Confirmation Theory (Continued)

- Administrative: let's talk about paper topics. Some of you already have some good, creative ideas for papers! Graduate students should see me about perhaps doing a longer paper . . .
- Finishing-Off Hempel's Account of Confirmation
  - The Ravens Paradox
  - The "Blite" (or "Grue") paradox
  - Final Remarks on Hempelian Confirmation
- Probabilistic Accounts of Confirmation, Part I
  - Probability Theory — Revisited
  - Bayesian Interpretation and Framework
  - Some Examples + Survey of Problems

Two Deeper Problems with Hempelian Confirmation

- Paradox of the Ravens: Consider the hypothesis that all ravens are black, H: $(\forall x)(Rx \to Bx)$. Which of these Hempel-confirms H?

E_1: $Ra \,\&\, Ba$	E_2: $\sim Rb$	E_3: Bc
E_4: $\sim Rd \,\&\, \sim Bd$	E_5: $\sim Re \,\&\, Be$	E_6: $Rf \,\&\, \sim Bf$

  Answer: *All but* E_6 *Hempel-confirm* H! So, *Red Herrings* confirm H?!

- Goodman's Grue Paradox: Consider the hypothesis (H') that all ravens are "blite", where the predicate "blite" (B) is defined as follows:

  x is blite iff *either* (i) x is examined before (the end of) today, and x is black *or* (ii) x is examined after today, and x is white.

  On Hempel's theory, $Ra \,\&\, Ba$ confirms H'. But, this means that a black raven observed today confirms the hypothesis that ravens observed tomorrow (and thereafter) will be white![a]

[a] Bayesian confirmation can say more about these problems. See Rosenkrantz's "Does the Philosophy of Induction Rest on a Mistake," and Skyrms' Chapter IV for insights.

Salmon & Earman on Conditional Probability I

- Unlike Skyrms, Salmon & Earman take conditional probability as primitive, and they axiomatize conditional probability as follows:
  1. $0 \leq \Pr(B \mid A) \leq 1$.
  2. If $A \vDash B$, then $\Pr(B \mid A) = 1$.
  3. If B and C are mutually exclusive, then
  $$\Pr(B \vee C \mid A) = \Pr(B \mid A) + \Pr(C \mid A).$$
  4. $\Pr(B \,\&\, C \mid A) = \Pr(B \mid A) \cdot \Pr(C \mid A \,\&\, B)$.
- After presenting these axioms for $\Pr(\cdot \mid \cdot)$, S & E proceed to show that it obeys the laws of the (ratio-defined) conditional probability of Skyrms.
- S & E neglect to say what happens to $\Pr(A \mid B)$ when B is *impossible* (the *undefined* case for Skyrms). Unfortunately, this causes their definition of conditional probability to be *logically inconsistent!*

Salmon & Earman on Conditional Probability II

- Proof that S & E's definition of $\Pr(Y \mid X)$ is *logically inconsistent*:
  - Assume X is *impossible* (hence, that $\Pr(Y \mid X)$ is *undefined* on Skyrms' ratio definition of conditional probability).
  - By S & E's axiom (2), we have $\Pr(B \mid X) = 1$ and $\Pr(C \mid X) = 1$ (for *any* B, C), since X entails *everything* (including B and C).
  - Now, assume that B and C are *mutually exclusive*.
  - Then, by axiom (3), we have:

    $$\Pr(B \vee C \mid X) = \Pr(B \mid X) + \Pr(C \mid X) = 1 + 1 = 2.$$

  - But, axiom (1) requires that $\Pr(B \vee C \mid X) \leq 1$. *Contradiction!*
- So, while S & E's $\Pr(X \mid Y)$ can be defined even in cases where $\Pr(Y \mid \Omega) = 0$, it is an *incoherent* concept (which is *even worse*!).
- This can be fixed by requiring that Y be *possible* in $\Pr(X \mid Y)$. This is a weaker requirement than Skyrms' $\Pr(Y) \neq 0$.

Bayesian Confirmation I

- Thomas Bayes (1701 – 1761) was one of the first people known to have used the following formula for computing the "inverse probability" $\Pr(H \mid E)$, from $\Pr(E \mid H)$, $\Pr(H)$, and $\Pr(E)$.[a]

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

- This is known as *Bayes' Theorem*. S *&* E prove BT from their axioms on page 71. And, Skyrms proves BT from his on page 131.
- Bayes' Theorem can be proved algebraically using our Venn Diagram technique (and the ratio definition of conditional prob.).

[a] It is somewhat controversial as to whether (and to what level of generality) Bayes used this formula explicitly (and for this purpose of "inversion"). See Dale.

Bayesian Confirmation II

- It is somewhat unclear as to why Bayesianism is called "*Bayes*ianism." Bayes' Theorem is just one of many important theorems in probability theory. Why not call it "probabilism?"
- Bayesian*ism* comes along with some substantive additional *epistemological* assumptions (see Sober's "Introduction").
- First, it is assumed that *a (epistemically) rational agent's degrees of belief (or credence) are probabilities* (in the technical sense).
- Second, it is assumed that *a (epistemically) rational agent learns by conditionalizing on their evidence* (up to that point in time).
- That is, at any given time t, a rational Bayesian agent a has degrees of belief that conform to a probability function $\Pr_t^a(\cdot)$.
- And, if a rational Bayesian agent a learns E between t_1 and t_2, then their degrees of belief at t_2 are given by $\Pr_{t_2}^a(\cdot) = \Pr_{t_1}^a(\cdot \mid E)$.

