Confirmation Theory

James Hawthorne

University of Oklahoma

Draft -- to be submitted to Prasanta S. Bandyopadhyay and Malcolm Forster (eds.), *Philosophy of Statistics*, *Handbook of the Philosophy of Science*, *Volume 7*, Elsevier.

CONTENTS

Introduction

Confirmation theory is the study of the logic by which scientific hypotheses may be confirmed or disconfirmed, or even refuted by evidence. A specific theory of confirmation is a proposal for such a logic. Presumably the epistemic evaluation of scientific hypotheses should largely depend on their empirical content – on what they *say* the evidentially accessible parts of the world are like, and on the extent to which they turn out to be right about that. Thus, all theories of confirmation rely on measures of how well various alternative hypotheses account for the evidence.[1] Most contemporary confirmation theories employ probability functions to provide such a measure. They measure how well the evidence fits what the hypothesis *says* about the world in terms of *how likely* it is that the evidence should occur were the hypothesis true. Such hypothesis-based probabilities of evidence claims are called *likelihoods*. Clearly, when the evidence is more likely on one hypothesis than on an alternative, that should redound to the credit of the former hypothesis and the discredit of the later. But various theories of confirmation diverge on precisely how this credit is to be measured?

A natural approach is to also employ a probabilistic measure to directly represent the degree to which the hypothesis is confirmed or disconfirmed on the evidence. The idea is to rate the degree to which a hypothesis is confirmed on a scale from 0 to 1, where tautologies are always assigned maximal confirmation (degree 1), and where the degree of confirmation of the disjunction of mutually incompatible hypotheses sum to the degrees of confirmation of each taken separately. This way of rating confirmation just recapitulates the standard axioms of probability theory, but applies them as a measure of degree-of-confirmation. Any theory of confirmation that employs such a measure is an *probabilistic confirmation theory*. However, confirmation functions of this sort will be of little value unless it can be shown that under reasonable conditions the accumulation of evidence tends to drive the confirmation values given by these functions towards 0 for false hypotheses and towards 1 for true hypotheses.

How should confirmation values be related to what hypotheses imply about evidence claims via the likelihoods? The most straightforward idea would be to have the confirmation function assign to a hypothesis whatever numerical value is had by the *likelihood* the hypothesis assigns to the evidence. However, this idea won't work. For one thing, in cases where the hypothesis logically entails the evidence, the likelihood is 1. But we cannot require the confirmational

---

[1] I make no distinction here between scientific hypotheses and scientific theories. For our purposes a theory is just a large, complex hypothesis. We will suppose (as a formal logic must) that scientific hypotheses are expressible as sentences of a language – e.g. English, or a mathematically sophisticated dialect thereof. This supposition need not be in opposition to a "semantic view" of theories, since presumably, if scientists can express a theory well enough to agree about what it says about the world (or at least about its testable empirical content), it must be expressible in some bit of language. For instance, if a theory is taken to be a set of models, presumably that set of models (of whatever kind) may be described in mathematical English, perhaps employing set theory. In that case the empirical content of the theory will consist of various hypotheses about what parts of the world are modeled (to what degree of approximation) by some particular model or set of models. Such hypotheses should be subject to empirical evaluation. So a theory of confirmation should apply to them.

probability of each hypothesis that entails the evidence to be 1. For, some alternative hypotheses may also entail the evidence, and a probabilistic confirmation measure cannot assign probability 1 to two (or more) alternative hypotheses based on the same evidence.[2]

If both *likelihoods* and the *degrees of confirmation* for hypotheses are to be measured probabilistically, it seems natural to represent both by a common probability function. In that case what relationship does the *degree to which a hypothesis is confirmed on evidence* have to the *likelihood of the evidence according to the hypothesis*? This is where Bayes' Theorem comes into play. Bayes' Theorem follows from the standard axioms for probabilities; and it explicitly shows how the probabilistic degree of confirmation for a hypothesis depends on the likelihoods of evidence claims. Thus, any confirmation measure that satisfies the standard axioms of probability theory and employs the same probability function to represent likelihoods will have to be a *Bayesian confirmation measure*. Any theory of confirmation that employs such a probabilistic measure of confirmation will thus be a *Bayesian confirmation theory*.

Various Bayesian approaches to confirmation primarily diverge with regard to how they understand the concept of probabilistic confirmation – i.e. with regard to how they interpret the notion of *probability* that is supposed to be captured by confirmational probability functions. Is the confirmation function supposed to represent the *warranted degree of confidence* (or *degree of belief*) an agent should have in hypotheses (based on the evidence), as the subjectivists and personalists would have it? Or, is the confirmation function some kind of *objective logical relationship* by which the evidence is supposed to *probabilistically entail* hypotheses, as many logical objectivists maintain? Or, might the confirmation function represent some other coherent conception of evidential support?

These days the most prominent strain of Bayesian confirmation theory takes probabilistic confirmation functions to represent the rationally ideal agent's subjective *degrees of confidence*, or *degrees of belief* in statements or propositions. This view has become so influential that the term 'Bayesian confirmation theory' is often taken to be synonymous with it. But to identify Bayesianism with the subjectivist view is a mistake. It tends to either mischaracterize or entirely disregard a host of non-subjectivist Bayesian accounts. Properly speaking, it is not the subjective interpretation of the confirmation function that makes a confirmation theory *Bayesian*. Rather, any confirmation theory that gives Bayes' Theorem a central role in the representation of how a hypothesis is confirmed or refuted, based on *what the hypothesis says about evidence*, is a *Bayesian* confirmation theory. And any account that employs a common function to represent both confirmation and the likelihoods will give Bayes' Theorem this central role.

Historically, a number of proposals for objective Bayesian confirmation theories have been developed – e.g., the theories of Keynes (1921), Carnap (1950, 1951, 1971, 1980), Jeffreys (1939), Jaynes (1968); and more recently Rosenkrantz (1981), Maher (1996, 2001, 2006), and Williamson (2006) have defended objectivist accounts. The founding proponents of subjectivist Bayesianism include Ramsey (1926), de Finetti (1937), and Savage (1954). Prominent

---

[2] The same argument holds whenever two alternative hypotheses both assign the evidence a likelihood greater than 1/2 – and similarly, whenever *n* alternative hypotheses each assign the evidence likelihoods greater than 1/*n*.

contemporary subjectivist treatments include those of Jeffrey (1965, 1992), Levi (1967, 1980), Lewis (1980), Skyrms (1984), Earman (1992), Howson and Urbach (1993), and Joyce (1998, 1999). It should be said that although all subjectivist approaches draw on a belief-strength notion of probability, there are important differences among some of them. For instance, some subjectivist's hold that belief-strength should be constrained by additional epistemic norms that go beyond mere probabilistic coherence.

In this article I'll spell out the probabilistic logic that underlies almost all accounts of probabilistic confirmation functions. I'll discuss what I take to be some important difficulties faced by objectivist and subjectivist views about the nature of these functions, and I'll describe a view about the nature of Bayesian confirmation theory that I think avoids or overcomes the interpretative problems. Here is how I'll proceed.

In section 1 I'll discuss the most basic features of this probabilistic logic. I'll set down axioms that characterize these functions, but without yet describing how the logic is supposed to apply to the evaluation of scientific hypotheses.

Section 2 will briefly describe two accounts of the nature of confirmation functions – views about what the confirmation functions are supposed to be, or what they are supposed to represent. We won't be able to delve too far into this until after we see how the logic of confirmation functions represents the evidential support of hypotheses, which we'll address in section 3. But the logic of confirmation described is section 3 will be more easily comprehended if we have some preliminary idea about what the confirmation functions represent.

Section 3 shows how the logic of confirmation functions represents the evidential support of hypotheses. I'll spell out several forms of Bayes' Theorem, showing how the Bayesian formulae represent the role of the likelihoods and another important factor (i.e. the *prior probabilities*) in the logic of evidential support.

In section 4 I'll return to the issue of what confirmation functions are supposed to be. I'll describe major problems with the two most prominent interpretations, and suggest an alternative view. In particular I'll address the issue of how confirmation functions are supposed to inform our beliefs about the truth or falsity of hypotheses.

