Author & Citation Info | Friends PDF Preview | InPho Search | PhilPapers Bibliography

Inductive Logic

First published Mon Sep 6, 2004; substantive revision Wed Feb 27, 2008; substantive revision Thurs Feb 10, 2011

An inductive logic is a system of evidential support that extends deductive logic to less-than-certain inferences. For valid deductive arguments the premises *logically entail* the conclusion, where the entailment means that the truth of the premises provides a *guarantee* of the truth of the conclusion. Similarly, in a good inductive argument the premises should provide some *degree of support* for the conclusion, where such support means that the truth of the premises indicates with some *degree of strength* that the conclusion is true. Presumably, if the logic of good inductive arguments is to be of any real value, the measure of support it articulates should meet the following condition:

> **Criterion of Adequacy (CoA)**:
> As evidence accumulates, the *degree* to which the collection of true evidence statements comes to *support* a hypothesis, as measured by the logic, should tend to indicate that false hypotheses are probably false and that true hypotheses are probably true.

This article will focus on the kind of the approach to inductive logic most widely studied by philosophers and logicians in recent years. These logics employ conditional probability functions to represent measures of the degree to which evidence statements support hypotheses. This kind of approach usually draws on Bayes' theorem, which is a theorem of probability theory, to articulate how the *implications of hypotheses about evidence claims* influences the degree to which hypotheses are supported by those evidence claims. We will examine the extent to which this kind of logic may pass muster as an adequate logic of evidential support, especially in regard to the testing of scientific hypotheses. In particular, we will see how such a logic may be shown to satisfy the Criterion of Adequacy.

Sections 1 through 3 present all of the main ideas behind the probabilistic logic of evidential support. For most readers these three sections will suffice to provide an adequate understanding of the subject. Those readers who want to know more about how the logic applies when the *implications of hypotheses about evidence claims* (called *likelihoods*) are vague or imprecise may, after reading sections 1-3, skip down to section 6.

Sections 4 and 5 are for the more advanced reader who wants a detailed understanding of some telling results about how this logic may bring about convergence to the truth. These results show that the Criterion of Adequacy is indeed satisfied — that as evidence accumulates, false hypotheses will very probably come to have evidential support values (as measured by their *posterior probabilities*) that approach 0; and as this happens, a true hypothesis will very probably acquire evidential support values (as measured by their *posterior probabilities*) that approach 1.

---

1. Inductive Arguments

Let us begin by considering examples of the kinds of arguments an inductive logic should explicate. Consider the following two arguments:

> **Example 1.**. Every raven in a random sample of 3200 ravens is black. This strongly supports the hypothesis that all ravens are black.

> **Example 2**. 62 percent of voters in a random sample of 400 registered voters (polled on February 20, 2004) said that they favor John Kerry over George W. Bush for President in the 2004 Presidential election. This supports with a probability of at least .95 the hypothesis that between 57 percent and 67 percent of all registered voters favor Kerry over Bush for President (at or around the time the poll was taken).

An argument of this kind is often called an *induction by enumeration* of cases. We may represent the logical form of such arguments semi-formally as follows:

> Premise: In random sample S consisting of n members of population B, the proportion of members that have attribute A is r.
>
> Therefore, with degree of support p,
>
> Conclusion: The proportion of all members of B that have attribute A is between $r-q$ and $r+q$ (i.e., is within *margin of errorq* of r).

Let's lay out this argument more formally. The Premise breaks down into three separate premises:[1]

	Semi-formalization	Formalization
Premise 1	The frequency (or proportion) of members with attribute A among the members of S is r.	$F[A,S] = r$
Premise 2	S is a random sample of B with respect to whether or not its members have A	$\mathrm{Rnd}[S,B,A]$
Premise 3	Sample S has exactly n members	$\mathrm{Size}[S] = n$
Therefore	with degree of support p	========{p
Conclusion	The proportion of all members of B that have attribute A is between $r-q$ and $r+q$ (i.e. is within *margin of errorq* of r)	$F[A,B] = r \pm q$

Any inductive logic that encompasses such arguments should address two challenges. (1) It should tell us which enumerative inductive arguments should count as *good* inductive arguments rather than as inductive fallacies. In particular, it should tell us how to determine the appropriate *degree p* to which such premises *inductively support* the conclusion, for a given margin of error *q*. (2) It should demonstrably satisfy the CoA. That is, it should be provable (as a metatheorem) that *if* a conclusion expressing the approximate proportion for an attribute in a population is true, *then* it is very likely that sufficiently numerous random samples of the population will provide true premises for *good* inductive arguments that confer *degrees of support p* approaching 1 for that true conclusion — where, on pain of triviality, these *sufficiently numerous* samples are only a tiny fraction of a large population. Later we will see how a probabilistic inductive logic may meet these two challenges.

Enumerative induction is rather limited in scope. This form of induction is only applicable to the support of claims involving simple universal conditionals (i.e., claims of form 'All *B*s are *A*s') and claims about the proportion of an attribute in a population (i.e., 'The frequency of *A*s among the *B*s is *r*'). And it applies only when the evidence for such claims consists of instances of *B*s observed to be either *A*s or non-*A*s. However, many important empirical hypotheses are not reducible to this simple form, and the evidence for hypotheses is often not composed of simple instances. Consider, for example, the Newtonian Theory of Mechanics:

> All objects remain at rest or in uniform motion unless acted upon by some external force. An object's acceleration (i.e., the rate at which its motion changes from rest or uniform motion) is in the same direction as the force exerted on it; and the rate at which the object accelerates due to a force is equal to the magnitude of the force divided by the object's mass. If an object exerts a force on another object, the second object exerts an equal amount of force on the first object, but in the opposite direction to the force exerted by the first object.

The evidence for (and against) this theory is not gotten by examining a randomly selected subset of objects and the forces acting upon them. Rather, the theory is tested by calculating observable phenomena entailed by it in a wide variety of specific situations — ranging from simple collisions between small bodies to the trajectories of planets and comets — and then seeing whether those phenomena really occur. This approach to testing hypotheses and theories is ubiquitous, and should be captured by an adequate inductive logic.

