1997

Towards a (Bayesian) Convergence?

Richard D. Friedman

University of Michigan Law School, rdfrdman@umich.edu

Available at: <https://repository.law.umich.edu/articles/729>

Follow this and additional works at: <https://repository.law.umich.edu/articles>

Part of the [Evidence Commons](#), and the [Jurisprudence Commons](#)

Recommended Citation

Friedman, Richard D. "Towards a (Bayesian) Convergence?" *Int'l J. Evidence & Proof* 1 (1997): 348-53.

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

Towards a (Bayesian) convergence?

By Richard D. Friedman

Professor of Law, University of Michigan Law School

I am grateful that my paper has been a subject of attention – even critical attention – by so many able scholars, and I welcome the opportunity to offer responsive commentary. It is impossible to comment in depth on all the contributions to this symposium, so I will pick my spots.

Analytical uses

If I understand them correctly, several leading Bayesioskeptics (Allen, Callen, Stein) acknowledge – with varying degrees of specificity and varying degrees of grudgingness – that standard probability theory can be useful as an analytical tool in considering evidentiary doctrines and the probative value of evidentiary items. Assuming I am not guilty of wishful perception, such an acknowledgment is important, because this description covers most of the uses that other evidentiary scholars and I have attempted to make of probability theory. To make the point more concrete, I offer here a very incomplete list of evidentiary matters, ranging from the very broad to the quite specific, on which I believe probabilistic analysis has been helpful:

1. Conceptualising and determining the standard of persuasion;
2. Addressing the so-called problem of conjunction;
3. Conceptualising how the probative value of an item of evidence should be measured, and assessing that value for miscellaneous types of evidence, such as:
 - (a) character impeachment evidence,
 - (b) hearsay evidence, and
 - (c) missing evidence;
4. Assessing how expert evidence should be presented, and re-examining the out-of-fashion ultimate issue rule;
5. Analysing two important fallacies – the ‘prosecutor’s fallacy’ and the ‘defendant’s fallacy’ – that often infect trial arguments about evidence; and
6. Assessing conditions under which evidence offered under the extended *Hillmon* doctrine, described below, has significant value without relying on the declarant’s perception and memory of the other person’s intentions.

If I am right that probability theory is a useful analytical tool in these contexts, or at least in some of them, then I wonder why we legal Bayesians have to spend so much effort defending what we do. If I am wrong, then I believe Bayesioskeptics should show why, with specific attention to specific applications. But that, despite the invitation I

issued in my main essay, they hardly even attempt. In this symposium, one might find Allen repeating and elaborating on some old arguments – to which I have responded in my main essay (written simultaneously with his) – regarding the standard of persuasion and the so-called problem of conjunction¹ – but no further engagement with the work of legal Bayesians addressed to specific and significant evidentiary issues.

It bears emphasis that Bayesian analysis often achieves results that are not obvious but that are sound, readily explainable and intuitively appealing. At least so I have claimed, and the Bayesioskeptics have made no effort to show otherwise. In my main essay, I offered one example, having to do with the ultimate issue rule. Here, I will offer another illustration, this one drawn from my own work (Friedman, 1987: 718) and concerning the extended *Hillmon* doctrine mentioned above.² Under that doctrine, a declarant's statement of intent to engage in an activity with someone else may be offered to prove that the two did indeed engage in that activity. Using probability theory, I have demonstrated that the declarant's statement has greatest probative value in establishing that they did engage in the activity, even without relying on the declarant's perception and memory of the other person's intentions, if, as assessed apart from the declaration, (a) there is a substantial probability that the two people did engage in the joint conduct, and (b) the declarant's willingness to engage in the conduct is not a foregone conclusion.

Now, though I regard most of the leading Bayesioskeptics as friends, I am not sure I have ever been able to persuade them of anything. But I suspect that given enough time I would be able to explain the basis for this insight and persuade them that it is correct. Certainly it has a good deal of intuitive appeal, and yet I believe it is far from obvious; it takes some thinking to get there. Accepting a pair of points made by Callen, I am willing to believe that one *could* get there without overt probabilistic analysis (Bring makes this point as well), and certainly a good deal more than understanding of probability theory is necessary. But the point is that overt use of probability theory is of great assistance in reaching insights like this – which, despite the barrels of ink spilled in discussing the *Hillmon* doctrine, had escaped previous detection – and in explaining them clearly and rigorously.

Examples like this have another significance as well. Bayesioskeptics seem to contend (some of them, I think, pretty much all of the time, and pretty much all of them some of the time) that a variety of problems, such as computational complexity, makes Bayesian analysis infeasible in all but the most exotic of settings. But if that were so, legal Bayesians could not persistently yield sound and intuitively appealing results.

Algorithms and consistency

Allen tries to brand legal Bayesianism with the pejorative, mechanical-sounding, label 'algorithmic'. One aspect of this attempt is Allen's persistent – and, as pointed out by Robertson and Vignaux, erroneous – equating of probability theory with Bayes's

1 In discussing this matter in my principal essay, I mentioned the idea of the average probability of the elements. Two of the commentators (Bring, Carriquiry) seem to think that I regard this as a useful concept. In fact, I do not, and that was the point I was trying to express.