Is there any good reason to think that, given a suitable amount of evidence, false hypothesis will become strongly disconfirmed and true hypotheses will become highly confirm? In section 5 I'll explicate a *Bayesian Convergence Theorem* that establishes this. It shows that under reasonable conditions, the accumulation of evidence should result in the near refutation of false hypotheses (i.e. in confirmational probability approaching 0), and should lead to a high degree of confirmation for the true alternative (i.e. to confirmational probability near 1).[3]

The discussion throughout sections 3 through 5 will suppose that the likelihoods for evidence

[3] This *Bayesian Convergence Theorem* avoids many of the usual objections to such theorems. It depends only on *likelihoods*; so those who are suspicious of *prior probabilities* may still find it of interest. And it provides explicit lower bounds on the likely rate of convergence.

statements on the various hypotheses, which express the empirical content of hypotheses, are either objective or inter-subjectively agreed to by the relevant scientific community – so all confirmation functions employed by a scientific community will agree on their numerical values. In section 6 I'll weaken this supposition, but argue that the important truth-acquiring features of probabilistic confirmation theory may be retained.

1. The Probabilistic Logic of Confirmation Functions

Confirmation functions represent a logic by which scientific theories are refuted or supported by evidence. So they should be defined on an *object language* rich enough to fully express any scientific theory. The language for first-order predicate logic should suffice.[4] All current scientific theories are expressible in such a language.[5] The standard logical connectives and quantifiers for this language are as follows: "not", '~'; "and", '·'; "or", '∨'; truth-functional "if-then", '⊃'; "if and only if", '≡'; the quantifiers "all", '∀', and "some", '∃'; and the identity relation, "is the same thing as", '='. These are the *logical terms* of the language. The meanings of all other terms, called *non-logical terms* (i.e., names, and predicate and relational expressions), are not specified in advance by the logic. Standard deductive logic neither depends on their meanings nor on the actual truth-values of sentences containing them. It only supposes that the non-logical terms are meaningful, and that sentences containing them can have truth-values.

A *degree of confirmation function* represents a relationship between *statements* (i.e. declarative sentences) that is somewhat analogous to the deductive *logical entailment relation*. However, the logic of confirmation will have to deviate from the deductive paradigm in several ways. For one thing, deductive logical entailment is an absolute relationship between sentences, but confirmation comes in degrees. Deductive logical entailment is monotonic: when *B logically entails A*, adding a premise *C* cannot undermine the logical entailment – i.e., $(C \cdot B)$ must entail *A* as well. But *confirmation* is *nonmonotonic*. Adding a new premise *C* to *B* may substantially raise the degree of to which *A* is confirmed, or may substantially lower it, or may leave it completely unchanged – i.e., for a confirmation function P_α, the value of $P_\alpha[A \mid C \cdot B]$ may be much larger than $P_\alpha[A \mid B]$ (for some statements *C*), while it may be much smaller (for some other *C*), and it may have the same, or nearly the same value (for some other statements *C*).

Arguably, another very significant difference is this. A given deductive logic specifies a unique *logical entailment relation*, and that relation depends *only* on the meanings of the logical connectives and quantifiers. Is there, similarly, a uniquely good confirmation function? And does it depend only on the meanings of logical terms, and not on the specifics of what the individual sentences mean (due to the meanings of the names and predicates they contain)? Most confirmation theorists would answer "no" to both questions. (I'll discuss some reasons for this

[4] Everything I'll say here also applies to languages for second-order and higher-order logics.
[5] If you resist the idea that scientific theories may be represented in such a language, then think of the logic of confirmation presented here as a logician's model of the real thing. Viewed this way, the logician's model is, like any model employed by the sciences, an approximation that is intended to capture the essential features of it's subject matter, which in this case is the nature of correct scientific inferences regarding hypotheses expressible in a scientific dialect of English.

later.) Rather, from the logician's point of view, confirmation functions are technically somewhat analogous to *truth-value assignments* to sentences of a formal language. That is, holding the meanings of the logical terms (connectives and quantifiers) fixed, there are lots of ways to assign meanings to names and predicate terms of a language; and for each such meaning assignment, each of the various ways the world might turn out to be specifies a corresponding truth-value assignment for each sentence of the language. Similarly, keeping the meanings of logical terms fixed, each way of assigning meanings to names and predicate expressions may give rise to a distinct confirmation function – and perhaps to more than one. So, from the point of view of formal logic, there should be a host of different possible confirmation functions.

In deductive logic the possible truth-value assignments to sentences of a formal language *L* are constrained by certain *semantic rules*, which are axioms regarding the meanings of the logical terms ('not', 'and', 'or', etc., the quantifiers, and the identity relation). The rules, or axioms, for confirmation functions play a similar role. They constrain each member of the family of possible confirmation functions, $\{P_\beta, P_\gamma, \ldots, P_\delta, \ldots\}$, to respect the meanings of the logical terms, but without regard for what the other terms of the language may mean. Although each confirmation function satisfies the same basic axioms, the further issue of which among them provides an appropriate measure of confirmation is not settled by these axioms alone. It presumably depends on additional factors, including the meanings of the non-logical terms of the language.

Here are semantic rules (or axioms) that constrain probabilistic confirmation functions.[6]

> Let *L* be a language for predicate logic with identity, and let '$\models$' be the standard logical entailment relation for that logic (where '$B \models A$' abbreviates '*B* logically entails *A*', and '$\models A$' abbreviates '*A* is a logical truth'). A confirmation function is a function P_α from pairs of sentences of *L* to real numbers between 0 and 1 that satisfies the following rules:
>
> 1. $P_\alpha[D \mid E] < 1$ for at least one pair of sentences *D* and *E*.
>
> For all sentence *A*, *B*, *C*,
>
> 2. if $B \models A$, then $P_\alpha[A \mid B] =1$;
> 3. If $\models (B \equiv C)$, then $P_\alpha[A \mid B] = P_\alpha[A \mid C]$;
> 4. If $C \models {\sim}(B \cdot A)$, then either $P_\alpha[(A \vee B) \mid C] = P_\alpha[A \mid C] + P_\alpha[B \mid C]$, or for every sentence *D*, $P_\alpha[D \mid C] = 1$;

---

[6] The language *L* in which scientific hypotheses and evidence claims are expressed is what logicians call the *object-language*. As logicians see it, confirmation functions are not themselves part of the object-language *L*. Rather, we consider confirmation functions, along with other semantic notions like *truth* and *logical entailment*, to be in the metalanguage, where properties of object-language expressions are treated. Logicians distinguish object-language from metalanguage in order to avoid certain kinds of paradoxes and logical incoherencies that can arise from applying semantic concepts to themselves.

5. $P_\alpha[(A{\cdot}B) \mid C] = P_\alpha[A \mid (B{\cdot}C)] \cdot P_\alpha[B \mid C]$.

This particular axiomatization takes conditional probability as basic. The conditional probability functions it characterizes agree with the more usual account of 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 a condition statement C has probability 0 (e.g., even when $P_\alpha[C \mid (D{\vee}{\sim}D)] = 0$). However, this feature is not the primary reason for taking conditional probability as primitive for confirmation functions.

The main reason is this. On the usual account, where unconditional probability is basic, conditional probability is defined as follows: $P[A \mid B] = P[(A{\cdot}B)] / P[B]$ if $P[B] > 0$, and is undefined otherwise. (This is closely related to axiom 5.) But if one takes conditional probability as defined in this way, the likelihood a hypothesis H assigns to some evidence statement E under experimental or observation conditions C, $P[E \mid H{\cdot}C]$, must be *defined* as follows: $P[E \mid H{\cdot}C] = P[E{\cdot}H{\cdot}C] / P[H{\cdot}C]$. However, in the context of confirmation functions it seems unnatural to take such likelihoods as *defined* like this. For, likelihoods often have very well-known, well-defined values all on their own, whereas the values of the probabilities in the numerator and denominator of '$P[E{\cdot}H{\cdot}C] / P[H{\cdot}C]$' are often only vaguely known or specified.

