Variation in Evidence and Simpson's Paradox

Corey Dethier

University of Notre Dame
Philosophy Department
corey.dethier@gmail.com

Jan. 11, 2020

Introduction

Motivation

There are a *lot* of different models of “variation in evidence” going under various different names: robustness, consilience, unification, coherence, focused correlation, triangulation...

Formal models include those offered by: Bovens and Hartmann (2003), Claveau (2013), Fitelson (2001), Heesen, Bright, and Zucker (2019), Lehtinen (2016, 2018), McGrew (2003), Myrvold (1996, 2003, 2017), Schlosshauer and Wheeler (2011), Schupbach (2005, 2018), Sober (1989), Staley (2004), Stegenga and Menon (2017), Wheeler (2009, 2012), and Wheeler and Scheines (2013), and that list doesn't include applications.

The project

The project in brief: provide a unified account (of unification).

This presentation in brief: weaken the assumptions of Bovens and Hartmann (2003), see what happens.

The project

The project in brief: provide a unified account (of unification).

This presentation in brief: weaken the assumptions of Bovens and Hartmann (2003), see what happens.

Initial reaction: avoiding Simpson's paradox is a sufficient condition on varied evidence confirming!

The project

The project in brief: provide a unified account (of unification).

This presentation in brief: weaken the assumptions of Bovens and Hartmann (2003), see what happens.

Initial reaction: avoiding Simpson's paradox is a sufficient condition on varied evidence confirming!

Present thought: the connection with Simpson's paradox shows why this sort of analysis is going to get into trouble.

The plan

1. Why you might want a reliability-based model.
2. The relationship between confirmation and Simpson's paradox.
3. Why this relationship is a problem and not a solution.

Reliability

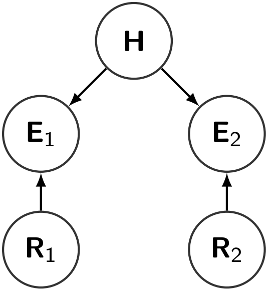
Sources of evidence

Consider:

- Witnesses testifying to the same fact.
- Multiple thermometers.
- Peterson (2003): study shows that global warming trend is robust across changes in location.

Crucial to these examples is that there's a difference between the *sources* of information.

The basic picture



H and **R** jointly control **E**; "varation" can be defined in terms of probabilistic relationships between **R** variables.

E.g.:

$$V = \frac{Pr(R_1 \vee R_2) - Pr(R_1 \& R_2)}{Pr(R_1 \vee R_2)}$$

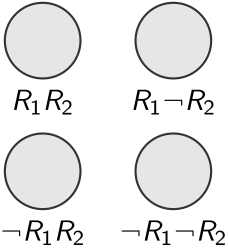
How does **E** affect **H**?

Suppose we learn E_1 and E_2.

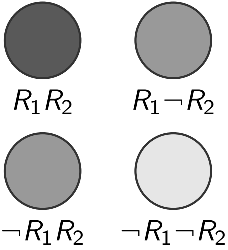
1. Direct effect: changes the probability of H given R_1 and/or R_2.
2. Indirect effect: changes the probability of R_1 and/or R_2.

The direct effect

Before learning $E_1 \& E_2$:

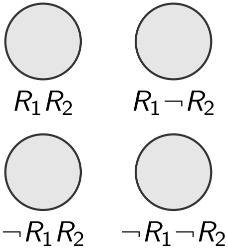


After learning $E_1 \& E_2$:

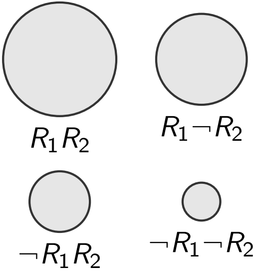


The indirect effect

Before learning $E_1 \& E_2$:



After learning $E_1 \& E_2$:



Three idealizations

IC1: H is probabilistically independent of the reliability of any source: $Pr(H) = Pr(H|R_i) = Pr(H|\neg R_i)$.

IC2: The posterior probability given by reliable evidence is not affected by the reliability of other sources of evidence: $Pr(H|E_i, R_i, R_j) = Pr(H|E_i, R_i, \neg R_j)$.

EC: there's no conditionalization on unreliable evidence: for all X, then $Pr(H|E_i, \neg R_i, X) = Pr(H|\neg R_i, X)$.

