Earman on Old Evidence and Measures of Confirmation

Branden Fitelson†
University of California–Berkeley

September 17, 2004

Abstract. In *Bayes or Bust?* John Earman quickly dismisses a possible resolution (or avoidance) of the problem of old evidence. In this note, I argue that his dismissal is premature, and that the proposed resolution (when charitably reconstructed) is reasonable.

†Thanks to Stephan Hartmann and Hannes Leitgeb for useful comments on previous drafts of this paper.

1 Earman's Dismissal

On page 121 of *Bayes or Bust?*, John Earman (1992) suggests, and then quickly dismisses, a possible way to avoid the problem of old evidence. Here is an excerpt from the relevant section of *Bayes or Bust?*.

> The original problem of old evidence would vanish for Bayesian personalists for whom $\Pr(E) \neq 1$, with Pr interpreted as personal degree of belief. ... However, denying that $\Pr(E) = 1$ only serves to trade one version of the old-evidence problem for another. Perhaps it was not certain in November 1915 that the true value of the anomalous advance was roughly 43" of arc per century, but most members of the scientific community were pretty darn sure, e.g., $Pr(E) = .999$. Assuming that Einstein's theory does entail E, we find that the confirmatory power $C(T, E)$ of E is $\Pr(T) \times .001/.999$, which is less than .001002. This is counterintuitive, since, to repeat, we want to say that the perihelion phenomenon did (and does) lend strong support to Einstein's theory.

Earman's claim is that, while the *qualitative* problem of old evidence would be avoided if Bayesians were to insist that $\Pr(E) < 1$, a *quantitative* problem of old evidence would still remain (at least, in the case of deductive evidence). In support of this claim, Earman gives an example which is an instance of the following general claim about incremental measures of degree of support $\mathfrak{c}$:

(1) If $H \vDash E$ and $\Pr(E) \approx 1$, then $\mathfrak{c}(H, E) \approx 0$.

Charitably, Earman is only claiming that (1) is true for the *difference measure* of degree of support $\mathfrak{c}(H, E) = d(H, E) =_{df} \Pr(H \mid E) - \Pr(H)$. Indeed, (1) is true if $\mathfrak{c} = d$, as the following theorem shows (see Appendix for proofs).

Theorem 1. *If $H \vDash E$ and $\Pr(E) = 1 - \epsilon$, then $d(H, E) \leq \frac{\epsilon}{1-\epsilon}$* (for small ϵ).[1]

In the footnote (#8) at the end of the passage quoted above, Earman claims that "similar problems arise if confirmatory power is measured in other ways."[2] It's not clear exactly what Earman means here, but, as we will see in the next section, claim (1) is not as robust as Earman might think.

[1] Similar theorems can be proven for several other measures of incremental support. For instance, the ratio measure $r(H, E) =_{df} \frac{\Pr(H \mid E)}{\Pr(H)}$ satisfies a similar theorem in which the upper-bound on $r(H, E)$ is $\frac{1}{1-\epsilon}$ (proof omitted). The ratio measure (or measures ordinally equivalent to r) have been used and/or defended by several philosophers, including Milne (1996). *Pace* Milne, there are many independent reasons to reject r as an adequate measure of incremental support. See Fitelson (1999) and (2001), and Eells and Fitelson (2001).

[2] In particular, Earman claims that Gaifman's (1985) measure of support $g(H, E) =_{df} (1 - \Pr(H))/(1 - \Pr(H \mid E))$ will suffer a similar fate as the difference measure does in such cases. Interestingly, this is incorrect, as $g(H, E)$ can be *arbitrarily large* in such cases. See the proof of Theorem 2 in the Appendix below for a demonstration of this.