For example, let H say "the coin is fair" (i.e. has a propensity to come up *heads* half the time when tossed in the usual way), let C say "the coin is tossed at present in the usual way", and let E say "the coin lands *heads* on the present toss". We may not be at all clear about the values of $P[E{\cdot}H{\cdot}C]$ (the probability that "the coin is fair, is presently tossed in the usual way, and lands *heads*) or of $P[H{\cdot}C]$ (the probability that "the coin is fair, and is presently tossed in the usual way"). Nevertheless, we still take the value of $P[E \mid H{\cdot}C]$ to be perfectly well-defined and well-known (it should clearly equal 1/2), due to what $H{\cdot}C$ says about E. Thus, because confirmation functions are supposed to represent such relationships between hypotheses and evidence statements, and because such *likelihoods* are often "better defined" or "better known" than the probabilities that would "define them" via the usual ratio definition, it seems more natural to axiomatize confirmational probabilities in a way that takes conditional probabilities as basic.[7]

One important special case where the present approach is especially helpful is this. Consider a statistical hypothesis that says that the chance (or measure) of an attribute X among systems in a state Y is 1. Formally, we might express such a hypothesis this way: '$Ch(X,Y)=1$'. Suppose it's known that a physical system g is in state Y (i.e., $g{\in}Y$). This gives rise to a likelihood: $P_\alpha[g{\in}X \mid Ch(X,Y)=1 \cdot g{\in}Y] = 1$. Now, adding certain kinds of additional information to the premise should lower this likelihood – e.g., $P_\alpha[g{\in}X \mid Ch(X,Y)=1 \cdot g{\in}Y \cdot {\sim}g{\in}X] = 0$. The

---

[7] See Hájek (2003a) for more reasons to take conditional probability as basic. Though this approach makes better sense conceptually, the account of Bayesian confirmation in the remainder of the paper does not (except where noted) essentially depend on taking conditional probability as basic. The more usual axiomatization will usually suffice. It goes like this: a confirmation function is a function P_α from sentences of L to real numbers between 0 and 1 that satisfies the following rules: (1) if $\models A$, then $P_\alpha[A] = 1$; (2) if $\models {\sim}(A{\cdot}B)$, then $P_\alpha[A{\vee}B] = P_\alpha[A] + P_\alpha[B]$; by definition $P_\alpha[A \mid B] = P_\alpha[A{\cdot}B] / P_\alpha[B]$ whenever $P_\alpha[B] > 0$.

confirmation functions permit this on our present approach, whereas on the usual approach, this "lowering of the likelihood" cannot happen when the likelihood has value 1.[8]

Let us now briefly consider each axiom, to see how plausible it is as a constraint on a measure of confirmation. First, notice that adopting a scale between 0 and 1 is merely a convenience. This scale is usual for probabilities, but another scale might do as well.

Rule (1) is a non-triviality requirement. It says that some sentences must be supported by others 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.

Rule (2) says that if *B* logically entails *A*, then *B* must maximally confirm *A*. This makes each probabilistic confirmation function a kind of generalization of the deductive logical entailment relation. Rule (3) adds the plausible constraint that whenever statements *B* and *C* are logically equivalent, each must provide precisely the same confirmational support to all other statements.

Rule (4) says that confirmational 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 the case where *C* acts like a contradiction and supports every sentence to degree 1.

To understand what Rule (5) says, think of a confirmation function P_α as describing a measure on possible states of affairs (or possible worlds): an expression of form '$P_\alpha[C \mid D] = r$' says that the proportion of states in which *C* is true among those where *D* is true is *r*. On this reading, Rule (5) says the following: suppose that *B* (together with *C*) is true in proportion *q* of all states where *C* is true; and suppose that *A* is true in fraction *r* of all those states where *B* and *C* are true together; then *A* and *B* (and *C*) will be true in fraction *r* of the proportion *q* (i.e. in $r\cdot q$) of all the states where *C* is true.

All of the usual theorems of probability theory are easily derived form these axioms. For example, logically equivalent sentences are always supported to the same degree: if $C \models (B\equiv A)$, then $P_\alpha[A \mid C] = P_\alpha[B \mid C]$. And the following generalization of the Addition Rule (4) is derivable: $P_\alpha[(A\vee B) \mid C] = P_\alpha[A \mid C] + P_\alpha[B \mid C] - P_\alpha[(A\cdot B) \mid C]$.

It also follows that if $\{B_1, \ldots, B_n, \ldots\}$ is any countable set sentences that are mutually exclusive, given *C* (in that, for each pair B_i and B_j, $C \models {\sim}(B_i\cdot B_j)$), 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*).[9]

---

[8] For, if by definition $P_\alpha[g{\in}X \mid Ch(X,Y){=}r \cdot g{\in}Y] = P_\alpha[g{\in}X \cdot Ch(X,Y){=}1 \cdot g{\in}Y] / P_\alpha[Ch(X,Y){=}r \cdot g{\in}Y]$, then for $P_\alpha[g{\in}X \mid Ch(X,Y){=}1 \cdot g{\in}Y] = 1$ we have $P_\alpha[g{\in}X \cdot Ch(X,Y){=}1 \cdot g{\in}Y] = P_\alpha[Ch(X,Y){=}1 \cdot g{\in}Y] = P_\alpha[(g{\in}X \vee {\sim}\, g{\in}X) \cdot Ch(X,Y){=}1 \cdot g{\in}Y] = P_\alpha[g{\in}X \cdot Ch(X,Y){=}1 \cdot g{\in}Y] + P_\alpha[{\sim}g{\in}X \cdot Ch(X,Y){=}1 \cdot g{\in}Y]$. So $P_\alpha[{\sim}g{\in}X \cdot Ch(X,Y){=}1 \cdot g{\in}Y] = 0$. Thus, $P_\alpha[g{\in}X \mid Ch(X,Y){=}1 \cdot g{\in}Y \cdot {\sim}g{\in}X]$ cannot be 0, but rather, it must be undefined because the term that would be in the denominator of the definition has value 0.

[9] This is not *countable additivity*. I have no objection to countable additivity in general. But it is not a natural part of a probabilistic logic defined on sentences of a formal language for predicate

In the context of the logic of confirmation it makes good sense to supplement the above rules with two more. Here's the first one:

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 true (given the meanings that P_α presupposes for the terms in L), then, for all C, $P_\alpha[A \mid C] = 1$.

The idea is that the logic of confirmation is about evidential support for contingent claims. Nothing can count as empirical evidence against or for non-contingent truths. They should be maximally confirmed given any and all possible statements.

An important respect in which the logic of confirmation functions should follow the deductive paradigm is in not presupposing the truth-values of contingent sentences. For the whole idea of a logic of confirmation is to provide a measure of the extent to which contingent premise sentences indicate the likely truth-values of contingent conclusion sentences. But this idea won't work

---

logic. In this context countable additivity would require the language to possess a means of expressing infinitely long disjunctions. Then the rule for countable additivity would say that whenever for all distinct B_i and B_j, $C \models {\sim}(B_i \cdot B_j)$, then $P_\alpha[B_1 \vee B_2 \vee \ldots \mid C] = \sum_{i=1}^{\infty} P_\alpha[B_i \mid C]$. The result stated in the main text may be derived without appeal to countable additivity, as follows:

Suppose that for each distinct B_i and B_j, $C \models {\sim}(B_i \cdot B_j)$, and suppose $P_\alpha[D \mid C] < 1$ for at least one sentence D. Notice that, for each i, $C \models ({\sim}(B_i \cdot B_{i+1}) \cdot \ldots \cdot {\sim}(B_i \cdot B_n))$. This implies $C \models {\sim}(B_i \cdot (B_{i+1} \vee \ldots \vee B_n))$. So for each finite list of the B_i, $P_\alpha[(B_1 \vee B_2 \vee \ldots \vee B_n) \mid C] = P_\alpha[B_1 \mid \mathrm{C}] + P_\alpha[(B_2 \vee \ldots \vee B_n) \mid C] = \ldots = \sum_{i=1}^{n} P_\alpha[B_i \mid C]$. By definition, $\sum_{i=1}^{\infty} P_\alpha[B_i \mid C] = \lim_n \sum_{i=1}^{n} P_\alpha[B_i \mid C]$. Thus, $\lim_n P_\alpha[(B_1 \vee B_2 \vee \ldots \vee B_n) \mid C] = \sum_{i=1}^{\infty} P_\alpha[B_i \mid C]$.

First-order logic does have a limited means to express infinite disjunctions via existential quantification. So one might consider adding a kind of countable additivity axiom, as follows:

for each open expression Fx, $P_\alpha[\exists x Fx \mid B] = \lim_n P[Fc_1 \vee \ldots \vee Fc_n \mid B]$, where the individual constants $c_1, \ldots, c_n, \ldots$, exhaust the countably infinite list of L's individual constants.

From this axiom the following form of countable additivity follows: if for each distinct c_i and c_j, $B \models {\sim}(Fc_i \cdot Fc_j)$, then $P_\alpha[\exists x Fx \mid B] = \sum_i P_\alpha[Fc_i \mid C]$. However, the proposed axiom seems overly strong, since it effectively assumes that every individual object gets named. That would mean there cannot be more than a countable number of objects under discussion.

