Chapter XYZ: Weight of Evidence

1 Introduction

The aim of legal-fact finding is, quite literally, to find out the facts. To find out the facts—for example, whether or not the defendant embezzled a large sum of money as the prosecution claims—it is necessary to gather, scrutinize and assess evidence from a variety of sources. One side will present evidence against the defendant, and the other side, the defense, will present evidence in favor of the defendant. But this adversarial process must come to an end at some point. A decision about the facts must be made based on the evidence presented up to that point. This decision is preceded by another, namely a decision about whether the process of presenting, scrutinizing and assessing evidence should come to an end.

In principle, more evidence could always be presented about a disputed factual issue: witnesses, forensic experts, telephone conversations, camera recordings, and so on. But, at some point, gathering more evidence will no longer bring value in addition to what the evidence already available could do because the costs of seeking more evidence may become too great.

1.1 Epistemic value or informativeness

What does it mean for new evidence to be valuable? Some terminology can help here. New evidence can be *epistemically* or *informationally* valuable. It provides novel information that existing, already available evidence did not provide, where this novel information can be consistent or inconsistent with the existing evidence. New evidence can be valuable even when it does not provide novel information but simply agrees with other evidence, for example, when two witnesses report having seen the same thing. So, new evidence is epistemically valuable if it adds new information or strengthens the credibility of existing evidence. But, at some point, seeking more evidence will become less and less epistemically valuable.

1.2 Decisional value

The epistemic value of new evidence need not be the same as its *decisional* value. A new piece of evidence is decisionally valuable if, when added to the existing body of evidence, it would change the trial decision, given whatever decision criterion is used. A new piece of evidence can be epistemically valuable without being decisionally valuable, but not the other way around. Perhaps a new piece of evidence can add novel information or strengthen the credibility of other evidence, but these changes might not be significant enough to change the decision.

1.3 When to stop seeking more evidence?

If new evidence is regarded as valuable in the sense specified, should the trial decision be postponed to seek it? Not necessarily. We should not assume that evidence that is informationally and decisionally valuable—if added, it would change the trial decision—should be *ipso facto* sought by investigators or that trial proceedings should be prolonged in search of this evidence. The evidence could be too expensive or even impossible to obtain. Having a detailed camera recording of what happened could have great informational and decisional value, but this does not mean a trial judgment must be postponed until the recording is obtained. The costs of obtaining it might be too great.

In general, the question of whether to stop seeking more evidence is a practical question, moral or political. It requires balancing the costs of seeking more evidence with the epistemic and decisional value of the new evidence. These costs can be monetary, but in time and delayed trial decisions. Before performing this balancing act, however, we need an account of when additional evidence is epistemically and decisionally valuable. We aim to provide an account of the epistemic value of new evidence here.

1.4 Prospective and retrospective questions

We identified three different types of questions:

- What is the epistemic or informational value of an item of evidence?
- What is its decisional value?
- Should the trial decision be postponed to seek more evidence that is valuable?

These three questions are typically viewed prospectively but they can also be formulated retrospectively. The first two questions about the epistemic and decisional value of evidence need not be confined to evidence that has not yet been obtained. They could also concern evidence that was obtained. Or they could be about evidence that is missing and could no longer be obtained. In this sense, they can be both prospective and retrospective.

The prospective questions inform what investigators, prosecutors or defense lawyers should do as they seek more evidence. But even if the decision about seeking more evidence is foreclosed, one can still ask retrospectively what the investigators should have done. The latter question, in turn, can guide the formulation of possible remedies, say whether the defendant should be acquitted because of missing evidence that the investigator should have gathered.

1.5 Stylized legal case

A stylized legal case will help fix ideas. A more detailed discussion of realistic legal cases is to be found at the end of this chapter.