Many less theoretical instances of inductive reasoning also fail to be captured by enumerative induction. Consider the kinds of inferences members of a jury are supposed to make based on the evidence presented at a murder trial. The inference to probable guilt or innocence is usually based on a patchwork of various sorts of evidence. It almost never involves consideration of a randomly selected sequences of past situations when people like the accused committed similar murders. Or, consider how a doctor diagnoses her patient on the basis of his symptoms. Although the frequency of occurrence of various diseases when similar symptoms were present may play a role, this is clearly not the whole story. Diagnosticians commonly employ a form of *hypothetical reasoning* — e.g., if the patient has a brain tumor, would that account for all of his symptoms?; or are these symptoms more likely the result of a minor stroke?; or is there another possible cause? The point is that a full account of inductive logic should not be limited to enumerative induction, but should also explicate the logic of *hypothetical reasoning* through which hypotheses and theories are tested on the basis of their predictions about specific observations. In Section 3 we will see how a kind of probabilistic inductive logic called "Bayesian Confirmation Theory" captures such reasoning.

2. Inductive Logic and Inductive Probabilities

Probability, and the equivalent notion *odds*, are the oldest and best understood ways of representing partial belief and uncertain inference. Probability has been studied by mathematicians for over 350 years, but the concept is certainly much older. In recent times a number of other related representations of uncertainty have emerged. Many of these have found useful application in computer based artificial intelligence systems that perform inductive inferences in expert domains such as medical diagnosis. This article will explicate the representation of inductive inferences in terms of *probability*. A brief comparative description of some of the most prominent alternative representations may be found in the following supplementary document:

Some Prominent Approaches to the Represention of Uncertain Inferences.

2.1 The Historical Origins of Probabilistic Logic

The mathematical study of probability originated with Blaise Pascal and Pierre de Fermat in the mid-17th century. From that time through the early 19th century, as the mathematical theory continued to develop, the theory was primarily applied to the assessment of risk in games of chance and to drawing simple statistical inferences about characteristics of large populations — e.g., to compute appropriate life insurance premiums based on mortality rates. In the early 19th century Pierre de Laplace made further theoretical advances and showed how to apply probabilistic reasoning to a much wider range of scientific and practical problems. Since that time probability has become an indispensable tool in the sciences, business, and many other areas of modern life.

Throughout its development various researchers appear to have thought of probability as a kind of logic. But the first extended treatment of probability as an explicit part of logic was George Boole's *The Laws of Thought* (1854). John Venn followed two decades later with an alternative empirical frequentist account of probability in *The Logic of Chance* (1876). Not long after that the whole discipline of logic was transformed by new developments in deductive logic.

In the late 19th and early 20th century Frege, followed by Russell and Whitehead, showed how deductive logic could be represented in the kind of rigorous formal system we now call *quantificational logic* or *predicate logic*. For the first time logicians had a fully formal deductive logic powerful enough to represent all valid deductive arguments in mathematics and the sciences — a logic in which the validity of deductive arguments depends only on the logical structure of the sentences involved. This development spurred some logicians to attempt to apply a similar approach to inductive reasoning. The idea was to extend the deductive entailment relation to a notion of *probabilistic entailment* for cases where premises provide less than conclusive support for conclusions. These *partial entailments* are expressed in terms of *conditional probabilities*, probabilities of the form $P[C \mid B] = r$ (read "the probability of C given B is r"), where P is a probability function, C is a conclusion sentence, B is a conjunction of premise sentences, and r is the probabilistic degree of support that B provides for C. Attempts to develop such a logic have varied widely in regard to precisely how the deductive model is emulated.

Some inductive logicians have tried to follow the deductive paradigm very closely by attempting to specify inductive support probabilities in terms of the syntactic structures of premise and conclusion sentences. In deductive logic the syntactic structure of the sentences involved completely determines whether premises logically entail a conclusion. So these logicians attempted to specify inductive support probabilities solely in terms of the syntactic structure of premise and conclusion sentences. In such a system each sentence confers a syntactically specified degree of support on each of the other sentences of the language. The inductive probabilities in such a system are *logical* in the sense that they depend on syntactic structure alone. This kind of conception was articulated to some extent by John Maynard Keynes in his *Treatise on Probability* (1921). Rudolf Carnap pursued this idea with greater rigor in his *Logical Foundations of Probability* (1950) and in several subsequent works (e.g., Carnap 1952). (For details of Carnap's approach see the section on logical probability in the entry on interpretations of the probability calculus, in this *Encyclopedia*.)

In the inductive logics of Keynes and Carnap, Bayes' theorem, which is a theorem of probability theory, plays a central role in expressing how evidence comes to bear on hypotheses. (We'll examine Bayes' theorem later.) So, such approaches might well be called *Bayesian logicist* inductive logics. Other well-known Bayesian logicist attempts to develop a probabilistic inductive logic include (Jeffreys, 1939), (Jaynes, 1968), and (Rosenkrantz, 1981).

It is now generally held that the core idea of Bayesian logicism is fatally flawed — that syntactic logical structure cannot be the sole determiner of the degree to which premises inductively support conclusions. 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. The difficulty is that in *any* probabilistic logic that satisfies the usual axioms for probabilities, the inductive support for a hypothesis must depend in part on its *prior probability*. This *prior probability* represents how plausible the hypothesis is supposed to be based on considerations other than the observational and experimental evidence (e.g. perhaps due to relevant plausibility arguments). A Bayesian logicist must tell us how to assign values to these pre-evidential *prior probabilities* of hypotheses, for each of the hypotheses or theories under consideration. Furthermore, this kind of Bayesian logicist must determine these *prior probability* values in a way that relies only on the syntactic logical structure of these hypotheses, perhaps based on some measure of their syntactic simplicities. There are severe technical problems with getting this idea to work. Moreover, various kinds of examples seem to show that such an approach must assign intuitively quite unreasonable prior probabilities to hypotheses in specific cases (see the footnote cited near the end of section 3.2 for details). Furthermore, for this idea to apply to the evidential support of real scientific theories, scientists would have to formalize theories in a way that makes their relevant syntactic structures apparent, and then evaluate theories solely on that syntactic basis (together with their syntactic relationships to evidence statements). Are we to evaluate alternative theories of gravitation (and alternative quantum theories) this way? This seems an extremely doubtful approach to the evaluation of real scientific theories and hypotheses. Thus, it seems that logical structure alone cannot suffice for the inductive evaluation of scientific hypotheses. (This issue will be treated in more detail in Section 3, after we first see how probabilistic logics employ Bayes' theorem to represent the evidential support for hypotheses as a function of *prior probabilities* together with their *evidential likelihoods*.)