If we don't assume that all individuals are named, the strongest claim we should want is this: $P_\alpha[\exists x Fx \mid B] \geq \lim_n P_\alpha[Fc_1 \vee \ldots \vee Fc_n \mid B]$. But this already follows from the other axioms:

$P_\alpha[\exists x Fx \mid B] \geq P_\alpha[\exists x Fx \mid B] \cdot P_\alpha[(Fc_1 \vee \ldots \vee Fc_n) \mid B \cdot \exists x Fx] = P_\alpha[(Fc_1 \vee \ldots \vee Fc_n) \cdot \exists x Fx \mid B] = P_\alpha[\exists x Fx \mid B \cdot (Fc_1 \vee \ldots \vee Fc_n)] \cdot P_\alpha[(Fc_1 \vee \ldots \vee Fc_n) \mid B] = P_\alpha[(Fc_1 \vee \ldots \vee Fc_n) \mid B]$ (since $P_\alpha[\exists x Fx \mid B \cdot (Fc_1 \vee \ldots \vee Fc_n)] = 1$, because $B \cdot (Fc_1 \vee \ldots \vee Fc_n) \models \exists x Fx$).

Thus, if for each pair of distinct c_i and c_j, $B \models {\sim}(Fc_i \cdot Fc_j)$, then $P_\alpha[\exists x Fx \mid B] \geq \sum_i P_\alpha[Fc_i \mid C]$.

One more point. Confirmation functions are in the meta-language, where logical relationships reside, rather than in the object-language L, where scientific hypotheses live. So, although countable additivity may not be a feature of confirmation functions, scientific hypotheses (expressed in the object-language) may themselves employ countably additive object-language probability functions to model chance processes in nature.

properly if the truth-values of some contingent sentences are presupposed by the confirmation function. Such presuppositions may hide significant premises, making the logic confirmation enthymematic. Thus, for example, no confirmation function P_α should permit a tautological premise to assign degree of confirmation 1 to a contingent claim: $P_\alpha[C \mid B \vee {\sim}B]$ should always be less than 1 when C is contingent.

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 functions should be considered mere abbreviations of proper, logically explicit, non-enthymematic, confirmational relationships. Thus, properly speaking, a confirmation function P_α should assign probability 1 to a sentence *relative to every possible premise* only if 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) the sentence is *analytic* according to the meanings of terms in the language presupposed by confirmation function P_α, and so outside the realm of evidential support. Thus, it is natural to adopt the following version of the so-called "axiom of regularity".

7. If A is not a consequence of set theory or some other piece of pure mathematics employed by the sciences, and is neither a logically nor an analytically true statement (given the meanings of the terms of L as represented in P_α), then $P_\alpha[A \mid {\sim}A] < 1$.[10]

Taken together with (6) it tells us that a confirmation function P_α counts as non-contingently true just those sentences that it assigns probability 1 on every possible premise.[11]

Bayesian logicists such as Keynes and Carnap thought that the logic of confirmation 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 do so with enough such axioms to reduce the number of possible confirmation functions to a single unique 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 that such rules can suffice to specify a single, uniquely qualified confirmation function based only on the formal structure of sentences. I'll say more about why this is doubtful a bit later, after we first see how confirmational probabilities capture the important relationships between hypotheses and evidence.

2. Two Conceptions of Confirmational Probability

Axioms 1-7 for conditional probability functions merely place formal constraints on what may

---

[10] It follows from the other axioms that $P_\alpha[A \mid {\sim}A] = 0$: for, ${\sim}A \models (A \vee {\sim}A)$, so $1 = P_\alpha[A \vee {\sim}A \mid {\sim}A] = P_\alpha[A \mid {\sim}A] + P_\alpha[{\sim}A \mid {\sim}A] = P_\alpha[A \mid {\sim}A] + 1$.

[11] Because, if $P_\alpha[A \mid C] = 1$ for all C, then $P_\alpha[A \mid {\sim}A] = 1$. Then 7 says that A must be either a consequence of set theory or some other piece of pure mathematics employed by the sciences, or it must be logically or analytically true.

properly count as a probabilistic confirmation function. Each function P_α that satisfies these rules may be viewed as a possible way of specifying a confirmation function 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 sentences of 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 these axioms for confirmation 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 confirmation functions are to be understood, may help by filling out our conception of what confirmational probability is really about. I'll describe two prominent views.[12]

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. Generally there will be more than one respectable way to define such a measure on possible worlds. Many of the various functions P_α, P_β, P_γ, …, etc., that satisfy the constraints imposed by axioms 1-7 may represent a viable measure of the *inferential import* of statements in a given language. This kind of *logicistist* reading of the Bayesian confirmation functions needs more fleshing out, of course. But the idea should be clear enough for present purposes.

Subjectivist Bayesians offer an alternative reading of the confirmation functions. First, they usually take unconditional probability as basic, and they take conditional probabilities as defined in terms of them. Furthermore, 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 being true. Roughly, the idea is this. Suppose that an ideally rational agent α is willing to accept a wager that would yield \$$u$ (but no less) if A turns out to be true and would lose him \$1 if A turns out to be false. Then, under reasonable assumptions about his desires for small amounts of 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 the usual axioms for unconditional probabilities.[13] 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. Such relationships between belief-strengths and the desirability of outcomes (e.g., gains in money or goods on bets) are at the core of *Bayesian decision theory*.[14] Subjectivist Bayesians usually take confirmational

[12] For a comprehensive treatment of the many interpretations of 'probability' see (Hájek 2003b).
[13] Note 7 lists these axioms.
[14] An alternative, but conceptually similar approach is to set down intuitively plausible constraints on the notion of *rational preference* among acts (or their outcomes), and then show

probability to just *be* this notion of probabilistic belief-strength.[15]

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 confirmational 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 taking confirmation functions to *be* Bayesian belief-strength functions, as we will see later. So, perhaps an agent's confirmation function is not simply identical to his belief function, and perhaps the relationship between confirmation and belief-strength is somewhat more complicated than the subjective Bayesian supposes.

In any case, some account of what confirmation functions are supposed to represent is clearly needed. The belief function account and the possible worlds account are two attempts to provide this. Let's put this interpretative issue aside until later (section 4). We'll try to get a better handle on what probabilistic confirmation functions *really are* after we take a careful look at how the logic that draws on them is supposed to work.

3. The Logical Structure of Evidential Support, and the role of Bayes' Theorem in that Logic

A theory of confirmation should explicate the notion of evidential support for all sorts of scientific hypotheses, ranging from simple diagnostic claims (e.g., the patient has pneumonia) to scientific theories about the fundamental nature of the world, like quantum mechanics or the theory of relativity. We'll now look at how the logic of probabilistic confirmation functions draws on Bayes' Theorem to bring evidence to bear, via the likelihoods, on the refutation or support of scientific hypotheses.

To begin with, consider some exhaustive set of mutually incompatible hypotheses or theories about some common subject matter, $\{h_1, h_2, \ldots\}$. The set of alternatives may consist of a simple pair of alternatives – e.g., {"the patient has pneumonia", "the patient doesn't have pneumonia"}. Or it may consist of a long list of possible alternatives, as is the case when the physician tries to determine which among a range of diseases is causing the patients symptoms. For the cosmologist the alternatives may consist of several alternative theories of the structure and dynamics of space-time, and may include various versions of the "same theory". Where confirmation theory is concerned, even a slightly different version of a given theory will count as a distinct theory, especially if it differs from the original in empirical import.

In principle there may be finitely or infinitely many alternative hypotheses under consideration. They may all be considered at once, or they may be constructed and assessed over many centuries. One may even take the set of alternative hypotheses to consist of all logically possible

---

that any such notion of preference can be represented (uniquely) by a probabilistic belief function together with a utility function, and that *preferred* acts (or their outcomes) are just those that maximized expected utility.

[15] For various versions of this sort of subjectivist approach see (Ramsey 1926), (de Finetti 1937), (Savage, 1954), (Jeffrey 1965), (Skyrms 1984), and (Joyce 1999).

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 logically possible alternative poses practical challenges, the logic works much the same way in the logically ideal case as it does in more practical cases.

The set of alternative hypotheses may contain a *catch-all hypothesis* h_K that says that none of the other hypotheses are true – e.g., "the patient has none of the known diseases". When only a finite number u of explicit alternative hypotheses is under consideration, h_K will be equivalent to the sentence that denies each definite alternative: $({\sim}h_1 \cdot \ldots \cdot {\sim}h_u)$.

