Logical Foundations of Evidential Support

Branden Fitelson

Department of Philosophy
University of California–Berkeley

branden@fitelson.org
http://fitelson.org/

Overview of Presentation

Preliminaries I: Some General Bayesian Background

- Bayesianism (*i.e.*, "Bayesian *epistemology*") assumes that the degrees of belief (or degrees of credence) of rational agents are *probabilities*.
- Let $\Pr(H)$ denote the (unconditional) degree of belief that a rational agent S assigns to H (at some time t). This is S's *prior* probability of H (at t).
- $\Pr(H \mid E)$ is the degree of belief S assigns to H (at t), *on the supposition that E* (*i.e.*, the d.o.b. that S *would* assign to H *upon learning E*). This is S's *posterior* probability of H (on E, at t). I omit "S"s and "t"s hereafter.
- A simple toy example (just to help fix our ideas): Let H be the hypothesis that a card (drawn at random from a standard deck) is a spade, and let E be the (evidential) proposition that the card is the ace of spades.
- Given standard assumptions about random card draws, $\Pr(H) = 1/4$ and $\Pr(H \mid E) = 1$. So, learning E raises the probability of (indeed, *verifies*) H.

Preliminaries II: Basics of Bayesian Confirmation

- According to contemporary Bayesianism, evidence E *confirms* (or *supports*) a hypothesis H if learning E *raises the probability of H*.
- If learning E *lowers* the probability of H, then E *dis*confirms (or *counter*-supports) H, and if learning E does not change the probability of H, then E is confirmationally *irrelevant* to (or neutral regarding) H.
- There are *many* logically equivalent ways of saying that E confirms H. Here are three of these (where $\sim X$ is the *logical negation* of X).
  - E confirms H if $\Pr(H \mid E) > \Pr(H)$. [$1 > \frac{1}{4}$ in card example]
  - E confirms H if $\Pr(E \mid H) > \Pr(E \mid \sim H)$. [$\frac{1}{13} > 0$ in card example]
  - E confirms H if $\Pr(H \mid E) > \Pr(H \mid \sim E)$. [$1 > \frac{12}{51}$ in card example]
- By taking differences, ratios, *etc.*, of the left/right sides of such inequalities, a *plethora* of possible *quantitative* Bayesian *relevance measures* $\mathfrak{c}(H, E)$ of the *degree* to which E confirms (or supports) H can be constructed.

Four Popular Bayesian Relevance Measures of Confirmation

- *Dozens* of Bayesian relevance measures of the degree to which E confirms H have been proposed & applied. Here are four popular and representative choices for $\mathfrak{c}$, generated using the three inequalities on the previous slide.[a]
  - The *Difference*: $d(H,E) = \Pr(H \mid E) - \Pr(H)$
  - The *Log-Ratio*: $r(H,E) = \log\left[\frac{\Pr(H \mid E)}{\Pr(H)}\right]$
  - The *Log-Likelihood-Ratio*: $l(H,E) = \log\left[\frac{\Pr(E \mid H)}{\Pr(E \mid {\sim}H)}\right]$
  - The "Normalized Difference":
$$s(H,E) = \Pr(H \mid E) - \Pr(H \mid {\sim}E) = \frac{1}{\Pr({\sim}E)} \cdot d(H,E)$$
- The first part of my presentation concerns the *disagreement* exhibited by these measures, and its ramifications for confirmation theory ...

[a] Logs of ratios are taken to ensure positive/negative/zero values when E confirms/disconfirms/is irrelevant to H. This is just a useful convention; it does not affect the ordinal structure of the measures.

Surprising and Radical Disagreement Between Measures

- *All* Bayesian confirmation measures (including our 4) agree on *qualitative* judgments of the form "E confirms/disconfirms/is irrelevant to H".
- Surprisingly, our four measures of degree of confirmation (and many others) can disagree radically on *comparative* or *ordinal* judgments of the form "E_1 confirms H_1 more strongly than E_2 confirms H_2."
- This *ordinal* disagreement between measures of confirmation has serious consequences for a wide variety of arguments in the literature (many of which have previously gone unnoticed and/or unappreciated).
- For instance, it is part of Bayesian Lore that the observation of a black raven (E_1) confirms the hypothesis that all ravens are black (H) more strongly than the observation of a red herring (E_2) does. But, this conclusion *depends sensitively on one's choice of confirmation measure*.
- *Almost all* comparative arguments are *sensitive to choice of measure*!

Tabular Summary of Some Measure-Sensitive Arguments

	Valid *wrt* relevance measure:			
Argument	d?	r?	l?	s?
Horwich *et al.* on Hempel's Ravens Paradox	Yes	Yes	Yes	No
Eells on Goodman's "Grue" Paradox	Yes	No	No	Yes
Sober on Goodman's "Grue" Paradox	Yes	No	Yes	Yes
Rosenkrantz on Irrelevant Conjunction	Yes	No	No	Yes
Earman on Irrelevant Conjunction	Yes	No	Yes	Yes
Horwich *et al.* on the Variety of Evidence	Yes	Yes	Yes	No
Christensen on the Old Evidence Problem	No	No	Yes	Yes
Popper-Miller's *Critique* of Bayesianism	Yes	No	No	Yes
Earman's Old Evidence *Critique* of Bayesianism	Yes	Yes	No	No

Inductive-Logical Foundations I

- I begin Part Two with some words of Carnap's (from *Logical Foundations of Probability*). The first quote gives us a sense of the basic idea behind inductive logic, as a quantitative analogue of deductive logic:

  > "Deductive logic may be regarded as the theory of the relation of logical consequence, and inductive logic as the theory of another concept [$\mathfrak{c}$] which is likewise objective and logical, *viz.*, ... degree of confirmation."

