Weight Chapter Outline

9/1/2022

1 Introduction

The probability we assign to a hypothesis of interest conveys our epistemic uncertainty about the hypothesis. But there is also a further uncertainty at play. It is the uncertainty about our assessment of the probability of the hypothesis. Suppose we test positive for a certain disease. The test isn't infallible. So, based on the test result, the probability we should assign to the hypothesis that we are actually positive is somewhere below 100%, say it is 75%. This number expresses our uncertainty about the hypothesis that we are positive. The 75% figure is the output of applying Bayes' theorem, which however requires as inputs the base rate of the disease and the error rates of the test. These numbers are themselves uncertain. So the 75% probability that we are positive carries its own uncertainty. It rests on numbers that are themselves uncertain. Or consider an example from the courtroom. A criminal defendant is standing trial. Traces that were found at the crime scene match the defendant's genetic profile. An expert testifies that the probability that a random person would coincidentally match is as low as 1 in 10 billion. This extremely low probability can serve, together with other prior information, to assess the probability that the defendant is actually the source. But the figure 1 in 10 billion could itself be disputed. Besides the uncertainty about whether or not the defendant was actually the source of the traces, there is uncertainty about the 1 in 10 billion figure itself.

Numbers carry with them an aura of objectivity. Their presentation by an expert might suggest they are not up for debate. But that would be a mistake. A skilled defense lawyer will rightly inquiry about the credentials of the expert, the source of the data, their reliability. At one extreme, the numbers presented in a courtroom can be made up. In the infamous People v. Collins, the prosecutor made up the frequencies of various identifying characteristics: interracial couple, driving a yellow convertible, women with ponytail, man with mustache, etc. But the numbers presented in a courtroom can also rest on well-established practices of data collection and genetic modeling, as in the case of random match probabilities. Other times these numbers, albeit not made up, might rest on less extensive research, as in the case frequencies about carpet fibers. All in all, our second-order uncertainty comes in degrees, just like our first-order uncertainty about a hypothesis of interest. What we need is a theory that can model both of them.

This chapter articulates a formal framework for modeling first-order uncertainty about hypotheses as well as uncertainty about the first-order uncertainty. The framework fits naturally with Bayesian statistics and can be implemented thorough Bayesian networks. We develop this framework as an improvement of imprecise probabilism in Bayesian epistemology. As the name suggests, imprecise probabilism posits that agents can be imprecise about the probabilities they assign. Instead of single probability measures, they can entertain multiple probability measures. The higher order framework we develop here brings imprecise probabilism one step further. Imprecise probabilism assumes that the multiple probability measures which agents can entertain are all equally likely. The higher order approach relaxes this assumption. The possible probability measures can themselves be more or less probable. The two-tiered framework for modeling uncertainty will then serve to articulate a theory of evidential weight.

Side comment: I think for expository clarity, we should emphasize two contributions here. One is the contribution to higher order uncertainty. This is an improvement on imprecise probabilism. This I think is already an important contribution that fares better than imprecise probabilism. The second contribution is the theory of weight that is built out of the higher order uncertainty framework. The two issues seems separate to me. One could be interested in question of higher order uncertainty and imprecise probabilism without being interested in questions of weight.

2 Precise, imprecise and higher order probabilism

This section (or perhaps multiple sections) will describe the three main frameworks of probabilism: precise, imprecise, and higher order probabilism. This section should do two things. First, to motivate why the move to imprecise probabilism is warranted but also show the limits of imprecise probabilism. Second, it should outline what higher order probabilism looks like and why it overcomes the problems of imprecise probabilism.

2.1 Motivating examples

(a1) Use example by Peirce (sampling 100 v. sampling 1000 times, with equal sample proportion). Balance of evidence for/against a certain hypothesis seems the same, but the informational basis (weight) is wider in the second case. (Brief comment –perhaps in a footnote– that this difference is modeled in standard statistics as the SE of the sample proportion– the SE decreases as the sample size increases. This is not what we are after though since we are not simply trying to estimate a parameter value such as population proportion, but assess the probability that a hypothesis is true.)

(a2) Give an example like Peirce's using varying sample sizes to estimate relative proportion of some identifying feature in match evidence, handwriting, genetic profile, fiber, etc. DNA evidence or a simpler form of match evidence (see e.g. the Georgia v. Wayne Williams case) should do here. This can connect back to DNA evidence example in the introduction.

2.2 Addditional motivating examples

(b) Another intuitive example is, one thing is absolute ignorance about an event (or suspension of judgment because of lack of knowledge), and another thing is having equal information for/against. Assigning in both cases a sharp probability of .5 fails to capture the intuitive difference. In one case the informational basis is very limited, or non-existent, while in the other the informal basis is wider, even though the balance might seem the same.

(c) A similar problem is raised by the "negation problem" by Cohen (little evidence in favor of H, so Pr(H) is low, cannot mean there is a lot of evidence in favor of not-H, Pr(not-H) is high).

What is now in sections 8, 9 and 10 should go in this section (or sections). I would remove all the stuff about Joyce (since this is about weight), but basically all the materials we need are in those sections.

