PHIL250 **Notes on James Hawthorne's "Inductive Logic"**
Wm Kallfelz

Recall from **<u>Lecture 2 Notes</u>** (September 4) A toy example:

"**Every** crow in a random sample of n crows is black. **Therefore**, this strongly supports the hypothesis that 'all crows are black.'

Inductive logic seeks to extend the principles of deductive reasoning to statements which aren't fully certain. The aim of inductive logic, using tools from **probability theory and statistics** is to derive such rules of inference that would provide an optimal support for a conclusion. Inductive logic essentially **involves** probability and statistics, but is **not merely** probability and statistics.

According to inductive logic, the above argument could be characterized by the following scheme:

1. S is a **random sample** of C (crows).
2. $|S| = n$ (i.e. S has n members)
3. The **frequency** (or proportion) of members of C having property B (i.e., being black) in sample S is r (= 1). (I.e. $f[B, C \cap S] = r = 1$)

==[p][1]

$\therefore$ [To a **margin of error** q conclude that] The **proportion** of **all** members of C having property B is r (= 1).

James Hawthorne's (2005) article[2] gives an overview which emphasize the major aspects such a logic should exhibit, what standards it should adhere to, as agreed on by those in the field.[3]

---

[1] Note: the ====[p] notation means: "Conclude with degree of support p."

[2] *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/logic-inductive

[3] Note the "circularity" here. Treating Inductive Logic as a *paradigm*, Hawthorne is describing its cognitive values. Recall from your Kuhn readings: the cognitive values internal to a paradigm are 'site specific' in the sense that they provide standards for puzzle-solving in the research problems as characterized by the paradigm. "When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular...The resulting circularity does not, of course, make the argument wrong…[however] the status of the circular argument is only that of persuasion. It cannot be made logically or even probabilistically compelling for those who refuse to step into the circle." (-*T. Kuhn,* **CC1998,** 88) So in the case of Hawthorne's article, *none* of these features or standards would sway or convince anyone who doesn't buy into the Inductive Logic paradigm. For instance, both Popper and Kuhn rejected inductive logic. For Popper, as you recall from the readings in the *The Logic of Scientific Discovery*, inductive logic suffers from *infinite regress*. Their conclusions are hedged by 'degrees of support,' yet such 'degrees of support' require further inductive procedures supporting *them*, which require further 'degrees of support,' etc. For Kuhn, as you recall in the "Objectivity, Value Judgment, and Theory Choice" article, the *cognitive values* governing theory-choice are *prior to*, and hence *cannot be* captured by any formal procedure(s), inductive or otherwise. Values aren't rules. So in the case of supporting a general conclusion, (based on particular claims) Kuhn would say that such (rational but not rule-based) values (whether cognitive or contextual) play a central role that cannot captured by the inductive logical machinery.

Every notion concerning 'degree of support' must meet the **CoA** (Condition of Adequacy), which (from the looks of it) appears *circular* (virtuous or vicious? It's in the eye of the beholder[4]). Basically the **CoA** states that:

> *In the long run* (i.e. as evidence accumulates) the degree of support (p) that an inductive logic assigns to a general hypothesis, based on collections of particular true evidence claims, "should tend to indicate that false hypotheses are probably false and true hypotheses are probably true." (1)

Note the circularity here. However, note a subtlety as well: Degree of support (denoted by p) **isn't the same notion as probability!** Degrees of support are parameters or functions that are internal to the logic of induction. Probability, on the other hand, is a general mathematical concept that is based on (at least) three different interpretations:

1. **Classical**: The ratio of the number of desired outcomes divided by the total number of outcomes. Stated more formally, $P(E) = {}^{|E|}/_{|X|}$, where X is the set of total number of outcomes (the *sample space*) and E is some desired event (a subset of the sample space). The " |…|" means "total number." For example, suppose I have a deck of 52 cards. Suppose I choose 1 card from that deck. Suppose I'm interested in whether or not that card is a King. Then $|X| = 52$, and $|E| = 4$. Then $P(E) = {}^{4}/_{52} = {}^{1}/_{13.}$

2. **Empirical:** The disadvantage of the classical definition is one must know (at least in principle) how many possible outcomes there are in advance. The sample space, in other words, must be *closed*. For *open* systems typical of empirical circumstances (whether scientific or ordinary) this simply won't do. Consider a some simple example: What is the expected lifetime of a lightbulb? What is the probability that I will live to age 85? What are the chances of getting into a serious car accident in the next twenty years? Etc. All these questions deal with sample spaces that are *open-ended*.[5] The empirical approach defines probability in terms of *frequency ratios*. That is to say, the ratio of the total possible number of *observed* instances of a desired even divided the total possible number of observed instances. For instance, consider the example of black ravens. Suppose I select a 'statistically significant'[6] sample S of 30 ravens, and observe 29 instances of black ravens. Then according to this interpretation, $P(E) = {}^{n(E)}/_{n(S)} = {}^{29}/_{30.}$ This simple example also reveals the underlying issues and tensions here, concerning the Problem of Induction and how Inductive Logic seeks to deal with it. For in the question of observed frequencies, one wonders how such

[4] See note 3. above. Popper would say "vicious."