At about the time the Bayesian logicist idea was developing, an alternative conception of probabilistic inductive reasoning was also emerging. This approach is now generally referred to as the Bayesian *subjectivist* or *personalist* approach to inductive reasoning (see, e.g., Ramsey, 1926; De Finetti, 1937; Savage 1954; Edwards, Lindman, Savage, 1963; Jeffrey, 1983, 1992; Howson, Urbach, 1993; Joyce 1999). It treats inductive probability as part of a larger normative theory of belief and action known as *Bayesian decision theory*. The principle idea is that the strength of an agent's desires for various possible outcomes should combine with her belief-strengths regarding claims about the world to produce optimally rational decisions. Bayesian subjectivists provide a logic that captures this idea, and they attempt to justify this logic by showing that in principle it leads to optimal decisions about which of various risky alternatives should be pursued. On the Bayesian subjectivist or personalist account of inductive probability, inductive probability functions represent the subjective (or personal) belief -strengths of ideally rational agents, the kind of belief strengths that figure into rational decision making. (See the section on subjective probability in the entry on interpretations of the probability calculus, in this *Encyclopedia*.)

Elements of the logicist conception of inductive logic live on today as part of the general approach called *Bayesian inductive logic*. However, among philosophers and statisticians the term ‘Bayesian’ is now most closely associated with the subjectivist or personalist account of belief and decision. And the term ‘Bayesian inductive logic’ has come to carry the connotation of a logic that involves purely subjective probabilities.

This current usage is misleading since for inductive logics the Bayesian/non-Bayesian distinction should really hang on whether the logic gives Bayes' theorem a prominent role, or whether the logic largely eschews the use of Bayes' theorem in inductive inferences (as do the *classical approaches* to statistical inference developed by R. A. Fisher (1922) and by Neyman and Pearson (1967)). Indeed, any inductive logic that employs the same probability functions to represent both the *probabilities of evidence claims due to hypotheses* and the *probabilities of hypotheses due to those evidence claims* must be a *Bayesian* inductive logic in this broader sense; because Bayes' theorem follows directly from the axioms that each probability function must satisfy, and Bayes' theorem expresses a necessary connection between the *probabilities of evidence claims due to hypotheses* and the *probabilities of hypotheses due to those evidence claims*.

In this article the *probabilistic inductive logic* we will examine is a *Bayesian* inductive logic in the broader sense. This logic will not presuppose the *subjectivist Bayesian theory* of belief and decision, and will avoid the objectionable features of Bayesian logicism. Later we will see that there are good reasons to distinguish *inductive probabilities* from Bayesian *degree-of-belief probabilities* and from *purely logical probabilities*. So, the probabilistic logic articulated in this article will be presented in a way that depends on neither of these conceptions of what the probability functions are. However, this version of the logic will be general enough that it may be fitted to a Bayesian subjectivist or Bayesian logicist program, if one desires to do that.

2.2 Probabilistic Logic: Axioms and Characteristics

All logics derive from the meanings of terms in sentences. What we now recognize as *formal deductive logic* rests on the meanings (i.e., the truth-functional properties) of the standard logical terms. These terms, and the symbols we will employ to represent them, are as follows: 'not', '~'; 'and', '·'; 'or', '∨'; truth-functional 'if-then', '⊃'; 'if and only if', '≡'; the quantifiers 'all', '∀', and 'some', '∃'; and the identity relation, '='. The meanings of all other terms (i.e., names, and predicate and relational expressions) are permitted to "float free". That is, the logic depends neither on their meanings nor on the truth-values of sentences containing them. It merely supposes that these other terms are meaningful, and that sentences containing them have truth-values. Deductive logic then tells us that the logical structures of some sentences — i.e., the syntactic arrangements of their logical terms — preclude them from being jointly true of any possible state of affairs. That is the notion of *logical inconsistency*. The notion of *logical entailment* is interdefinable with it. A collection of premise sentences *logically entails* a conclusion sentence just when the negation of the conclusion is *logically inconsistent* with those premises.

An inductive logic must, it seems, deviate from this paradigm in several significant ways. For one thing, logical entailment is an absolute, all-or-nothing relationship between sentences, whereas inductive support comes in degrees of strength. For another, although the notion of *inductive support* is analogous to the deductive notion of *logical entailment*, and is arguably an extension of it, there seems to be no inductive logic extension of the notion of *logical inconsistency* — at least none that is inter-definable with *inductive support* in the way that *logical inconsistency* is inter-definable with *logical entailment*. That is, *B logically entails A* just when $(B\cdot{\sim}A)$ is *logically inconsistent*. However, it turns out that when the unconditional probability of $(B\cdot{\sim}A)$ is very nearly 0 (i.e., when $(B\cdot{\sim}A)$ is "nearly inconsistent"), the degree to which *B inductively supports A*, $P[A \mid B]$, may range anywhere between 0 and 1.

Another notable difference is that when *B logically entails A*, adding a premise *C* cannot undermine the entailment — i.e., $(C\cdot B)$ must entail *A* as well. This property of *logical entailment* is called *monotonicity*. But *inductive support* is *nonmonotonic*. In general, depending on what *A*, *B*, and *C* mean, adding a premise *C* to *B* may substantially raise the degree of support for *A*, or may substantially lower it, or may leave it completely unchanged — i.e., $P[A \mid C\cdot B]$ may have a value much larger than

$P[A \mid B]$, or a much smaller value, or it may have the same, or nearly the same value as $P[A \mid B]$.

In a formal treatment of probabilistic inductive logic, inductive support is represented by conditional probability functions defined on sentences of a formal language *L*. These conditional probability functions are constrained by certain rules or axioms that are sensitive to the meanings of the logical terms (i.e., 'not', 'and', 'or', etc., the quantifiers 'all' and 'some', and the identity relation). The axioms apply without regard for what the other terms of the language may mean. In essence the axioms specify a family of *possible support functions*, $\{P_\beta, P_\gamma, \ldots, P_\delta, \ldots\}$ for a given language *L*. Although each support function satisfies these same axioms, the further issue of which among them provides an appropriate measure of *inductive support* is not settled by the axioms alone. That may depend on additional factors, such as the meanings of the non-logical terms in the language.