- The next two quotes give Carnap's *intuitive* characterizations of the terms "logical" and "objective" as they apply to the confirmation relation $\mathfrak{c}$.

  > "The principal common characteristic of the statements in both fields is their independence of the contingency of facts. This characteristic justifies the application of the common term 'logic' to both fields."

  > "That $\mathfrak{c}$ is an objective concept means this: if a certain $\mathfrak{c}$ value holds for a certain hypothesis with respect to a certain evidence, then this value is entirely independent of what any person may happen to think about these sentences."

Inductive-Logical Foundations II: Three Historical Desiderata

- Abstracting away from the idiosyncrasies of specific attempts to explicate a *logical* $\mathfrak{c}$ (*e.g.*, Carnap), we can extract the following three (nearly universal) central desiderata from the historical literature on inductive logic:

($\mathcal{D}_1$) Inductive logic should provide a *quantitative generalization of deductive logic*. That is, the relations of deductive entailment and deductive refutation should be captured as limiting (extreme) cases of confirmation (disconfirmation) with cases of 'partial entailment' and 'partial refutation' lying somewhere on a $\mathfrak{c}$-continuum (or range) between these extremes.

($\mathcal{D}_2$) Inductive logic should use *probability* (in its technical, modern sense) as its central conceptual building block. [Mainly *historical* vs *material*.]

($\mathcal{D}_3$) Inductive logic (*i.e.*, the *non*-deductive relations between propositions that are characterized by inductive logic) should be *objective* and *logical*.

- In what follows, I will adopt Carnap's *intuitive*, *pre-theoretical* usage (above) of the terms 'logical' and 'objective' (I'll *diverge* from his *theoretical* usage).

Inductive Logic: A Naive Proposal

- According to classical deductive logic, $\{P_1, \ldots, P_n\}$ *entails* C iff the material conditional $P \supset C$ is (logically) *necessarily* true (P abbreviates $P_1 \,\&\, \cdots \,\&\, P_n$).
- Naively, then, one might try to quantitatively generalize deductive entailment by saying that the argument from $\{P_1, \ldots, P_n\}$ to C is *inductively strong* if the material conditional $P \supset C$ is (logically?) *probably* true. This suggests:

(NIL) $\mathfrak{c}(C, P) = \Pr(P \supset C) = \Pr(\text{either} \sim\!P \text{ or } C)$

- But, as Skyrms explains, (NIL) will not do. Consider the following argument:

  (P) There is a man in Cleveland who is 1999 years and 11-months-old and in good health. Therefore, (C) No man will live to be 2000 years old.

- Intuitively, this argument is *not* strong (P seems to *dis*confirm C). But, intuitively, $\Pr(P \supset C) = \Pr(\text{either} \sim\!P \text{ or } C)$ is high, since $\Pr(C)$ is high. Thus, $\Pr(P \supset C)$ does *not* capture the *confirmation relation* between P and C.[a]

[a] Note that (NIL) also violates ($\mathcal{D}_1$). If P *refutes* C, then $\Pr(P \supset C) = \Pr(\sim\!P)$, which is *not* minimal.

Inductive Logic: "The Received View"

- According to Skyrms, (NIL) conflates the *probability of the material conditional* $P \supset C$ with *the conditional probability of* C, *given* P; and, it is the latter that captures the *confirmation relation*. This is "The Received View":

(TRV) $\mathfrak{c}(C, P) = \Pr(C \mid P) = \dfrac{\Pr(P \,\&\, C)}{\Pr(P)}$

- It is clear that (TRV) copes better with Skyrms' example than (NIL) does. Intuitively, *given that* there is a healthy 1999 year + 11-month-old man in Cleveland, the probability that no man will live to be 2000 is *not* high.
- It also seems clear that (TRV) satisfies the first two of our historical desiderata.

$\mathcal{D}_1$: If P entails C, then $\Pr(C \mid P) = 1$ (maximal), and if P refutes C (*i.e.*, if P entails $\sim\!C$), then $\Pr(C \mid P) = 0$ (minimal), with a continuum in between.

$\mathcal{D}_2$: (TRV) *identifies* $\mathfrak{c}$ with *conditional probability* (in its technical sense).

- What about desideratum ($\mathcal{D}_3$)? Does (TRV) supply a conception of $\mathfrak{c}$ that is *logical* and *objective*? This depends on how (TRV) *interprets* "$\Pr(C \mid P)$" …

"Logical" Probability?

- If one accepts (TRV), then it *seems* that one must provide a "logical" (or *a priori*) probability function $\Pr(\cdot \mid \cdot)$ in order to satisfy desideratum ($\mathcal{D}_3$).
- Many philosophers — including Leibniz, Wittgenstein, Keynes, Waismann, Carnap, and many others — have tried to explicate something like this:

$$\Pr(C \mid P) = \frac{\Pr(P \,\&\, C)}{\Pr(P)} = \frac{\text{The proportion of logically possible worlds in which } P \,\&\, C \text{ is true}}{\text{The proportion of logically possible worlds in which } P \text{ is true}}$$

- Unfortunately, no adequate, general account of "logical" or *a priori* probability has ever been found (the consensus now is that no such account is forthcoming). Here, the *Principle of Indifference* rears its ugly head …
- It seems that there is no "logical aether" with respect to which propositions receive their "absolute probabilities". Probability seems to be an *inherently relational* property of propositions. Most philosophers (*e.g.*, Bayesians) have come to view this relativity as *epistemic*, but to do so is to abandon ($\mathcal{D}_3$).
- As such, it seems that we need to take a different approach in order to satisfy desideratum ($\mathcal{D}_3$) using probability. I propose a "Neo-Carnapian" solution …