[5] Which is not the same thing as saying that they're *infinite.* Simply that it becomes difficult if not impossible to specify in advance what their boundaries are. In the cases listed above, the sample spaces are even difficult to *define.* Take, for example, the case of the second question (what is the probability of ym living to the age of 85)? What does it *mean* to specify all possible 'outcomes'? There's only one actual system (myself)! One can speak, for instance, of 'possible worlds' or carbon-copies of myself under varying possible circumstances (i.e., all things being equal, if I grew up in a different country, etc.) This is known as the *ensemble* approach to specifying a sample space X for such an open-ended question. *Or* I could specify the whole of humanity in this present epoch, by sampling a 'statistically significant' number of individuals from this population and doing a 'longitudinal study' tracking their life-lines, etc.

[6] As we'll see, this is a notion that is not free of controversy either.

probabilities can be *generalized*? Let's take an even simpler example. Suppose I have an unbiased coin. According to the classical interpretation of probability, the chances of getting heads = $^1/_2$, since $X = \{ H , T\}$ (the sample space consists of two outcomes: H or T, heads or tails.) Suppose I flip the coin $N = 30$ times, and discover that the # of heads = 13. Then according to the empirical interpretation $P(E) = {}^{n(H)}/_N = {}^{13}/_{30}$. Naturally, we'd expect that as $N \rightarrow \infty$, the empirical definition and the classical definition of probability should give the same answer. But this is only because the sample space is completely specified unambiguously in advance. Unfortunately, in the case of open-ended empirical cases, the issue is much cloudier. Aside from the point that in such circumstances it's probably not even *possible* to specify a sample space *X*, (hence one cannot ascribe a classical probability in these cases), there remains the central problem of 'generalizing' a notion of empirical probability (how can one extrapolate beyond some concrete observed frequency ratio?) These are the issues that inductive logic grapples with.

3. **Subjective:** Closely allied with the problems specified in the case of empirical probability, is this interpretation with ascribes a measure of *belief* (or *degree of certainty*) to certain claims. For instance, one speaks of being 100% certain (in typical cases involving *analytical* claims[7]) but one usually qualifies *a posteriori* claims (i.e. claims whose truth-conditions depend on experience) with less than 100% certainty. The whole *Bayesian* paradigm[8] seeks to hook up this notion of probability with the empirical notion, in attempt to render some long-term empirical beliefs with a high measure of subjective certainty some empirically objective plausibility. Bayesians also claim to show how empirical probabilities can be extrapolated beyond their specific contexts of observation, to given them degrees of objective validity, qualified accordingly. Of course, the whole point is *how* such 'qualifications' can be appropriately systematized to the extent that they can be precisely quantified. The Bayesian advertises their a strategy which presumably solves this problem (again, though, not without controversy). Obviously, Inductive Logic draws much from the Bayesian paradigm.

O.K., now that the notion of probability has been clarified, and distinguished from the notion of degree of support, note also the heavy reliance of the article placed on *conditional* probabilities. It's seldom (in 'real world' empirical cases) we utter statements concerning probabilities that are *unqualified*, or pertain to the entire sample space.[9] Basically a conditional probability is one which is qualified in the following manner:

Given *Y to be the case, what is the probability of E?*

---

[7] An analytical claim is one whose meaning in the predicate is contained in thee meaning of the subject. For instance, "A **triangle** is a **three-sided polygon**" is one such example. On the other hand, a *synthetic* claim in one whose meaning in the predicate is *not* contained in the meaning of the subject. (For example: "Bob has a cold today.") Recall Sept 20 class: Many (but not all) philosophers want to believe that all analytic claims a *a priori* (i.e., claims whose truth-conditions don't depend on experience), and that all synthetic claims are *a posteriori* (i.e., claims whose truth-conditions depend on experience.) Again, many, but not *all.* The late 18th century philosopher Immanuel Kant was one such exception.

[8] A school of thought developed after Bayes, which we'll examine in greater detail in this course. Recall Kuhn's as well as Popper's critical remarks against Bayesianism.

[9] See footnote 5 above. In many cases, it may be impossible to specify the sample space *in principle.*

This is expressed in shorthand form as: $P(E \mid Y)$.

For example, in the case of the deck of cards, as mentioned in the definition of classical probability, suppose we drew a hand of 5 cards, and we're interested in the chances of there being a King, *given that 5 cards were drawn,* This is expressed as $P(E \mid Y)$, where Y is the event that 5 cards were drawn.[10]

Aside from the essential point that degree of support shouldn't be interpreted as identical to the concept of probability (though it involves this concept), keep in mind the more general point that **inductive logic does not reduce to probability theory! (Though it certainly involves this notion.) Why? Recall <u>Lecture 2</u> (September 4): "**What distinguishes logic from any old rule-based formal system, is that in logic, the rules must be truth-preserving." (p. 6, lecture notes) 'Truth' refers to *content* of propositions, not just *form.* Such content "cannot help [but be] subjected to the **'variable' conditions of thinking."** (Agazzi's quote, p.3 September 4 notes **<u>Lecture 2</u>**.)