A good way to specify the rules or axioms of the logic of inductive support functions is as follows. Let *L* be a language for predicate logic with identity, and let '⊨' be the standard *logical entailment* relation — i.e. the expression '$B \vDash A$' says "*B logically entails A*" and the expression '$\vDash A$' says "*A* is a tautology".

> A support function is a function P_α from pairs of sentences of *L* to real numbers between 0 and 1 that satisfies the following rules or axioms:
>
> 1. $P_\alpha[D \mid E] < 1$ for at least one pair of sentences *D* and *E*.
>
> For all sentence *A*, *B*, and *C*,
>
> 2. If $B \vDash A$, then $P_\alpha[A \mid B] = 1$;
> 3. If $\vDash (B \equiv C)$, then $P_\alpha[A \mid B] = P_\alpha[A \mid C]$;
> 4. If $C \vDash {\sim}(B\cdot A)$, then $P_\alpha[(A \vee B) \mid C] = P_\alpha[A \mid C] + P_\alpha[B \mid C]$ or $P_\alpha[D \mid C] = 1$ for every *D*;
> 5. $P_\alpha[(A\cdot B) \mid C] = P_\alpha[A \mid (B\cdot C)] \times P_\alpha[B \mid C]$.

This axiomatization takes conditional probability as basic, as seems appropriate for *evidential support functions*. These functions agree with the usual unconditional probability functions when the latter are defined — just let $P_\alpha[A] = P_\alpha[A \mid (D \vee {\sim}D)]$. However, these axioms permit conditional probabilities $P_\alpha[A \mid C]$ to remain defined even when condition statement C has probability 0 (i.e., even when $P_\alpha[C \mid (D \vee {\sim}D)] = 0$).

Notice that conditional probability functions apply only to pairs of sentences, a conclusion sentence and a premise sentence. So in probabilistic inductive logic we represent finite collections of premises by conjoining them into a single sentence. Rather than say, 'A is supported to degree r by the set of premises $\{B_1, B_2, B_3,\ldots,B_n\}$', we say '$A$ is supported to degree r by the premise $(\ldots((B_1 \cdot B_2) \cdot B_3) \cdot \ldots \cdot B_n)$', and write this as '$P[A \mid (\ldots((B_1 \cdot B_2) \cdot B_3) \cdot \ldots \cdot B_n)] = r$'.

Let us briefly consider each axiom, 1-5, to see how plausible it is as a constraint on a quantitative measure of inductive support, and how it extends the notion of deductive entailment. First, notice that adopting an inductive support scale between 0 and 1 is merely a convenience. This scale is usual for probabilities; but any other scale might do as well.

Rule (1) is a non-triviality requirement. It says that at least one sentence must be supported by another to degree less than 1. We might instead have required that $P_\alpha[(A \cdot {\sim}A) \mid (A \vee {\sim}A)] < 1$; but this turns out to be derivable from Rule (1) together with the other rules.

Each *degree-of-support* function P_α on L measures *support strength* with numerical values between 0 and 1, with maximal support at 1. When B *logically entail* A, the support of A based on B is maximal. This is what Rule (2) asserts. It comports with the idea that an inductive support function is a generalization of the deductive entailment relation.

Rule (3) is equally obvious. It says that whenever B is logically equivalent to C, as premises each must provide precisely the same amount of support to every conclusion.

Rule (4) says that inductive support "adds up" in a plausible way. When C logically entails the incompatibility of A and B, the support C provides each separately must sum to the support it provides for their disjunction. The only exception is in cases where C acts like a contradiction and supports all sentences to degree 1.

To understand what Rule (5) says, think of a support function P_α as describing a measure on possible worlds or possible states of affairs. '$P_\alpha[C \mid D] = r$' says that the proportion of worlds in which C is true among those where D is true is r. Rule (5) then

says the following: if A is true in fraction r of worlds where B and C are true together, and if B (together with C) is true in proportion q of all the C-worlds, then A and B (and C) should be true together in fraction r of that proportion q of B (and C) worlds among the C-worlds.[2]

From these five rules all of the usual theorems of probability theory are easily derived. For example, logically equivalent sentences are always supported to the same degree: if $C \vDash (B \equiv A)$, then $P_\alpha[A \mid C] = P_\alpha[B \mid C]$. The following generalizations of the Addition Rule (4) may be proved as well:

$$P_\alpha[(A \vee B) \mid C] = P_\alpha[A \mid C] + P_\alpha[B \mid C] - P_\alpha[(A \cdot B) \mid C].$$

If $\{B_1, \ldots, B_n\}$ is any finite set of sentences such that for each pair B_i and B_j, $C \vDash {\sim}(B_i \cdot B_j)$ (i.e., the members of the set are mutually exclusive, given C), then

$$P_\alpha[((B_1 \vee B_2) \vee \ldots \vee B_n) \mid C] = \sum_{i=1}^{n} P_\alpha[B_i \mid C],$$

unless $P_\alpha[D \mid C] = 1$ for every sentence D.

If $\{B_1, \ldots, B_n, \ldots\}$ is any countably infinite set of sentences such that for each pair B_i and B_j, $C \vDash {\sim}(B_i \cdot B_j)$, then

$$\lim_n P_\alpha[((B_1 \vee B_2) \vee \ldots \vee B_n) \mid C] = \sum_{i=1}^{\infty} P_\alpha[B_i \mid C],$$

unless $P_\alpha[D \mid C] = 1$ for every sentence D.[3]

In the context of inductive logic it makes good sense to supplement the above rules with two additional rules. One is this:

6. If A is an axiom of set theory or any other piece of pure mathematics employed by the sciences, or if A is analytically truth (given the meanings of terms in L associated with support function P_α), then, for all C, $P_\alpha[A \mid C] = 1$.

The idea is that inductive logic is about evidential support for contingent claims. Nothing can count as empirical evidence against non-contingent truths. They should be maximally supported by all claims C.

One important respect in which inductive logic should follow the deductive paradigm is in not presupposing the truth-values of contingent sentences. No *inductive support function* P_α should permit a tautological premise to assign degree of support 1 to a contingent claim — i.e., $P_\alpha[C \mid (B \vee {\sim}B)]$ should always be less than 1 when C is contingent. For, the whole idea of inductive logic is to provide a measure of the extent to which contingent premise sentences indicate the likely truth-values of contingent conclusion sentences. This idea won't work properly if the truth-values of some contingent sentences are presupposed. Such presuppositions would make inductive logic enthymematic. It may hide significant premises in inductive support relationships.