Evidence for scientific hypotheses comes from specific experiments or observations. For a given experiment or observation, let 'c' represent a description of the relevant experimental or observational *conditions* under which the evidence is obtained, and let 'e' represent a description of the *evidential outcome* that comes about under 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 auxiliary hypotheses 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 $h_i \cdot b \cdot c \models e$ or $h_i \cdot b \cdot c \models {\sim}e$.

For example, h_i might be Newtonian Theory of Gravitation. A test of the theory may involve conditions described by a statement c about how measurements of Jupiter's position are made at various times; the outcome description e states the results of each position measurement; and the background information (or auxiliary hypotheses) b may state some trustworthy (already well confirmed) theory about the workings and accuracy of the devices used to make the position measurements. If outcome e can be calculated from the theory h_i together with b and c, we have that $h_i \cdot b \cdot c \models e$. The so-called *hypothetico-deductive* account of confirmation says that in such cases, if $(c \cdot e)$ actually occurs, this may be considered good evidence for h_i, given b. On the other hand, if from $h_i \cdot b \cdot c$ we calculate some outcome incompatible with e, then we have $h_i \cdot b \cdot c \models {\sim}e$. In that case, from deductive logic alone we get that $b \cdot c \cdot e \models {\sim}h_i$, and h_i is said to be *falsified* by $b \cdot c \cdot e$.

Duhem (1906) and Quine (1953) are generally credited for 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. So, generally speaking, the evidence can only falsify hypotheses relative to background or auxiliary hypotheses that tie them to that evidence. However, the auxiliaries themselves will usually be testable on some separate body of evidence in much the same way that the hypotheses $\{h_1, h_2, \ldots\}$ are tested. Furthermore, when we are not simply interested in assessing the hypotheses $\{h_1, h_2, \ldots\}$ relative to a specific package of auxiliaries b, but instead want to consider various alternative packages of auxiliary hypotheses, $\{b_1, b_2, \ldots\}$, as well, the set of *alternative hypotheses* to which the logic of confirmation applies should be the various possible combinations of original hypotheses in conjunction with the possible alternative auxiliaries, $\{h_1 \cdot b_1, h_1 \cdot b_2, ..., h_2 \cdot b_1, h_2 \cdot b_2, ..., h_3 \cdot b_1, h_3 \cdot b_2, \ldots\}$. When this is the case, the logic of confirmation will remain the same. The only difference is that the hypotheses 'h_i' in our later discussion should be taken to stand for the complex conjunctive hypotheses of form $(h_k \cdot b_v)$, and 'b' in our later discussion should stand for whatever remaining, common auxiliary hypotheses are not at issue. In the most extreme case, where each hypothesis at issue includes within itself all relevant auxiliaries, the term '*b*' may be empty – i.e. we may take it to be some simple tautology.

In probabilistic confirmation theory the degree to which a hypothesis h_i is supported or confirmed on evidence $c \cdot e$, 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 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 (give that b and c are true), $P[e \mid h_i \cdot b \cdot c]$ – together with the likelihoods of the outcomes according to h_i's competitors h_j, $P[e \mid h_j \cdot b \cdot c]$, etc.... I'll now examine each of these two kinds of factors more closely. Then I'll discuss how the values of posterior probabilities depend on them.

3.1 Likelihoods

Hypotheses express their empirical import via *likelihoods*, which are confirmation function probabilities of form $P[e \mid h_i \cdot b \cdot c]$. A likelihood expresses how likely it is that outcome e will occur according hypothesis h_i under conditions c, supposing that auxiliaries b hold.[16] If a hypothesis together with auxiliaries and observation conditions deductively entails an evidence claim, the probability axioms make the corresponding likelihood objective in the sense that every confirmation function must agree on its values: i.e., for all confirmation functions P, $P[e \mid h_i \cdot b \cdot c] = 1$ if $h_i \cdot b \cdot c \models e$, and $P[e \mid h_i \cdot b \cdot c] = 0$ if $h_i \cdot b \cdot c \models {\sim}e$. 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 there may be an auxiliary statistical hypothesis in the background b that connects h_i to the evidence. For the sake of clarity 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 about 5% of all cases where no HIV is present. And suppose the true-positive rate is .99 – i.e., the test correctly shows the blood sample to be positive for HIV in about 99% 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 play the role of 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

---

[16] Bayesians often refer to the probability of an evidence statement on a hypothesis, $P[e \mid h \cdot b \cdot c]$, as the *likelihood of the hypothesis*. This can be a somewhat confusing convention since it is clearly the evidence that is made likely to whatever degree by the hypothesis. So, I'll disregard the usual convention here. Also, presentations of the logic of confirmation often suppress c and b, and simply write '$P[e \mid h]$'. But c and b are important parts of the logic of the likelihoods. So I'll continue to make them explicit.

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 may relay on 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, used to assess the strength of the gravitational force between test masses. In that case *b* may say that for this kind of device 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 consists of hypotheses about the chance-characteristics of coin-tossing. Let $h_{[r]}$ be a hypothesis that says a specific coin has a propensity *r* for coming up *heads* on normal tosses, and that all such tosses are probabilistically independent of one another. Let *c* state that the coin is tossed *n* times in the usual way; and let *e* state that on these specific *n* 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 well-known: $P[e \mid h_{[r]} \cdot b \cdot c] = [n!/(m!(n-m)!)]\ r^m\ (1-r)^{n-m}$.

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 $P[e \mid h \cdot b \cdot c]$ of possible outcomes of the experimental arrangement.

Likelihoods that arise from explicit statistical claims – either within the hypotheses being tested, or from statistical background claims that tie the hypotheses to the evidence – are sometimes called *direct inference likelihoods*. Such likelihoods are completely objective. So it seems reasonable to suppose that all confirmation functions should agree on their values, just as all confirmation functions agree on likelihoods when evidence is logically entailed. Direct inference likelihoods are *logical* in an extended, non-deductive sense. Some logicians have attempted to spell out the logic of *direct inferences* in terms of the logical form of the sentences involved.[17] If that can be made to work in a Bayesian context, we'll need to supplement the axioms for probabilistic confirmation (in section 1) with additional axioms that capture the logic of the direct inference likelihoods. But regardless of whether that project can be made to succeed, it

[17] Attempts to do so in a Bayesian context have not been wholly satisfactory thus far, but research continues. For an illuminating discussion of the logic of direct inference and the difficulties involved in providing a formal account, see the series of papers (Levi, 1977), (Kyburg, 1978) and (Levi, 1978).

seems reasonable to take likelihoods that derive from explicit statistical claims to have 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 we may need to draw on a range of different confirmation functions, P_α, P_β ,…, P_γ, etc., to represent the different confirmation functions employed by various members of a scientific community (due to the different values they assign to prior probabilities), 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. 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 whereas α takes competing theory h_2 to probabilistically imply that *e* is quite unlikely, his colleague β reads h_2 as saying that *e* is very likely. So, for α the outcome *e* supplies strong support for h_1 over h_2, because $P_\alpha[e \mid h_1{\cdot}b{\cdot}c^n] >> P_\alpha[e \mid h_2{\cdot}b{\cdot}c]$. But his colleague β takes outcome *e* to show just the opposite – that h_2 is strongly supported over h_1 – because $P_\beta[e \mid h_1{\cdot}b{\cdot}c] << P_\beta[e \mid h_2{\cdot}b{\cdot}c]$. 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. The empirical testability of its hypotheses and theories would be completely undermined. Under such circumstances, although each scientist employs the same *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 very different theory than *h* as understood by β.[18] Thus, the empirical objectivity of the sciences requires that experts should be in close agreement about the values of the likelihoods.

For now we will suppose that the likelihoods have objective or intersubjectively agreed values, common to all agents in a scientific community, and shared by all confirmation functions they employ. 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

---

[18] This idea should not be confused with the view called *logical positivism*. *Positivism* holds that if two theories assign the same likelihood values to all possible evidence claims, then they are essentially the same theory, though they may be couched in different words. The slogan is: *same likelihoods* implies *same theory*. The point being defended here, however, is not the *positivistic* claim, but its inverse, which should be much less controversial: *different likelihoods* implies *different theories*. That is, given that all of the relevant background and auxiliaries are made explicit (represented in '*b*'), if two scientists disagree significantly about the likelihoods of important evidence claims on a given hypothesis, they must understand the empirical content of that hypothesis quite differently. To that extent, though they may be using the same sentences – the same syntactic expressions – they use them to express empirically distinct hypotheses.

