Confirmation Theory as a Branch of Inductive Logic: Historical and Philosophical Reflections

Branden Fitelson

Department of Philosophy
Group in Logic and the Methodology of Science
&
Institute for Cognitive and Brain Sciences

University of California–Berkeley

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

1. Overview
2. Prelude: Inductive Support and Confirmation
3. Hempel on Confirmation
   - Confirmation via deductive relations
4. Carnap on Confirmation
   - Confirmation via logical probabilities
5. Bayesian Confirmation
   - Confirmation via subjective probabilities
6. Steps Toward an Alternative Approach
   - None of the above

- Traditionally, confirmation theory has aimed to give accounts (or explications) of three kinds of informal "inductive support" relations between propositions:
  - **Qualitative**. E inductively supports H.
  - **Comparative**. E supports H *more strongly than* E' supports H'.
  - **Quantitative**. E inductively supports H *to degree* r.
- My focus will be on the *qualitative* and *quantitative* kinds.
- I'll discuss both logical and epistemic "support" notions.
- By *inductive logic* (IL) I mean a theory of "argument strength" which (in some sense) *generalizes* deductive logic.
- I see confirmation theory as a branch of IL (*i.e.*, "argument strength"/"inductive support" as almost interchangeable).
- Of course, it is controversial as to whether there *is* such a thing as (logical) "inductive strength" or whether "degree of inductive support" is best explicated as a *logical* concept.
- Two related themes: the role/logical status of *probability* in IL, and the relationship between IL and *epistemology*.

- Hempel (generalizing Nicod's earlier naïve *instantial* account) offered a *formal, logical* theory of confirmation.
- Hempel's explication characterized confirmation as a logical relation between sentences in simple, first-order languages.

  Qualitative. E confirms H if $E \vdash Z$, where Z is as sentence constructed syntactically from E and H in a certain way. [Hempel also gave **quantitative** and **comparative** accounts of confirmation, but I won't be discussing those today.]
- Hempel's confirmation relation has various properties, *e.g.*,

  (EC) If $E \vdash H$, then E confirms H.

  (SCC) If E confirms H and $H \vdash H'$, then E confirms H'.

  (M) If E confirms H, then so does $E \,\&\, K$, for any K (provided K mentions no individuals not mentioned in either E or H).
- (EC) is an uncontroversial confirmation-theoretic principle. On the other hand, (SCC) and (M) are quite controversial.
- *Hempel's own intuitions* about inductive support ran counter to (M). In his discussion of the raven paradox, he warns us not to conflate the following two inductive support claims.

(PC) $\sim Ra \,\&\, \sim Ba$ supports $(\forall x)(Rx \supset Bx)$, *relative to* $\top$.
(PC*) $\sim Ra \,\&\, \sim Ba$ supports $(\forall x)(Rx \supset Bx)$, *relative to* $\sim Ra$.

- But, Hempel's *confirmation* relation is *monotonic* (M), and there is no (classical) distinction between the following two:
  - $E \,\&\, K$ entails Z, *relative to* (given) tautological corpus $\top$
  - E entails Z, *relative to* (given) corpus K
- Thus, Hempel's *theory* implies that it is *impossible* for (PC) to be true while (PC*) is false — *contradicting* his *intuitions*.
- Hempel's *intuitions* about inductive support seem *right* here. It seems that inductive support is *non-monotonic*.
- Also, Hempel's discussion of (PC) and (PC*) suggests there may be *more than two relata* (K?) involved in "support."
- So, Hempel's theory of confirmation seems inadequate (by his own lights, it seems "dissimilar to the explicandum").
- Carnap (and others) gave *probabilistic* explications that are able to capture these sorts of features of "support."