---

[10] Which the answer is $P(E \mid Y) = {}^{C(4,1)C(48,4)}/_{C(52,5)}$ where $C(n, k)$ is a **combination,** i.e. $C(n, k) = {}^{n!}/_{k!(n-k)!}$. In other words, what this states (using the multiplication principle) is the ratio of selecting 1 card from a total of 4 different Kings (which can be accomplished in $C(4,1) = {}^{4!}/_{1!3!} = 4$ ways) AND (read 'multiply') the number of ways to choose the remaining 4 cards from the other 48 non-Kings which can be accomplished in $C(48,4) = {}^{48!}/_{4!44!} = 194{,}580$ ways) divided by the total # of outcomes (which happens to be the number of ways to choose 5 cards from a collection of 52 cards, (i.e. $C(52,5) = {}^{52!}/_{5!47!} = 2{,}598{,}960$ ways). Working this out: $P(E \mid Y) = {}^{5\cdot4\cdot(47!\cdot48!)}/_{(52!\cdot44!)}$ $=$ **$0.29946 \approx 30$ %**

Note that this answer can be arrived at using more elementary means. For instance, one can use the multiplication and addition principle (of outcomes). Here's how it works:
Suppose I choose a hand 5 five, and my first card is a King. What are the chances of that occurring?
Answer: $({}^{4}/_{52})\cdot({}^{48}/_{51})\cdot({}^{47}/_{50})\cdot({}^{46}/_{49})\cdot({}^{45}/_{48})$ ways, which represent the successive probabilities of obtaining non-Kings the rest of the time. (The second case is $({}^{48}/_{51})$, since there are 48 non-Kings remaining and there are 51 cards remaining. The third case is $({}^{47}/_{50})$, since there are 47 non-Kings remaining and there are 50 cards remaining, etc.) Now, this is just **one** outcome. Consider another outcome in which the 2nd card happens to be a King. The probability of this occurring is: $({}^{48}/_{52})\cdot({}^{4}/_{51})\cdot({}^{47}/_{50})\cdot({}^{46}/_{49})\cdot({}^{45}/_{48})$, for the same reasoning. To make a long story short, if you consider all the outcome (there are 5 independent cases), each has the same probability of occurring. Therefore using the addition principle, the chances of obtaining one King in a hand of 5 is: $5({}^{4}/_{52})\cdot({}^{48}/_{51})\cdot({}^{47}/_{50})\cdot({}^{46}/_{49})\cdot({}^{45}/_{48}) =$ **$0.29946 \approx 30$ %.**

The connection between the two approaches resides is the formula :

$$P(E \mid Y) = {}^{P(E\&Y)}/_{P(Y)}$$

…which is the general definition of conditional probability. In this case, Y refers to the event in which I've chosen a hand of 5 cards from a deck of 52 cards, and E refers to the event that I've chosen one King from a set of 4. This toy example can be reduced to an expression involving total number of outcomes (obtained via the combination, which is a mathematical expression computing total # of cases in which the order is independent), by virtue of the fixity of the sample space (hence the denominator expressions are all equal and can be canceled). Therefore: $P(E \mid Y) = {}^{n(E\&Y)}/_{n(Y)} = {}^{C(4,1)C(48,4)}/_{C(52,5)}$ (where n stands for number of outcomes.)

As Hawthorne states:

> "Any inductive logic…should address two challenges. (1) It should tell us which enumerative inductive arguments should count as *good*…rather than inductive fallacies….(2) [I]t should demonstrably satisfy the **CoA."** (2)

By 'enumerative induction' Hawthorne simply means the kind of reasons by which we form a conclusion by aggregating over observed instances. Enumerative inductive fallacies are easy to come by (and fall under the category of **False Generaliaztions** [see posted notes on **Argument Fallacies.])** A simple example would be one in which I (falsely) conclude "it never rains in Southern California" because since I've been in Southern California, it hasn't rained.

As Hawthorne is quick to point out, enumerative induction is by no means the only form of *ampliative* (or non-deductive) reasoning, it's only the simplest. More complicated cases which are non-enumerative include the examples of the 'universal' inference of Newton's laws[11], cases of medical diagnosis, a jury deciding the innocence/guilt of a triee, etc.

"[A] full account of inductive logic should not be limited to enumerative induction, but should also explicate the logic of *hypothetical reasoning*.." (3)

- **Possible question to address in your paper** : Would you consider, based on the complexities alone of the attempt to formalize enumerative induction as discussed in the assigned reading (pages 1-13) that Inductive Logic could address such challenges posed by more general cases of hypothetical reasoning? Or would the simple scheme of falsificationism prove a more reliable alternative?

