The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity

Branden Fitelson

Department of Philosophy

University of Wisconsin–Madison

October 22, 1998

Introduction

- Preliminaries: Some Bayesian Background & Our Framework
- Abstract Overview of the Problem of Measure Sensitivity
- Concrete Examples of the Problem of Measure Sensitivity
- Some Existing Attempts to Resolve the Problem
- Tabular Summary of Key Results
- Conclusion: Where do we go from here?

Preliminaries I: Some Bayesian Background

- E confirms H iff E is probabilistically correlated with H.
- More formally, E confirms H iff $\Pr(H \mid E) > \Pr(H)$.
- There are *many* logically equivalent ways to say this, *e.g.*,
  - E confirms H iff $\Pr(E \mid H) > \Pr(E \mid \bar{H})$.
  - E confirms H iff $\Pr(H \,\&\, E) > \Pr(H) \cdot \Pr(E)$.
- This leads to a *plethora* of possible *relevance measures*.
- Various differences, ratios, *etc.*, can be generated ...

Preliminaries II: Four Relevance Measures

- The following four measures have been proposed & defended
  - The *Difference* Measure: $d(H,E) =_{df} \Pr(H \mid E) - \Pr(H)$
  - The Log-*Ratio*: $r(H,E) =_{df} \log\left[\frac{\Pr(H \mid E)}{\Pr(H)}\right]$
  - The Log-*Likelihood-Ratio*: $l(H,E) =_{df} \log\left[\frac{\Pr(E \mid H)}{\Pr(E \mid \bar{H})}\right]$
  - Carnap's *Covariance* Measure:
    $\mathfrak{r}(H,E) =_{df} \Pr(H \,\&\, E) - \Pr(H) \cdot \Pr(E) = \Pr(E) \cdot d(H,E)$
- It is known that these measures are *not* ordinally equivalent.
- Does this *technical* non-equivalence affect any *actual* arguments in Bayesian Confirmation Theory?

The Problem of Measure Sensitivity I: Overview

- **Definition:** An argument $\mathcal{A}$ is *sensitive to choice of measure* if the validity of $\mathcal{A}$ varies, depending on which of the four measures d, r, l, or $\mathfrak{r}$ is used in $\mathcal{A}$. Otherwise, $\mathcal{A}$ is said to be *in*sensitive to choice of measure (or, more simply, *robust*).
- *Many* well-known arguments are sensitive to choice of measure.
- Why accept the conclusions of such arguments, without some reason to use certain measures rather than others?
- This makes many arguments in the field *enthymatic* — such logical gaps constitute *the problem of measure sensitivity.*

The Problem of Measure Sensitivity II: Some Examples

- Gillies's Version of the Popper-Miller Argument
- Rosenkrantz and Earman on "Irrelevant Conjunction"
- Eells on the Grue Paradox
- Horwich *et al.* on the Ravens Paradox
- Horwich *et al.* on the Variety of Evidence

Gillies's Version of the Popper-Miller Argument

- Gillies (and Popper-Miller) use the following property of the difference measure d to argue against Bayesianism:

(1) $$d(H, E) = d(H \vee E, E) + d(H \vee \bar{E}, E).$$

- As it turns out, neither the log-ratio measure r, nor the log-likelihood-ratio measure l has property (1).
- $\therefore$ Gillies's argument is *sensitive to choice of measure.*
- Gillies gives some compelling reasons to prefer d over r, but (as far as I know) he gives no reason to prefer d over l.
- So, as it stands, Gillies's argument is *enthymatic*.

Rosenkrantz on Irrelevant Conjunction

- Rosenkrantz provides a Bayesian resolution of the problem of Irrelevant Conjunction (*a.k.a.*, the Tacking Problem) which trades on the following property of the difference measure:

(2) If $H \models E$, then $d(H \,\&\, X, E) = \Pr(X \mid H) \cdot d(H, E)$.

- Neither r nor l has property (2).
- Like Gillies, Rosenkrantz gives some good reasons to reject r. However, he explicitly admits that he knows of "no compelling considerations that adjudicate between" d and l.
- This makes it unclear as to how one might consistently fill the gap in Rosenkrantz's argument.