A "Neo-Carnapian" Approach to Inductive Logic I

- When it came to *de*ductive logic, Carnap adopted a *Principle of Tolerance*, which said (roughly) that the deductive logician is not in the business of telling us which deductive-logical frameworks we should adopt ("external").
- The deductive logician's role is simply to tell us what the deductive relations are *within* specified deductive-logical frameworks ("internal" to logic).
- I suggest that we make a *similar* kind of move in the case of inductive logic. I submit that the inductive logician is not in the business of telling people which probability models they should use (*that*, I submit, is *not* a *logical* question).
- The inductive logician's job is to explicate the inductive-logical (*confirmation*) relations between statements *within* (*arbitrary*!) specified probability models.
- I am *not* endorsing *logical pluralism*. I (and Carnap) think there's "*one true* $\mathfrak{c}$."
- Carnap's inductive-logical tolerance was (if anything) *too weak*, since it couldn't tolerate many important probability models (*e.g.*, analogical ones).

A "Neo-Carnapian" Approach to Inductive Logic II

- Here's my proposal. Confirmation is a *three-place* relation: between premises, conclusion, and *a probability model* $\mathcal{M}$ (*i.e.*, an algebra of propositions – *with numbers assigned to them* – satisfying your favorite axiomatization of probability). This proposal has its advantages:
  - It cleanly separates the *logical* confirmation relation from *epistemic* concepts like "*a priori*" or "epistemically rational" credence (often seen in "logical" accounts of probability — *via* the *Principle of Indifference*).
  - It guarantees the *transparent* satisfaction of all three historical desiderata (*esp.*, $\mathcal{D}_3$) — *without* requiring a "logical interpretation" of probability.
- This is "neo-Carnapian" (*not* Carnapian), since it is a *logical* conception of confirmation (in Carnap's sense), but the confirmation function is 3-place (*not* 2-), and it is *not* constructed syntactically out of the underlying languages.
- Next, I will propose a fourth and final desideratum for the ternary $\mathfrak{c}(C, P, \mathcal{M})$. This is a *relevance* desideratum, borrowed from Bayesian Confirmation Th.

A Fourth Desideratum & A New Foundation I

- Consider the following argument:

  (P) Dennis Rodman has been taking birth control pills for the past year.
  Therefore, (C) Dennis Rodman is not pregnant.

- Intuitively, $\Pr_{\mathcal{M}}(C \mid P)$ is high (*i.e.*, assuming an $\mathcal{M}$ reflecting our background knowledge about D.R.'s gender & human biology, *etc*.). But, do we want to say that there is a *strong confirmation relation* between P and C (in $\mathcal{M}$)?
- According to (TRV), we should say just that. This seems wrong. Intuitively, $\Pr_{\mathcal{M}}(C \mid P) = \Pr_{\mathcal{M}}(C)$. That is, $\Pr_{\mathcal{M}}(C \mid P)$ is high *solely because* $\Pr_{\mathcal{M}}(C)$ is high, and *not* because of any *confirmation relation* between P and C (in $\mathcal{M}$).
- This is the same kind of complaint Skyrms makes about (NIL). It seems just as compelling here. The problem here is that P is *probabilistically irrelevant to* C (in $\mathcal{M}$). This leads to a fourth desideratum for the ternary $\mathfrak{c}(C, P, \mathcal{M})$:

  ($\mathcal{D}_4$) $\mathfrak{c}(C, P, \mathcal{M})$ should be a $\Pr_{\mathcal{M}}$-*relevance measure*.

A Fourth Desideratum & A New Foundation II

$$(\mathcal{D}) \quad \mathfrak{c}(C, P, \mathcal{M}) \text{ should be } \begin{cases} \text{Maximal } (> 0, \text{ constant}) & \text{if } P \text{ entails } C. \\ > 0 \text{ (confirmation)} & \text{if } \Pr_{\mathcal{M}}(C \mid P) > \Pr_{\mathcal{M}}(C). \\ = 0 \text{ (irrelevance)} & \text{if } \Pr_{\mathcal{M}}(C \mid P) = \Pr_{\mathcal{M}}(C). \\ < 0 \text{ (disconfirmation)} & \text{if } \Pr_{\mathcal{M}}(C \mid P) < \Pr_{\mathcal{M}}(C). \\ \text{Minimal } (< 0, \text{ constant}) & \text{if } P \text{ entails } {\sim}C. \end{cases}$$

- Of the many relevance measures in the literature, *the only one* (up to ordinal equivalence) that satisfies all of our desiderata ($\mathcal{D}$) is the *log-likelihood-ratio*:

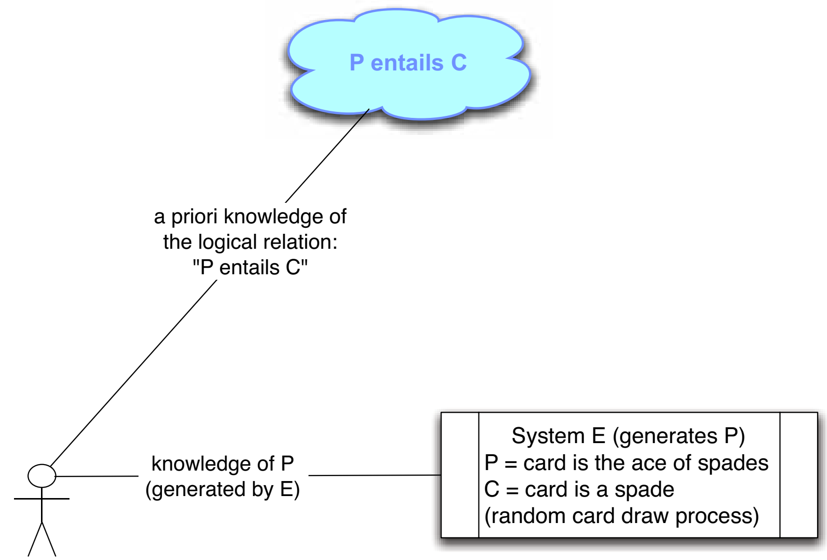
$$l(C, P, \mathcal{M}) = \log\left[\frac{\Pr_{\mathcal{M}}(P \mid C)}{\Pr_{\mathcal{M}}(P \mid {\sim}C)}\right]$$

- This provides a very elegant solution to the problem of measure sensitivity in Bayesian confirmation theory (the problem with which we began).
- Moreover, it provides a general logical foundation for both Bayesian and non-Bayesian Pr-accounts of evidential support — *via* likelihood ratios.

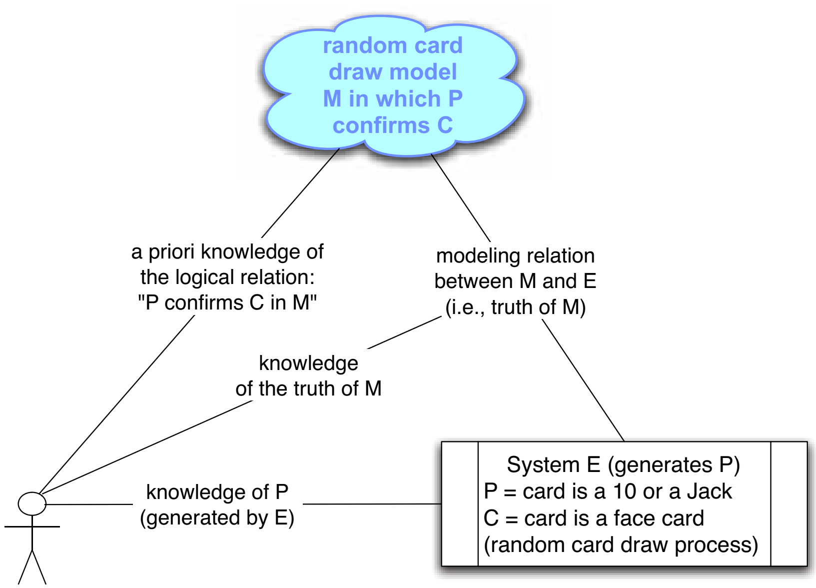
Bayesian & Non-Bayesian Notions of Evidential Support

- For a Bayesian, "E_1 is *better evidence* for H_1 than E_2 is for H_2" is a *biographical remark* concerning relations among *S 's degrees of belief at t*. This has little *epistemological significance* (since *any* Pr function will do!).
- Non-Bayesians in statistics use *objective* statistical models $\mathcal{M}$ (*e.g.*, models of *causal-statistical regularities in experimental set-ups*). Claims about *statistical* evidential support are thus meant to be objective, not personalistic.
- In either case, such claims should (from a logical point of view) be understood as expressions of *likelihood-ratio* comparisons in the salient models (as opposed to comparisons of *probabilities* or *other* relevance measures).
- You can plug any $\mathcal{M}$ you like into our framework. Whether the resulting *logical* confirmation relations (in $\mathcal{M}$) will be of any *philosophical* significance will require additional analysis [What is $\mathcal{M}$ a model *of*? Is $\mathcal{M}$ *true*? *etc*.].
- Now, we're going beyond inductive logic, and into the E & M of Pr-models.

Epistemological Connections 1: Entailment & Conclusive Evidence



Epistemological Connections 2: Confirmation & Non-Conclusive Evidence



Carnap on the Principle of Indifference

. . . the statement of equiprobability to which the principle of indifference leads is, like all other statements of inductive probability, not a factual but a logical statement. If the knowledge of the observer does not favor any of the possible events, then with respect to this knowledge as evidence they are equiprobable.

- Carnap deduces (PI) from logic alone, by *assuming* that "K does not favor any s_i over any other s_j" just *means* "$\mathfrak{c}(s_i, K) = \mathfrak{c}(s_j, K)$", for all i and j.
- Since Carnap also assumes that $\mathfrak{c}(x, y) = \Pr(x \mid y)$, it then follows (by logic alone!) that if K does not favor any s_i over any other s_j, then $\Pr(s_i \mid K) = \Pr(s_j \mid K)$, for all i, j, which is just what the (PI) says.
- This argument is cogent only if we grant that $\mathfrak{c}(x, y) = \Pr(x \mid y)$, which I think is incorrect. Interestingly, Carnap *himself* sometimes appeals to *relevance*!
- He appeals to the absence of *correlations* to favor certain applications of (PI) over others. And, his discussions of principles like (CCC) employ *relevance*.

