(*final draft*) In A. Hájek and V. F. Hendricks, (eds.), *Probability and Statistics: 5 Questions,* Chapter 6, 2009, Automatic Press, pp. 65-74.

6

James Hawthorne

Associate Professor, Philosophy

University of Oklahoma, USA

1. Why were you initially drawn to probability theory and/or statistics?

As an undergraduate at the University of Kansas I was initially a math and physics major. In my sophomore year I took a Philosophy of Science course from Don Marquis. The two books for the course were Ernest Nagel's *The Structure of Science* and Wesley Salmon's *The Foundations of Scientific Inference*. I was immediately enthralled by the Bayesian-frequentist approach to confirmation theory presented by Salmon. At the same time I was taking an advanced course in symbolic logic. And it seemed to me that inductive logic, including confirmation, should be formalizable as well. Salmon made it clear that the Carnapian approach didn't work. But I thought that some sort of Bayesian approach looked like the most promising way to go. I then took Philosophy of Mathematics and more Philosophy of Science with Richard Cole, and a course on logical theory with Art Skidmore. I ended up dropping the physics major (loved the theory, hated the lab work), and majored in math and philosophy instead.

I then went to the University of Minnesota to do graduate work in philosophy. I spent most of my time there studying philosophy of science. The Minnesota Center for Philosophy of Science was thriving, and an exciting place to be – always lots of top philosophers coming through and giving talks. The Center held weekly seminars (not for credit) that focused on a different topic each semester (the Mind-Body problem, the nature of laws, confirmation theory, etc.). Early on Herbert Feigl still attended many of the meetings, and always had interesting and provocative things to say. I studied logic and philosophy of science, including Bayesian approaches to the epistemology of the sciences, with Tony Anderson, John Earman, Bill Hanson, Geoff Hellman, Grover Maxwell, and Paul Meehl. All of these guys were on my dissertation committee, although Grover Maxwell died quite a while before I completed the dissertation.

My dissertation advisor was Bill Hanson, who is a magnificent logician. I wrote my dissertation on Bayesian confirmation. It was an attempt to logicize confirmation as much as possible. I think that Carnap is right about the likelihoods (often) being logical, subject to a treatment in terms of logical form, but that prior probabilities cannot depend on logical form alone. The dissertation was an attempt to work this out. The main idea is to formalize the logic of direct inference likelihoods in a Bayesian context (on an object language rich enough to represent any scientific theory), and then employ Bayesian convergence as a way of dealing with the prior probabilities, by washing them out.

With regard to the formalization of direct inference, I was heavily influenced by a series of articles in the Journal of Philosophy by Kyburg and Levi, with a follow-up by Seidenfeld. I wanted to construct, within a Bayesian context, a logic for likelihoods from statistical hypotheses to instances based on logical form alone – i.e. to construct a logic like Kyburg's, but that works for Bayesian likelihoods. I approached this by starting with the Popper functions extended to full first-order logic, as developed by Hartry Field. I supplemented the object language with ZF set theory, so that it's rich enough to represent any scientific theory, and added a way to express contingent chance claims. The idea then was to add enough restrictions to these Popper-Field confirmation functions that they would all agree on the direct inference likelihoods from theories containing chance claims to conjunctions of their instances, even in cases where those chance claims are embedded in very rich, very complex scientific theories . The difficulty in doing this properly is that in a Bayesian context direct inferences are easily defeated when the chance statements on which they are based are accompanied by other statements; and trying to spell out exactly when the direct inference is defeated, and when not, in terms of logical structure alone turns out to be really complicated. (Those familiar with Kyburg's successive non-Bayesian attempts to do this sort of thing know what I mean about how hard it is to get this kind of thing right. Levi has proposed a way to do this in a Bayesian context, but I'm not convinced that his proposal works in full generality in terms of logical structure alone.) Anyway, my attempt at carrying this project out in the dissertation was not entirely successful, but I learned a lot from working on it. I've since done additional work on this on and off, and believe I've found a better approach, but I'm not yet satisfied with it, so haven't published on it. In any case, I had much better luck with the part on Bayesian convergence. I've refined that since, and have published on it.

