Chapter 1: Against Legal Probabilism

We present the theory of legal probabilism, discuss several objections against it and outline a number of responses available to the legal probabilist.

1 A probability threshold

Witnesses are called to testify in court about questions relevant to the defendant's civil or criminal liability. They are examined and cross-examined by the lawyers of the two parties. The purpose of this lenghty and elaborate process is to ascertain whether the defendant engaged in behaviour or committed acts that are prohibited by the applicable law. In other words, the purpose of the examination and cross-exmamination of the evidence is to answer the question, did the defendant did it or not? Only if the evidence is strong enough to establish that the defendant did it, the defendant should be found liable.

The evidence is strong enough when it meets the governing burden of proof. This burden is usually different in civil or criminal cases. In civil cases, the burden of proof is 'balance of probabilities'; in criminal cases, the burden of proof is 'proof beyond a reasonable doubt'. The latter is meant to be more stringent than the former.

According to legal probabilism, the burden of proof is a probability threshold applied to the probability of liability. This is the probability, based on the evidence presented in court, that the defendant committed the unlawful acts or engaged in the unlawful behavior they are accused of. If this probability is sufficiently high, the decision should be against the defendant, and otherwise it should favor the defendant.

How stringent the threshold should be—51 percent, 99 percent or what?—depends on the relative costs of false positive and false negative decisions, and the relative benefits of true positive and true negative decisions. Typically, since the cost of a false positive—judging the defendant liable while they are not—is greater in criminal than civil trials, the probability threshold is set higher in criminal trials. In general, the stringency of the threhsold is a function of maximizing expected utility and minimizing expected disutility.

This probabilistic picture of the burden of proof is not uncontroversial, as we shall soon seen. But it has some plausibility, especially in civil trials. So it is worth exploring it more closely.

According to a simple version of legal probabilism, the defendant's liability is established by the balance of probabilities—the burden of proof governing civil trials—provided the defendant's liability, based on the evidence presented, is greater than 50 percent or greater than .5. That is, if L stands for 'the defendant is liable' and E stans for the total evidence presented via examination and cross-examination, the burden of proof is formulated as follows:

find against the defendant if $P(L|E) > .5$

The word 'liability' can seem unspecific. The defendant is accused of having committed a specific set of acts or behaviors that, according to the applicable law, count as impermissible. For example, the act of driving and the act of driking alchool, in close temporal sucession one after the other, make one liable of driving under the influence. To represent liability at a more fine-grained level, let H_A denote the theory or hypothesis against the defendant, the accusation theory, and $\neg H_A$ its negation. The accusation theory should have some degree of specificity. It should not simply be the assertion that, say, the defendant drove under the influence. It should also say when, where and how the defendant drove under the influence. How specific the accusation theory should be is a question we will investigate later.

So, if E is the total evidence presented in court and H_A is the accusation theory, the burden of proof in civil cases can be formulated as follows:

find against the defendant if $P(H_A|E) > .5$

This formulation is equivalent to:

find against the defendant if $P(H_A|E) > P(\neg H_A|E)$.

The two formulations are equivalent because, if $P(H_A|E) > .5$, then $P(\neg H_A|E) < .5$ and thus also $P(H_A|E) > P(\neg H_A|E)$.

2 Challenge: Where do the numbers come from?

How is 'probability' to be understood in speaking of the probability of liability or the probability that the defendant did this and that? It cannot be a long-run frequency because the acts the defendant committed or not, are not repeatable events. It cannot be an objective chance either becuase, as a matter of fact, the defendant either committed those acts or did not. They lie somewhere in the past. Either they occurred or they did not. So the probability of liability must reflect the extent to which the evidence presented in court supports the claim that the defendant committed the unlawful acts they are accused of. The probability of liability must be evidence-relative. It must be an epistemic probability of some kind.