Note how Hawthorne echoes Agazzi's more general points. (Agazzi was quoted in the notes for **Lecture 2).** Agazzi claims that logic, as understood to be a truth-preserving systematic discipline, *cannot* be reduced to mere formal syntax alone. (The notion of truth inevitably involves *semantics*, or meaning, i.e. the *content* and not just the form of our propositions). In a similar manner, writes Hawthorne:

> "[T]he core idea of Bayesian[12] logicism is fatally flawed—syntactical logical structure cannot be the sole determiner of the degree to which premises inductively support conclusions. [13] A crucial facet of the problem faced by Bayesian logicism involves how the logic is supposed to apply to scientific contexts, where the conclusion sentence is some hypothesis or theory, and the premises are evidence claims." (4)

---

[11] Of course, fallibilists like Popper or in general hypothetico-deductivists have an easy answer for that Recall Popper: *All* bona fide scientific laws should be expressible in a universal manner, and belong to a falsifiable theory. ("It will rain or it will not rain tomorrow" is *not* a law, since it's a tautology. On the other hand, "all ravens are black" is a scientific law, since (a) it's expressed in universal form, (b) it can be falsified.

[12] See above remark in the description of subjective probability.

[13] And of course, for people like Kuhn, and Longino, they'd argue that this is *precisely* where *values* (whether cognitive or contextual) play a crucial role. For Kuhn, such values (recall **Lecture VII**) *cannot* be reduced to rules, therefore defy being characterized by *any* algorithmic procedure (whether inductive logical or otherwise). Kuhn would retort: "Just because it's not *logical* [in the sense of being characterized by logic] does *not* mean it's not *rational*."

The key problem (and precisely *why* Popper accused such inductive logical strategies as sliding into infinite regress—see footnote 3 above) is that inductive support (p) depends in a crucial way on *prior probability.* (Recall from the remarks here: though inductive support cannot be reduced to mere probability, nevertheless the notion depends on the notion of probability: No probability theory, then no inductive logic!)

Apart from there being "severe technical problems" (ibid.) with securing the prior probabilities for inductive support "various counter-examples seem to show that such an approach must assign intuitively quite unreasonable prior probabilities to many hypotheses…it appears that logical structure alone cannot distinguish good inductive inferences from bad ones." (4)

Consider, for instance, (again!) the toy example

> "**Every** crow in a random sample of n crows is black. **Therefore**, this strongly supports the hypothesis that 'all crows are black.'
>
> Which according to Inductive Logic can be depicted according to the following scheme:
>
> **P1** S is a **random sample** of C (crows).
> **P2** $|S| = n$ (i.e. S has n members)
> **P3** The **frequency** (or proportion) of members of C having property B (i.e., being black) in sample S is r (= 1). (I.e. $f[B, C \cap S] = r = 1$)
> ---
> ---[p]
> $\therefore$ [To a **margin of error** q conclude that] The **proportion** of **all** members of C having property B is r (= 1).

Now everything hinges on *how* one can systematically pin down p (the degree of support). For a Bayesian, p depends on prior probabilities, i.e. probabilities (either empirical or subjective) established *before* the evidence was gathered in the experiment, as depicted in the premises **P1, P2, P3.** Circular worries aside[14] what Hawthorne is mentioning here is how formal approaches at best provide a *constraint* to classes of hypotheses. Furthermore, the consensus is that the Bayesian strategy for establishing constraints to hypotheses is far too broad…admitting too many ridiculous hypotheses into the background by and through which such prior probabilities can be established. Based on these criticisms, Hawthorne mentions (page 6) that Bayesians generally retreated to a *subjectivist* position, i.e. interpreting probability in a subjective manner (in terms of degrees of belief; recall remark in page 3 above.)

Hawthorne (unfortunately) uses a somewhat atypical logical formalism (representing negation via "~" , conjunction by "·", and implication by "⊃"). For the sake of consistency with more generic uses, I will use the formalism I introduced in **<u>Lecture 2</u>** (September 4th notes), i.e. I'll represent negation, conjunction, and implication with "¬", "∧", "→", respectively.

---

[14] I.e, Popper's worry: Other prior probabilities are required to establish these prior probabilities, and other prior probabilities are needed to establish the basis of those other prior probabilities, etc. .Bayesians claim to have a strategy circumventing such a worry, or at least rendering the circulariry virtuous,.

By "entailment" assume without loss of generality that Hawthorne is describing implication.[15] So, for instance (recall **Lecture II** Sept 4 notes): for P to entail Q: $P \rightarrow Q \leftrightarrow Q \vee \neg P$. Negating this entailment[16]: $\neg(P \rightarrow Q) \leftrightarrow \neg(Q \vee \neg P) \leftrightarrow \neg Q \wedge \neg\neg P \leftrightarrow \neg Q \wedge P \leftrightarrow \boldsymbol{P \wedge \neg Q}$.