However, it is common practice for probabilistic logicians to sweep provisionally accepted contingent claims under the rug by assigning them probability 1. This saves the trouble of repeatedly writing a given contingent sentence B as a premise, since $P_\gamma[A \mid B \cdot C]$ will just equal $P_\gamma[A \mid C]$ whenever $P_\gamma[B \mid C] = 1$. Although this device is useful, such probability functions should be considered mere abbreviations of proper, logically explicit, non-enthymematic, inductive support functions. Thus, properly speaking, an inductive support function $P_\alpha$should not assign probability 1 to a sentence *relative to all possible premises* unless that sentence is either (i) logically true, or (ii) an axiom of set theory or some other piece of pure mathematics employed by the sciences, or (iii) unless according to the interpretation of the language that $P_\alpha$presupposes, the sentence is *analytic*, and so outside the realm of evidential support. Thus, we adopt the following version of the so-called "axiom of regularity".

7. If, for all C, $P_\alpha[A \mid C] = 1$, then A is a logical truth or an axiom of set theory or some other piece of pure mathematics employed by the sciences, or A is analytically true (according to the meanings of the terms of L as represented in P_α).

This is more a convention than an axiom. Taken together with (6) it tells us that a support function P_α counts as non-contingently true just those sentences that it assigns probability 1 on every premise.

Some Bayesian logicists (e.g. Carnap) thought that inductive logic might be made to depend solely on the logical form of sentences, just like deductive logic. The idea was, effectively, to supplement axioms 1–7 with additional axioms that depend only on the logical structures of sentences, and to introduce enough such axioms to reduce the number of possible support functions to a single uniquely best confirmation function. It is now widely agreed that this project cannot be carried out in a plausible way. Perhaps there are additional rules that should be added to 1–7. But it is doubtful such rules can suffice to specify a single, uniquely qualified support function based only on logical structure. We will se why in Section 3, but only after first seeing how inductive probabilities capture the relationship between hypotheses and evidence.

2.3 Two Conceptions of Inductive Probability

Axioms 1–7 for conditional probability functions merely place formal constraints on what may properly count as a *degree of support function*. Each function P_α satisfying these rules may be viewed as a possible way of applying the notion of *inductive support* to a language L that respects the meanings of the logical terms, much as each possible *truth-value assignment* for a language represents a possible way of assigning truth-values to its sentences in a way that respects the semantic rules expressing the meanings of the logical terms. The issue of which of the *possible* truth-value assignments to a language represents the *actual* truth or falsehood of its sentences depends on more than this — it depends on the meanings of the non-logical terms and on the state of the actual world. 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. It must, at least, rely on what the sentences of the language mean, and perhaps on much more besides. But, what more? Various "interpretations of probability", which offer accounts of how support functions are to be understood, may help by filling out our conception of what *inductive support* is really about. There are two prominent views.

One reading is to take each P_α as a measure on possible worlds, or possible states of affairs. The idea is that, given a fully meaningful language (and, perhaps relative to the inferential inclinations of a particular agent, α) '$P_\alpha[A \mid B] = r$' says that among the worlds in which B is true, A is true in proportion r of them. There will generally not be a single privileged way to define such a measure on possible worlds. Rather, it may be that each of a number of functions P_α, P_β, P_γ, …, etc., satisfying the constraints imposed by axioms 1-7 can represent a viable measure of the *inferential import* of propositions expressed by sentences of the language. This idea needs more fleshing out, of course. The next section will give some indication of how that might go.

Subjectivist Bayesians offer an alternative reading of the support functions. First, they usually take unconditional probability as basic, and they take conditional probabilities as defined in terms of them: the conditional probability '$P_\alpha[A \mid B]$'is defined as a ratio of unconditional probabilities, $P_\alpha[A{\cdot}B]/P_\alpha[B]$.*Subjectivist Bayesians* take each unconditional probability function P_α to represent the belief-strengths or confidence-strengths of an ideally rational agent, α. On this understanding '$P_\alpha[A] = r$' says, "the strength of α's belief (or confidence) that *A* is truth is *r*." Subjectivist Bayesians usually tie such belief strengths to what the agent would be willing to bet on *A* turning out to be true. Roughly, the idea is this. Suppose that an ideally rational agent α would be willing to accept a wager that would yield (no less than) $*u* if *A* turns out to be true and would lose him $1 if *A* turns out to be false. Then, under reasonable assumptions about how much he desires money, it can be shown that his belief strength that *A* is true should be $P_\alpha[A] = 1/(u+1)$. And it can further be shown that any function P_α that expresses such betting-related belief-strengths on all statements in agent α's language must satisfy axioms for unconditional probabilities analogous to axioms 1–5.[4] Moreover, it can be shown that any function P_β that satisfies these axioms is a possible rational belief function for some ideally rational agent β. These relationships between belief-strengths and the desirability of outcomes (e.g., gaining money or goods on bets) are at the core of *subjectivist Bayesian decision theory*. *Subjectivist Bayesians* usually take *inductive probability* to just *be* this notion of *probabilistic belief-strength*.

Undoubtedly real agents do believe some claims more strongly than others. And, arguably, the belief strengths of real agents can be measured on a probabilistic scale between 0 and 1, at least approximately. And clearly the inductive support of evidence for hypotheses should influence the strength of an agent's belief in those hypotheses. However, there is good reason for caution about viewing *inductive support functions* as Bayesian belief-strength functions, as we will see a bit later. So, perhaps an agent's support function is not simply *identical to* his belief function, and perhaps the relationship between *inductive support* and *belief-strength* is somewhat more complicated.

In any case, some account of what support functions are supposed to represent is clearly needed. The belief function account and the possible worlds account are two attempts to provide this. Let us put this interpretative issue aside for now. One may be

able to get a better handle on what inductive support functions *really are* after one sees how the inductive logic that draws on them is supposed to work.

3. The Application of Inductive Probabilities to the Evaluation of Scientific Hypotheses

One of the most important applications of a formal inductive logic is to the confirmation or refutation of scientific hypotheses. The logic should explicate the notion of evidential support for all sorts of hypotheses, ranging from simple diagnostic claims (e.g., "the patient is infected with the HIV") to scientific theories about the fundamental nature of the world, like quantum mechanics or the theory of relativity. We'll now look into how support functions (a.k.a. confirmation functions) represent the logic of hypothesis confirmation. This kind of inductive logic is often referred to as *Bayesian Confirmation Theory*.

