes (the set of all proper subsets of the sample space) of an example where H denotes 'Patient has disease' and E denotes 'Patient has symptoms'. Note that a negation is shown by $\neg$ and so $\neg H$ means not H. Note also that evidence in practical decision making comes in many forms including natural language. This means that the precision (or fuzziness) with which H and E are defined may vary depending on the nature and level of the modelling in the FIR space.

In many problems of this kind fortunately we have access to a lot of precisely defined data. We can therefore use these data to determine *a priori* (before the fact) the background chance of having the disease within the population at large $p(H)$. Also from many trials we can determine that the chance of having a positive test for the symptoms if you have the disease $p(E/H)$ and the chance of having a false positive result i.e., a positive test for the symptoms if you do not have the disease $p(E/\neg H)$ as well as a false negative $p(\neg E/H)$. We can then use these data to find out your chance of having the disease if you test positive. To do this we use Bayes rule effectively to update the *a priori* probability in the light of the evidence E.

$$p(H/E)=p(E/H)p(H)/p(E)$$

where $p(E)=p(E/H)p(H)+p(E/\neg H)p(\neg H)$

If we have sufficient data then the probability measure can be interpreted as a frequency as in classical statistics. However all too often sufficient data is lacking, hypotheses and evidence may be quite imprecise (or fuzzy) and judgement is required. In Bayesian probability theory the probability measure is interpreted as a subjective degree of belief about the incomplete data and our judgements concerning it. The theory is used to update *a priori* beliefs in the light of new evidence. Any difficulties dealing with contextual unknown uncertainties are handled by making it clear that Bayesian probabilities should be always considered as conditional on context and other unknowns. In that sense even an *a priori* probability is conditional on context.

However it is difficult to make judgements about a background probability for H where data is sparse, the sample space is unclear and evidence comes from many different and possibly conflicting diverse sources. In these circumstances we need an alternative formulation. Evidence theory as developed by Dempster and Shafer [19] is an extension of classical probability theory and its Bayesian interpretation as a degree of belief. Whereas Bayesian inference relies on a complete model of the problem domain evidence theory recognises that in many complex modelling situations it is very difficult indeed to construct a complete model. In evidence theory we infer $p(H)$ and not $p(H/E)$. In other words the inference is defined on the universe of discourse and is not necessarily conditioned on the evidence—though it is of course conditional on the assumptions of the inference process. We also recognise that it is very hard to estimate the likelihood of every individual event E occurring (or sub-process being successful) given the super-event H occurs (or super-process is successful). Instead the reverse approach of estimating the relevance of the sub-events (sub-processes) to the super-event (process) $p(H/E)$ is more natural. In this way the evidence theory approach functions even if the domain is incomplete. So if, for example, H is 'there is global warming due to human activity' we may have no *a priori* belief but we do have evidence which is more or less convincing. So if E is 'a significant number of glaciers are melting due to human activity' we can express a belief in $p(H/E)$ i.e., the degree to which we believe that there is global warming due to human activity if a significant number of glaciers are indeed melting' and in $p(\neg H/\neg E)$ it is not the case that 'there is global warming due to human activity' if it is not the case that 'a significant number of glaciers are indeed melting due to human activity' and in $p(E)$ ' a significant number of glaciers are melting due to human activity'. We then can calculate

$$p(H)=p(H/E)p(E)+[1-p(\neg H/\neg E)][1-p(E)]$$

Table 1

	H	***¬H***
E	*E & H*	*E &¬H*
¬E	*¬E & H*	*¬E &¬H*

However if we assert that $p(E)=p$ then in classical probability it must be the case that $p(\neg E)=(1-p)$. The reality is much more subtle. We may have positive evidence that the glaciers are melting but we may also want to be able to capture the negative evidence that we are not totally convinced that the warming is man made. In other words we wish to admit some uncertainty of incompleteness into our judgements.

The net result is that evidence theory challenges the notion of completeness. Shafer [19] argues that if one believes a statement *A* to some degree of belief *Bel* (*A*), then $Bel\,(\neg A)\neq(1-Bel\,(A))$. Like Shafer we will use a measure of probability over an enhanced set of all proper subsets of the sample space. In Table 1 the sample space is simply (*E, H*) with the possibilities shown. This is extended in Table 2 where *Hu* expresses the idea of an uncertain *H* and *Eu* is an uncertain *E*. In other words the reference set has been expanded to include measures that enable us to accommodate incompleteness as *Eu* and *Hu*.

In Table 3 we allocate, simply for illustration, specific but arbitrary example probability measures (Shafer calls them probability masses *m*) to each element of the reference set. χ represents the set of all possibilities. In effect the probabilities are now summed over the enhanced reference set. Of course any negation of *H* or of *E* contains all other uncertainties and is therefore inherently problematic.