Hence we arrive via these simple transformations that the negation of $P \rightarrow Q$ is $P \wedge \neg Q$. Hence what Hawthorne means when he says that if we're certain that P entails Q, then we're certain that $P \wedge \neg Q$ cannot be the case. (page 5) To give a simple example: "x is an integer implies x is a real number" is a certain statement (i.e. deductively true). Its negation reads: "x is not a real number and x is an integer", which of course is always false.

Ordinary deductive exhibits *monotonicity*. That is to say, if any premise set $\Gamma = \{p,q,\ldots,s\}$ entails a conclusion Q, then adding extra premises to $\Gamma = \{p,q,\ldots,s\}$ won't undermine the truth of the entailment. In the simple case involving implication[17]: $P \rightarrow Q \vdash (P \wedge R) \rightarrow Q$. To name a simple example: "x is an integer implies x is a real number" guarantees also the truth of the statement: "x is an integer *and*…implies x is a real number."[18]

Unfortunately for the case of inductive reasoning, if one associates conditional statements $p \rightarrow q$ with conditional probabilities $P(q \mid p)$ then monotonicity no longer holds. For instance, suppose I'm 99% certain that if I miss the 7:00AM (p) train, then I'll be late for work (q). I.e., $P(q \mid p) = 0.99$. I *cannot* tack on some additional premise/condition and expect my probability to increase (or at least stay constant). For instance $P(q \mid p \wedge r) < 0.99$ depending on r. Here's an example, suppose r stands for "I caught a taxicab and the traffic wasn't heavy." In other words, I'm now *less than 99% certain* I'll be late for work if I missed the 7:00AM train *and* caught a taxi.

The distinguishing factor of Inductive Logic from mere probability alone is the notion that "support" is defined by a *class* of functions (not, in other words, some unique function). These functions should obey **Axioms 1. – 5.,** as listed on page 6. Note that the support functions are represented by conditional probabilities. Using the notation for simpler predicate logic (recall my remark above and the notation used in **Lecture 2** notes, they read:

---

[15] Entailment is actually a *semantic* notion, i.e. truth-finctional. Entailment is typically defined in more advanced logic as follows: "P entails Q" if and only if in every possible world where P holds true, Q holds true as well. Whereas implication is *syntactic*, i.e. defined via logical structure alone. (For instance, $P \rightarrow Q \leftrightarrow Q \vee \neg P$, recall **Lecture 2**.) For simple cases, at any rate, as demonstrated by the use of truth-tables for $\rightarrow$, one is in no danger to assume the equivalence of entailment with ordinary implications. For more subtle and complicated instances, however, one must distinguish these two concepts.

[16] As discussed in **Lecture 2**, these transformations for distributing negation ($\neg$) include *DeMorgan's Rule* (the negation of $\vee$ is $\wedge$ and vice versa), double negation: (P can be inferred from $\neg\neg P$), and commutativity of $\vee$ and $\wedge$: $(P \vee Q) \leftrightarrow (Q \vee P)$, $(P \wedge Q) \leftrightarrow (Q \wedge P)$.

[17] Recall (page 6, **Lecture 2 notes**) that $\vdash$ means "it logically follows that" (i.e. is guaranteed by the rules of inference).

[18] At first blush, this may seem a little bit strange. Suppose I add the obviously false premise like "1+1=3". **The conditional "*x*** is an integer and 1+1=3 implies x is a real number." **is still true!** (In other words, conditional statements are automatically true if their premises are *false*). To see why this is true, consider another absurd (but true) conditional: "If pigs fly in the midnight sun then all circles are squares." This case seems even more shocking because both the premise and the conclusion are false. **But this conditional statement is true!** To see why, recall: $P \rightarrow Q \leftrightarrow Q \vee \neg P$, so the right hand side in this case states: "All circles are squares or pigs don't fly in the midnight sun." This is true, since one part of the of the disjuction "pigs don't fly…" is true. In a nutshell, the only time a conditional statement is ever false occurs when the antecedent P is true but the consequent Q is false. ("If 1+1=2 then pigs fly" is false)

1. $0 < P_\alpha(q \mid p) < 1$ (for any sentences p, q that aren't deductively certain
2. If $p \rightarrow q$ then $P_\alpha(q \mid p) = 1$
3. If $p \leftrightarrow q$ then $P_\alpha(r \mid p) = P_\alpha(r \mid q)$
4. If $p \rightarrow \neg (r \wedge q)$ then either $P_\alpha(r \vee q \mid p) = P_\alpha(r \mid p) + P_\alpha(q \mid p)$ or $P_\alpha(s \mid p)$ for any sentence s
5. $P_\alpha(q \wedge r \mid p) = P_\alpha(q \mid p \wedge r)$ & $P_\alpha(r \mid p)$