Earman on Irrelevant Conjunction

- Earman gives a more robust resolution of the tacking problem which requires only the following logically weaker cousin of (2):

(2′) If $H \models E$, then $d(H \,\&\, X, E) < d(H, E)$.

- r violates even this weaker condition (2′), but l *satisfies* (2′).
- In this sense, Earman's account is *less* sensitive to choice of measure (*i.e.*, more robust) than Rosenkrantz's is.
- Nonetheless, it would still be nice to hear some *independent* reasons why we should prefer d over r.

Eells on the Grue Paradox

- Eells offers a Bayesian account of the Grue paradox which trades on the following property of the difference measure (where $\beta =_{df} \Pr(H_1 \,\&\, E) - \Pr(H_2 \,\&\, E)$, and $\delta =_{df} \Pr(H_1 \,\&\, \bar{E}) - \Pr(H_2 \,\&\, \bar{E})$):

(3) If $\beta > \delta$ and $\Pr(E) < \frac{1}{2}$, then $d(H_1, E) > d(H_2, E)$.

- Neither r nor l has property (3).
- Eells has offered (personal communication) an argument against the log-ratio measure r, but (as far as I know) he does not provide any (*independent*) reasons to prefer d over l.
- Pending such reasons, Eells's argument remains enthymatic.

Horwich *et al.* on Ravens & Variety of Evidence

- The vast majority of Bayesian explications of both the Ravens Paradox and the confirmational value of varied evidence presuppose the following (where $\mathfrak{c}$ is some relevance measure):

(4) If $\Pr(H \mid E_1) > \Pr(H \mid E_2)$, then $\mathfrak{c}(H, E_1) > \mathfrak{c}(H, E_2)$.

- Interestingly, Carnap's covariance measure $\mathfrak{r}$ *violates* (4).
- Typically, the advocates of such arguments have used either d or r in their arguments (note: d, r, and l *all* satisfy (4)).
- As far as I know, none of these commentators have given (*independent*) reasons to prefer their measures over Carnap's $\mathfrak{r}$ (or, over any other measures that violate (4) — see slide 13).

Some Existing Attempts to Resolve the Problem

- There do exist a few general arguments in the literature which rule-out all but a small class of ordinally equivalent measures (*e.g.*, Milne, Good, Carnap, and Heckerman).
- Others have given “piecemeal” arguments which attack a *particular* class of measures, but fail to rule-out other competing measures (*e.g.*, Rosenkrantz, Gillies, and Eells).
- Most notably, I have heard *no* compelling reasons to prefer the difference measure d over either l or $\mathfrak{r}$.
- Until such reasons are provided, the arguments of Gillies, Rosenkrantz, Eells, Horwich *et al.* will remain *enthymatic*.

Tabular Summary of Key Results

	Is $\mathcal{A}$ valid *wrt* the measure:			
Name of Argument $\mathcal{A}$	d?	r?	l?	$\mathfrak{r}$?
Rosenkrantz on Irrelevant Conjunction	YES	NO	NO	YES
Earman on Irrelevant Conjunction	YES	NO	YES	YES
Eells on the Grue Paradox	YES	NO	NO	YES
Horwich *et al.* on Ravens & Variety	YES	YES	YES	NO[a]
Gillies's Popper-Miller Argument	YES	NO	NO	YES

[a]There are other relevance measures *wrt* which the standard Ravens/Variety arguments do *not* go through (*e.g.*, Mortimer's measure $\Pr(E \mid H) - \Pr(E)$).

Conclusion: Where do we go from here?

- It seems to me that there are two viable general strategies for coping with the problem of measure sensitivity:
  1. Avoid the problem entirely, by making sure that all of one's confirmation-theoretic arguments are *robust*.
  2. Or, if there is some argument $\mathcal{A}$ which one cannot make *robust*, then one should give some *independent reasons* why those measures with respect to which $\mathcal{A}$ is *valid* should be preferred over other measures which render $\mathcal{A}$ *in*valid.
- In particular, those current defenders of the difference measure d should (where necessary) either seek robust arguments or explain why d should be preferred over both l and $\mathfrak{r}$.