The sum of the probability masses in the second column of Table 3 gives us $p(H)=0.3$. We will follow Blockley and Godfrey [5] and label this as *g* for green. Likewise from the third column we can deduce that $p(\neg H)=0.3$ which we will label as *r* for red. From the fourth column we see that $p(Hu)=0.4=w$ with *w* for white. We can express this as an interval measure such that $p(H)=[g,\ 1-r]=[0.3,\ 0.7]$. Note that in this formulation $p(H)\neq(1-p(\neg H))$. If we draw this interval number using the green white and red colours then we have an Italian Flag as in Fig. 2.

The way in which the numerical values are distributed over the enhanced reference sets of Tables 2 and 3 has to be decided. Shafer [19] uses Dempster's rule of combination which allocates probability mass assignments from input values by multiplication. Thus if we assume the inputs we calculated for *p*(*H*), *p*(*E*) in Table 3 i.e., $p(H)=[0.3,\ 0.7]$ and $p(E)=[0.1,\ 0.5]$ then the assignments to each element are as in Table 4. For example $m(H\ \&\ E)=p(H)\times p(E)=0.3\times 0.1=0.03$ and $m(Hu\ \&\ Eu)=p(Hu)\times p(Eu)=0.4\times 0.4=0.16$.

Clearly the distributions of Tables 3 and 4 are very different. The assumption by Shafer of the Dempster rule of combination is, in effect an assumption of independence ([4], [21]) and is not warranted except where independence is known to be the case. In general one has to assess a degree of dependence as suggested by Cui and Blockley [11] or Walley [21]. However assessing dependence between multiple pieces of evidence is difficult if not impossible in all but a few specific cases. Blockley [6] has suggested a practical but approximate way of overcoming this difficulty. The method, which under certain circumstances is exact, is based on pair wise comparisons as will be illustrated by example in the next section.

As a summary Table 5 sets out the differences between a Bayesian approach and that of evidence based on an Italian Flag.

Walley [21] defines an interval probability as an imprecise probability. Using the FIR analysis presented earlier it would be more accurately termed an incomplete probability. Bradley [9] calls it an ambiguity presumably because different bets for and against an event could imply more than one interpretation. Both of these authors interpret a probability as a

Table 2

	H	**¬*H***	***Hu***
E	*E & H*	*E &¬H*	*E & Hu*
¬*E*	*¬E & H*	*¬E &¬H*	*¬E & Hu*
Eu	*Eu & H*	*Eu &¬H*	*Eu & Hu*

Table 3

	$p(H)$	$p(\neg H)$	$p(Hu)$	Totals
$p(E)$	$m(H\&E)=0.1$	$m(\neg H\&E)=0$	$m(Hu\&E)=0$	$p(E)=0.1$
$p(\neg E)$	$m(H\&\neg E)=0.2$	$m(\neg H\&\neg E)=0.3$	$m(Hu\&\neg E)=0$	$p(\neg E)=0.5$
$p(Eu)$	$m(H\&Eu)=0$	$m(\neg H\&Eu)=0$	$m(Hu\&Eu)=0.4$	$p(Eu)=0.4$
Totals	$p(H)=0.3$	$p(\neg H)=0.3$	$p(Hu)=0.4$	$p(\chi)=1.0$

Table 4

	$p(H)$	$p(\neg H)$	$p(Hu)$	Totals
$p(E)$	$m(H\&E)=0.03$	$m(\neg H\&E)=0.03$	$m(Hu\&E)=0.04$	$p(E)=0.1$
$p(\neg E)$	$m(H\&\neg E)=0.15$	$m(\neg H\&\neg E)=0.15$	$m(Hu\&\neg E)=0.20$	$p(\neg E)=0.5$
$p(Eu)$	$m(H\&Eu)=0.12$	$m(\neg H\&Eu)=0.12$	$m(Hu\&Eu)=0.16$	$p(Eu)=0.4$
Totals	$p(H)=0.3$	$p(\neg H)=0.3$	$p(Hu)=0.4$	$p(\chi)=1.0$