Axioms, of course, mean *postulates*, or (hopefully) *reasonable* assumptions that undergird any formal system. One cannot *derive* such axioms, since they're established as *givens* (in order to make the combinations formed by the formal machinery possible).[19] These axioms seem reasonable enough and as Hawthorne shows, one can *deduce* the ordinary axioms of probability from 1.-5. By and large, the technical details here aren't important. Hawthorne attempts to explain what 1.-5. in ordinary parlance (pages 6-7), in somewhat obscure prose. Briefly, the underlying gist of these axioms state:

1. In the case of any two statements that aren't connected by deductive certainty, then we're never 100% certain of the truth or falsity of their relation.
2. On the other hand, if *p* (deductively) implies *q*, then we *are* 100% certain that "q, given p."
3. Moreover, if *p* and *q* are *coextensive* (i.e. logically equivalent[20]) , then our degree of support for "*r*, given *p*" should be equal to our degree of support of "*r*, given *q*."
4. See footnote 18 here. Note that If $p \rightarrow \neg (r \wedge q) \leftrightarrow \neg p \vee \neg (r \wedge q) \leftrightarrow \neg p \vee \neg r \vee \neg q$ (I.e. "If I'm at work on time today then it's not true that I skipped the meeting and hung out with my friends and watched the Superbowl" is equivalent to "I'm late for work or I didn't skip the meeting or I didn't watch the Superbowl with my friends". Given this disjunction it's easy to see how we can break up this into independent non-overlapping cases: My degree of certainty that I missed the meeting or I watched the superbowl with my friends, given that I'm at work on time is partitioned according to my degree of support that I watched the Superbowl with my friends, given that I went to work on time OR (exclusively, read "+") my degree of certainty that I skipped the meeting, given that I went to work on time.
5. Suppose I believe that $p \rightarrow (r \wedge q)$ some degree of support <1. Then it's reasonable that my belief is broken down to my belief that *q* given *p* and *r* AND my belief that *r*, given *p*. For instance, suppose I'm reasonably sure that if the match didn't strike, then the box is wet and chemical is ruined. Then, given that match didn't light when struck, my belief that it's wet and that there's something wrong with the chemical is equal to my belief that the match is wet, given that it didn't light when struck and the chemical is ruined & my belief that the chemical is ruined given that the match didn't light when struck.

Hawthorne discusses two extra provisional rules (6. and 7.) in pages 7-8. The point of this section is to show that Inductive Logic is an open research program, hence its controversial nature. Don't worry about the technical details here.

Hawthorne writes in the next section:

[19] Think, for example, of the case of ordinary Euclidean geometry. From four axioms + the fifth paralell line axiom one derives the entire formal system.
[20] Recall **Lecture II**)

> Each function P_α satisfying these rules [Axioms 1.-5., and to a lesser extent 6., 7.] may be viewed as a possible way of applying the notion of *inductive support* to a language L[21]…The issue [however] of which of the *possible* truth-value assignments to a language represents the *actual* truth or falsehood of its sentence depends…on the meaning of the non-logical terms and on the state of the actual world.[22] **Similarly, the degree to which some sentences *actually* support others in a fully meaningful language must rely on something more than merely satisfying the axioms for support functions**. (8)

Recall the problem about specifying a sample space for open-ended situations, as mentioned in footnote 5 above (page 2). A similar issue arises here. As mentioned also in page 6 above, concerning the problems with Bayesianism, subjective Bayesianist interpret P_α as analogous to subjective probability; i.e. in terms of the agent's *belief strength.* Non-subjective Bayesians interpret it in semi-empirical terms, as a measure of frequency over an ensembles of possible worlds (or possibly alternate sets of circumstances). For example, the inductive support that the match will light when struck depends on the frequency of cases of lit matches in which alternate set of counterfactual circumstances are varied in a *ceteris paribus* (all [other] things being equal) fashion. For example, would the match have lit if: (at possible 'world'1) *Ceteris paribus,* the match didn't light because it was damp, (at possible 'world'2) *Ceteris paribus,* the match didn't light because the air is too thin, etc..

There are problems with both interpretations. The problems with Bayesianism was alluded to earlier (page 6). Whether or not the subjective interpretation takes care of Bayesianism major contentious issues is a point we'll re-visit later on in the course. To note in passing here, however, a **simple** aspect[23] of the subjective interpretation is that "[*s*]*ubjectivist Bayesians* usually take *inductive probability* to just *be* th[e] notion of *probabilistic belief-strength.*" (8)

The 'possible worlds' interpretation is applied in the form of possible alternative hypotheses, in the case seeking to apply Inductive Logic to science.

Suppose, for instance, one has a *finite*[24] set of mutually incompatible **alternative** hypotheses/theories $\mathbb{A} = \{H_1, H_{2,\dots,} H_N\}$ alongside the hypothesis/theory H being tested. For example, consider the observed effect E of two magnetic stuck together. H could be the hypothesis that 'the two magnets are stuck together due to their magnetic force.' $\mathbb{A} = \{H_1, H_{2,\dots,} H_N\}$ on the other hand could consist of hypotheses like: 'the magnets stuck together because they're glued together,' 'the magnets stuck together because of gravity,' 'the magnetic stuck together because of some hitherto-unknown force X', etc.