The metaphor of a scale can be helpful. Evidence may tip the scale in one direction or the other. Evidence can point against the defendant, making it more probable that that the accusation theory is true. Or it can point in favor of the defendant, making it less probable that the accusation theory is true. The overall balance of the scale, based on the total evidence, is the probability that the accusation theory is true.

The metaphor of scale tilting on one side or othe other is good only up to a point, however. How do we move past the metaphor?

The posterior probability of liability, assessed on the basis of the evidence presented, is determined starting from a prior or initial value, assessed prior to considering the evidence. Typically, the prior probability is equated to $P(L)$, while the posterior probability, given evidence E, is equated to $P(L|E)$. The relation between prior and posterior is set by the well-known formula of Bayes' theorem. Its simplest formulation goes, as follows:

$$P(L|E) = \frac{P(E|L)}{P(E)}P(L) = \frac{P(E|L)}{P(L)P(E|L) + P(\neg L)P(E|\neg L)}P(L)$$

The generic L could be replaced by the more spec ific accuasatio theory H_A, but to keep the notation easier and more readable, we will stick to L.

Here we face the first major challange for legal probabilism. *Where do the numbers come from?* The prior probability $P(L)$ must be set somewhere, but where? Should the prior be $1/n$, where n is the number of individuals who could have committed the unlawful acts in question? Setting $P(L) = 1/n$ makes sense in a criminal case in which the identity of the perpertrator is disputed, even though what happaned is clear. If n are the possible perpetrators and we have no other information to distinguish them—prior to considering the trial evidence E, that is—it makes senses to set the prior probability $P(L)$ to $1/n$ since any of the suspects could be the perpetrator. But $1/n$ does not make sense in other contexts, criminal or civil, in which how the events occurred is disputed, while the identity of the person who partcipated in them is not. In such cases, the prior could be $1/s$, where s is number of possible ways in which the events could have unfolded. But while n is a well-defined number, that is not so with s. It is clear how to count n possible suspects. It is less clear how to count s possible the ways in which the events could have unfolded.

The challange does not stop at assessing prior probabilities. What numbers should be assigned to $P(E|L)$ and $P(E|\neg L)$? The answer is far from clear. As a first step, it can be helpful to break down L into smaller level statements, say, whether the defendant visited the crime scene or left a blood stain at the scene. It can also be helpful to break down E into smaller pieces: fingerprint evidence, witness testimonies, genetic matches, expert reports, etc. Once the statements are broken down this way, assigning probabilities to them becomes more manageable.

For example, let M stand for 'the defendant's genetic profile matches the crime traces' and S stand for 'the defendant is the source of the crime trace'. The original formula reduces to the more manageable:

$$P(S|M) = \frac{P(M|S)}{P(M)}P(S) = \frac{P(M|S)}{P(C)P(M|S) + P(\neg C)P(M|\neg S)}P(C),$$

where the general level statement L is replaced by C and E by M. As we will see later, $P(M|S)$ can be set to one and $P(M|\neg S)$ to the genotype probability, the expected frequency of finding the matching genotype in a reference population. The details do not matter for now, but the important point is that these numbers can be assigned following well-established procedures. Finally, after setting $P(C) = 1/n$, where n is number of possible contributors, the posterior probability $P(C|M)$ can be obtained by easy calculations.

The strategy of focusing on smaller level statements makes the numerical assignments possible, but also pushes the problem elsewhere. Granted, probabilities can be assigned to smaller level statements such as M and S, but what about general level statements such as L or H_A? The general level statements should still be connected to the smaller level ones. The point in a trial is not just to ascertain whether the defendant is a source of the traces found at the scene, but whether the defendant is ultimately liable.

Bayesian networks are used to map out complex cases involving multiple propositions and multiple pieces of evidence. We will see how they work and how they are built in later chapters. To fix ideas, a simple criminal case could be represented as follows: $W \leftarrow L \rightarrow S \rightarrow M$, where the letters L, S and M are interpreted as before. In addition, let W stand for an incriminmating eyewitness testimony, say the testimony that they saw the defendant run away from from the crime scene at the relevant time.