Table 5

	Bayesian	Evidence theory and the Italian Flag
	The probability measure is a single number e.g., $p(H)=0.3$.	The probability measure is an interval number e.g., $p(H)=[0.2, 0.5]$. The interval 0.0 to 0.2 is green. The interval 0.2 to 0.5 is white. The interval 0.5 to 1.0 is red.
Input	$p(H)$ is the background *a priori* probability such as the chance of having a disease *H*. $p(\neg H)=1-p(H)$.	$p(E)$ is the belief that the evidence *E* is dependable for our purpose. $p(\neg E)$ is assessed separately and is the belief that the evidence *E* is not dependable for our purpose. $p(\neg E)\neq(1-p\ (E))$.
	$p(E/H)$ is the chance of having symptoms *E* if you have the disease *H*. $p(E/\neg H)$ is the chance of having the symptoms *S* if you do not have the disease.	$p(H/E)$ is the belief that the process *H* will eventually be successful if we depend on evidence *E*. $p(\neg H/\neg E)$ is the belief that the process *H* will eventually fail if the evidence *E* cannot be depended on.
	$p(E)$ is the chance of having symptoms *E* and is calculable from the inputs as $p(E)=P(E/H).p(H)+p(E/\neg H).\ p(\neg H)$	
Output	$p(H/E)$ is the chance of having the disease *H* if you show the symptoms *E*. $p(H/E)=p(E/H).\ p(H)/p(E)$.	$p(H)$ is the chance that the process *H* will eventually be successful. $p(H)=p(H/E).\ p(E)+(1-p(\neg H/\neg E).\ p(\neg E)$.

gamble in which the lower bound is the price you would be prepared to pay to receive 1 unit of value (say £1) if an event is true and the upper bound is the amount for which you would sell a gamble. Thus if the gamble is expressed as (A, g) where g is the lower bound probability and A is the event which if it turns out to be true (your horse wins) then you would receive £1 from the bookie. In other words you are paying £g for the chance to win £1. The bookie is effectively taking on the bet $(\neg A, 1-g)$. Bradley shows that if you have the preference that for all (A, g) either you prefer (A, g) over $(\neg A, 1-g)$ or the other way round then this gamble is effectively the betting analogue of the law of the excluded middle. Dropping this preference means that a bet of (A, g) does not imply you are willing to accept a bet of $(\neg A, 1-g)$. Bradley concludes that the Dutch book argument is too strong in examples of epistemic vagueness (fuzziness). He writes that probability theory is a nice mathematical theory but does not do justice to all people's observed behaviour. Both Bradley and Walley are concerned to create a theory that is coherent, consistent and avoids sure loss (a Dutch book). Walley argues that it may never be possible to completely formalise real decision making because intelligence, intuition, experience and judgement required cannot be captured. Nevertheless formal methods can provide a check on some mistakes, can check logical consistency, coherence and the avoidance of decisions that entail sure loss (for example as in a gamble). Therefore a formal mathematical calculation can support decision making as long as it is done with an adequate understanding of the limitations of context and application.

4.2. Interval probabilities: An example

In general there may be a large number of disparate sources of evidence concerning some hypothesis *H*. For example the rather fuzzy assertion *H* that 'Serious changes in the climate (such as an average rise in global temperature of between 2 and 6 degrees C) over the next 50 years are anthropogenic' may be supported by rather fuzzy assertions such as *A* 'The hole in the ozone layer over the south pole has been caused by CFC gases and is reputedly causing a rise in skin cancers' and evidence *B* that 'the measured average temperature for the last 50 years is rising and is due to greenhouse gas emissions' and evidence *C* that 'Glaciers are melting at an alarming rate due to greenhouse gas emissions'.

We can express the support for such evidential assertions as $p(A)$, $p(B)$, $p(C)$. It is then theoretically possible to do exact calculations to deduce $p(H)$ from $p(A)$, $p(B)$, $p(C)$ but this would require many totally consistent judgements about the bearing of the evidence *A*, *B*, *C* on the hypothesis *H* and about the complex interdependencies between *A*, *B* and *C*. This is rarely possible in practice. For example for 5 pieces of evidence (*A*, *B*, *C*, *D*, *E*) then 5 assessments of $p(H/A)$, $p(H/B)$, $p(H/C)$, $p(H/D)$, $p(H/E)$ and a total of 32 assessments about combinations of evidence including $p(A\cap B\cap\neg C\cap D\cap\neg E)$ and $p(A\cap\neg B\cap C\cap\neg D\cap E)$ would be needed. As we shall see in an example this latter kind of detail is possible for three pieces of evidence which are either independent or totally dependent (i.e., one piece of evidence implies another) even when there are inconsistencies between the conditional evidence for *H* expressed as $p(H/A)$, $p(H/B)$, $p(H/C)$. However in most practical inferences the evidence will be partially dependent or assumptions about partial dependence will lead to conflicts or inconsistencies and an exact result is not a practical option. It is for this reason that Blockley [6] has proposed a method of pair wise comparisons which is practically feasible and under certain conditions is exact. The method is exact where conditional judgements are consistent and the evidence is entirely self consistent (no conflicts) and the structure of the interdependencies is such that pair wise comparisons do not over specify the constraints. However in general the method over specifies the constraints on the distribution of the probability masses and hence is an approximation.