- Carnap clarifies his explicandum (what I call "inductive support") in various ways, including [from page 21 of LFP]:
  Qualitative. ($\star$) E gives some (positive) evidence for H.
- In the 1st ed. of LFP, Carnap characterizes "the degree to which E confirms H" as $\mathfrak{c}(H, E) = \Pr(H \mid E)$, which leads to:
  Quantitative. $\Pr(H \mid E) = r$.
  Comparative. $\Pr(H \mid E) > \Pr(H' \mid E')$.
  Qualitative. $\Pr(H \mid E) > t$ (for some "threshold value" t).
  - Doesn't sound like ($\star$). More on this dissonance below.
- Like Hempel, Carnap wanted a *logical* explication of "support" (as a relation between sentences in FOLs).
- For Carnap, this meant that the conditional probability functions of IL must *themselves* be "logical", which implied that its Pr-claims must be *analytic* and knowable *a priori*.
- This leads naturally to the Carnapian project of providing a "logical explication" of conditional probability $\Pr(\cdot \mid \cdot)$ itself.
- I'm skeptical about *that* project. But, this issue is orthogonal to my talk today, so I'll leave this one to Jim and Patrick.

- I want to focus on two *other* questions that are raised by (my reading of) Carnap's discussions about confirmation.
  - **Logicality of Pr**. Must Pr *itself* be logical if (probabilistic) *confirmation* claims are to be logical? Must the confirmation theorist *qua inductive-logician* tell us *which* Pr*-assignment is salient* for the assessment of the strength of an argument?
  - **Epistemology and Inductive Logic**. Carnap presumed (as do I) that there are *some* connections between deductive logic and epistemology. Are there analogous connections between confirmation theory/IL and epistemology?
- Carnap thought the answer to the logicality of Pr question was "YES." I disagree. Below, I will describe an alternative approach that does not require the logicality of Pr.
- Carnap also thought the answer to the epistemology & IL question was "YES." With this, I agree. I will try to gesture toward something positive here. First, a key distinction.
- In LFP, Carnap gives a counterexample to Hempel's (SCC), which presupposes a more ($\star$)-like **qualitative** conception.

Qualitative. E confirms H iff $\Pr(H \mid E) > \Pr(H)$.

- This *probabilistic relevance* conception *violates* (SCC), whereas the previous Pr-threshold conception *implies* (SCC).
- The 2nd ed. of LFP includes a preface which acknowledges an "*ambiguity*" in LFP_1, and concedes that the (**qualitative**) relevance conception is the "more interesting" of the two.
  - **Firmness**. The degree to which E confirms$_f$ H:
    $$\mathfrak{c}_f(H, E) = \Pr(H \mid E).$$
  - **Increase in Firmness**. The degree to which E confirms$_i$ H:
    $$\mathfrak{c}_i(H, E) = \mathfrak{f}[\Pr(H \mid E), \Pr(H)]$$
    $\mathfrak{f}$ measures "the degree to which E *increases* the Pr of H."
- The 1st ed. of LFP was mainly about firmness (except the chapter on relevance, which has the discussion of Hempel).
- The 2nd edition only adds the preface, which says very little about $\mathfrak{c}_i$. Specifically, no function $\mathfrak{f}$ is rigorously defended.
- My positive proposal for confirmation theory can be seen as an attempt to extend *some* of Carnap's ideas about $\mathfrak{c}_f$ to $\mathfrak{c}_i$.

- *Many* candidate functions $\mathfrak{f}$ satisfy the *relevance* constraint:

  ($\mathcal{R}$) $\mathfrak{f}[\Pr(H \mid E), \Pr(H)] \gtreqless 0$ iff $\Pr(H \mid E) \gtreqless \Pr(H)$

- From an inductive-logical point of view, confirmation measures should *quantitatively generalize* entailment:

  ($\mathcal{D}$) Provided that both E and H are *contingent* claims[1] $\mathfrak{c}_i(H, E)$ should be *maximal* if $E \vdash H$, and *minimal* if $E \vdash {\sim}H$. [Note: $\Pr(H \mid E)$ satisfies *this*, but *not* $\mathcal{R}$.]

- Kemeny & Oppenheim used this consideration (and others) to argue that the best explication of $\mathfrak{c}_i(H, E)$ is given by:

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

- F can be expressed as a function $\mathfrak{f}$ of $\Pr(H \mid E)$ and $\Pr(H)$, and it satisfies $\mathcal{R}$, $\mathcal{D}$, and various other IL desiderata.
- One can use F to define **comparative** $[F(H, E) > F(H', E')]$ and **qualitative** $[F(H, E) > 0]$ confirmation$_i$ concepts.

[1] Here, I'm bracketing the "paradox of entailment" cases, which are tricky.