Possible addition: To give the reader an intuitive picture, we could provide three Bayesian networks for the diagnostic example from the introduction section. The first network has the shape with $D \to T$, and is just how one would do things in the standard way with sharp probabilities. The second network contains multiple probability measures about (a) the prior probability of D and second-order uncertainty about the error rates that go into the conditional table for $P(T|D)$. This is the network using imprecise probabilism. This is also something like "sensitivity analysis." Finally, the third network contains distributions over multiple probability distributions—the higher order approach. The same could be done for the DNA match example. This comparison would convey succinctly the first major contribution of the chapter. The three Bayesian networks will also nicely connect with the two motivating examples right at the start of the chapter.

3 Weight of evidence

The chapter now turns to the weight of evidence and its formalization. I think it is conceptually important to separate the discussion about precise, imprecise, and higher order probabilism from this further idea. Weight of evidence is one way in which higher order probabilism can be put to use. It can be confusing to run the discussion of higher order probabilism together with weight of evidence. Higher order probabilism can still make perfect sense even if no theory of weight cen be worked out.

This section should contain illustrative examples of why "balance alone" isn't enough to model the evidential uncertainty relative to a hypothesis of interest. These examples should be chosen carefully. We can use legal and non-legal examples. The driving intuition is given by Keynes with the weight/balance distinction. Some of the examples we saw earlier in talking about imprecise probabilism can be mentioned here again, such as (a1) and (a2), and perhaps also (b) and (c).

Upshot is that uncertainty cannot be captured by balance of evidence alone. There is a further dimension to uncertainty. So we need a theory that can accommodate this further level of uncertainty. This theory is essentially the higher order probabilism introduced before.

Question: There seems to be a nice symmetry here. Take probabilism theory (sharp probabilism). We can use sharp probability theory to offer a theory of the value of the evidence (i.e. likelihood ratio). Actually, I think that the likelihood ratio model the idea of balance of the evidence. What Keynes distinction weight/balance shows is that likelihood ratio are not, by themselves, enough to model the value of the evidence. The straightforward move here seems to just have higher order likelihood ratios. Wouldn't higher order likelihood ratio be essentially your formal model of the weight of the evidence?

3.1 Completeness and resilience

Note sure exactly, but we should mention completeness and resilience at some point, maybe in this section or maybe elsewhere. We should think where it might be best to discuss them.

- Give an example using completeness of evidence (pick one or more court cases). The court case we can use is Porter v. City of San Francisco (see file with Marcello's notes).[1] The jury is given an instruction that a call recording is missing, but no instruction whether the call should be assumed to be favorable or not. What is the jury supposed to do with this information? If the call could contain information that is favorable or not, shouldn't the jury simply ignore the fact that the call recording is missing (Hamer's claim)? Modelling with Bayesian network might turn out useful. Cite also David Kaye on the issue of completeness. His claim is that when evidence is known to be missing, then this informatio should simply be added as part of the evidence, which is precisely what the court in Porter does. But again, once wer add the fact that the evidence is missing what is the evidentiary significance of that?
  (I am thinking that incompleteness is modeled by adding an evidence node to a Bayesian network but without setting a precise value for that node, and then see if the updated network yields a different probability than the previous network without the missing evidence node. The missing evidence node could be added in different places and this might changes things. In the Porter case the missing evidence seems to affect the credibility of the other evidence in the case, we would have a network like this: $H \rightarrow E \leftarrow C$, where C is the missing evidence node and E is the available evidence nose.)
- Also, give an example using resilience (pick another court case). Since the discussion about completeness seems alreayd quite extensive, we might just stop at that and go no further.

Both completeness and resilience suggest there is a further level of uncertainty besides balance of evidence.

3.2 Desiderata

Here we can discuss monotonicity, completeness, strong increase, etc. We can list the intuitive properties (based on the example we presented in both philosophy and law) that any theory of weight/completeness/resilience should be able to capture. We should try to keep these requirements as simple as possible and leave complications to footnotes.

3.3 Limits of our contribution

Work by Nance of Dahlman suggests that "weight" should play a role in the standard of proof. We do not take a position on that. Weight could be regulated by legal rules at the level of rules of decision, rule

[1] This is a wrongful death case in which victim was committed to a hospital facility, but escaped and then died under unclear circumstances. So the nurses and other hospital workers—actually, the city of San Francisco—are accused of contributing to this person's death. Need to check exact accusation—this is not a criminal case. A phone call was made to social services shortly after the person disappeared, but its content was erased from hospital records. Court agrees that content of phone call would be helpful to understand what happened and to assess the credibility of hospital's workers ("The Okupnik call is the only contemporaneous record of what information was reported to the SFSD about Nuriddin's disappearance, and could contain facts not otherwise known about her disappearance and CCSF's response. Additionally, the call is relevant to a jury's assessment of Okupnik's credibility"). The court thought that the hospital should have kept records of that call. But court did not think the hospital acted in bad faith or intentionally, so it did NOT issue an "adverse inference instruction" (=the missing evidence was favorable to the party that should have preserved it, but failed to do it).

of evidence, admissibility, sanctions at the appellate level. All that matters to us is that, in general, legal decision-making is sensitive to these further levels of uncertainty (quantity, completeness, resilience), but whether this should be codified at the level of the standard of proof or somewhere else, we are not going to take a stance on that.