We will now demonstrate these ideas for our relatively simple example with three propositions. In this example, as we shall see, the initial assessments of the evidence will be inconsistent but the pairwise combination is exact because the structure of the interdependencies does not result in an over specification of the constraints on the distribution of the probability masses. Imagine that, for the purposes of this example, we agree that *A* and *B* are independent and so are *A* and *C*. However *B* and *C* are totally dependent so that *C* implies *B* and *C* is wholly contained in *B*. We will allocate degrees of belief purposely only to illustrate the methodology rather than be drawn into a debate about what the actual judgements should be. Therefore we will use $p(A)=0.5$, $p(B)=0.3$, $p(C)=0.2$ and we will assert (again for simplicity here) that $p(H/A)=p(H/B)=p(H/C)=1.0$ and $p(H/\neg A)=p(H/\neg B)=p(H/\neg C)=0$. Note that these conditional assessments concern the support for *H* from each of the propositions *A*, *B*, *C* separately and will normally lead to different estimates of *H*. In this example the simple estimates of $p(H/A)$ etc imply that *H* is equivalent to *A*, that *H* is equivalent to *B* and that *H* is equivalent to *C*—clearly these judgements are inconsistent but they are the best we can do based on the opinions of those professionals whose duty of care is to make them. Using them we can calculate

$p(H)=p(H/A)p(A)+p(H/\neg A)p(\neg A)$ so that for these special assumed values $p(H)=p(A)$ with similar results for $p(B)$ and $p(C)$. Therefore we now have three different estimates of $p(H)$ i.e., $p(A)$, $p(B)$, $p(C)$. Again it is important to stress that these estimates are inconsistent because of inconsistencies in the judgements on which they are based.

In order to make the best practical use of them we now work with pairs (*A*,*B*), (*A*,*C*) and (*B*,*C*). We do this because we assume that it is possible for a decision maker to make pair wise comparisons but impossible to assign probability masses across many disparate sources of evidence. However for only three pairs we can actually compare the results from a pair wise method with an exact result on a 3 dimensional space.

So for example Table 6 sets out the possibilities for *A* and *B* as $A\times B$ or the Cartesian product of *A* and *B*. The four subsets of $A\times B$ are $(A\cap B)$, $(A\cap\neg B)$, $(\neg A\cap B)$, $(\neg A\cap\neg B)$. Along the top of Table 6 are listed the set *A* and the set $\neg A$ and then in the second row $\neg A$ is expanded to show the possible subsets of $\neg A$ with *B*, *C* i.e., $(\neg A\cap B\cap C)$, $(\neg A\cap B\cap\neg C)$, $(\neg A\cap\neg B\cap C)$, $(\neg A\cap\neg B\cap\neg C$. Down the side of Table 6 are *B* and $\neg B$ and the possible subsets of $\neg B$ with *A*, *C*. An entry in Table 6 is the probability mass allocated to the set which is intersection of the row and column. Thus the entry in the top left corner is $p(A\cap B)$ and in the third row of the left hand column is $p(A\cap(A\cap\neg B\cap\neg C))$. We can allocate the probability masses to Table 6 using the initial probabilities and logical relations between *A*, *B*, *C*. Fig. 3 shows the relationships as a Venn diagram. So to illustrate the method we note that *A* and *B* are independent so $p(A\cap B)=0.5\times 0.3=0.15$. Likewise we know *A* and *C* are

Table 6

$A\times B$		A	$\neg A$				
		A	$\neg A\cap B\cap C$	$\neg A\cap B\cap\neg C$	$\neg A\cap\neg B\cap C$	$\neg A\cap\neg B\cap\neg C$	Σ
B	B	0.15	0.10	0.05	–	–	0.3
$\neg B$	$A\cap\neg B\cap C$	0	–	–	–	–	0
	$A\cap\neg B\cap\neg C$	0.35	–	–	–	–	0.35
	$\neg A\cap\neg B\cap C$	–	–	–	0	–	0
	$\neg A\cap\neg B\cap\neg C$	–	–	–	–	0.35	0.35
	Σ	0.5	0.1	0.05	0	0.35	1.0

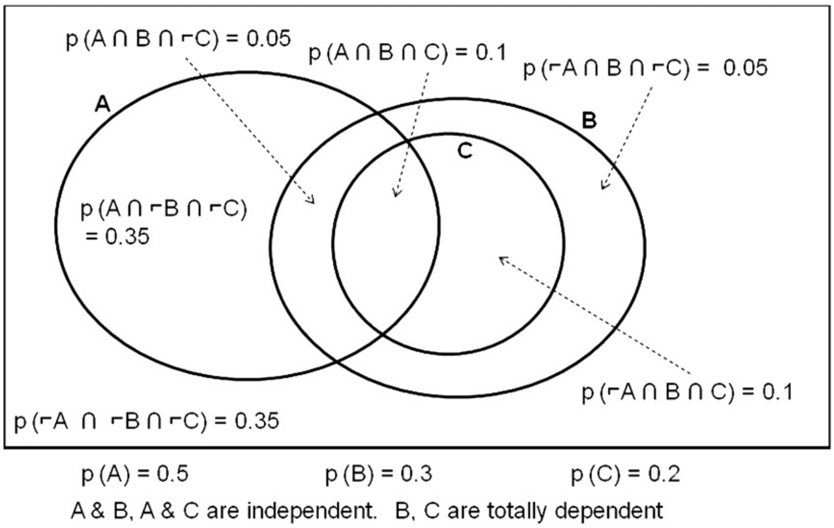

Fig. 3. An allocation of probability masses.