Consider some exhaustive set of mutually incompatible hypotheses or theories about some subject matter, $\{h_1, h_2, \ldots\}$. The set of alternatives may be very simple, e.g., {"the patient has HIV", "the patient is free of HIV"}. Or, when the physician is trying to determine which among a range of diseases is causing the patient's symptoms, the alternative hypotheses may consist of a long list of possible diseases. For the cosmologist the alternatives may be a list of several alternative gravitational theories, or several versions of the "same theory". Where inductive logic is concerned, even a slightly different version of a given theory will count as a distinct theory if it differs from the original in empirical import. (This should not be confused with the converse claim, which is the positivistic assertion that theories with the same empirical content are really the same theory. Inductive logic doesn't require you to buy that!)

In general there may be finitely or infinitely many such alternatives under consideration. They may all be considered at once, or they may be constructed and compared over a long historical period. One may even think of the set of alternative hypotheses as consisting of all logically possible alternatives expressible in a given language about a given subject matter — e.g., all possible theories of the origin and evolution of the universe expressible in English and mathematics. Although testing every possible alternative may pose practical challenges, it turns out that the logic works much the same way in the logically ideal case as it does in realistic cases.

If the set of alternative hypotheses is finite, it may contain a *catch-all hypothesis* h_K that says that none of the other hypotheses are true (e.g., "none of the other known diseases is present"). When only some number u of explicit alternative hypotheses is under consideration, h_K is just the sentence $(\sim h_1 \cdot \ldots \cdot \sim h_u)$.

Evidence for scientific hypotheses consists of the results of specific experiments or observations. For a given experiment or observation, let 'c' represent a description of the relevant *conditions* under which it is performed, and let 'e' represent a description of the result, the *evidential outcome* of conditions c.

Scientific hypotheses often require the mediation of background knowledge and auxiliary hypotheses to help them express claims about evidence. Let 'b' represent all background and auxilliary hypothese not at issue in the assessment of the hypotheses h_i, but that mediate their implications about evidence. In cases where a hypothesis is deductively related to evidence, either $h_i \cdot b \cdot c \vDash e$ or $h_i \cdot b \cdot c \vDash \sim e$.

For example, h_i might be the Newtonian Theory of Gravitation. A test of the theory might involve a condition statement c describing the results of some earlier measurements of Jupiter's position, and describing the means by which the next position measurement will be made; the outcome description e states the result of this additional position measurement; and the background information (or auxiliary hypotheses) b might state some already well confirmed theory about the workings and accuracy of the devices used to make the position measurements. Thus, if $(c \cdot e)$ occurs, this may be considered good evidence for h_i, given b, as the *hypothetico-deductive* account of confirmation maintains. On the other hand, if from $h_i \cdot b \cdot c$ we calculate some outcome incompatible with e, then $h_i \cdot b \cdot c \vDash \sim e$. In that case from deductive logic alone we get that $b \cdot c \cdot e \vDash \sim h_i$, and h_i is said to be *falsified* by $b \cdot c \cdot e$.

Duhem (1906) and Quine (1953) are generally credited with alerting inductive logicians to the importance of auxiliary hypotheses. They point out that scientific hypotheses often make little contact with evidence claims on their own. Rather, most scientific hypotheses only make testable predictions relative to background claims or auxiliary hypotheses that tie them to that evidence. Typically auxiliaries are highly confirmed hypotheses from other scientific domains. They often describe the operating characteristics of various devices (e.g. measuring instruments) used to make observations or conduct experiments. They are usually not at issue in the testing of h_i against its competitors, because h_i and its alternatives usually rely on the same

auxiliary hypotheses to tie them to the evidence. But even when an auxiliary hypothesis is already well-confirmed, we cannot simply assume that it is unproblematic, or just *known to be true*. Rather, the evidential support or refutation of a hypothesis h is *relative to* whatever auxiliaries and background information (in b) is being supposed. In other contexts the auxiliary hypotheses used to test h_i may themselves be among a collection of alternative hypotheses that are themselves subject to evidential support or refutation. (Furthermore, to the extent that competing hypotheses employ different auxiliary hypotheses in accounting for evidence, the evidence only tests each such hypothesis in conjunction with its distinct auxiliaries against alternative hypotheses packaged with their distinct auxiliaries.) Thus, what counts as a *hypothesis to be tested*, h_i, and what counts as auxiliary hypotheses and background information, b, and even to some extent what counts as the conditions c for an experiment or observation, will always depend on the epistemic context — on what alternative hypotheses are being tested by the same experiments or observations, and on what claims are being presupposed or held fixed for present purposes, and on what claims are considered to be immediate preconditions for the evidential outcome e. No statement is intrinsically a *hypotheis*, or intrinsically an *auxiliary* (or a *background condition*), or intrinsically an *evidential condition*. Rather, those are roles statements may play in an epistemic context, and the very same statement may play different roles in different confirmational contexts.

In a probabilistic inductive logic the degree to which evidence $c \cdot e$ supports a hypothesis h_i relative to background b is represented by the *posterior probability* of h_i, $P_\alpha[h_i \mid b \cdot c^n \cdot e^n]$. It turns out that the *posterior probability* of a hypothesis depends on just two kinds of factors: (1) its *prior probability*, $P_\alpha[h_i \mid b]$, together with the prior probabilities of its competitors, $P_\alpha[h_j \mid b]$, etc.; and (2) the *likelihood* of evidential outcomes e according to h_i, given that b and c are true, $P[e \mid h_i \cdot b \cdot c]$, together with the likelihoods of outcomes according to its competitors, $P[e \mid h_j \cdot b \cdot c]$, etc. In this section we will first examine each of these two kinds of factors in some detail, and then see precisely how the values of posterior probabilities depend on them.

3.1 Likelihoods

In probabilistic inductive logic *the likelihoods* carry the empirical import of hypotheses. A *likelihood* is a support function probability of form $P[e \mid h_i{\cdot}b{\cdot}c]$. It expresses how likely it is that outcome e will occur according to hypothesis h_i.[5] If a hypothesis together with auxiliaries and observation conditions deductively entails an evidence claim, the axioms of probability make the corresponding likelihood objective in the sense that every support function must agree on its values: i.e., $P[e \mid h_i{\cdot}b{\cdot}c] = 1$ if $h_i{\cdot}b{\cdot}c \vDash e$; $P[e \mid h_i{\cdot}b{\cdot}c] = 0$ if $h_i{\cdot}b{\cdot}c \vDash {\sim}e$. However, in many cases the hypothesis h_i will not be deductively related to the evidence, but will only imply it probabilistically. There are (at least) two ways this might happen. Either h_i may itself be an explicitly probabilistic or statistical hypothesis, or it may be that an auxiliary statistical hypothesis, as part of background b, connects h_i to the evidence. Let's briefly consider examples of each.