DRAW BAYES NET HERE

Even this simple Bayesian network will need several conditional probabilities to get the calculations going. It will need, for example, $P(W|L)$ and $P(W|\neg G)$, as well as $P(S|L)$ and $P(S|\neg L)$. It is not at clear what these numbers should be. If the defendant is liable (or not liable), how probable that they would be leaving traces at the crime scene? If they are liable (or not liable), how probable that they would be seen running away from the crime scene? Presumably, $P(S|L) > P(S|\neg L)$ and $P(W|L) > P(W|\neg G)$, but what else besides thase inequalities?

So Bayesian networks do not solve the problem of how to assign the numbers. If anything, they make the problem even more apparent. It is difficult to find all the numbers required by the probability tables of a Bayesian network even in very simple cases, with just few nodes making up the network. When network becomes more complex, involving many more propositions and items of evidence, the problem becomes even more serious. So, then, the required numbers are often inserted as educated guesses or simply because they cannot be left blank.

2.1 Possible responses

Legal probabilists could respond in a number of different ways. Here are some of them:

- *Bite the bullet:* Yes, we do not currently have all the numbers we need, and that is a good reason to figure out what they are and collect relevant data. If the numbers are missing, then we do not currently have a good, precise way to quantify uncertainty. This isn't a problem for legal probabilism; it is a problem for any procedure that attempts to ascertain disputed matters of fact. Legal probabilism has the merit to tell us what is missing.

- *Localize:* We can focus on those domains, propositions or forms of evidence for which the required numbers are available, for example, match genetic evidence or other forms of scientific evidence. Legal probabilism need not aspire to model the evidence of an entire legal case, but only evidence amenable to probabilistic quantification.
- *Qualitative:* While precise numbers can be helpful, they can also be a distraction. Even without delivering precise numbers, legal probabilism is still valuable. It forces us to focus on the logic of reasoning under uncertainty. Probability theory imposes coherence constraints on our evidence-based beliefs about uncertain events. Similarly, legal probabilism imposes coherence constraints on evidence-based beliefs about civil and criminal liability. Legal probabilism is not primarily concerned with the task of arriving at precise probabilities. They can still be useful for illustrative purposes, but should not be taken as a basis for making decisions.

So, legal probabilism is committed to at least one of the following tenets:

- practically, we can determine the probability of liability with some degree of precision
- theorethically, the probability of liability is a good measure of the overall uncertanity about the disputed factual issue

The first tenet is questionable, as we have just seenb, and the qualitative version of legal probabilism outlined above respond by giving it up. The second tenet seems more defensible. In fact, most versions of legal probabilism out there are committed to the second tenet. But this second tenet is also questionable. This is our next topic.

3 Challenge: not just high probability

Besides high probability, other dimensions (should) guide decision-making and they might not be reducible to the probability of liability. Some of these other dimensions are:

- How certain are we about the probability of liability? (Higher-order uncertainty)
- How good (specific, coherent, plausible, explanatory powerful) is the story presented?
- Did the defense challenged the other party's story? Did the story survive the challenges?
- Is any evidence missing? Is the evidence presented representative of both sides or was the evidence collected in a biased or skewed manner?

A more sophisticated version of legal probabilism, then, should be able to do at least two things: first, formally model these additional dimension using the language of probability (or determine to what extent they fall outside the scope of probability theory and cognate theories); and second, show why relying on these additional dimensions in decision-making does foster important values, such as the accuracy and fairness of trial decisions.