The direct effect

Let $\delta(H,E) = Pr(H|E) - Pr(H)$. Then:

$$\begin{aligned}
\delta(H, E_1 \& E_2) = {} & Pr(R_1, R_2) && \times \delta(H, E_1 \& E_2 | R_1, R_2) \\
& + Pr(R_1, \neg R_2) && \times \delta(H, E_1 | R_1, \neg R_2) \\
& + Pr(\neg R_1, R_2) && \times \delta(H, E_2 | \neg R_1, R_2)
\end{aligned}$$

The direct effect

Let $\delta(H,E) = Pr(H|E) - Pr(H)$. Then:

$$\begin{aligned}
\delta(H, E_1\&E_2) = {} & Pr(R_1, R_2) && \times\delta(H, E_1\&E_2|R_1, R_2)\\
& + Pr(R_1, \neg R_2) && \times\delta(H, E_1|R_1, \neg R_2)\\
& + Pr(\neg R_1, R_2) && \times\delta(H, E_2|\neg R_1, R_2)
\end{aligned}$$

The only value that can be negative is $\delta(H, E_1\&E_2|R_1, R_2)$ (compare Mayo-Wilson 2011, 2014; Stegenga and Menon 2017).

The direct results

Result 1: Sufficient condition on confirmation:

$$-\delta(H, E_1 \& E_2 | R_1, R_2) < \frac{V(R_1, R_2)}{1 - V(R_1, R_2)} \delta(H, E|R)$$

Result 2: (Assuming that the sufficient condition holds:) increasing $Pr(R_1 \vee R_2)$ increases the degree of confirmation, *ceteris paribus*.

Simpson's Paradox

The indirect effect

Recall: learning E_1 and E_2 has two effects.

1. Direct effect: changes the probability of H given R_1 and/or R_2.
2. Indirect effect: changes the probability of R_1 and/or R_2.

We've only discussed the direct effect. How does considering the indirect effect change things?

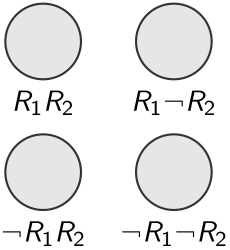
More complexity!

This is what $\delta(H, E_1 \& E_2)$ looks like (**EC** enforced):

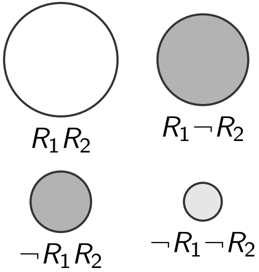
$$
\begin{aligned}
&= Pr(R_1, R_2|E_1, E_2)Pr(H|E_1, E_2, R_1, R_2) - Pr(R_1, R_2)Pr(H|R_1, R_2) \\
&+ Pr(R_1, \neg R_2|E_1, E_2)Pr(H|E_1, R_1, \neg R_2) - Pr(R_1, \neg R_2)Pr(H|R_1, \neg R_2) \\
&+ Pr(\neg R_1, R_2|E_1, E_2)Pr(H|E_2, \neg R_1, R_2) - Pr(\neg R_1, R_2)Pr(H|\neg R_1, R_2) \\
&+ Pr(\neg R_1, \neg R_2|E_1, E_2)Pr(H|\neg R_1, \neg R_2) - Pr(\neg R_1, \neg R_2)Pr(H|\neg R_1, \neg R_2)
\end{aligned}
$$

The same condition identified earlier

Before learning $E_1 \& E_2$:



After learning $E_1 \& E_2$:



Not quite that simple

Recall **IC1**: H is probabilistically independent of the reliability of any source: $Pr(H) = Pr(H|R_i) = Pr(H|\neg R_i)$.

What happens if we relax this assumption?

Not quite that simple

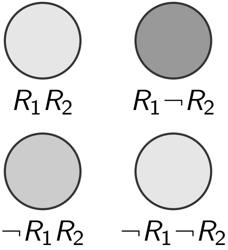
Recall **IC1**: H is probabilistically independent of the reliability of any source: $Pr(H) = Pr(H|R_i) = Pr(H|\neg R_i)$.

What happens if we relax this assumption? Same result for $\delta(H, E_1 \& E_2)$:

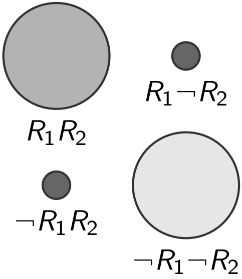
$$\begin{aligned}
&= Pr(R_1, R_2|E_1, E_2)Pr(H|E_1, E_2, R_1, R_2) - Pr(R_1, R_2)Pr(H|R_1, R_2) \\
&+ Pr(R_1, \neg R_2|E_1, E_2)Pr(H|E_1, R_1, \neg R_2) - Pr(R_1, \neg R_2)Pr(H|R_1, \neg R_2) \\
&+ Pr(\neg R_1, R_2|E_1, E_2)Pr(H|E_2, \neg R_1, R_2) - Pr(\neg R_1, R_2)Pr(H|\neg R_1, R_2) \\
&+ Pr(\neg R_1, \neg R_2|E_1, E_2)Pr(H|\neg R_1, \neg R_2) - Pr(\neg R_1, \neg R_2)Pr(H|\neg R_1, \neg R_2)
\end{aligned}$$

A new problem emerges

Before learning $E_1 \& E_2$:



After learning $E_1 \& E_2$:



Simpson's paradox

Pearl (2014): "Simpson's paradox refers to a phenomena whereby the association between a pair of variables (X, Y) reverses sign upon conditioning of a third variable, Z, regardless of the value taken by Z. If we partition the data into subpopulations, each representing a specific value of the third variable, the phenomena appears as a sign reversal between the associations measured in the disaggregated subpopulations relative to the aggregated data, which describes the population as a whole."