A blood test for HIV has a known false-positive rate and a known true-positive rate. Suppose the false positive rate is .05 — i.e., the test incorrectly shows the blood sample to be positive for HIV in 5% of all cases where HIV is not present. And suppose the true-positive rate is .99 — i.e., the test correctly shows the blood sample to be positive for HIV in 99% of all cases where HIV really is present.When a particular patient's blood is tested, the hypotheses under consideration are 'the patient is infected with HIV', h, and 'the patient is not infected with HIV', ${\sim}h$. In this context the known test characteristics function as background information, b. The experimental condition c merely states that this patient was subjected to a blood test for HIV, which was processed by the lab in the usual way. Let us suppose that the outcome e states that the result is positive for HIV. The relevant likelihoods, then, are $P[e \mid h{\cdot}b{\cdot}c] = .99$ and $P[e \mid {\sim}h{\cdot}b{\cdot}c] = .05$.

In this example the values of the likelihoods are entirely due to the statistical characteristics of the accuracy of the test, which is carried by the background information b. The hypothesis h being tested is not itself statistical.

This kind of situation may, of course, arise for much more complex hypotheses. The hypothesis of interest may be some deterministic physical theory, say Newtonian Gravitation Theory. Some of the experiments that test this theory relay on somewhat imprecise measurements that have known statistical error characteristics, which are

expressed as part of the background or auxiliary hypotheses b. For example, the auxiliary b may describe the error characteristics of a device that measures the torque imparted to a quartz fiber, where the measured torque is used to assess the strength of the gravitational force between test masses. In that case b may say that for this kind of device the measurement errors are normally distributed about whatever value a given gravitational theory predicts, with some specified standard deviation that is characteristic of the device. This results in specific values r_i for the likelihoods, $P[e \mid h_i{\cdot}b{\cdot}c] = r_i$, for each of the various alternative gravitational theories h_i being tested.

On the other hand, the hypotheses being tested may themselves be statistical in nature. One of the simplest examples of statistical hypotheses and their role in likelihoods are hypotheses about the chance characteristic of coin-tossing. Let $h_{[r]}$ be a hypothesis that says a specific coin has a propensity r (e.g., 1/2) for coming up heads on normal tosses, and that such tosses are probabilistically independent of one another. Let c state that the coin is tossed n times in the normal way; and let e say that on these tosses the coin comes up heads m times. In cases like this the value of the likelihood of the outcome e on hypothesis h for condition c is given by the well-known binomial term:

$$P[e \mid h_{[r]}{\cdot}b{\cdot}c] = \frac{n!}{m! \times (n-m)!} \times r^m\,(1-r)^{n-m}.$$

There are, of course, more complex cases of likelihoods involving statistical hypotheses. Consider, for example, the hypothesis that plutonium 233 nuclei have a half-life of 20 minutes — i.e., the propensity for a Pu-233 nucleus to decay within a 20 minute period is 1/2. This hypothesis, h, together with background b about decay products and the efficiency of the equipment used to detect them (which may itself be an auxiliary statistical hypothesis), yields precisely calculable values for likelihoods $P[e_k \mid h{\cdot}b{\cdot}c]$ of possible outcomes of the experimental arrangement.

Likelihoods that arise from explicit statistical claims — either within the hypotheses being tested, or from explicit statistical background claims that tie the hypotheses to the evidence — 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, just as all support functions agree on likelihoods when evidence is logically entailed. Direct inference likelihoods are *logical* in an extended, non-deductive sense. Indeed, some logicians have attempted to spell out the logic of *direct inferences* in terms of the logical form of the sentences involved.[6] But regardless of whether that project succeeds, it seems reasonable to take likelihoods of this sort to have highly objective or intersubjectively agreed values.

Not all likelihoods of interest in confirmational contexts are warranted deductively or by explicitly stated statistical claims. Nevertheless, the likelihoods that relate hypotheses to evidence in scientific contexts should often have objective or intersubjectively agreed values. So, although a variety of different support functions P_α, P_β ,…, P_γ, etc., may be needed to represent the differing "inductive proclivities" of the various members of a scientific community, all should agree, at least approximately, on the values of the likelihoods. For, likelihoods represent the empirical content of a hypothesis, what the hypothesis (together with background b) *probabilistically implies* about the evidence. Thus, the empirical objectivity of a science relies on a high degree of objectivity or intersubjective agreement among scientists on the numerical values of likelihoods.

To see the point more vividly, imagine what a science would be like if scientists disagreed widely about the values of likelihoods. Each practitioner interprets a theory to *say* quite different things about how likely it is that various possible evidence statements will turn out to be true. Whereas scientist α takes theory h_1 to probabilistically imply that event e is highly likely, his colleague β understands the empirical import of h_1 to say that e is very unlikely. And, conversely, α takes competing theory h_2 to probabilistically imply that e is quite unlikely, whereas β reads h_2 to say that e is very likely. So, for α the evidential outcome e supplies strong support for h_1 over h_2, because $P_\alpha[e \mid h_1 \cdot b \cdot c] \gg P_\alpha[e \mid h_2 \cdot b \cdot c]$. But his colleague β takes outcome e to show just the opposite — that h_2 is strongly supported over h_1 — because $P_\beta[e \mid h_1 \cdot b \cdot c] \ll P_\beta[e \mid h_2 \cdot b \cdot c]$. If this kind of thing were to occur often or for significant evidence claims in a scientific domain, it would make a shambles of the empirical objectivity of that science. It would completely undermine the empirical testability of its hypotheses and theories. Under such circumstances, although each scientist employs the same *theoretical sentences* to express a given theory h, each understands the *empirical import* of these sentences so differently that h as understood by α is an empirically different theory than h as understood by β. Thus, the empirical objectivity of the sciences requires that experts should be in close agreement about the values of the likelihoods.[7]

For now we will suppose that the likelihoods have objective or intersubjectively agreed values, common to all agents in a scientific community. Let us mark this agreement by dropping the subscript 'α', 'β', etc., from expressions that represent likelihoods. One might worry that this supposition is overly strong. There are many legitimate scientific contexts where, although scientists should have enough of a common understanding of the empirical import of hypotheses to assign quite similar values to likelihoods, precise agreement on the numerical values is unrealistic. This point is right. So later we will see how to relax the supposition that likelihood values agree precisely. But for now the main ideas behind probabilistic inductive logic will be more easily explained if we focus on those contexts were objective or intersubjectively agreed likelihoods are available. Later we will see that much the same logic continues to apply in contexts where the values of likelihoods may be somewhat vague, or where members of the scientific community disagree to some extent about their values.