The defendant confessed to having killed the victim when he burglarized the victim's apartment and stole cash and other valuables kept in a tin box. The confession was the prosecution's key incriminating evidence, but at trial, the defense raised doubts that the defendant was covering up for his son. These doubts were corroborated by the fact that the defendant was rather short and it would take a tall person to kill the victim. And the defendant's son was much taller. On the other hand, the defendant confessed and provided abundant details all consistent with the findings of the investigators. One piece of evidence is missing: because the defendant confessed right away, the investigators did not examine the tin box for fingerprints. The box was later wiped clean. Had they found the son's fingerprints on it, that would be evidence of the defendant's innocence.

This case raises several questions:

- First, questions on the epistemic or informational value of the evidence. Retrospectively—knowing what we know now—what is the value of the missing fingerprint evidence? Prospectively—knowing what the investigators knew back them—what was the value of the fingerprint evidence?
- Second, questions about the decision. Retrospectively, would the fingerprint evidence—if it had been obtained during the investigation—have changed the trial decision?
- Third, questions about what to do. Prospectively or retrospectively, what should the investigators have done? This question is relevant to another, related yet different question: Should the defendant be acquitted because of the missing fingerprint evidence?

1.6 Weight v. balance of evidence

Another bit of terminology is the distinction between 'weight of evidence' and 'balance of evidence'. This distinction goes back to Keynes. The balance tracks whether the evidence presented tips for or against a hypothesis of interest. Weight, for Keynes, tracks the quantity of the evidence. Here both balance and weight are properties of the *total* body of evidence.

Instead, the notion of epistemic or informational value of evidence we are after is a property of a piece of evidence relative to an existing body of other evidence.

Section 2

2 First pass

2.1 Absolute probability difference

Let's start with the simplest account of the value of new evidence, mostly in terms of a probability change. Formally, the epistemic value of new evidence E_n, against the background of other evidence E_b, can be measured by the probability change the new evidence brings about relative to a hypothesis H of interest. So the epistemic value of E_n given background evidence E_b relative to hypothesis H is the absolute difference

$$|P(H|E_b) - P(H|E_b \wedge E_n)|.$$

The greater this difference, positive or negative, the greater the epistemic or infomational value of the new evidence. The smaller the difference, the lesser the value of the new evidence.

Dahlman's account of informativeness closely follows this definition. Nance offers a similar account. For him, the informational value (or weight) of the evidence is the absolute value of the log likelihood ratio

$$log(\frac{P(E_n|H \wedge E_b)}{P(E_n|\neg H \wedge E_b)}).$$

As we know, likelihood ratios mirror changes in probabilities: the greater the likelihood ratio, the greater the change upwards, and the lower the likelihood ratio, the greater the change downwards.

2.2 Objection

This account, however, is not completely satisfactory, as it may count additional evidence as worthless, while it is instead quite valuable. Suppose two witnesses testify one in favor and another against the defendant, one says they were with the defendant when the crime occurred away from the crime scene, while the other says they saw the defendant around the crime scene at the relevant time. Absent any reason to believe one witness more than the other, assume the chances are 50/50, that is $P(G|W1 \wedge W2) = .5$ Next, we consult two other witnesses, and again we find one favoring the defendant and the other incriminating the defendant. So, again, $P(G|W1 \wedge W2 \wedge W2 \wedge W5) = .5$. The probabilities have not changed, but the testimonies of the two additional witnesses are not worthless.

2.3 Weighted probability difference

Here is a quick fix. True, after listening to the two additional witnesses, we realize their input is worthless. But *before* listening to them, we did not know what they would have said. They could have spoken both against (or in favor) of the defendant, and in that case, the probability of G would have changed quite dramatically, either upwards or downwards. So, there are three possible scenarios: (S_1) they both speak against the defendant; (S_2) they both speak in favor; and (S_3) they speak one in favor and the other against. If each scenario is 1/3 likely, there is a 2/3 chance the new evidence would be epistemically valuable and only a 1/3 chance it would not be. So, overall, the new evidence is valuable.