2 The Measure-Sensitivity of Earman's Dismissal

As it turns out, Earman's dismissal is strongly sensitive to his choice $\mathfrak{c} = d$ of measure of support [in the sense of Fitelson (1999)]. That is, there are many other measures of incremental support (or confirmatory power) which *violate* (1). As Christensen (1999) points out, the measure $s(H, E) =_{df} \Pr(H \mid E) - \Pr(H \mid \neg E)$ does not satisfy (1), since $s(H, E)$ can be *arbitrarily large* in cases involving highly probable, deductive evidence. However, as Ellery Eells and I have argued [see Eells & Fitelson (2000) and (2001), and Fitelson (2001)], Christensen's measure s is woefully inadequate as a measure of inductive support, for many independent reasons. Interestingly, there is another, perfectly good Bayesian measure of inductive support which violates (1): the likelihood-ratio measure $l(H, E) =_{df} \frac{\Pr(E \mid H)}{\Pr(E \mid \neg H)}$, as the following theorem shows:

Theorem 2. *Even if* $H \vDash E$ *and* $\Pr(E) \approx 1$, $l(H, E)$ *can be arbitrarily large.*

There have been many ardent supporters of l as a measure of incremental support or confirmation, including Kemeny & Oppenheim (1952), Alan Turing, I.J. Good [see Good (1985) for a nice historical survey concerning l], and, more recently, Heckerman (1988), Schum (1994), and Fitelson (2001).

3 The Proposed Resolution — Revisited

I know of no arguments in the literature which aim to show that the difference measure d (used in Earman's reconstruction of the resolution) should be preferred to the likelihood-ratio measure l.[3] And, since the use of the likelihood-ratio measure avoids the quantitative problem of old evidence described by Earman, it seems to me that Earman's dismissal of this way of avoiding the problem of old evidence is somewhat premature. A more charitable reading of the proposed resolution seems only to depend on the following two claims:

(2) A rational Bayesain agent should not assign probability 1 to contingent empirical propositions (*e.g.*, evidential propositions E).

and

(3) The likelihood-ratio measure l is a perfectly adequate way for rational Bayesian agents to gauge confirmational power.

On page 121, Earman (1992) comments that if (2) is pushed to "its logical conclusion, we will eventually reach the position that no 'thing language' proposition of the sort useful in confirming scientific theories is ever learned for certain, and the strict conditionalization will collapse." But, Earman continues,

[3] Rosenkrantz (1981, Exercise 3.6) remarks that he knows of "no compelling considerations that adjudicate between" the difference measure and the likelihood ratio measure. The fact that l avoids Earman's quantitative problem of old evidence (but d does not) seems to count as evidence in favor of l. Christensen (1999) makes a similar point regarding his measure s. James Joyce (1999) also defends s (in contrast to d) by appealing to similar considerations.

"Bayesians are hardly at a loss here, since Jeffrey (1983) has proposed a replacement for strict conditionalization that allows for uncertain learning." So, it seems that Earman doesn't have any compelling argument against (2). Indeed, Jeffrey's 'radical probabilism' seems to provide a perfectly sensible underpinning for (2). Moreover, Earman does not give us any reason to doubt (3). And, as I mentioned above, there are many independent reasons to think (3) is true. This suggests that the above proposal for avoiding the problem of old evidence should probably not be dismissed so quickly.

Appendix: Proofs of Theorems

A Proof of Theorem 1

Theorem 1. *If* $H \vDash E$ *and* $\Pr(E) = 1 - \epsilon$*, then* $d(H, E) \leq \frac{\epsilon}{1-\epsilon}$ (for small ϵ).

Proof.

$$
\begin{aligned}
d(H, E) &= \Pr(H \mid E) - \Pr(H) \qquad \text{[def. of } d\text{]} \\
&= \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E)} - \Pr(H) \qquad \text{[Bayes' Theorem]} \\
&= \Pr(H) \cdot \frac{1 - \Pr(E)}{\Pr(E)} \qquad [H \vDash E, \text{ algebra}] \\
&= \Pr(H) \cdot \frac{\epsilon}{1 - \epsilon} \qquad [\Pr(E) = 1 - \epsilon] \\
&\leq \frac{\epsilon}{1 - \epsilon} \qquad [\Pr(H) \in [0, 1], \text{ small (non-negative) } \epsilon]
\end{aligned}
$$