2 The doctrine takes its name from the famous case of the *Mutual Life Insurance Co. v Hillmon* 145 US 285 (1892). There, one Walters wrote that he intended to travel with Hillmon. The insurance company, attempting to prove that Hillmon had murdered Walters, offered this letter to prove that the two had indeed travelled together.

theorem, which is merely a consequence, albeit a particularly important one, of the theory.

In fact, probability theory does not provide an algorithm – a procedure, rule or mechanical set of steps – by which even an ideal fact finder should attempt to assess the probability of an uncertain factual proposition. In my principal essay, I have emphasised how flexible probability theory is, and how it may be applied to any given problem in a multitude of different ways.

What probability theory does do is to provide a set of consistency constraints on an evaluator's probability assessments. Notwithstanding Ligertwood's assertion that Bayesian logic is 'concerned with numerical accuracy', the legal Bayesian does not usually need to be worried about numerification at all. Thus, an informal, and totally non-numerical, statement of some of these constraints might be as follows:

1. All other things being equal, the more probable a proposition appears without consideration of a given body of evidence, the more probable it will appear upon consideration of that evidence.
2. All other things being equal, the more probable it appears that a given body of evidence would arise given the truth of a proposition, the more probable the proposition will appear given the body of evidence.
3. All other things being equal, the less probable it appears that a given body of evidence would arise given the falsity of a proposition, the more probable the proposition will appear given the body of evidence. ^

The Bayesioskeptics presumably would not deny the logical validity of any of these principles. Which of them, I wonder, do they think a rational fact finder can violate in any substantial way and still achieve sensible, mutually consistent results?

Incompleteness

Another kind of criticism of legal Bayesianism made repeatedly in this symposium (Callen; Ligertwood; Stein) is that probability theory is incomplete, in that it depends on values and judgments about which the theory has little or nothing to say. Certainly the theory is dependent in this way; probability theory is a set of constraints applied to assessments, and not a source of the elemental assessments themselves. But this consideration does not undermine the value of the theory – any more than the laws of physics are rendered useless because they do not reveal the mass of a given object but only indicate what happens in prescribed conditions to an object of a given mass.

Bayesians in this symposium (e.g. Carriquiry) have noted the importance of judgment in using probability theory to analyse evidentiary issues. Judgment is necessary in assessing not only the elemental probabilities (which, as Allen points out, include not only the prior probabilities but also conditional probabilities). Judgment is further necessary in aggregating possibilities to make use of the theory tractable. As Donnelly emphasises, judgments of this type – determining what types of approximation are useful, simplifying a task sufficiently at a tolerable cost in loss of information – are common in many contexts. There is no reason why they may not be made in conjunction with probability theory.

Judgment and intuition are also necessary, as Robertson and Vignaux usefully acknowledge, in forming hypotheses: Bayesian reasoning is 'a method of reasoning about hypotheses in the light of evidence' and 'does not formally provide a process for

generating new hypotheses', though it may stimulate such generation. But this consideration, as they further state, is not an argument against a probabilistic view of fact finding; it is merely an acknowledgment that probability theory provides only part of what a fact finder needs. Robertson and Vignaux have previously offered a description of hypothesis refinement in terms of probability theory, and Carriquiry offers another in this symposium. Probability theory may be of only limited *assistance* in generating new hypotheses, but the generation process does not demand *inconsistency* with the theory.

The standard of persuasion

One significant matter that arguably cannot be expressed solely in terms of probability theory is the standard of persuasion. In my principal essay I suggested that, along with probability of the proposition at issue, another component, representing quality or completeness of the evidence, may play a role in the standard. Kaye takes me to task for simply 'tossing a few algebraic symbols into a formula'. I certainly did present this idea in skeletal form – not because I believe I have provided a final answer but because I am not sure at this point how this idea might be elaborated. I do believe that concerns over quality or completeness account for the resistance many people, scholars and others, feel against a purely probabilistic definition of the standard of persuasion. Kaye has raised doubts about the need for an extra component. He relies on the unquestionably sound point that gaps in the evidence may themselves affect the probability that a fact finder assigns to a material proposition. I certainly agree that the 'missing evidence' inference is forceful, but I suspect that something else is at work as well. I believe many people have a more visceral feeling that, if too little is known – if, in Allen's terms, however much uncertainty there is, there is too much ignorance – then the court should not disturb the status quo, even though the best guess favours the plaintiff. In any event, like Kaye, I would welcome further thought on this matter.

Another complexity associated with the standard of persuasion is identifying the goal to be achieved, or the measure to be maximised. Probability theory itself cannot, of course, perform this function. The model I present in my principal essay assumes that the goal is to maximise expected utility, treating the utility of the outcome of each case as independent of the utility of the outcome of any other. This, as I understand it, is also the goal that DeKay (1996) favours; as Kaye points out, it is ironic that Allen cites this article in his favour. It is possible that the goal should be articulated differently; perhaps, for example, if a given standard of persuasion yields errors disproportionately favouring one type of litigant as compared to another (product manufacturers as compared to consumers, for example), that is a material factor. But, as Kaye suggests, whether or not this is so, the analytical framework provided by probability theory is necessary to consider the matter sensibly.