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 well taken. Later (in section 6) we will see how to relax the supposition that likelihood values agree precisely. But for now, the main ideas behind probabilistic confirmation theory 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 moderately disagree about their values.

An adequate treatment of the likelihoods requires 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, …, c_n, describe the conditions under which a sequence of experiments or observations are conducted. And let the corresponding outcomes of these conditions be represented by the respective sentences $e_1, e_2,…,e_n$. 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 the 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 is equal to the product of the likelihoods of the individual outcomes: $P[e^n \mid h_i{\cdot}b{\cdot}c^n] = P[e_1 \mid h_i{\cdot}b{\cdot}c_1] \cdot\ldots\cdot P[e_n \mid h_i{\cdot}b{\cdot}c_n]$ – that is, the individual bits of evidence are *probabilistically independent given the hypothesis*. I'll discuss this kind of probabilistic independence in more detail later. For present purposes we'll not assume it.

3.2 Posterior Probabilities and Prior Probabilities

In a probabilistic confirmation theory the degree to which a hypothesis is supported on the evidence is represented by the *posterior probability* of the hypothesis, $P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n]$. The posterior probability represents the net plausibility of the hypothesis. It depends on the combined influence of evidence and non-evidential plausibility considerations. The likelihoods are the means through which evidence makes its contribution. But another factor, the *prior probability* of the hypothesis (on background b), $P_\alpha[h_i \mid b]$, represents the weight of all non-evidential plausibility considerations on which posterior probabilities (i.e. *posterior plausibilities*) may depend. Bayes' Theorem shows how the posterior probabilities of hypotheses depend 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 earlier. What the physician and patient want to know is the posterior probability $P_\alpha[h \mid b{\cdot}c{\cdot}e]$ of the hypothesis h that the patient has HIV, given the evidence of the positive test, $c{\cdot}e$, and given the error rates of the test, described by 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 posterior probability depends on how plausible it is that the patient has HIV before taking the test results into account, $P_\alpha[h \mid b]$. In the context of medical diagnosis this prior probability is sometimes called the *base rate*. It is the plausibility 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 may be explicitly stated in the background, *b*. To see its importance, consider the following numerical results (which may be calculated using Bayes' Theorem, as presented in the next subsection). 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 plausibility 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 shouldn't worry the patient too much. The positive 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 expensive test with better error characteristics.)

In more theoretical disciplines prior probabilities may represent assessments 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 the evaluation of hypotheses to be highly problematic. Bayesian confirmation theorists counter that such assessments often play an important role in the sciences, especially when there is insufficient evidence to distinguish among some of the alternatives. And they point out that 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 kinds of plausibility considerations brought to bear in assessing the various interpretations of quantum theory (e.g., those related to the measurement problem). Many of these considerations go to the heart of conceptual issues that were central to the development of the theory. Indeed, many of these issues were first explored by those 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 the relative plausibility of alternative views, especially when distinguishing evidence has yet to be found, or is far from definitive.

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 with regard to the implausibility of some "logically possible" alternatives. This seems to be the primary epistemic role of the thought experiment. It is arguably a virtue of probabilistic confirmation theory that it provides a place for such assessments to figure into the net evaluation of hypotheses. Prior probabilities are subjective in the sense that agents may continue to disagree on the relative merits of plausibility arguments – and so disagree on the prior plausibilities of various hypotheses. But assessments of priors are far from being *mere subjective whims*. Moreover, it can be shown that when sufficient empirical evidence becomes available, such plausibility assessments may be "washed out" or overridden by the evidence. We'll see how this works in the next subsection.

Our discussion of the nature of prior probabilities isn't over yet. We will return to it a bit later. But let's now see precisely how the logic of confirmation is supposed to work – how the likelihoods combine with prior probabilities to yield posterior probabilities for hypotheses.

3.3 Bayes' Theorem and the Logic of Hypothesis Evaluation

Any *probabilistic logic of confirmation* that draws on the usual axioms of probability theory to represent how evidence supports hypotheses must be a *Bayesian inductive logic* in the broad sense. For, Bayes' Theorem is just a simple theorem of probability theory. Its importance derives from the way it shows how evidence, through the likelihoods, combines with prior probabilities of hypotheses to produce assessments of their posterior probabilities.

We will examine several forms of Bayes' Theorem, each derivable solely from the axioms of probability theory. The simplest is this:

Bayes' Theorem: Simple Form:

$$(8) \qquad P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n] \; = \; \frac{P[e^n \mid h_i{\cdot}b{\cdot}c^n] \cdot P_\alpha[h_i \mid b] \cdot P_\alpha[c^n \mid h_i{\cdot}b]}{P_\alpha[e^n \mid b{\cdot}c^n] \cdot P_\alpha[c^n \mid b]}$$

$$= \frac{P[e^n \mid h_i{\cdot}b{\cdot}c^n] \cdot P_\alpha[h_i \mid b]}{P_\alpha[e^n \mid b{\cdot}c^n]} \quad \text{if } P_\alpha[c^n \mid h_i{\cdot}b] = P_\alpha[c^n \mid b].$$

This equation expresses the posterior probability of h_i, $P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n]$, in terms of the *likelihood* of the evidence on the hypothesis (together with background and observation conditions), $P[e^n \mid h_i{\cdot}b{\cdot}c^n]$, the *prior probability* of the hypothesis (given background conditions), $P_\alpha[h_i \mid b]$, and the *simple probability* of the evidence (given background and observation conditions), $P_\alpha[e^n \mid b{\cdot}c^n]$. This latter probability is sometimes called the *expectedness of the evidence*. This version of Bayes' Theorem also includes a term, ($P_\alpha[c^n \mid h_i{\cdot}b]$ / $P_\alpha[c^n \mid b]$), that represents the ratio of the *likelihood of the experimental conditions* on the hypothesis and background, to the *"likelihood" of the experimental conditions* on the background alone. Bayes' Theorem is usually expressed in a way that suppresses this factor, perhaps by building conditions c^n into the background b. However, if c^n is built into b, then technically b must change as new evidence is accumulated. It is better to make the factor explicit, and see how to deal with it logically. Arguably the term ($P_\alpha[c^n \mid h_i{\cdot}b]$ / $P_\alpha[c^n \mid b]$) should be 1, or nearly 1, since the truth of the hypothesis at issue should not significantly affect how likely it is that the experimental conditions themselves are satisfied. If various alternative hypotheses assign significantly different likelihoods to the experimental conditions themselves, then such conditions should more properly be included as part of the evidential outcomes e^n.

Both the *prior probability* of the hypothesis and the *expectedness* tend to be "subjective". That is, various agents from the same scientific community may legitimately disagree on what values these factors should take. Bayesian logicians usually accept the subjectivity of the prior probabilities for hypotheses, but they find the subjectivity of the *expectedness* more troubling. (How likely is the evidence supposed to be, given only the background – how are we supposed to assess this value?) However, this problem is easily finessed.

One way to circumvent the subjective *expectedness* of the evidence is to consider a ratio form of

Bayes' Theorem, a form that compares hypotheses one pair at a time:

Bayes' Theorem: Ratio Form:

$$(9) \quad \frac{P_\alpha[h_j \mid b \cdot c^n \cdot e^n]}{P_\alpha[h_i \mid b \cdot c^n \cdot e^n]} = \frac{P[e^n \mid h_j \cdot b \cdot c^n]}{P[e^n \mid h_i \cdot b \cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]} \cdot \frac{P_\alpha[c^n \mid h_j \cdot b]}{P_\alpha[c^n \mid h_i \cdot b]}$$

$$= \frac{P[e^n \mid h_j \cdot b \cdot c^n]}{P[e^n \mid h_i \cdot b \cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]} \quad \text{if } P_\alpha[c^n \mid h_j \cdot b] = P_\alpha[c^n \mid h_i \cdot b].$$

The second line follows if c^n is no more likely on $h_i \cdot b$ than on $h_j \cdot b$ – i.e., if neither hypothesis makes the occurrence of experimental or observation conditions more likely than the other.[19]

This ratio form of Bayes' Theorem expresses how much more plausible one hypothesis is than another, based on the evidence and on their non-evidential plausibilities. Notice that the only subjective element affecting the ratio of posterior probabilities is the ratio of prior probabilities. We see from this equation that the *likelihood ratios* carry the full import of the evidence. The evidence influences the evaluation of hypotheses in no other way.