Carnapian Monadic Predicate Logical Probability 1

- Generalizing on Wittgenstein, Carnap defined his logical measure functions over sentences in monadic predicate logical languages $\mathfrak{L}_Q^{m,n}$ containing n monadic predicates ($F, G, H, \ldots$) and m individual constants ($a, b, c, \ldots$).
- To fix ideas, consider the language $\mathfrak{L}_Q^{2,2}$, which contains two monadic predicates F and G and two individual constants a and b.
- In $\mathfrak{L}_Q^{2,2}$, we can describe 16 states, using the 16 *state descriptions* of $\mathfrak{L}_Q^{2,2}$:

$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	$Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$
$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$
$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$
${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$
${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$
	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	

Carnapian Monadic Predicate Logical Probability 2

- Following (PI), Carnap's first measure function $\mathfrak{m}^\dagger$ assigns *equal probability* to each state description s_i of $\mathfrak{L}_Q^{m,n}$. In our example $\mathfrak{L}_Q^{2,2}$, $\mathfrak{m}^\dagger(s_i) = \frac{1}{16}$.
- To extend $\mathfrak{m}^\dagger$ to all of $\mathfrak{L}_Q^{m,n}$, we stipulate that the probability of a disjunction of mutually exclusive sentences is the sum of the probabilities of its disjuncts.
- And, since every sentence in $\mathfrak{L}_Q^{m,n}$ is equivalent to some disjunction of state descriptions, and every pair of state descriptions is mutually exclusive, this gives us a complete unconditional probability function $\Pr^\dagger(\cdot)$ over $\mathfrak{L}_Q^{m,n}$.
- Finally, we define the conditional probability function $\Pr^\dagger(q \mid p)$ over pairs of sentences in $\mathfrak{L}_Q^{m,n}$ (in the standard way) as the following *ratio*: $\frac{\Pr^\dagger(p \& q)}{\Pr^\dagger(p)}$.
- Claims of the form $\ulcorner \Pr^\dagger(q \mid p) = x \urcorner$ are *analytic in* $\mathfrak{L}_Q^{m,n}$ since their truth-values are determined solely by the syntactical structure of the logical language $\mathfrak{L}_Q^{m,n}$. But, why is *this* choice of measure function $\mathfrak{m}^\dagger$ *logical*? Logicality is ensured by the application of (PI) to state descriptions. Or is it?

Carnapian Monadic Predicate Logical Probability 3

- As it turns out, Carnap ultimately *rejects* the measure function $\mathfrak{m}^\dagger$ in favor of an alternative measure function $\mathfrak{m}^*$, for epistemic-sounding reasons.
- Carnap notes that $\mathfrak{m}^\dagger$ causes $\Pr^\dagger$ to have the following property ($b \neq a$):

$$(*) \qquad \Pr^\dagger(Fb \mid Fa) = \frac{\Pr^\dagger(Fb \,\&\, Fa)}{\Pr^\dagger(Fa)} = \frac{4 \cdot \frac{1}{16}}{8 \cdot \frac{1}{16}} = \frac{1}{2} = 8 \cdot \frac{1}{16} = \Pr^\dagger(Fb)$$

- In other words, (*) says that one object a's having property F can never raise the probability that another object b also has F. As Kyburg shows, this *generalizes* to *any* number of Fs: $\Pr^\dagger(Fa_1 \mid Fa_2 \,\&\, \cdots \,\&\, Fa_m) = \Pr^\dagger(Fa_1)$.
- Carnap (*et al*) characterize this (in epistemic terms) as $\mathfrak{m}^\dagger$ leading to a logical probability function $\Pr^\dagger$ that fails to allow for "learning from experience".
- Carnap views this consequence of applying (PI) to the state descriptions of $\mathfrak{L}_Q^{m,n}$ as unacceptable. Presumably, then, he must think that $K_\top$ "favors" some state descriptions over others. But, which ones? Enter the $\mathfrak{m}^*$ measure …

Carnapian Monadic Predicate Logical Probability 4

- Two state descriptions s_i and s_j in $\mathfrak{L}_Q^{m,n}$ are *permutations* of each other if one can be obtained from the other by a mere permutation of constant symbols.
- "$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$" can be obtained from "${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$" by permuting "$a$" and "$b$". Thus, "$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$" and "${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$" are permutations of each other (in $\mathfrak{L}_Q^{2,2}$).
- A *structure description* in $\mathfrak{L}_Q^{m,n}$ is a disjunction of state descriptions, each of which is a permutation of the others. $\mathfrak{L}_Q^{2,2}$ has 10 structure descriptions:

$(Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb) \vee ({\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$
$(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb) \vee (Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb)$	$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$
$(Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb) \vee ({\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb)$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$
$(Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb)$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$
$(Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb)$	
$({\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb)$	

- The measure $\mathfrak{m}^*$ [$\therefore \Pr^*$] assigns equal probability to *structure descriptions*. $\mathfrak{m}^*$ takes individuals to be *indistinguishable* (Bose-Einsten statistics), and $\mathfrak{m}^\dagger$ does not (Fermi-Dirac statistics). Otherwise, they operate in the same way.