NEED TO EXPLAIN EACH POINT MORE CLEARLY

4 Comparative thresholds

We pause our examination of legal probabilism and outline an alternative theory, relative plausibility. This is can be used a contrast for the discussion of legal probabilism. The starting point of relative plausiblility is that at trial the two parties put forward competing explanations of the evidence, and then these competing explanations are tested against the evidence. The explanation that is more plausible in light of the evidence prevails. So the judgment of liability should follow the explanation that prevails. Think of the two competing explanations as the accusation theory H_A and a defebnse theory H_D. Instead of assessing the probability of H_D or H_D in light of the evidence E, the theory of relative plausibility submit that the point of legal fact-find to assess the plausibility of H_A relative to H_D in light of evidence E. Plausibility is a multidimensional notion, comprising considerations of fit and consistency with the evidence, predictive power, logical coherence, coverage of the evidence, etc. The more plausible explanation is one that prevails along a weighted combination of these criteria.

To articulate the idea of plausible more precise woud us too far afield. But one aspect of relative plausibility can be immediuately grasped: instead of focusing on just the acusation H_A, focus on the comparison of the accusation theory against its alternative H_D. This comparative idea can be adopted by legal probabilism. It is worth examining what comes out.

Legal probabilism is a flexible theory. What if $\neg H_a$ is replaced by a more specific hypothesis alternative to H_A, call it H_D, say the theory put foward by the defense? While H_D entails $\neg H_A$, becuase H_D and H_A must be incompatible, the converse does not hold. $\neg H_A$ does not entail H_D becuase H_D is just one particular way in which H_A can fail to hold.

So, the burden of proof in civil cases can now be formulated as follows (call it the comparative formulation):

> find against the defendant if $P(H_A\ vertE) > P(H_D|E)$.

In other words, to establish the defendant's liability by the balance of probabilities, the accusation theory H_A should be more probable than the defense theory. Crucially, the condition $P(H_A\ vertE) > P(H_D|E)$ is not equivalent to $P(H_A\ vertE) > .5$ seen earlier. It could be that both H_A and H_D have probability below .5, even though H_A is more probable than H_D. So, following this comparative formulation, a defendant could be found liable even though the probability of liability is below .5. This result seems counterintuitive, and perhpas it is a reason to favor the earlier, non-comparative formulation of the burden of proof.

A related reason to be cautious of the comparative formulation is that the resultig decision rule depends on the choice of H_A and H_D. It is possible that, given the same stock of evidence E, in one case the probability of H_A exeecds that of H_D, while in another case, given a different framing of the two theories, the probability of H_D execeeds that of H_A. This result signals a worrisome level of sujbectivity in the decision rule.

4.1 Challenge 2: Evidence is evaluated holistically.

The chapter on story coherence should address this challange.

4.2 Challenge 3: Learning isn't updating

Ronald Allen complains that Bayesian updating isn't an adequate model of what goes on in the courtroom when evidence is presented. The decision-makers do not start from priors and update them based on the pieces of evidence presented. What happens is more complicated and cannot be modeled by Bayesian updating. *The chapter on cross-examination and arguments should address this challange.*

4.3 Challenge 4: Trials are adversarial

Trials are often adversarial. Evidence is examined and cross-examined. How can this adversarial process be modeled probabilistically? *The chapter on cross-examination and arguments should address this challange.*

4.4 Challenge 5: No evidence that probability reduces errors

It is clear that people make probabilistic mistakes in reasoning, but does this show that mistaken convictions are caused by these probabilistic mistakes? There is no evidence of that. In what way does probability actually improve the accuracy of legal decisions? *Discussion about accuracy and fairness should address this challange*

5 Structure

So we can envision four central chapters:

Chapter: Higher-order probability See existing chapter and paper on higher-order legal probabilism.

Chapter: Narratives, specificity, coherence etc. See Rafal's paper on coherence.

Chapter: Cross-examination and arguments See Marcello's paper on cross-examination and Bayesian networks, and also paper on awareness growth and Bayesian networks.

Chapter: Gaps in Evidence See existing paper on gaps in the evidence.

This more sophisticated version of legal probabilism should answer some of existing challenges to simple legal probabilism.