The **catch-all hypothesis** merely denies all the other alternative hypotheses, i.e., stating that 'none of the alternative hypotheses are true': $H_K \leftrightarrow \neg H_1 \wedge \neg H_2 \wedge \ldots \wedge \neg H_N$.

---

21 In this context, you can assume the basic terms and connectives of L can be characterized by FOPL (First Order Predicate Logic, recall **Lecture 2** notes.)
22 I.e., recall Agazzi's remark quoted earlier here. Logic isn't simply a matter of form, but also of content. The 'non-logical terms' are the words in the sentences, whose semantic content (i.e. meaning) *refer* to (extra-logical) notions and things. For example, "there will be a sea-battle tomorrow" refers to some singularly possible future contingent state of affairs.
23 Recall Kuhn's criteria (**Lecture VII**)
24 Note the restriction here. **Possible paper question to address**: Does Popper's falsificationism require such resitrictions? Are there significant cases in science when the list of hypotheses is opn-ended? Conversely, is presupposing such a finite list analogous to fixing a sample space in advance (as in the case of classical probability?) Is this a realistic maneuver?

In the hypothetico-deductive [25]conception of science, a hypothesis can be deduced from evidence E and "all background and auxiliaries not at issue in the assessment of the hypothesis" (9), B, and the relevant conditions C under which the test is performed. I.e., adopting (without loss of generality—recall footnote 15, page 7 above) the notion of 'implies' a hypothetico-deductivist would say that for a given hypothesis H: $H\wedge B\wedge C \rightarrow E$ or $H\wedge B\wedge C \rightarrow \neg E$ [26]

Suppose $H\wedge B\wedge C \rightarrow \neg E$. Then[27] $E\wedge B\wedge C \rightarrow \neg H$.

By *posterior probability* $P_\alpha(H| B\wedge C^n\wedge E^n)$ this means the probability of H being true given that B is fixed and I performed n trials with confirming evidence E and conditions C. I.e.:

$C^n\wedge E^n = (C\wedge E)\wedge (C\wedge E)\wedge\ldots\wedge (C\wedge E)$ (n times)

For example suppose under suitably equivalent[28] conditions C given fixed (non-H dependent) background assumptions B I sample n=10 statistically significant samples of 30 ravens, and in each of these samples I observe that the ravens are all black (i.e. in each of these 10 trials of samples of 30 ravens, E holds). I seek to quantify the posterior probability $P_\alpha(H| B\wedge C^n\wedge E^n)$, i.e. the probability that H: "All ravens are black" is **true**, given that $B\wedge C^n\wedge E^n$ where $n = 10$.

In Inductive Logic, it turns out that $P_\alpha(H| B\wedge C^n\wedge E^n)$ depends on:

1. Prior probability $P_\alpha(H| B)$, i.e. the probability of H being true given non-H dependent background assumptions.
2. Likelihood of E being true, given H and B and C, i.e. $P_\alpha(E| H\wedge C\wedge B)$
3. Likelihoods of E being true, given all the other alternative hypotheses H_k, where k=1,…,N and B and C, i.e. $P_\alpha(E| H_k \wedge C\wedge B)$

If you've had a course in probability and statistics, then you know the formula and theorem by and through which the posterior probability $P_\alpha(H| B\wedge C^n\wedge E^n)$ is related to the prior probabilities mentioned in 1., 2., 3. above. (It's of course *Bayes* Theorem, or formula). If you haven't seen this before, don't worry about it. We'll be *conceptually* analyzing the Bayesian paradigm in further detail later in the course. The point for now is to recognize that Bayesianism is an example of a Kuhnian paradigm or a Lakatosian research programme. And no, it's not universally accepted. (Though Bayesianism undergirds most of *actuarial* sciences and risk-assessment, i.e. what insurance companies do when they try to figure out rates to charge

[25] Recall from class discussion/lecture. Hypothetico-deductivists (like Popper and others) argue that one *deduces* hypotheses/theories from premises which are fallible (i.e. could be proven wrong). **Note: hypothetico-deductivism isn't equivalent to falsificationism!** Other hypothetico-deductivists have different approaches for testing theories besides Popper's approach. As is apparent in this article, ofr instance, inductive logicians can be hypothetico-deductivists as well.

[26] For example: H : "All ravens are black", B: All background assumptions not theory-related, C: conditions of empirical testing (selected 30 ravens). E : "All 30 ravens were black". ¬ "Some of the 30 weren't black."

[27] Since $H\wedge C\wedge B\rightarrow\neg E \leftrightarrow \neg(H\wedge C\wedge B)\vee\neg E \leftrightarrow \neg H \vee \neg C \vee\neg B \vee \neg E \leftrightarrow \neg C \vee\neg B \vee \neg E \vee\neg H \leftrightarrow$ $\neg(E\wedge C\wedge B)\vee\neg H \leftrightarrow E\wedge C\wedge B\rightarrow\neg H$. For instance, if in the example described in footnote 26 above, we get evidence that some of the 30 ravens aren't black, then the evidence that all 30 ravens observed were black & background assumptions & conditions entail that one should *not* conclude that all ravesn are black.