Carnapian Monadic Predicate Logical Probability 5

- We can then define $\Pr^*(\bullet)$ and $\Pr^*(\bullet \mid \bullet)$ in terms of $\mathfrak{m}^*$, by assuming equiprobability of *states within* structure descriptions. Let's compare the $\mathfrak{m}^\dagger$ and $\mathfrak{m}^*$ distributions over the 16 state descriptions of our toy language $\mathfrak{L}_Q^{2,2}$:

Fa	Ga	Fb	Gb	State Descriptions (s_i)	$\mathfrak{m}^\dagger(s_i)$	$\mathfrak{m}^*(s_i)$
$\top$	$\top$	$\top$	$\top$	$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/10
$\top$	$\top$	$\top$	$\bot$	$Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
$\top$	$\top$	$\bot$	$\top$	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
$\top$	$\top$	$\bot$	$\bot$	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
$\top$	$\bot$	$\top$	$\top$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
$\top$	$\bot$	$\top$	$\bot$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/10
$\top$	$\bot$	$\bot$	$\top$	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
$\top$	$\bot$	$\bot$	$\bot$	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
$\bot$	$\top$	$\top$	$\top$	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
$\bot$	$\top$	$\top$	$\bot$	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
$\bot$	$\top$	$\bot$	$\top$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/10
$\bot$	$\top$	$\bot$	$\bot$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
$\bot$	$\bot$	$\top$	$\top$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
$\bot$	$\bot$	$\top$	$\bot$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
$\bot$	$\bot$	$\bot$	$\top$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
$\bot$	$\bot$	$\bot$	$\bot$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/10

- Now, m^* does *not* have the "no learning from experience" property (*), since:

$$\Pr^*(Fb \mid Fa) = \frac{\Pr^*(Fb \,\&\, Fa)}{\Pr^*(Fa)} = \frac{2\cdot\frac{1}{10} + 2\cdot\frac{1}{20}}{2\cdot\frac{1}{10} + 6\cdot\frac{1}{20}} = \frac{3}{5} > 2\cdot\frac{1}{10} + 6\cdot\frac{1}{20} = \frac{1}{2} = \Pr^*(Fb)$$

- Generally, $\Pr^*(\bullet \mid \bullet)$ says that the more objects that are assumed to have F, the more probable it is that other objects will also have F. This is called *instantial relevance*. Carnap prefers $\Pr^*$ because it has this property, but $\Pr^\dagger$ does not.
- Is Carnap committed to the view that $K_\top$ favors certain state descriptions over others? He prefers $\mathfrak{m}^*(Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb) > \mathfrak{m}^*(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$ to $\mathfrak{m}^\dagger(Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb) = \mathfrak{m}^\dagger(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$. But, *why*?
- Carnap realizes that if a theory of "logical probability" is going to provide a logical foundation for claims about *evidence*, then it *must* be able to furnish probability models exhibiting *correlations* between *contingent claims*.
- This is just the tip of the iceberg of a much more general problem for linguistic theories of "logical probability" like Wittgenstein's and Carnap's. Such theories are unable to emulate a broad enough range of probability functions, so as to be generally applicable to intuitive models of evidential relations.

Carnapian Monadic Predicate Logical Probability 6

- $\Pr^*$ also has the advantage of supporting certain analogical correlations, such as:

$$\Pr^*(Gb \mid Fa \,\&\, Ga \,\&\, Fb) > \Pr^*(Gb \mid Fa \,\&\, Ga \,\&\, {\sim}Fb)$$

- Carnap sees this as another virtue of $\Pr^*$ (as compared with $\Pr^\dagger$). However, there are other kinds of analogical correlations that $\Pr^*$ cannot support.
- Hesse (in "Analogy in Confirmation Theory") discusses more general principles of "analogical inference". In particular, she argues for the following principles:
  - *a* and *b differing* regarding F should *weaken* the inference from Ga to Gb, *without completely undermining it*: $\Pr(Gb \mid Ga) > \Pr(Gb \mid Fa \,\&\, Ga \,\&\, {\sim}Fb) > \Pr(Gb)$. Neither $\Pr^*$ nor $\Pr^\dagger$ satisfies this condition. [$\Pr^*(Gb \mid Fa \,\&\, Ga \,\&\, {\sim}Fb) = \Pr^*(Gb)$]
  - Generally: if a has properties $P_1 \ldots P_n$, and b has $P_1 \ldots P_{n-2}$, but lacks P_{n-1}, that should be of *some* relevance to b's having P_n. In the $n = 3$ case, we should have: $\Pr(Hb \mid Ha) > \Pr(Hb \mid Fa \,\&\, Ga \,\&\, Ha \,\&\, Fb \,\&\, {\sim}Gb)$

    $> \Pr(Hb \mid Fa \,\&\, Ga \,\&\, Ha \,\&\, {\sim}Fb \,\&\, {\sim}Gb) > \Pr(Hb)$
  - *I.e.*, Differing on 2 properties should be worse than 1, and neither should completely undermine instantial relevance. $\Pr^*$ violates this. Exercise: Prove this!

Some Details on the Limiting (Deductive) Case

- I claimed that, when E *entails* H, $l(H,E)$ is maximal ($+\infty$), and does not depend on the prior probability of H.
- But, if E *entails* H, then $\Pr(E \mid {\sim}H) = 0$. Shouldn't we say that $l(H,E)$ is *undefined* in such cases, since it has a zero denominator?
- There are two ways to handle this. First, one could maintain that, *in the limit as* $\Pr(E \,\&\, {\sim}H)$ *approaches zero*, $l(H,E)$ diverges ($+\infty$). So, $l(H,E)$ is maximal and doesn't depend on $\Pr(H)$ in such cases.
- Or, more satisfyingly, one could use the alternative measure:

$$l'(H,E) = \frac{\Pr(E \mid H) - \Pr(E \mid {\sim}H)}{\Pr(E \mid H) + \Pr(E \mid {\sim}H)}$$

  It is easy to show that (*i*) l' *is ordinally equivalent to*[a] l, and (*ii*) l' takes on the values $+1/{-1}$ in cases where E entails/refutes H.

[a] This is because l' is a monotone increasing function of l [*viz.*, $l' = \tanh(l/2)$].