Let's consider a simple example of how this form of Bayes' Theorem applies. Suppose you possess a warped coin and want to determine its propensity for *heads*. You may compare two hypotheses, $h_{[q]}$ and $h_{[r]}$, that propose the propensity for *heads* is q and r, respectively. Let c^n report that the coin is tossed n times in the usual way, and let e^n report a total m *heads*. Equation (9) then yields:

$$\frac{P_\alpha[h_{[q]} \mid b \cdot c^n \cdot e^n]}{P_\alpha[h_{[r]} \mid b \cdot c^n \cdot e^n]} = \frac{q^{\mathrm{m}}\,(1-q)^{n-m}}{r^{\mathrm{m}}\,(1-r)^{n-m}} \cdot \frac{P_\alpha[h_{[q]} \mid b]}{P_\alpha[h_{[r]} \mid b]}$$

When, for instance, the coin is tossed $n = 100$ times and comes up *heads* $m = 72$ times, the evidence for hypothesis $h_{[1/2]}$ as compared to $h_{[3/4]}$ is given by the likelihood ratio $[(1/2)^{72}(1/2)^{28}]/[(3/4)^{72}(1/4)^{28}] = .000056269$. So, even if prior to taking account of the evidence, one considers it 100 times more plausible that the coin is fair than that it is warped towards heads with propensity 3/4 – i.e., even if $P_\alpha[h_{[1/2]} \mid b] \,/\, P_\alpha[h_{[3/4]} \mid b] = 100$ – the evidence provided by these tosses makes the posterior plausibility that the coin is fair only about 56/10,000$^{\text{th}}$ as plausible as the hypothesis that it is warped towards heads with propensity 3/4 – i.e., $P_\alpha[h_{[1/2]} \mid b \cdot c^n \cdot e^n] \,/\, P_\alpha[h_{[3/4]} \mid b \cdot c^n \cdot e^n] = .0056269$. Thus, such evidence *strongly refutes* the

[19] This assumption may be substantially relaxed without affecting the analysis below; we might instead only suppose that the ratios $P_\alpha[c^n \mid h_j \cdot b]/P_\alpha[c^n \mid h_i \cdot b]$ are bounded so as not to get exceptionally far from 1. If *that* supposition were to fail, then the mere occurrence of the experimental conditions would count as very strong evidence for or against hypotheses, which is a highly implausible effect. Our analysis could include such bounded condition-ratios, but this would only add inessential complexity.

"fairness hypothesis" relative to the "3/4-*heads*-propensity hypothesis", provided the assessment of prior plausibilities doesn't make the latter hypothesis *too extremely implausible* to begin with. Notice, however, that *strong refutation* is not *absolute refutation*. Additional evidence could reverse the trend against the fairness hypothesis.

This example employs repetitions of the same kind of experiment – repeated tosses of a coin. But the point holds more generally. If, as the evidence increases, the *likelihood ratios* $P[e^n \mid h_j{\cdot}b{\cdot}c^n] / P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ approach 0, then the Ratio Form of Bayes' Theorem (Equation 9) shows that the posterior probability of h_j must approach 0 as well. The evidence comes to strongly refute h_j with little regard for its prior plausibility value. Indeed, Bayesian induction turns out to be a version of *eliminative induction*, and Equation 9 begins to illustrate this. For, suppose that h_i is the true hypothesis, and consider what happens to *each* of its false competitors, h_j. If enough evidence becomes available to drive each of the likelihood ratios $P[e^n \mid h_j{\cdot}b{\cdot}c^n] / P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ toward 0 (as n increases), then Equation 9 says that each false h_j will become effectively refuted – each of their posterior probabilities approaches 0. As a result, the posterior probability of h_i must approach 1. The next two equations make this clear.

If we sum the ratio versions of Bayes' Theorem in Equation 9 over all alternatives to hypothesis h_i (including the catch-all h_K, if there is one), we get the Odds Form of Bayes' Theorem. The *odds against A* given *B* is defined as $\Omega_\alpha[{\sim}A \mid B] = P_\alpha[{\sim}A \mid B] / P_\alpha[A \mid B]$. So, we have:

Bayes' Theorem: Odds Form

$$(10)\quad \Omega_\alpha[{\sim}h_i \mid b{\cdot}c^n{\cdot}e^n] \;=\; \sum_{j\neq i} \frac{P[e^n \mid h_j{\cdot}b{\cdot}c^n]}{P[e^n \mid h_i{\cdot}b{\cdot}c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]} \;+\; \frac{P_\alpha[e^n \mid h_K{\cdot}b{\cdot}c^n]}{P[e^n \mid h_i{\cdot}b{\cdot}c^n]} \cdot \frac{P_\alpha[h_K \mid b]}{P_\alpha[h_i \mid b]}\,.$$

If the catch-all alternative isn't needed, just drop the expression after the '+' sign (i.e. let $P_\alpha[h_K \mid b] = 0$). We represent the term for the catch-all hypothesis (if needed) separately because the likelihood of evidence relative to the catch-all hypothesis will not generally enjoy the same kind of objectivity as the likelihoods for *specific* hypotheses. We leave the subscript α on the likelihood for the catch-all to indicate this lack of objectivity.

The catch-all hypothesis may lack objective likelihoods, but the influence of the catch-all term in Bayes' theorem diminishes as additional *specific* hypotheses become articulated. That is, when a new hypothesis h_{u+1} is formulated and made explicit, the old catch-all h_K is replaced by a new catch-all, h_{K*}, of form $({\sim}h_1{\cdot}\ldots{\cdot}{\sim}h_u{\cdot}{\sim}h_{u+1})$. The prior probability for the new catch-all hypothesis comes from diminishing the prior of the old catch-all: $P_\alpha[h_{K*} \mid b] = P_\alpha[h_K \mid b] - P_\alpha[h_{u+1} \mid b]$. Thus, the influence of the catch-all term diminishes towards 0 as new alternative hypotheses are made explicit and "peeled off" of previous catch-all terms.[20]

[20] For example, when a new disease is discovered, a new hypothesis h_{u+1} about that disease being a possible cause of patients' symptoms is made explicit. The old catch-all was, "the symptoms are caused by some unknown disease – some disease other than $h_1,\ldots, h_u$". The new catch-all hypothesis states that "the symptoms are caused by one of the remaining unknown diseases – some disease other than $h_1,\ldots, h_u, h_{u+1}$". Then, $P_\alpha[h_K \mid b] = P_\alpha[{\sim}h_1{\cdot}\ldots{\cdot}{\sim}h_u \mid b] =$

If increasing evidence drives the likelihood ratios comparing h_i with each competitor towards 0, then the odds against h_i, $\Omega_\alpha[{\sim}h_i \mid b{\cdot}c^n{\cdot}e^n]$, will approach 0 (provided that priors of catch-all terms, if needed, approach 0 too, as new alternative hypotheses are made explicit and peeled off). And, as $\Omega_\alpha[{\sim}h_i \mid b{\cdot}c^n{\cdot}e^n]$ approaches 0, the posterior probability of h_i goes to 1. The relationship between the odds against h_i and its posterior probability is this:

Bayes' Theorem: From Posterior Odds to Posterior Probability

(11) $P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n] \;=\; 1/(1 + \Omega_\alpha[{\sim}h_i \mid b{\cdot}c^n{\cdot}e^n])$.

A confirmation function provides a probabilistic index of the net support a hypothesis receives *all-things-considered*. It explicitly divides considerations that bear on the evaluation of hypotheses into two kinds of components: non-evidential comparative plausibility considerations, represented by ratios of prior probabilities; and evidential support that derives from what hypotheses imply about the evidence, represented by ratios of likelihoods. Equation 9 shows precisely how these two kinds of considerations are brought to bear. In particular, it makes clear that if the series of likelihood ratios $P[e^n \mid h_j{\cdot}b{\cdot}c^n] / P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ approaches 0 on increasing evidence (i.e., as n increases), the posterior probability of hypothesis h_j must approach 0 as well, with little regard for the value of its prior probability $P_\alpha[h_j \mid b]$. As this happens to each of h_i's competitors, Equations 10 and 11 show that the posterior probability of hypothesis h_i approaches 1, as evidence increases.