What's happened: each worldly "subpopulation" observes an increase in confirmation while confirmation decreases overall.

A cool result?

Potential upshot: for confirmation from varied evidence, all we need is to (a) avoid Simpson's paradox situations and (b) avoid the reversals discussed by Stegenga and Menon (2017).

And that result would hold in a general setting, with relatively few idealizations.

A problem, not a solution

Moving forward

What's the next step for a theory of variation in evidence?

Based on the above, an account of how **R** is affected by **E**—i.e., how our the probability of reliability changes with multiple confirming reports.

(That's essentially what Bovens and Hartmann (2003) and Claveau (2013) are both doing.)

The problem

Notice, however, that **E** will have both direct and indirect (through **H**) effects on **R**.

Bovens and Hartmann (2003) and Claveau (2013) both avoid this problem with IC1.

But IC1 is horribly unrealistic.

The problem, then

Claim: There's no interesting general relationship between the hypotheses that we're interested in testing and the (un)reliability of our tools.

Means that we're unlikely to be able to use the present tools to say anything more interesting about when these weird reversals occur.

Thank you

Thank you!

Bovens, Luc and Stephan Hartmann (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

Claveau, François (2013). The Independence Condition in the Variety-of-Evidence Thesis. *Philosophy of Science* 80.1: 94–118.

Fitelson, Branden (2001). A Bayesian Account of Independent Evidence with Applications. *Philosophy of Science* 68.3: 123–40.

Heesen, Remco, Liam Kofi Bright, and Andrew Zucker (2019). Vindicating Methodological Triangulation. *Synthese* 196.8: 3067–81.

Lehtinen, Aki (2016). Allocating Confirmation with Derivational Robustness. *Philosophical Studies* 173.9: 2487–509.

– (2018). Derivational Robustness and Indirect Confirmation. *Erkenntnis* 83.3: 539–76.

Mayo-Wilson, Conor (2011). The Problem of Piecemeal Induction. *Philosophy of Science* 78.5: 864–74.

– (2014). The Limits of Piecemeal Causal Inference. *The British Journal for the Philosophy of Science* 65.2: 213–49.

McGrew, Timothy (2003). Confirmation, Heuristics, and Explanatory Reasoning. *The British Journal for the Philosophy of Science* 54.4: 553–67.

Myrvold, Wayne (1996). Bayesianism and Diverse Evidence: A Reply to Andrew Wayne. *Philosophy of Science* 63.4: 661–65.

– (2003). A Bayesian Account of the Virtue of Unification. *Philosophy of Science* 70.2: 399–423.

– (2017). On the Evidential Import of Unification. *Philosophy of Science* 84.1: 92–114.

Pearl, Judea (2014). Comment: Understanding Simpson's Paradox. *The American Statistician* 68.1: 8–13.

Peterson, Thomas C. (2003). Assessment of Urban Versus Rural In Situ Surface Temperatures in the Contiguous United States: No Difference Found. *Journal of Climate* 16.18: 2941–59.

Schlosshauer, Maximilian and Gregory Wheeler (2011). Focused Correlation, Confirmation, and the Jigsaw Puzzle of Variable Evidence. *Philosophy of Science* 78.3: 376–92.

- Schupbach, Jonah (2005). On a Bayesian Analysis of the Virtue of Unification. *Philosophy of Science* 72.4: 594–607.
- – (2018). Robustness Analysis as Explanatory Reasoning. *British Journal for the Philosophy of Science* 69.1: 275–300.
- Sober, Elliot (1989). Independent Evidence About a Common Cause. *Philosophy of Science* 56.2: 275–87.
- Staley, Kent W. (2004). Robust Evidence and Secure Evidence Claims. *Philosophy of Science* 71.4: 467–88.
- Stegenga, Jacob and Tarun Menon (2017). Robustness and Independent Evidence. *Philosophy of Science* 84.3: 414–35.
- Wheeler, Gregory (2009). Focused Correlation and Confirmation. *The British Journal for the Philosophy of Science* 60.1: 79–100.
- – (2012). Explaining the Limits of Olsson's Impossibility Result. *Southern Journal of Philosophy* 50.1: 136–50.
- Wheeler, Gregory and Richard Scheines (2013). Coherence and Confirmation through Causation. *Mind* 122.485: 135–70.