Making the Measure l More Intuitive: The Odds Form of Bayes' Theorem

- It might seem odd to compare $\Pr(E \mid H)$ and $\Pr(E \mid {\sim}H)$ for the purposes of measuring "how much E raises the probability of H."
- It might help "pump your intuitions" to note that:

$$\frac{\Pr(E \mid H)}{\Pr(E \mid {\sim}H)} = \frac{\text{Odds}(H \mid E)}{\text{Odds}(H)}$$

  where $\text{Odds}(H) = \frac{\Pr(H)}{\Pr({\sim}H)}$, and $\text{Odds}(H \mid E) = \frac{\Pr(H \mid E)}{\Pr({\sim}H \mid E)}$.
- In other words, comparing $\Pr(E \mid H)$ and $\Pr(E \mid {\sim}H)$ (in a ratio) is equivalent to comparing the *posterior and prior odds of H*.
- This is because of the following "Odds Form" of Bayes' Theorem:

$$\frac{\Pr(H \mid E)}{\Pr({\sim}H \mid E)} = \frac{\Pr(E \mid H)}{\Pr(E \mid {\sim}H)} \cdot \frac{\Pr(H)}{\Pr({\sim}H)}$$

- Note: $\text{Odds}(\cdot)$ is on a $[0, \infty)$ scale, but $\Pr(\cdot)$ is on a $[0, 1]$ scale. Thus, $\Pr(H)$ being "close to 1" is an *artifact* of the (*conventional*) Pr scale.

The Problem of "Irrelevant Conjunction"

- According to deductive accounts of confirmation (*e.g.*, Hempel's H-D account), E confirms H (roughly) iff H entails E.
- Such accounts of confirmation have the following consequence:

  (1) If E confirms H, then E confirms H & X, for *any* X.
- But, the X's in (1) can be *utterly irrelevant* to H (and E).
- While (1) is *not* a consequence of Bayesian confirmation, the following *is*:

  (2) If H entails E, then E confirms H & X, for *any* X.
- Bayesians try to mitigate the effects of (2), by arguing that:

  (3) If H entails E, then $\mathfrak{c}(H \,\&\, X, E) < \mathfrak{c}(H, E)$, for *any* X.
- (3) is sensitive to choice of measure, and it makes no appeal to the *irrelevance* of X. I have recently published a new account.

Rosenkrantz on the Problem of "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:

  (4) If H entails E, then $d(H \,\&\, X, E) = \Pr(X \mid H) \cdot d(H, E)$.
- Neither r nor l satisfies property (4).
- Rosenkrantz does provide some (pretty good) reasons to reject r. However, he explicitly admits that he knows of "no compelling considerations that adjudicate between" d and l.
- So, it is (at best) unclear how one might consistently complete Rosenkrantz's enthymematic treatment of the tacking problem.
- What's worse, as I have explained, I think there are compelling reasons to *favor l over d* as a measure of confirmation.

Earman on the Problem of "Irrelevant Conjunction"

- Earman gives a more robust resolution of the tacking problem which requires only the following logically weaker cousin of (4):

  (5) If H entails E, then $d(H \,\&\, X, E) < d(H, E)$.
- r violates even this weaker condition, but l *satisfies* (5).
- In this sense, Earman's account is *less* sensitive to choice of measure (*i.e.*, more robust) than Rosenkrantz's is.
- Earman's account can be bolstered by providing compelling independent reasons to favor d (or l) over r (as I have done).
- Unfortunately, even the bolstered version of Earman's account is not completely satisfying. See for a new approach (next two slides).

A New Approach to Irrelevant Conjunction

- First, we need to say what it *means* for a conjunct X to be *irrelevant* in a confirmational context involving hypothesis H and evidence E.[a]

  Definition. *X is an irrelevant conjunct with respect to the pair $\langle H, E\rangle$ iff X is confirmationally irrelevant to H, E, and H & E.*

- Do we even *have* a *Bayesian* problem of *irrelevant* conjunction? Yes.

  Theorem 1. *If E confirms H, and X is an irrelevant conjunct with respect to the pair $\langle H, E\rangle$, then E also confirms H & X.*

- Can we say anything general and interesting concerning *deleterious effects* of tacking irrelevant conjuncts onto hypotheses H? Yes.

  Theorem 2. *If E confirms H, and X is an irrelevant conjunct with respect to $\langle H, E\rangle$, then $\mathfrak{c}(H \,\&\, X, E) < \mathfrak{c}(H, E)$, where $\mathfrak{c}$ may be* ***any*** *measure of degree of confirmation, except r $[r(H \,\&\, X, E) = r(H, E)]$.*

[a] This is a *stronger* definition of irrelevant conjunct than we will actually need for our present purposes. Hawthorne and I have now shown that all we *need* is $\Pr(E \mid H \,\&\, X \,\&\, K) = \Pr(E \mid H \,\&\, K)$.

Goodman's "Grue" Paradox: The Standard Bayesian Representation

- Goodman famously presents an example involving the following two hypotheses (H and H') and observation report ($\mathcal{E}$):

  H: All emeralds are green. $[(\forall x)(Ex \supset Gx)]$

  H': All emeralds are grue. $[(\forall x)(Ex \supset \mathcal{G}x)]$

  $\mathcal{E}$: An object a has been observed to be a green emerald $[Ea \,\&\, Ga]$.

- Where, the predicate "grue" ($\mathcal{G}x$) is defined as follows:

  x is grue if and only if either (*i*) x has been observed and x is green, or (*ii*) x has not been observed and x is not green.