Prior probability and the presumption of innocence

Bayesioskeptics regard the concept of prior probability as a weak spot for a probabilistic approach to evidence. I think the term is probably unfortunate. The prior probability merely means the probability of a proposition as assessed without a given body of evidence. There is no requirement that it be assessed prior to what is called the posterior probability, though that of course is usually what occurs. It seems to me that concepts of prior and posterior probability are required of any system that would purport to help answer the fundamental question, as articulated in Federal Rule of

Evidence 401, of whether a given body of evidence makes a proposition of interest 'more probable or less probable than it would be without the evidence'.

There is no requirement that the prior remain fixed: if the prior and posterior probabilities and the likelihood ratio as first assessed by an observer are not mutually consistent according to probability theory, the observer can adjust one or more of them. What, then, is the point of the whole exercise, asks Bring. The answer is consistency. I do not believe that probability *assumes* consistency, as Callen puts it; rather, I would say that the theory *shows how to be* consistent.

The presumption of innocence may be thought of as a rule providing that, before the presentation of evidence, the fact finder in a criminal case must treat as low the probability of the facts underlying guilt. But how low? I do not think the solution offered by Fienberg (Fienberg and Kadane, 1983; Fienberg and Schervish, 1986) – essentially, as I understand it, that the prior probability need only be below the standard of persuasion – suffices, for that would allow a fact finder disposed to guilt to begin with a very high prior. In my principal essay, I offer some reflections on the matter. It is, I believe, a complication but not a serious theoretical problem, and not much at all of a practical problem.

Empiricism, complexity, and the *as if* world

As I have already made clear in my principal essay, I do not contend that fact finders in litigation ordinarily ought to be required or encouraged to make explicit or conscious use of Bayesian methods at trial. Even Robertson and Vignaux, perhaps the most unyielding Bayesians of all in the evidentiary debate, make clear in this symposium that they agree. Thus, it seems to me that many of the points offered by Allen, Ligertwood and others among the Bayesioskeptics about the difficulties of 'implementing' probability theory at trial are simply beside the point.

I certainly agree with the comments made by numerous commentators in this symposium (Lempert, Madansky, Michon) that it is important for some purposes to understand how fact finders actually do reason. And I further agree that most people are not, as Lempert says, 'natural Bayesians' in the sense of being felicitous with the conscious use of probability theory. Indeed, if one needed an illustration of this last point, one need look no further than Stein's contribution to this symposium. Stein is a person of superb intelligence and profound discernment. And yet, when he attempts to demonstrate Bayesian methods in a hypothetical involving medical malpractice, he makes a terrible bungle of the job. (Explaining this admittedly blunt assertion requires some space. I have already offered a full explanation as part of an essay I have written in response to an article by Stein and Porat, both soon to be published in volume 18 of the *Cardozo Law Review*, and so I will not repeat the argument here.)

Lempert explains that much of evidentiary discourse is about an ideal *as if* world. But why does that world matter? Think of the balance of probative value and prejudicial impact that, as exemplified by Federal Rule of Evidence 403, is a cornerstone of much of evidentiary law. To assess prejudicial potential – the bad impact the evidence might have on fact finders – we must know how fact finders will actually respond. But probative value, it seems to me, is the *good* impact the evidence might have, and so assessing it posits a fact finder dealing well with information. Different fact finders might make different assessments of the evidence and be equally considered rational. But in assessing probative value, the adjudicative system is entitled to posit a fact finder acting consistently, as indicated by probability theory.

I am therefore not sure how much it matters for present purposes whether the results achieved by fact finders tend to be roughly in accordance with probability theory. (I have, of course, acknowledged that people do not usually *apply* the theory or consciously go through anything like the calculations it suggests.) But I do believe that, when people are thinking well, this condition tends to hold, using as a baseline a stripped down application of probability theory that achieves simplicity at a tolerable cost of information loss. In my principal essay, I have sketched a few ways of thinking about such a stripped down system. Donnelly makes an additional useful point, that determination of whether a probability is greater than a given threshold – all that is necessary in this context – is far simpler than making an exact assessment. Michon's 'satisficing' suggestion may fit along the same lines.

And so too, ironically, may Allen's 'relative plausibility' approach. This approach resembles probability theory in at least one crucial respect,³ and the attempt to aggregate numerous slightly different possibilities into one hypothesis is just the type of simplification I am suggesting. Like Lempert, I wonder: Ron Allen as a Bayesian? How intriguing. What a delicious thought on which to end.

3 Allen seems to have acknowledged elsewhere (1991: 413) that this approach requires cardinality to account for standards of persuasion like 'clear and convincing'. But once cardinality is introduced, plausibility begins to sound a lot like probability.