independent so $p(A\cap C)=0.5\times 0.2=0.1$ and $p(\neg A\cap C)=p(C)–p(A\cap C)=0.2–0.1=0.1$. Also $p(A\cap B\cap\neg C)=p(A\cap B)–p(A\cap C)=0.15–0.10=0.05$. We can go on to calculate the rest of the numbers in Fig. 3 and Table 6 in like manner for example $p(A\cap\neg B\cap\neg C)=p(A)-p(A\cap B\cap\neg C)-p(A\cap B\cap C)=0.5–0.05–0.1=0.35$. However we also note that some of the intersections in Table 6 are logical contradictions—for example $(\neg A\cap B\cap C)\cap(A\cap\neg B\cap C)$ contains $A\cap\neg A$. An entry of—in Table 6 means there would be a logical contradiction if any probability mass were to be allocated to it.

Table 6 reduces to Table 7 when we collect all subsets of $\neg A$ together and all subsets of $\neg B$ together.

From Table 7 we can deduce that $p(A\cap B)=0.15$ and $p(A\cup B)=(1-0.35)=0.65$.

Now $p(A)$ and $p(B)$ are both measures of evidence for H but taken from different sources with the relationships assumed between A and B. Therefore we can think of $p(A\cap B)$ as an upper bound on the value of the positive evidence supporting H. It is only one possible upper bound because as far as these data are concerned the intersection represents common evidential support. However there may be other possible upper bounds based on other pairwise comparisons. Likewise $p(A\cup B)$ is the upper bound of the possible evidence supporting H. This is because it is one bound on the value of the positive evidence for $\neg H$ which is the negative evidence against H. However there is a problem. As A and B are estimates of H then there is still an irreducible logical conflict between $\neg A$ and B of 0.15 and between $\neg B$ and A of 0.35 and we will have to decide how to deal with it.

Now let us calculate the same quantities for $A\times C$ (Table 8).

This becomes Table 9.

From Table 9 we can deduce that $p(A\cap C)=0.1$ and $p(A\cup C)=(1-0.4)=0.6$ but also that there is a logically irreducible conflict of 0.5.

Likewise for B and C we get Tables 10 and 11.

From which the logical conflict is 0.1 and $p(B\cap C)=0.2$ and $p(B\cup C)=(1-0.7)=0.3$.

Table 7

$A\times B$	A	$\neg A$	Σ
B	0.15	0.15	0.3
$\neg B$	0.35	0.35	0.7
Σ	0.5	0.5	1.0

Table 8

$A\times C$		A	$\neg A$				
		A	$\neg A\cap B\cap C$	$\neg A\cap B\cap\neg C$	$\neg A\cap\neg B\cap C$	$\neg A\cap\neg B\cap\neg C$	Σ
C	C	0.10	0.10	–	0	–	0.2
$\neg C$	$A\cap B\cap\neg C$	0.05	–	–	–	–	0
	$A\cap\neg B\cap\neg C$	0.35	–	–	–	–	0.35
	$\neg A\cap B\cap\neg C$	–	0.05	0	–	–	0
	$\neg A\cap\neg B\cap\neg C$	–	–	–	–	0.35	0.35
	Σ	0.5	0.15	0	0	0.35	1.0

Table 9

$A\times C$	A	$\neg A$	Σ
C	0.1	0.1	0.2
$\neg C$	0.4	0.4	0.8
Σ	0.5	0.5	1.0

Table 10

$B\times C$		B	$\neg B$				
		B	$A\cap\neg B\cap C$	$A\cap\neg B\cap\neg C$	$\neg A\cap\neg B\cap C$	$\neg A\cap\neg B\cap\neg C$	Σ
C	C	0.2	0	–	0	–	0.2
$\neg C$	$A\cap B\cap\neg C$	0.05	–	–	–	–	0.05
	$A\cap\neg B\cap\neg C$	–	–	0.35	–	–	0.35
	$\neg A\cap B\cap\neg C$	0.05	–	–	0	–	0.05
	$\neg A\cap\neg B\cap\neg C$	–	–	–	–	0.35	0.35
	Σ	0.3	0	0.35	0	0.35	1.0

Table 11

$B \times C$	B	$\neg B$	Σ
C	0.2	0	0.2
$\neg C$	0.1	0.7	0.8
Σ	0.3	0.7	1.0

Table 12

$A \times B$	A	$\neg A$	Au	Σ
B	0.15	0.15	0	0.3
$\neg B$	0.35	0.35	0	0.7
Bu	0	0	0	0
Σ	0.5	0.5	0	1.0

Summarising

$$p(A \cap B) = 0.15, p(A \cup B) = 0.65$$
$$p(A \cap C) = 0.10, p(A \cup C) = 0.60$$
$$p(B \cap C) = 0.20, p(B \cup C) = 0.30$$

We can compare these solutions with the exact result for the combinations of the evidence for A, B, C shown in Fig. 3. In this example this exact result is possible because the structure of this example means that there are six proper subsets and six equations relating the probability masses.