- Thus, $\mathcal{E}$ is equivalent to $Ea \,\&\, \mathcal{G}a$, and so $\mathcal{E}$ is a positive instance of both H and H'. So, by (NC), $\mathcal{E}$ confirms both H and H' – seems paradoxical.

- Bayesian answers to Goodman's "new riddle of induction" have aimed to establish that H is better confirmed by $\mathcal{E}$ than H' is, relative to our *actual* K_α. That is, Bayesians have tried to show: $\mathfrak{c}(H, \mathcal{E} \mid K_\alpha) > \mathfrak{c}(H', \mathcal{E} \mid K_\alpha)$.

Eells on Goodman's "Grue" Paradox

- Eells offers a Bayesian account of the Grue paradox (*a.k.a.*, Goodman's "new riddle of induction") which trades on the following property of the difference measure [where β, δ are defined as follows: $\beta =_{df} \Pr(H_1 \,\&\, E) - \Pr(H_2 \,\&\, E)$, and $\delta =_{df} \Pr(H_1 \,\&\, \bar{E}) - \Pr(H_2 \,\&\, \bar{E})$]:

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

- Neither r nor l satisfies property (6).
- Eells does provide reasons (as reported in a paper by Sober, see below) to prefer the difference measure d over the ratio measure r, but he does not supply reasons to prefer d over l.
- Pending such reasons, Eells's argument remains *enthymematic*.
- Moreover, I have provided reasons to favor *l over d*.

Sober on Goodman's "Grue" Paradox

- Sober describes a more robust Bayesian account of the Grue paradox which exploits the following weaker property of d:

  (7) If H_1, H_2 entail E and $\Pr(H_1) > \Pr(H_2)$, then $d(H_1, E) > d(H_2, E)$.

- r violates even this weaker condition, but *l satisfies* (7).
- In this sense, Sober's resolution of Goodman's "Grue" paradox is *less* sensitive to choice of measure (*i.e.*, more robust) than Eells's is.
- And, like Eells, Sober does provide *some* reasons to prefer d to r.
- However, as I explain in my dissertation, these reasons (which are borrowed from Eells) are not very good reasons to prefer d to r.
- Aside from being measure-sensitive, the standard Bayesian approaches to Goodman's grue paradox don't even capture its full logical structure. As a result, they can't be considered adequate resolutions. We can do better …

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

- Almost all Bayesian accounts of both the Ravens Paradox and the value of "varied" evidence (*i.e.*, why more "varied" evidence E_1 is more confirmationally powerful then less "varied" evidence E_2) presuppose:

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

- The "normalized" difference measure s *violates* (8).[a]
- Typically, the advocates of such arguments have used either d or r in their arguments (as it turns out, d, r, and l *all* satisfy (8)).
- None of these authors seems to provide (*independent*) reasons to prefer their measures over s (or other measures which violate (8)).
- In my dissertation, I propose a novel Bayesian explication of the confirmational value of *independent* evidence, based on l.

[a] So do Carnap's relevance measure $\mathfrak{r}(H, E) = \Pr(H \,\&\, E) - \Pr(H) \cdot \Pr(E)$, Mortimer's measure $\Pr(E \mid H) - \Pr(E)$, and Nozick's measure $\Pr(E \mid H) - \Pr(E \mid {\sim}H)$.

The Popper-Miller Argument *Against* Bayesianism

- It isn't just arguments/accounts *within* Bayesian confirmation theory that are sensitive to choice of measure. Some well-known *criticisms* of Bayesianism also rest on measure sensitive arguments.
- Most famously, Popper and Miller use the following property of the difference measure d to argue *against* Bayesianism (generally):

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

- As it turns out, neither of the measures r or l satisfies (9). Therefore, the Popper-Miller argument is *sensitive to choice of measure*.
- In the absence of reasons to think that d is a more accurate (and charitable) reconstruction of Bayesian confirmation theory than either r or l, the Popper-Miller argument remains *enthymematic* (*i.e.*, a *straw man*).
- Popper-Miller give no such reasons. I have argued *there are none*.

Earman on the Quantitative Problem of Old Evidence

- Earman argues that quantitative Bayesian confirmation theory, together with the "radical probabilism" of Jeffrey does not suffice to avoid Glymour's problem of old evidence.
- His argument presupposes that Bayesians use d to measure degree of confirmation, and it rests on the following fact about d:

(10) If H entails E and $\Pr(E) \approx 1$, then $d(H, E) \approx 0$.

- This argument has two flaws. First, (10) does hold for d and r, but it does *not* hold for l or s (contrary to what Earman suggests). Second, this argument only applies to the case of *deductive evidence* (H entails E).
- As it turns out, we can avoid Earman's objections, by using our l instead of d:

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

- As Joyce and Christensen point out, s *also* satisfies (11).

What About Pluralism?

- My proposed solution to the problem of measure sensitivity involves arguing for "one true measure" of confirmation. Why not be a pluralist?
- *I.e.*, Why not say instead that each of the many proposals is the "true measure" of *something*, and that which measure we should use in a context will depend on what we aim to measure in that context?
- I guess if I were willing to be a pluralist about *logic*, then I might be willing to go along. Don't we think some arguments are (logically) good?
- Note: one can accommodate *any* confirmational ordering one likes, simply by cleverly choosing the "right" measure of confirmation.
- So, in order to avoid *ad hoc* "reverse engineering," we'll need *legislation* on *precisely how* contextuality (or aims) should effect choices of measure.
- I'm open to the *possibility* of a "pluralistic" account with normative force. But, until I see one, I'll stick with "monism." Besides, I think l is *right*!