Bayesian Confirmation III

- Historically, there have been two kinds of probabilistic confirmation:
  1. **Absolute**: E confirms H (relative to K) if $\Pr(H \mid E \,\&\, K) > \tau$, for some "threshold value" τ (*i.e.*, if $\dfrac{E \,\&\, K}{\therefore H}$ is "*inductively strong*").
  2. **Incremental**: E confirms H (rel. to K) if $\Pr(H \mid E \,\&\, K) > \Pr(H \mid K)$ (*i.e.*, if E is *positively stochastically relevant to* H, given K).

  Incremental confirmation has become more popular in recent years. It will be the main focus of our discussion of "Bayesian" confirmation.

- Such an account becomes *Bayesian* only if the probabilities are interpreted as *degrees of belief of rational agents*.

- This makes Bayesian confirmation somewhat *psychologistic* (or *subjective*). But, it is still *normative*, because it is "rational" agents.

- Contrast: Deductive confirmation was *not* defined in terms of rational agents' *beliefs* about deductions (no psychologism in deductive logic!).

Bayesian Confirmation IV

- There is an important and subtle ambiguity in the incremental notion of Bayesian confirmation. First, let's introduce some terminology:
- The *prior* probability of a hypothesis H (for an agent a) is the degree of belief a has in H *before* observing evidence E.
- The *posterior* probability of a hypothesis H (for an agent a) is the degree of belief a has in H *after* observing evidence E.
- The fact that we are talking about an agent a's degrees of belief can be captured by letting K include the *background knowledge of the agent*.
- So, we can write the *prior* of H (for a) as $\Pr(H \mid K)$, and the *posterior* of H as $\Pr(H \mid E \,\&\, K)$ (K is a's knowledge *prior to learning* E).
- The subtle ambiguity is between *synchronic* and *diachronic* senses of incremental confirmation. [The absolute notion avoids this ambiguity!]

Bayesian Confirmation V

- We will come back to the synchronic/diachronic issue later (when we talk about the problem of old evidence — a central problem).
- So far, the incremental theory just gives us a *qualitative* account of confirmation (analogous to the previous, *deductive* accounts).
  - E *confirms* H (relative to K) if $\Pr(H \mid E \,\&\, K) > \Pr(H \mid K)$
  - E *dis*confirms H (relative to K) if $\Pr(H \mid E \,\&\, K) < \Pr(H \mid K)$
  - E is *irrelevant to* H (rel. to K) if $\Pr(H \mid E \,\&\, K) = \Pr(H \mid K)$
- The basic idea behind the incremental account is that *confirmation is probabilistic relevance*. How can we *quantitatively* generalize this?
- Each of these says "E incrementally confirms H relative to K":
  - $\Pr(H \mid E \,\&\, K) > \Pr(H \mid K)$
  - $\Pr(H \,\&\, E \mid K) > \Pr(H \mid K) \cdot \Pr(E \mid K)$
  - $\Pr(E \mid H \,\&\, K) > \Pr(E \mid \bar{H} \,\&\, K)$

Bayesian Confirmation VI

- *Prima Facie*, any of these inequalities (or any other equivalent inequality!) could be used to generate a (initially plausible) *quantitative measure* of *degree* of (incremental) confirmation.
- For instance, we might adopt one of the following three measures:

$$
\begin{aligned}
d(H, E \mid K) &=_{df} \Pr(H \mid E \,\&\, K) - \Pr(H \mid K) \\
r(H, E \mid K) &=_{df} \log\left[\frac{\Pr(H \mid E \,\&\, K)}{\Pr(H \mid K)}\right] \\
l(H, E \mid K) &=_{df} \log\left[\frac{\Pr(E \mid H \,\&\, K)}{\Pr(E \mid \bar{H} \,\&\, K)}\right] \\
&= \log\left[\frac{\Pr(H \mid E \,\&\, K) \cdot [1 - \Pr(H \mid K)]}{[1 - \Pr(H \mid E \,\&\, K)] \cdot \Pr(H \mid K)}\right].
\end{aligned}
$$

- These and many other measures have been proposed and defended in the literature on (incremental) Bayesian confirmation theory.

Bayesian Confirmation VII

- The only reason we take logarithms in r and l is to ensure that (like d) the measures r and l are (i) *positive* for *confirmation*, (ii) *zero* for *irrelevance*, and (iii) *negative* for *dis*confirmation.
- In my dissertation (and in several publications), I survey the many measures that have been proposed and show that a wide variety of arguments in the literature depend on which measure one chooses.
- It can be shown that both the incremental and the absolute notions of confirmation differ significantly from previous, deductive accounts.
- For instance, they both (largely) avoid the problem of alternative hypotheses. Moreover, they both avoid the problem of statistical hypotheses, as well as the problem of quantitative generalization. And, they also *can* resolve "irrelevant conjunction," "ravens," and "grue."
- However, both the absolute and incremental notions of confirmation have potential problems (*e.g.*, *subjectivity* of Pr). Here are a few more . . .

Bayesian Confirmation VIII

- The absolute account of confirmation has (potential) problems:
  - The problem of the threshold value τ (what should τ be?)
  - The problem of insensitivity to *relevance* (Fred Fox, Skyrms)
- The incremental account of confirmation has (potential) problems:
  - The problem of symmetry (incremental confirmation is symmetric)
  - The problem of measure sensitivity (which I have made famous!)
  - The problem of old evidence (another famous problem)
- The problem of old evidence has caused a great stir in the literature. I think it trades on a *synchronic/diachronic* confusion or equivocation:
- It seems to me that the definition of confirmation is *synchronic*:
  - E confirms H (relative to K) for a at t if $\Pr_t^a(H \mid E \,\&\, K) > \Pr_t^a(H \mid K)$.
- Is t *prior* to (or *at*) the time when a learns E, or is t is *after* a has learned E? If t is *after* a learns E, then confirmation is *impossible*!