[28] See note 26 above. Is this restriction too idealized or is it reasonable? Does Popper require a similar condition? Why or why not?

customers, that sort of thing.) Nevertheless, as Kuhn pointed out (recall the Sept 20th reading) such approaches seem to be non-starters (in his opinion) when dealing with the complex situation involving theory-choice in science.

Writes Hawthorne:

> Likelihoods..aris[ing] from explicit statistical background claims…are often called *direct inference likelihoods.* Such likelihoods are completely objective. So it seems reasonable to suppose that all support functions should agree on their values [i.e., these likelihoods]…[d]irect inference likelihoods are *logical* in an extended, non-deductive sense. (11)

Such an example of a direct inference likelihood (DIL) would be the posterior probability of testing the hypothesis that 'all ravens are black' discussed above, i.e. $P_\alpha(H| B\wedge C^n\wedge E^n)$, since the n =10 trials are all statistically sampled.

Question for your paper (possible) How often do these DILs occur in the 'hard sciences' like physics? In the life sciences, hypothesis testing via statistical sampling is of course indispensable. But consider the case of physicist testing some very abstract mathematically based hypothesis/theory H that may in principle not even be statistically testable. For instance cosmologists don't use statistics, sine their theories are all based on *one* system (the Universe itself!) Are there therefore limits to Inductive Logic in terms of what sciences the logic is applicable toward? Does Popper's falsificationism suffer similar restrictions? (For instance, by the same token many theories in physics are often difficult to falsify in Popper's elegant sense. String theory is one such example.)

"[T]he empirical probability of a science relies on a high degree or objectivity or intersubjective agreement among scientists on the numerical values of likelihoods…we will suppose [therefore] suppose that the likelihoods have objective or intersubjectively agreed values, common to all agents in a scientific community." (11)

For instance (back to raven example!) $P_\alpha(E^n \mid H\wedge C^n\wedge E^n) = r$ (where $0 < r < 1$) represents such a DIL, in the case of n=10 trials yielding confirming evidence E (all in that sample consisting of 30 ravens were black), given the assumption that H ('all ravens are black') holds true. The above claims concerning objectivity and intersubjectivity enable us to fix some constant value r to this expression.

So going full circle, trying to quantify a degree of support for the theory H that 'all ravens are black' given agreed-upon (non-theory dependent) background assumptions B, along with n trials consisting of evidence E and conditions of testing C, amounts to trying to ascertain the posterior probability $P_\alpha(H| B\wedge C^n\wedge E^n)$. As mentioned above (points 1., 2., 3. on page 10) this depends on the associated prior probabilities $P_\alpha(H| B)$, etc.

"In the evidential evaluation of scientific theories, prior probabilities often represent **assessment by agents of non-evidential, conceptually motivated *plausibility weightings* among hypotheses."** (12) Oh really? Bayesians argue that such 'subjective' factors can be rendered systematically regular and are drawn on "forceful conceptual considerations." (Kuhn, on the other hand, would beg to differ!—At least insofar as buying into the notion that such factors can be algorithmically based, much less quantifiable.)

> “Although prior probabilities may be subjective in the sense that agents may disagree on the relative strengths of plausibility arguments—and so disagree on the plausibilities of various hypotheses—priors are far from being *mere subjective whims.*” (13)

Nevertheless:,

> “[M]ost logicians now take the project [of trying to fix prior probabilities by logical form alone] to have failed…semantic content should matter.” (13)

And therein lies the rub. Kuhn would argue that the semantic content is almost completely determined (in this case) by cognitive and contextual values that defy systematization and quantification. Popper rejects the very premises of Bayesian assumptions of convergence (recall from your readings of the *Logic of Scientific Discovery*). The question then becomes, based on you acquaintance with some of the notions of inductive logic presented in your reading and in these notes here, is inductive logic still a worthy program to embark on, when considering the testing of hypotheses? For on the flip side, though Popper’s falsificationism appears *prima facie* far simpler, it leaves us with no *positive* measure by which we can assess a theory’s worth. Inductive logicians, on the other hand, propose that inductive *support* provides just such a positive measure. Yet, as discussed here, their program is riddled with idealizations and assumptions. Do you think these assumptions miss some essential aspect concerning actual scientific theory-choice that Kuhn complains they do? Do you think Popper’s falsificationism does or doesn’t? All things considered, which, do you think, is in the final analysis the ‘lesser of the two evils’? Or more optimistically stated, given the reasonable degree of ‘success’ Bayesian approach seem to enjoy in other fields (actuarial sciences), controversies notwithstanding, do you Inductive logic or falsificationism as more productive research programmes?