However following the pair wise methodology if $(A \cap B)$, $(A \cap C)$, $(B \cap C)$ are all upper bound estimates of positive support for H then

$$\min\{p(A \cap B), p(A \cap C), p(B \cap C)\}$$

is the lowest upper bound on H which we will call $p(H_l)$.

Likewise if $(A \cup B)$, $(A \cup C)$, $(B \cup C)$ are all upper bound estimates of possible support for H then max $\{p(A \cup B), p(A \cup C), p(A \cup C)\}$ is the maximum upper bound on H and the lowest upper bound on $\neg H$. We can call this bound $p(H_u)$.

So now we have $p(H) = [p(H_l), p(H_u)] = [0.1, 0.65]$ which accords with the exact result.

Unfortunately we still have the severe problem that the results contain irreducible logical conflicting evidence through the probability masses assigned to $(A \cap \neg B)$, $(A \cap \neg C)$, $(B \cap \neg C)$ since A, B, C are actually all estimates of H but based on different sources of evidence.

5. Incompleteness

To deal with this conflict we can introduce and recognise explicitly further unknowns U. These can even include unknown unknowns. So now we have a problem space of $A \times B \times C \times U$ which is unknowable but we can say something about $A \times B$, $A \times C$ and $B \times C$. However now for $A \times B$ we need to set out Table 2 differently as in Table 12 so that we can explicitly recognise the unknowns U.

We can also extend Table 12 to include situations where A and B are incompletely known to start with. For example if $p(A) = [0.5, 0.6]$ and $p(B) = [0.3, 0.4]$ to allow for the unknowns within our estimates of A and B then we get Table 13.

The distribution of probability masses in Table 13 follow the Dempster rule i.e., they are obtained by multiplying the sums of the rows and columns e.g., $m(A \cap B) = 0.5 \times 0.3 = 0.15$ and $m(\neg A \cap \neg B) = 0.24$.

You will recall that in our example of the previous section A and B are estimates of H so this means that $\neg A$ and B are in conflict as are A and $\neg B$. The best we can do is to relocate the probability masses to $m(Au \cap Bu)$.

As previously $p(A \cap B)$ and $p(A \cup B)$ are bounded estimates of $p(H)$ where

$$p(A \cap B) = [m(A \cap B), \{m(A \cap B) + m(A \cap Bu) + m(Au \cap B) + m(Au \cap Bu)\}]$$
$$= [0.15, 0.15 + 0.05 + 0.03 + 0.01] = [0.15, 0.24]$$
$$p(A \cup B) = [1 - \{m(\neg A \cap \neg B) + m(\neg A \cap Bu) + m(Au \cap \neg B) + m(Au \cap Bu)\},$$
$$1 - \{m(\neg A \cap \neg B)\}] = [1 - (0.24 + 0.04 + 0.06 + 0.01), (1 - 0.24)] = [0.65, 0.76]$$

We can perform a similar analysis for $A \times C$ as in Table 14.

From Table 14 we get a conflict of $0.08 + 0.35 = 0.42$ and

$$p(A \cap C) = [0.1, 0.18]; \quad p(A \cup C) = [0.6, 0.72]$$

For B and C Dempsters rule does not apply since they are totally dependent and the distribution of probability masses is as in Table 15.

Table 13

$A \times B$	A	$\neg A$	Au	Σ
B	0.15	0.12	0.03	0.3
$\neg B$	0.30	0.24	0.06	0.6
Bu	0.05	0.04	0.01	0.1
Σ	0.5	0.4	0.1	1.0

Table 14

A × C	A	¬A	Au	Σ
C	0.1	0.08	0.02	0.2
¬C	0.35	0.28	0.07	0.7
Cu	0.05	0.04	0.01	0.1
Σ	0.5	0.4	0.1	1.0

Table 15

$B \times C$	B	$\neg B$	Bu	Σ
C	0.2	0	0	0.2
$\neg C$	0	0.6	0.1	0.7
Cu	0.1	0	0	0.1
Σ	0.3	0.6	0.1	1.0