❑

B Proof of Theorem 2

Theorem 2. *Even if* $H \vDash E$ *and* $\Pr(E) \approx 1$*,* $l(H, E)$ *can be arbitrarily large.*

Proof. We will provide an algorithm for generating probability models in which $H \vDash E$, and $\Pr(E)$ is as close to 1 as you like, but in which $l(H, E)$ is as large as you like. The algorithm for generating such probability spaces is as follows. Assume $H \vDash E$. Thus, $\Pr(H \,\&\, \neg E) = 0$. Now, set $\Pr(H \,\&\, E) = \frac{2\epsilon - 1}{\epsilon - 1}$, $\Pr(\neg H \,\&\, E) = \frac{\epsilon^2}{1-\epsilon}$, and $\Pr(\neg H \,\&\, \neg E) = \epsilon$. By letting $\epsilon \in (0, \frac{1}{2})$ be as close to zero as you like, $\Pr(E)$ will be as close to 1 as you like $(1 - \epsilon)$, and at the same time, $l(H, E)$ will be as large as you like $(\frac{1}{\epsilon})$. All spaces generated in this way will be *probability* spaces satisfying the conditions of the theorem.

The following proof shows that, when the conditions of the theorem are satisfied, the likelihood-ratio measure is approximately equal to Gaifman's (1985) measure $g(H,E) =_{df} \frac{\Pr(\neg H)}{\Pr(\neg H \mid E)}$:

$$\begin{aligned} l(H,E) &= \frac{\Pr(E \mid H)}{\Pr(E \mid \neg H)} && \text{[def. of } l\text{]} \\ &= \frac{1}{\Pr(E \mid \neg H)} && [H \vDash E] \\ &= \frac{Pr(\neg H)}{\Pr(\neg H \mid E) \cdot \Pr(E)} && \text{[Bayes' Theorem]} \\ &\approx \frac{Pr(\neg H)}{\Pr(\neg H \mid E)} && [\Pr(E) \approx 1] \end{aligned}$$

This, together with the proof of Theorem 2, shows that Earman's claim about g in footnote 8 on page 121 of *Bayes or Bust?* is incorrect. ❑

References

Christensen, D. (1999). Measuring confirmation. *Journal of Philosophy XCVI*, 437–61.

Earman, J. (1992). *Bayes or Bust?: A Critical Examination of Bayesian Confirmation Theory*. Cambridge: MIT Press.

Eells, E. and B. Fitelson (2000). Measuring confirmation and evidence. *Journal of Philosophy XCVII*, 663–672.

Eells, E. and B. Fitelson (2001). Symmetries and asymmetries in evidential support. *Philosophical Studies* (to appear).

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science 66*, S362–S378.

Fitelson, B. (2001). A Bayesian account of independent evidence with applications. *Philosophy of Science* (to appear).

Gaifman, H. (1985). On inductive support and some recent tricks. *Erkenntnis 22*, 5–21.

Good, I. (1985). Weight of evidence: a brief survey. In *Bayesian Statistics, 2* (Valencia, 1983), pp. 249–269. Amsterdam: North-Holland.

Heckerman, D. (1988). An axiomatic framework for belief updates. In L. Kanal and J. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 2*, pp. 11–22. New York: Elsevier Science Publishers.

Jeffrey, R. C. (1983). *The logic of decision* (Second ed.). Chicago, IL: University of Chicago Press.

Joyce, J. (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Kemeny, J. and P. Oppenheim (1952). Degrees of factual support. *Philosophy of Science 19*, 307–324.

Milne, P. (1996). $\log[p(h/eb)/p(h/b)]$ is the one true measure of confirmation. *Philosophy of Science 63*, 21–26.

Rosenkrantz, R. (1981). *Foundations and Applications of Inductive Probability*. Atascadero, Calif.: Ridgeview.

Schum, D. (1994). *The Evidential Foundations of Probabilistic Reasoning*. New York: John Wiley & Sons.