- I agree with K&O that F has the right *form* for a $\mathfrak{c}_i$ measure, but since K&O (like Carnap) assume that Pr must *itself* be "logical", their account of $\mathfrak{c}_i$ is *too* Carnapian for my taste.
- I'll return to this issue later. But, first, some remarks on Carnap's views on connections between IL & epistemology.
- Carnap (*p*. 201 of LFP) endorsed the following *epistemic bridge principle* for entailment ($\vdash$) and *knowledge*:

  If E is known by person X at time t, then H is likewise known by X at t [provided $E \vdash H$ is also known by X at t].
- For confirmation as firmness ($\mathfrak{c}_f$) and *credence*, Carnap endorsed the following, which he saw as *analogous*:

  If E *and nothing else* is known by person X at time t, then the degree of belief justified by the knowledge of X at t is r [provided that $\mathfrak{c}_f(H,E) = r$ is also known by X at t].
- The "*and nothing else*" is required *because* $\mathfrak{c}_f$ *violates* (M)! Adding more information besides E can *change* the degree of $\mathfrak{c}_f$. I'll return to this "requirement of total evidence."
- Before sketching my account, I'll glance at *Bayesianism*...

- Most modern Bayesians don't believe there are "logical" probabilities. I'm inclined to agree, but I'll let Jim field that.
- As a result, most modern Bayesians simply *give up on* the traditional project of confirmation theory *as a branch of IL*.
- Instead, they set their sights on explicating an explicitly *epistemic* (and subjective) notion of "inductive support":

  Qualitative. E confirms H *for agent* X *at time* t iff E and H are positively correlated under X's credence function at t.

- This is *formally* similar to the inductive-logical concept $\mathfrak{c}_i$. But, it is *subjective* and *epistemic*, *not* objective and logical.
- This forges a connection between confirmation theory & epistemology — by making it *a branch of* epistemology.
- Moreover, like Carnap, Bayesians assume *all inductive support relations supervene on* ***one kind*** *of probability*.
- There is also controversy about **quantitative/comparative** *Bayesian* $\mathfrak{c}_i$. We've written on that (not on today's agenda).

- *Epistemically*, it is "*all things considered*" assessments of the "goodness" of arguments *from E to H in contexts C* that count. These have *both* logical *and non*-logical components.
- Assessing (in C) the "goodness" (*soundness*) of a *deductive* argument from E to H requires the determination (in C) of:
  - Whether the argument is valid. [logical]
  - Whether E is true. [(generally) non-logical]
- How do we *generalize* this to include the *inductive* case?
- Of course, we *still* have to determine (in C) whether E is true.
- What *else*? I guess Carnap would say we need to determine (in C) "the degree to which E confirms$_i$ H." But, again, this determination must be made in accordance with [LFP, 211]:

  The Requirement of Total Evidence. In the application of IL to a given knowledge situation, the total evidence available must be taken as a basis for determining the degree of confirmation.
- Let K_C express "everything we (assessor) know in C." Then, if we know (in C) that E is true, $\Pr(H \mid E \,\&\, K_C) = \Pr(H \mid K_C)$.

- Therefore, if we know (in C) that E is true, we *cannot* — on a Carnapian approach — determine (in C) that E confirms$_i$ H.
- This is problematic. The non-logical component of our all things considered assessment of the argument's "goodness" has *interfered* with its (Carnapian) logical component!
- This is *not* a problem for *firmness*, since it doesn't prevent the logical probability $\Pr(H \mid E \,\&\, K_C) = \Pr(H \mid K_C)$ from being greater than *a threshold value*. This is *just* a problem for $\mathfrak{c}_i$.
- Bayesians face a similar problem, which is called "the problem of old evidence." Problem: *no* Pr*-assignment such that* $\Pr(E) = 1$ *can reflect a correlation between* E *and* H*.*
- So, how in the world can anyone who knows (in C) that E is true determine (in C) that an inductive argument from E to H is "all things considered good" in *some* "$\mathfrak{c}_i$ sense"?
- *That* is one of the central challenges motivating my project. I'll look at a concrete example of this phenomenon below.
- First, I will outline the basic ideas behind my approach.