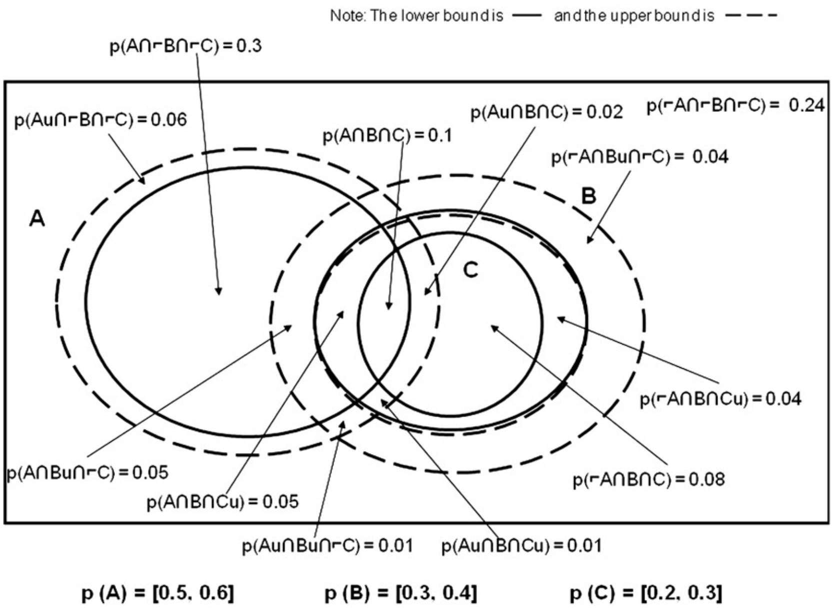

Fig. 4. An example of the allocation of probability masses.

Thus from Table 15 we get

$$p(B \cap C) = [0.2,0.3]; p(B \cup C) = [0.3,0.4]$$

Again we can obtain an exact solution because the structure of this problem is as described previously but with 12 proper subsets and 12 equations relating the probability masses—see Fig. 4.

Note that we can express ignorance or a 'don't know' verdict about unknown–unknowns using [0,1]. Norton [16] is unhappy with this formulation since he interprets the 0 in [0,1] as ignorance whereas he sees it as disbelief in [0,0]. Here we interpret the interval pairs together. Hence the first number (0) is the degree of evidence available for the proposition and the second (1) means there is (0) evidence for the negation. Therefore [0,1] obeys the principles of indifference and of

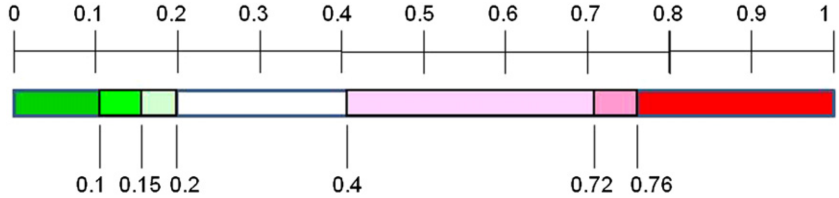

Fig. 5. A 'fuzzy' Italian Flag.

the invariance of ignorance because we have no evidential grounds for preferring one proposition or its negation. The pair [0,0] is interpreted as there being total evidence for the negation.

In summary the pair wise comparisons provide us with three estimates of $p(H)$. The lower bounds derive from our understanding and judgements about $A \times B$ to give us [0.15, 0.24], one from $A \times C$ to give us [0.1, 0.18] and one from $B \times C$ to give us [0.2, 0.3]. The upper bounds derive from our understanding and judgements about $A \times B$ to give us [0.65, 0.76], one from $A \times C$ to give us [0.6, 0.72] and one from $B \times C$ to give us [0.3, 0.4].

From Fig. 4 the correct result is

$$p(H) = [p(A \cap B \cap C), p(A \cup B \cup C)] = [0.1, 0.76]$$

The choice of $p(A \cap B \cap C)$ as a lower bound for H is severe since the intersection of more than two pieces of evidence will be small if one of them is small. In effect the smaller evidence will compromise the impact of the larger evidence. In the pair wise approximation we neglect intersections of more than two pieces of evidence and hence we have

$$p(H) = [\min\{p(A \cap B), p(A \cap C), p(B \cap C)\}, \max\{p(A \cup B), p(A \cup C), p(B \cup C)\}]$$

i.e., $p(H)=[0.1, 0.76]$

However by using only the minimum and maximum values we are missing some richness in the evidence available to us from the pair wise comparisons. We can enrich our result by taking each lower bound and each upper bound as shown in Fig. 5. In other words the Italian Flag becomes fuzzy with colours representing shades of membership in the set of probabilities of H i.e., $p(H)$. The Italian Flag has gradations of colour from green through lighter shades of green to white and then from lighter red through to red. In effect such a flag indicates how the evidence builds from the centre outwards—left towards positive confirming evidence and right towards negative falsifying evidence.