It took me a number of years to complete the dissertation. The project was way too ambitious for a dissertation, but I wouldn't be dissuaded from trying. (I really owe a lot to Bill Hanson for his unwavering support , which permitted me to pursue my project to its end.)

While continuing to work on the dissertation, I got a job at the Honeywell Systems Research Center, which is located in Minneapolis. Jan Wald (who has a philosophy Ph.D.) was putting together an artificial intelligence group there, and hired me because of my training in formal logic. At Honeywell I had the opportunity to experiment with various AI techniques, including Bayesian networks. We worked on diagnostic, decision support, and data fusion systems, among others. My work on these applications reinforced my respect for Bayesian approaches to uncertain inference. I ended up working at Honeywell for about nine years. I left in 1989 to take a job as a faculty member in philosophy at the University of Oklahoma, where I've been ever since.

2. What is distinctive about your work in the foundations of probability or its applications?

I consider myself an "objectivist Bayesian" with regard to inductive inference and confirmation theory. I've got nothing against subjective degree-of-belief functions. They play an important role in decision theory. But I think that Bayesian confirmation functions have to be distinct from Bayesian belief-strength functions. There should, of course, be a connection between degree-of-confirmation and degree-of-belief. For, presumably the whole point of evaluating the degree-of-confirmation for a hypothesis is to use that to influence how strongly the agent should legitimately come to believe that the hypothesis is true. The easiest way to get that connection would be to have the confirmation function itself *just be* the belief function. But, it turns out, that won't work.

One of the central points of Bayesian confirmation is to bring objective likelihoods to bear in the evaluation of hypotheses and theories – to evaluate hypotheses based on what they "say" the evidence will be like, which is what the likelihoods represent. The reason that confirmation functions have to be distinct from belief functions is that an agent's belief-function likelihoods cannot easily maintain the objective values that confirmation function likelihoods are supposed to have. That's the lesson we should learn from Glymour's *problem of old evidence*. And, indeed, this problem turns out to be much worse for the objectivity of likelihoods than is generally recognized. An agent's belief function likelihoods have to represent the probability of the evidence when the hypothesis is added to *everything else* the agent already holds. But generally other (partial) beliefs the agent holds *must* end up interfering with the objective values that the likelihoods are supposed to have for confirmational purposes. For instance, when the hypothesis is statistical and the agent has *any* additional information (even an opinion or hunch) that involves how the particular evidential outcome at issue might turn out, it can be shown that this information (or hunch) *has to* interfere with the objective value that the (belief function) likelihood is supposed to have.

For example, suppose 'E' is some bit of evidence (e.g. the coin lands heads on the next toss) statistically implied to degree r (e.g. 1/2) by a hypothesis H (e.g. the coin is fair, and is tossed in the usual unbiased way on the next toss), so that the confirmational likelihood is $P[E \mid H] = r$ (e.g, for $r = 1/2$). Let F be any statement at all, say some statement that is intuitively not relevant in any way to how likely E should be on H (e.g. let F say "Jim will be pleased with the outcome for that next toss"). Now suppose an agent becomes certain that 'either E, or else not E but F' (either the coin lands heads on the next toss, or it doesn't land heads but Jim will be pleased with the outcome for the next toss). Let B represent the agent's belief function before she becomes convinced that 'either E, or else not E but F'. Presumably her belief function likelihood may possess the objective likelihood value for E on H: $B[E \mid H] = P[E \mid H]$. (If not, then my point that her belief function disagrees with the objective likelihood has already been granted right up front.). It will turn out that on our analysis, below, the degree of the agent's belief that F holds conditional on $\sim E\&H$ will be a relevant factor; so let her degree of belief in that regard have any value s at all other than 1 (i.e., $B[F \mid \sim E\&H] = s < 1$ – i.e., $B[F\&\sim E\&H] = s \times B[\sim E\&H] < B[\sim E\&H]$ – e.g. let s = 1/2). Now the agent learns in a completely convincing way (e.g. I seriously tell her so, and she believes me completely) that 'either E, or else not E but F', and she updates to a new belief function in the usual Bayesian way: for all statements S, $B_{\text{new}}[S] = \mathrm{B}[S \mid E\vee(\sim E\&F)]$. But that *has to* screw up the objectivity of her belief function likelihood for E on H, because:

$$
\begin{aligned}
B_{\text{new}}[E \mid H] &= \frac{B_{\text{new}}[E\&H]}{B_{\text{new}}[H]} \\[1ex]
&= \frac{B[E\&H \mid E\vee(\sim E\&F)]}{B[H \mid E\vee(\sim E\&F)]} \\[1ex]
&= \frac{B[E\&H\&(E\vee(\sim E\&F))]}{B[H\&(E\vee(\sim E\&F))]} \\[1ex]
&= \frac{B[E\&H]}{(B[H\&E] + B[H\&\sim E\&F])} \\[1ex]
&= \frac{B[E \mid H]}{(B[E \mid H] + \mathrm{B}[F \mid \sim E\&H]\, B[\sim E \mid H])} \\[1ex]
&= 1/(1 + [(1-r)/r]\, s)
\end{aligned}
$$

which cannot possibly be the objective value for the likelihood, r. (I.e., $r = 1/(1 + [(1-r)/r]\, s)$ if and only if either s=1 or r=1 – e.g. if r=1/2 and s=1/2, then $\mathrm{B}_{\text{new}}[\mathrm{E} \mid \mathrm{H}] = 2/3 \neq r$.) It turns out that a similar analysis applies whenever the agent's belief strength for *E*, $B[E]$, is altered in any way. The alteration need not make $B_{new}[E] = 1$ (as in the traditional problem of old evidence), and need not be due to becoming certain of a disjunction involving *E* (as in the example just given).

The point is that even the most trivial bit of "information" that involves *E* can completely undermine the objectivity of the direct inference likelihoods for Bayesian belief functions. And any real agent will very often possess some such trivial "information". Thus, if confirmation-function likelihoods are supposed to have objective (or intersubjectively agreed) values, then an agent's belief function likelihoods cannot generally be those employed by confirmation functions. So confirmation functions *must* be distinct from belief functions. To reiterate, confirmation-function likelihoods are supposed to represent what hypotheses "say" about the evidence, not what the agent would believe if the hypothesis were added to everything else she holds. So, a full Bayesian account of *confirmation and belief* will require confirmation functions that are distinct from belief functions, and some account of how the degrees-of-confirmation are supposed to be employed to inform an agent's degrees-of-belief.

It's also worth noting that on the sort of account I'm pushing, where we take confirmation functions to express a *logic of confirmation*, rather than as expressing *belief-strengths of ideal agents*, there is no "logical omniscience problem" for confirmation functions. The agent who employs confirmation functions to evaluate hypotheses is supposed to use them to inform her beliefs, but only to the extent that she is cognizant of the logical relationships involved, much as she might use logical entailments she knows about to inform her beliefs..

So, for confirmation functions the likelihoods represent the empirical import of scientific hypotheses and theories – what hypotheses "say" about what the evidential part of the world is like. These should be highly objective, or intersubjectively agreed to by the appropriate scientific community. To the extent that likelihoods aren't objective in this way, to that extent the hypothesis makes no empirically testable scientific claim. The objective likelihoods are supposed to represent the testable empirical import of scientific claims.