- What we need is a way to be able to know (in C) that E is true *and at the same time* determine (in C) that — in *some* sense that is *epistemically salient* — E confirms$_i$ H (in C).
- **Step 1.** Confirmation$_i$ is a *three*-place relation [$F_{\Pr}(H,E)$] between E, H, *and a probability assignment* Pr, where Pr is a *non-logical parameter*, which gets fixed *contextually* (in C).
- Note: On this account, *confirmation$_i$* claims $F_{\Pr}(H,E) = r$ are *logical* and *objective*, *i.e.*, *analytic*, knowable *a priori*, and *quantitative generalizations of* $E \vdash \pm H$, *without* "logical" Pr.
- **Step 2.** When assessing (in C) the "all things considered goodness" of an argument from E to H (*i.e.*, whether E *evidentially supports* H in C), one must determine (in C):
  - Whether E is true. [(generally) non-logical]
  - The (most?) *evidentially salient* probability assignment with respect to E and H in C (Pr). [(generally) *non*-logical]
  - Whether $F_{\Pr}(H,E) > 0$. [logical component]
- Now, let's look at a concrete "problematic context."

- In some contexts (where E is known), there may be *some* (*prima facie*) salient assignments Pr relative to which E confirms$_i$ H and *others* Pr′ relative to which it does *not*.
- Here's an example, which also illustrates the "old evidence problem" (for Bayesians and Carnap) described above:
  - Mary administers a pregnancy test. It comes out positive (E). Let (H) be the hypothesis that Mary is pregnant. John knows that E is true (I will assume *with certainty*), and that the test is highly reliable, administered properly, *etc*.
  - John has available to him a (*prima facie*) evidentially salient Pr relative to which E *is* positively correlated with H [$F_{\Pr}(H,E) > 0$]: the *statistical* model of the test's reliability.
  - John also has available to him a (*prima facie*) evidentially salient Pr′ relative to which E is *not* positively correlated with H [$F_{\Pr'}(H,E) \ngtr 0$]: his *current credence function*.
  - Similarly for a "logical" Pr, *conditioned on* J's *total evidence*.
- My intuition is that John's assessment in this context should be that E provides some (positive) evidence for H.

- This suggests the following (contextual) "bridge principle":

  If X knows that E is true in C and that Pr is the (most?) evidentially salient Pr-assignment wrt E and H in C, then X knows that E provides some (positive) evidence for H in C [provided that X also knows $F_{\Pr}(H, E) > 0$].

- In this **qualitative** case, *any relevance measure* will do in place of F, since they all agree on such ">0" judgments.
- But, if we extend this to **comparative** relations such as "E favors H_1 over H_2" [$F_{\Pr}(H_1, E) > F_{\Pr}(H_2, E)$], things get more interesting, since those *depend on* the *form* of our chosen F.
- Two remarks. First, I'm *not* claiming that *all* (rational) evidential judgments can be explicated in this way. It may be that some are (inherently) probabilistically inexplicable.
- Second, in some C's the evidentially salient Pr *does* assign $\Pr(E) = 1$. *E.g.*, if the pregnancy test is *known* to *always* give positive results in C. Then, $F_{\Pr}(H, E)$ *shouldn't be* > 0.
- The $64 question is: What *determines evidential salience*? Of course, I have no *theory* of this. But, I'll say a few things.

- First Thing: "Pr is (most?) *evidentially salient* with respect to E and H in C" *is importantly distinct from* "It would (in C) be epistemically rational to adopt Pr as a *credence function*."
- This follows straightaway from the consideration of "old evidence" cases, in which the agent's credence function in C is *not evidentially salient* with respect to E and H in C.
- Second Thing: *Accuracy*. If we think (other things being equal) that it's the reliability of the test that is salient, then we'll want a model of its reliability (in C) that is *accurate*.
- Third Thing: Which Pr-assignments one takes to be salient (in which contexts) will depend on one's *epistemological commitments* concerning *the nature of evidential relations*.
- If one is an *externalist*, then one might tend to see *objective* probability models as *more salient* than *subjective* ones.
- Final Thing: The inductive logician *qua logician* needn't resolve such issues. Their job is to determine the *form* of the $\mathfrak{c}_i$-function, *not* to determine "evidential salience in C."