An adequate treatment of the likelihoods calls for the introduction of one additional notational device. Scientific hypotheses are generally tested by a sequence of experiments or observations conducted over a period of time. To explicitly represent the accumulation of evidence, let the series of sentences $c_1, c_2, \ldots, c_n$, describe the conditions under which a sequence of experiments or observations are conducted. And let the corresponding outcomes of these observations be represented by sentences $e_1, e_2, \ldots, e_n$. We will abbreviate the conjunction of the first n descriptions of experimental or observation conditions as 'c^n', and abbreviate the conjunction of descriptions of their outcomes as 'e^n'. Then, for a stream of n observations or experiments and their outcomes, the likelihoods take form $P[e^n \mid h_i \cdot b \cdot c^n] = r$, for appropriate r between 0 and 1. In many cases the likelihood of the evidence stream will be equal to the product of the likelihoods of the individual outcomes:

$$P[e^n \mid h_i \cdot b \cdot c^n] = P[e_1 \mid h_i \cdot b \cdot c_1] \times \ldots \times P[e_n \mid h_i \cdot b \cdot c_n].$$

When this equality holds the individual bits of evidence are said to be *probabilistically independent on the hypothesis*. In what follows such *independence* will only be assumed in those places where it is explicitly invoked.

3.2 Posterior Probabilities and Prior Probabilities

In the probabilistic logic of evidential support the evaluation of a hypothesis on evidence is represented by its *posterior probability*, $P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n]$. The posterior probability represents the net plausibility of the hypothesis resulting from the combination of the evidence together with any relevant non-evidential plausibility considerations (which should be packaged within b). The likelihoods are the means through which evidence contributes to posterior probabilities. But another factor, the *prior probability* of the hypothesis based on considerations expressed within b, $P_\alpha[h_i \mid b]$, also makes a contribution. It represents the weight of all non-evidential plausibility considerations on which posterior probabilities may depend. It turns out that posterior probabilities depend *only* on the values of (ratios of) likelihoods *and* on the values of (ratios of) prior probabilities.

To understand the role of *prior probabilities*, consider the *HIV test* example described in the previous section. What the physician and patient want to know is the value of the posterior probability $P_\alpha[h \mid b{\cdot}c{\cdot}e]$ that the patient has HIV, h, given the evidence of the positive test, $c{\cdot}e$, and given the error rates of the test, described within b. The value of this posterior probability depends on the likelihood (due to the error rates) of this patient obtaining a true-positive result, $P[e \mid h{\cdot}b{\cdot}c] = .99$, and of obtaining a false positive result, $P[e \mid {\sim}h{\cdot}b{\cdot}c] = .05$. In addition, the value of the of the posterior probability depends on how plausible it is that the patient has HIV before the test results are taken into account, $P_\alpha[h \mid b]$. In the context of medical diagnosis this prior probability is sometimes called the *base rate*. It represents the probability that the patient may have contracted HIV based on his risk group (i.e., whether he is an IV drug user, has unprotected sex with multiple partners, etc.). Such information should be explicitly stated, and represented within b as well. To see its importance, consider the following numerical results (which may be calculated using the formula called Bayes' Theorem, presented in the next section). If the base rate for the patient's risk group is relatively high, say $P_\alpha[h \mid b] = .10$, then the positive test result yields a probability for his having HIV of $P_\alpha[h \mid b{\cdot}c{\cdot}e] = .69$. However, if the patient is in a very low risk group, $P_\alpha[h \mid b] = .001$, then a positive test only raises the probability of HIV infection to $P_\alpha[h \mid b{\cdot}c{\cdot}e] = .02$. This posterior probability is much higher than the prior probability of .001, but should not worry the patient too much. This positive test result is more likely due to the false-positive rate of the test than to the presence of HIV. (This sort of test, with such a large false-positive rate, .05, is best used as a screening test; a positive result should lead to a second, more rigorous, more expensive test.)

In the evidential evaluation of scientific theories, prior probabilities often represent assessments by agents of non-evidential, conceptually motivated *plausibility weightings* among hypotheses. However, because such plausibility assessments tend to vary among agents, critics often brand them as *merely subjective*, and take their role in probabilistic induction to be highly problematic. Bayesian inductivists counter that such assessments often play an important role in the sciences, especially when there is insufficient evidence to distinguish among some of the alternative hypotheses. And, they argue, the epithet *merely subjective* is unwarranted. Such plausibility assessments are often backed by extensive arguments that may draw on forceful conceptual considerations.

Consider, for example, the kind of plausibility arguments that have been brought to bear on the various interpretations of quantum theory (e.g., those related to the measurement problem). These arguments go to the heart of conceptual issues that were central to the development of the theory. Indeed, many of these issues were first raised by the scientists who made the greatest contributions to the theory's development, in the attempt to get a conceptual hold on the theory and its implications. Such arguments seem to play a legitimate role in the assessment of alternative views when distinguishing evidence has yet to be found.

More generally, scientists often bring plausibility arguments to bear in assessing their views. Although such arguments are seldom decisive, they may bring the scientific community into widely shared agreement, especially regarding the *implausibility* of some logically possible alternatives. This seems to be the primary epistemic role of the thought experiment. Thus, 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 comparative plausibilities of various hypotheses — the priors used in scientific contexts should not represent *mere subjective whims*. Rather, they should be supported (or at least be supportable) by explicit arguments regarding how much more plausible one hypothesis is than another. The important role of plausibility assessments is apparent in such received bits of scientific wisdom as the old saw that *extraordinary claims require extraordinary evidence*. That is, it takes especially strong evidence, in the form of extremely high values for ratios of likelihoods, to overcome the extremely