In a Bayesian confirmation theory the posterior probabilities represent the net confirmational plausibility of hypotheses after the evidence is taken into account. But posterior probabilities depend not only on likelihoods, but also on values for prior probabilities. Prior probabilities represent how plausible hypotheses are taken to be on the basis of *non-evidential* considerations. Such considerations need not be wholly *a priori*. They may well include both conceptual and broadly empirical considerations not captured by the likelihoods. However, because such plausibility assessments tend to vary among agents, critics often brand them as *merely subjective*, and take their role in the evaluation of hypotheses to be highly problematic. Bayesian confirmation theorists should counter that such assessments often do play an important role in the sciences, especially when there is insufficient evidence to distinguish among some of the alternatives. And, it should be pointed out, the epithet "*merely subjective*" is unwarranted. Such plausibility assessments are often backed by extensive arguments that may draw on forceful conceptual and empirical (but *non-evidential*) considerations (that go beyond the likelihoods). This seems to be the primary epistemic role of the thought experiment.

Consider, for example, the kinds of plausibility considerations brought to bear in assessing the various interpretations of quantum theory. Many of these considerations go to the heart of conceptual issues that were central to the development of the theory in the first place, and were originally introduced by those scientists who've made the greatest contributions to the theory's development, in their attempts to get a conceptual hold on the theory and its implications. Such arguments seem to play a legitimate role in the assessment of the relative plausibility of alternative views, especially when "distinguishing evidence" has yet to be found, or is far from definitive. We may often have good reasons besides the evidence to strongly reject some logically possible alternatives as *just too implausible*, or at least as much less plausible than some *better conceived* candidates. In fact, in evaluating hypotheses, we always do bring such considerations to bear, at least implicitly. For, given any hypothesis, logicians can always cook up numerous alternatives that agree with it on all the evidence thus far. Any reasonable scientist will reject most of these inventions immediately, because they look *ad hoc*, contrived, or “just foolish”. Such reasons for rejection appeal to neither purely logical characteristics of these hypotheses, nor to evidential considerations. All such reasons ultimately rest on plausibility assessments (at least implicitly) that are not part of the evidence itself (as represented by the likelihoods).

Although scientists often bring plausibility arguments to bear in assessing their views, such arguments are seldom decisive, though they may bring the scientific community into widely shared agreement with regard to the implausibility of some “logically possible” alternatives. It is arguably a virtue of Bayesian confirmation theory that it provides a place for such assessments to figure into the net evaluation of hypotheses. Prior probabilities remain “subjective” in the sense that agents may continue to disagree on the relative merits of plausibility arguments – and so disagree on the prior plausibilities of various hypotheses. But assessments of priors are far from being *mere subjective whims*. Moreover, it can be shown that when sufficient empirical evidence becomes available, much of the disagreement due to such plausibility assessments may be “washed out” or overridden by the evidence.

When hypotheses are empirically (evidentially) distinct, the influence of the prior probabilities can be effectively “washed out” (unless they are set extremely close to 0 or 1). There is an especially striking Bayesian convergence result that establishes this. The result I have in mind is not subject to the usual criticisms of Bayesian convergence results. This result doesn’t depend on prior probabilities at all – only on ratios of likelihoods. (So it’s a result that even a non-Bayesian likelihoodist should love.) I call it the “Likelihood Ratio Convergence Theorem”. It shows that for empirically distinct hypotheses, if a decent body of experiments or observations is conducted, it is highly likely that the resulting stream of evidential outcomes will be such as to drive the ratio of the likelihood for a false competitor as compared to the likelihood for the true hypothesis to approach zero. This result doesn’t suppose that the evidence is “identically distributed” – so it applies to almost any pair of empirically distinct hypotheses. It’s a “weak law of large numbers” type result that gives explicit lower bounds on the rate of convergence – there’s no need to wait for the infinite long run. It is a “convergence to truth” result (not “merely convergence to agreement”). It permits the non-evidential prior probabilities to be reassessed and changed at will – e.g., as new conceptual and “broadly empirical” considerations are introduced. And it doesn’t depend on countable additivity (though I personally have no problem with countable additivity, where appropriate). It follows from this result about likelihood ratios that the posterior probabilities of false hypotheses (when compared to a true hypothesis) will be driven ever closer to zero. As this happens, the posterior probability of the true hypothesis (or its disjunction with empirically equivalent rivals) approaches 1. The Likelihood Ratio Convergence Theorem itself depends only on the workings of the likelihoods -- they do all the heavy lifting in “washing out “ priors, and bringing about the convergence of posterior probabilities.