The proposal would be this. The epistemic value of new evidence E_n given E_b and relative to H is the weighted sum of the difference $P(H|E_b) - P(H|E_b \wedge E_n^i)$, where the weights are the probabilities of the scenarios $P(S_i)$ that could materialize and $i \in \{1, 2, 3, ...k\}$:

$$\sum_{1}^{k} P(S_i) \times |P(H|E_b) - P(H|E_b \wedge E_n^i)|$$

2.4 Another objection (or recall higher-order probabilities)

This quick fix is still not satisfactory. In the previous chapter, we argued that precise probabilities are not enough because they fail to capture higher-order uncertainty. If precise probabilities are not enough to model uncertainty, then an account of the epistemic or informational value of new evidence that is based on precise probabilities will fall short for similar reasons.

Suppose an expert who listened to a telephone conversation testifies that the voice of the caller matches the defendant's voice. To quantify the uncertainty associated with the voice match, the expert introduces as is customary the random match probability, a number that expresses the likelihood that a random person would still be a match, simply by coincidence. This number is based on a probability model and sample data, so it is itself subject to uncertainty.[1]

Compare the voice expert with a genetic expert who testifies that the defendant genetically matches the crime traces. Suppose the genetic expert and the voice expert both testify, in different trials, that the defendant is a match, and they both give roughly similar random match probability. Should we say that both forms of evidence have the same value? After all, they change the guilt probability to the same extent. But the genetic match is better evidence than the voice match: better sample, better models, etc. Sticking to first-order probabilities neglects this additional layer of uncertainty and the difference between genetic match and voice

[1] This uncertainty is higher-order and is best modeled by a higher probability, a probability of a probability. What is the probability that the random match probability is what the expert says it is? Since the random match probability can take any value between 0 and 1, each of these values will have a higher-order probability, or in other words, there will be a distribution over the first-order probabilities (in this case, a distribution over the possible random match probabilities).

match. The next section outlines the right account of the epistemic or informational value of the evidence. This account incorporates higher-order uncertainty.

3 Our proposal

To model higher-order uncertainty, we propose using a probability density function f on a probability function $p(x)$ that captures the probability of hypotheses. This can be mathematically represented as $f(p(x))$. This construct effectively maps weights to their respective hypotheses by adding another layer of probabilities in the form of a probability density function, rather than using less modest approaches like extending the set of hypotheses with a variable representing the weight of the hypotheses.

A higher-order model for obtaining new evidence, which represents the weight of that evidence, can be described as follows:

- The set of possible hypotheses θ_n
- Prior probability distribution over the hypotheses $p(\theta_n)$
- A piece of evidence E
- Posterior probability distribution over the hypotheses $q(\theta_n)$

The update process follows Bayes' theorem:

$$q(\theta_n) = \frac{p(E \mid \theta_n) \times p(\theta_n)}{p(E)}$$

Consider a case where a DNA test plays a crucial role, and we are assessing the strength of the evidence given the random match probability (RMP). The RMP was calculated three times, as there were some doubts about the first test on a random sample. These pieces of evidence are represented as follows:

- 1 match in 200: $E_1 = \text{Beta}(201, 2)$
- 1 match in 2000: $E_2 = \text{Beta}(2001, 2)$
- 1 match in 20000: $E_3 = \text{Beta}(20001, 2)$

The sample sizes differ significantly, and intuitively, the weight of the evidence for E_3 is the highest.

Assuming a grid approximation and defining our set of hypotheses as $\theta_1, \ldots, \theta_{1000}$, we use a uniform prior defined with a beta distribution $p(\theta_n) = \text{Beta}(1, 1)$.

To compare the weights of these pieces of evidence, we can use information entropy $H = -\sum(p \log(p))$:

$$\text{Weight}(q) = 1 - \frac{H(q)}{H(p)}$$

The weights of the three posteriors achieved from learning from these three pieces of evidence are as follows:

- $\text{Weight}(q_1) \sim 0.5428$
- $\text{Weight}(q_2) \sim 0.8964$
- $\text{Weight}(q_3) \sim 1.0$

This clearly demonstrates the significance of the number of observations that the evidence introduces.

4 Legal cases

4.1 Tin box

4.2 Missing fingers