Thus with three pieces of evidence we have 3 lower bounds $p(H_l)=0.1$, $p(H_l)=0.15$, $p(H_l)=0.2$ and three upper bounds $p(H_u)=0.4$, $p(H_u)=0.72$, $p(H_u)=0.76$.

6. Conclusions

1. An analysis of uncertainty that contrasts definitions in natural languages with those of formal languages reveals that the categorisation of uncertainty as either aleatoric or epistemic is unsatisfactory for practical application.
2. Uncertainty, it is argued, can be captured by three orthogonal attributes—FIR i.e., fuzziness, incompleteness and randomness. A number of characterisations of uncertainty in natural language such as ambiguity, dubiety and conflict are complex mixes of interactions between FIR.
3. A Bayesian approach which uses a probability measure either as a frequency or as a subjective judgement about a degree of belief can be conditional on unknown variables but such an approach does not allow the decision maker to acknowledge incompleteness explicitly. For the management of risk in complex systems it is considered important to recognise the extent to which we 'don't know' about uncertain unintended and unwanted consequences so that we may be alert to unexpected hazards.
4. The Bayesian approach has been compared with an interval probability approach to show how conflict due to incomplete information can be managed.
5. A Fuzzy Italian Flag based on pair wise comparisons of more than two pieces of evidence can reflect the richness of that evidence to provide decision makers with a richer picture of the evidence for a particular hypothesis.

References

[1] T. Aven, On the need for restricting the probabilistic analysis in risk assessments to variability, Risk Anal. 30 (3) (2010) 354–360.
[2] D.I. Blockley, The Nature of Structural Design and Safety, Ellis Horwood, Chichester, 1980.
[3] D.I. Blockley, Reliability or responsibility, Struct. Saf. 2 (4) (1985) 273–280.
[4] D.I. Blockley, Fuzziness and probability: a discussion of Gaines' axioms, Civ. Eng. Environ. Syst. 2 (4) (1985) 195–200.
[5] D.I. Blockley, P S Godfrey, Doing it Differently, Thomas Telford, London, 2000.
[6] D.I. Blockley, Managing risks to structures, Proc. Inst. Civ. Eng., Struct. Build. 161 (SB4) (2008) 231–237.
[7] D.I. Blockley, Uncertainty, prediction or control, Int. J. Eng. Under Uncertainty: Hazard., Assess. Mitig. 1 (1–2) (2009) 73–80.
[8] D.I. Blockley, The importance of being process, Civ. Eng. Environ. Syst. 27 (3) (2010) 189–199.

[9] S. Bradley, Dutch Books and Ambiguity, Pluralism in the Foundations of Statistics, University of Kent, September, See 〈http://www.kent.ac.uk/secl/philosophy/jw/2010/plurastats/Bradley-handout.pdf〉. Also S. Bradley, Dutch book arguments and imprecise probabilities, in: D. Dieks, W.J. Gonzalez, S. Hartmann, M. Stolzner, M. Weber (Eds.), Probabilities, Laws and Structures, Springer, 2012. See 〈http://www.seamusbradley.net/Papers/dba-ip.pdf〉, 2010, 2011.

[10] R. Corotis, Risk communication with generalised uncertainty and linguistics, Struct. Saf. 31 (2009) 113–117.

[11] W.C. Cui, D.I. Blockley, Interval probability theory for evidential support, Int. J. Intell. Syst. 5 (2) (1990) 183–192.

[12] A. Der Kiureghian, O. Ditlevsen, Aleatory or epistemic? Does it matter? Struct. Saf. 31 (2009) 105–112.

[13] D. Dubois, Representation, propagation and decision issues in risk analysis under incomplete probabilistic information, Risk Anal. 30 (3) (2010) 361–368.

[14] I. Hacking, The Emergence of Probability, Cambridge University Press, Cambridge, 1975.

[15] W.A. Huber, Ignorance is not probability, Risk Anal. 30 (3) (2010) 371–376.

[16] J. D. Norton, Ignorance and Indifference. See 〈http://www.pitt.edu/~jdnorton〉, 2011.

[17] C. Perrow, Normal Accidents: Living with High Risk Technologies, Basic Books, New York, 1984.

[18] J. Reason, Human Error, Cambridge University Press, New York, 1990.

[19] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[20] B. A. Turner, N. F. Pidgeon, Man-Made Disasters, 1998.

[21] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991.

[22] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst. Man Cybern. SMC-3 (1) (1973) 28–44.

[23] L.A. Zadeh, The role of fuzzy logic in modelling, identification and control, in: L.A. Zadeh, G.J. Klir, B. Yuan (Eds.), Fuzzy Sets, Fuzzy Logic and Fuzzy Systems: Selected Papers, World Scientific, 1996, pp. 783–795.