3. How do you conceive of the relationship between probability theory and/or statistics and other disciplines?

Probability plays an essential role in most of the sciences. It provides the logic of hypothesis confirmation (usually via likelihoods – priors and posteriors are often not made explicit). Quantum theory is essentially probabilistic. There I take the relevant notion to be that of “objective chance”, or “propensity”, or “causal probability”. In higher level sciences probability plays an important role in modeling phenomena. There the models again tend to draw on something like the notion of “objective chance”, or “propensity”, or “causal probability”, although these notions are (to a large extent) employed by the models as a way of abstracting away form an unmanageable multitude of specifics and details. In the social sciences the notion of subjective or personal probability is quite useful for modeling the preferences and actions of agents.

4. What do you consider the most neglected topics and/or contributions in probability theory and/or statistics?

Bayesian convergence results – i.e. results on the washing out of prior probabilities – have gotten a bum rap. In particular, the Likelihood Ratio Convergence Theorem overcomes all of the usual objections to Bayesian convergence results. But it has gotten almost no attention. One still inevitably hears the same old obsolete objections to Bayesian convergence repeated.

5. What do you consider the most important open problems in probability theory and/or statistics and what are the prospects for progress?

I think we need a better foundation for decision theory. One sort of improvement would be to disentangle preference from belief and confidence. That is, traditional accounts, like those of Savage and Jeffrey, set down axioms for preference relations, and then via representation theorems show that the notion of preference can be represented in terms of probabilities and utilities, such that expected utilities recover the preference relations. Some of the axioms these accounts draw on don't look very plausible as constraints on "preference". That may be due to the fact that the "preference axioms" have to be strong enough to implicitly encode comparative confidence relations as well (which in turn give rise to probabilistic belief strengths via the representation theorems). Thus, the notion of belief or confidence only arises "pragmatically", in the representation of preferences via probabilities (i.e. degrees of confidence) together with utilities. Comparative confidence is not a basic notion within the theory of preference.

Philosophically this approach was originally motivated by the idea that "belief and confidence" have to be operationalized in terms of their role in preference and choice behavior. These days most of us are no longer so squeamish about attributing real mental/cognitive states like belief and confidence to agents. Radical behaviorism has gone by the way. So it makes good sense to axiomatize comparative confidence on its own. And, in fact, we already know how to do that in an intuitively plausible way. The idea, then, is to add to the comparative confidence axioms a set of really plausible axioms for preference, and then prove an appropriate representation result for that account of confidence and preference that shows how they are captured by probabilities and utilities. Jim Joyce's approach in his book on Causal Decision Theory is an important move in this direction. But I think that better axioms, that disentangle preference from confidence more clearly, should be available. And perhaps such a reworking, if done right, would generalize decision theory so as to handle various paradoxical cases – e.g., cases where the expected utility is infinite, as in the old St. Petersburg game, or is not well-defined, as in the Pasadena game introduced by Nover and Hájek.

I also think that that something might be learned from working out the details of a Bayesian account of direct inference likelihoods in terms of the logical form of the statements involved. Practitioners would no more employ such a logic than do mathematician employ formal deductive logic to prove mathematical theorems. But formal deductive logic provides a "standard of rigor" for the mathematician in the sense that any mathematical "proof" of a theorem that could not in principle be reconstructed in formal logic is not "really a proof" after all. A proper account of direct inference likelihoods should provide a similar standard for inference from complex statistical theories to collections of their instances. In the deductive case, having such a standard has taught us important things about the nature of logic and mathematics. Perhaps such a standard for direct inference likelihoods would tell us important new things about the nature of statistical inferences (e.g., about the nature and role of the notion of "randomness" needed to warrant at least some such inferences).