Is there any good reason for thinking that when h_i is true, the series of likelihood ratios $P[e^n \mid h_j{\cdot}b{\cdot}c^n] / P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ will indeed favor it by heading towards 0 for empirically distinct alternatives, as evidence accumulates? There is a result, a kind of *Bayesian Convergence Theorem*, that speaks directly to this issue. It says: given that h_i (together with $b{\cdot}c^n$) is true, it's *very likely* that these likelihood ratios will approach 0 as evidence accumulates. The theorem expresses this in terms of a likelihood (given $h_i{\cdot}b{\cdot}c^n$). It supposes nothing about the values of the prior probabilities. Let's call this result the *Likelihood Ratio Convergence Theorem*. When it applies, putting it together with Equation 9 shows that the posterior probability of false competitor h_j is *very likely* to approach 0 as evidence accumulates, with little regard for its prior probability. And as this happens to each of h_i's false competitors, Equations 10 and 11 show that the posterior probability of the true hypothesis, h_i, is *very likely* to approach 1.[21] Thus, Bayesian

---

$P_\alpha[{\sim}h_1{\cdot}\ldots{\cdot}{\sim}h_u{\cdot}(h_{u+1}\vee{\sim}h_{u+1}) \mid b] = P_\alpha[{\sim}h_1{\cdot}\ldots{\cdot}{\sim}h_u{\cdot}{\sim}h_{u+1} \mid b] + P_\alpha[h_{u+1} \mid \mathrm{b}] = P_\alpha[h_{K^*} \mid b] + P_\alpha[h_{u+1} \mid b]$.

[21] This claim depends, of course, on h_i being empirically distinct from each alternative h_j. I.e., there must be conditions c_k with possible outcomes o_{ku} on which the likelihoods differ: $P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] \neq P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k]$. Otherwise h_i and h_j are empirically equivalent, and no amount of evidence can distinguish one from the other. If the true hypothesis has empirically equivalent rivals, then convergence just implies that the odds against *the disjunction* of the true hypothesis with these rivals very probably goes to 0, and so the posterior probability of this *disjunction* goes to 1. Among empirically equivalent hypotheses the ratio of their posterior probabilities equals the ratio of their priors: $P_\alpha[h_j \mid b{\cdot}c^n{\cdot}e^n] / P_\alpha[h_i \mid b{\cdot}c^n{\cdot}e^n] \;=\; P_\alpha[h_j \mid b] / P_\alpha[h_i \mid b]$. So the true hypothesis will have a posterior probability near 1 (after evidence drives the posteriors of empirically distinct rivals near 0) *just in case* non-evidential considerations make its evidence-independent

confirmation is at bottom a version of *induction by elimination*, where the elimination of false alternatives comes by way of likelihood ratios approaching 0 as evidence accumulates. We will examine the *Likelihood Ratio Convergence Theorem* in detail in section 5.

3.4 Likelihood Ratios, Likelihoodism, and the Law of Likelihood

The versions of Bayes' Theorem expressed by Equations 9-11 show that for probabilistic confirmation, the influence of empirical evidence on posterior probabilities of hypotheses is completely captured by the ratios of likelihoods, $P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n]$. The evidence $(c^n \cdot e^n)$ influences the posterior probabilities in no other way.[22] So, the following "Law" is a consequence of the logic of confirmation functions.

> General Law of Likelihood:
> Given any pair of incompatible hypotheses h_i and h_j, whenever the likelihoods $P_\alpha[e^n \mid h_j \cdot b \cdot c^n]$ and $P_\alpha[e^n \mid h_i \cdot b \cdot c^n]$ are defined, the evidence $(c^n \cdot e^n)$ supports h_i over h_j, given b, *if and only if* $P_\alpha[e^n \mid h_i \cdot b \cdot c^n] > P_\alpha[e^n \mid h_j \cdot b \cdot c^n]$. The ratio of likelihoods $P_\alpha[e^n \mid h_i \cdot b \cdot c^n] / P_\alpha[e^n \mid h_j \cdot b \cdot c^n]$ measures the *strength of the evidence* for h_i over h_j given b.

Two features of this law require some explanation. As stated, the General Law of Likelihood does not presuppose that likelihoods of form $P_\alpha[e^n \mid h_j \cdot b \cdot c^n]$ and $P_\alpha[e^n \mid h_i \cdot b \cdot c^n]$ are always *defined*. This qualification is introduced to accommodate a conception of evidential support called *Likelihoodism*, which I'll say more about in a moment. Also, the likelihoods in the law are expressed with the subscript α attached to indicate that the law holds for each confirmation function P_α, even when the values of the likelihoods are not completely objective or agreed on by all agents in a given scientific community. These two features of the law are closely related.

Each probabilistic confirmation function satisfies the axioms of section 1. According to these axioms the conditional probability of one sentence on another is always defined. So, in the context of probabilistic confirmation theory, the likelihoods are always defined, and the qualifying clause about this in the General Law of Likelihood is automatically satisfied. Furthermore, it should be noted that for confirmation functions all of the versions of Bayes' theorem (Equations 8-11) continue to hold even when the likelihoods are not objective or intersubjectively agreed on. Although in many scientific contexts there is agreement on the values of likelihoods, whenever such agreement fails, the subscripts α, β, etc. must remain attached to the support function likelihoods to indicate this. Even so, the General Law of Likelihood continues to hold for probabilistic confirmation functions.

There is a view, or family of views, called *likelihoodism* that maintains that confirmation theory should only concern itself only with the impact of the evidence in support of one hypothesis over another, and only in cases where this evaluation involves ratios of *completely objective* likelihoods. When the likelihoods involved are objective, the ratios $P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n]$ provide a *pure, objective measure* of how strongly the evidence supports h_i as compared to h_j,

---

plausibility much higher than the sum of the plausibility ratings of empirically equivalent rivals.
[22] However, there may be other useful ways to measure evidential import. Fitelson (1999) provides a penetrating comparison of a number of such measures.

"untainted" by prior plausibility considerations. According to likelihoodists, only this *pure measure* of the impact of evidence is scientifically appropriate for the assessment of hypotheses.

More specifically, likelihoodists maintain that it is not appropriate for scientific inference to incorporate prior probabilities of hypotheses into the assessment of evidential support. They say it is not the place of those engaged in hypothesis testing (e.g., statisticians) to compute *recommended values* of posterior probabilities for the scientific community. When the results of experiments are made public, say in scientific journals, only objective likelihoods should be reported. The evaluation of the impact of the objective likelihoods on agents' posterior probabilities must depend on each agent's individual *subjective* prior probability, which represents plausibility considerations that have nothing to do with the evidence. So, posterior probabilities should be left to individuals to compute, if they wish to do so.

Furthermore, the conditional probabilities for most pairs of sentences fail to be objectively defined in a way that suits *likelihoodists*. So, for them, the general *logic of confirmation functions* (captured by the axioms of section 1) cannot represent an *objective logic* of evidential support for hypotheses. Thus, they eschew the logic of confirmation functions. Because of this *likelihoodists* do not have Bayes' theorem available to them, and so cannot derive the Law of Likelihood from it. Rather, they must state this Law as an *axiom* of their logic of evidential support, and it applies only when the likelihoods have *well-defined* objective values.

Likelihoodists tend to have a very strict conception of what it takes for likelihoods to be *well-defined*. They consider a likelihood to be well-defined only when it is what we referred to earlier as a *direct inference likelihood* – i.e., only when either, (1) the hypothesis (together with background and experimental conditions) logically entails the data, or (2) the hypothesis (together with background and conditions) logically entails an explicit *simple statistical hypothesis* that (together with experimental conditions) specifies precise probabilities for the each of the events that make up the evidence.

Likelihoodists make a point of contrasting *simple statistical hypotheses* with *composite statistical hypotheses*, which only entail vague, or imprecise, or *directional* claims about the statistical probabilities of evidential events. Whereas a *simple statistical hypothesis* might say, for example, "the chance of *heads* on tosses of the coin is precisely .65", a composite statistical hypothesis might say, "the chance of *heads* on tosses is either .65 or .75", or it may be a *directional hypothesis* that says, "the chance of *heads* on tosses is greater than .65". *Likelihoodists* maintain that *composite hypotheses* are not an appropriate basis for well-defined likelihoods because such hypotheses represent a kind of disjunction of simple statistical hypotheses, and so must depend on non-objective factors. The *direction hypothesis*, for instance, is essentially a disjunction of the various *simple statistical hypotheses* that assign specific values above .65 to the chances of heads on tosses. Likelihoods based on such hypotheses are not *appropriately objective* for likelihoodist because they must in effect depend on factors that represent the degree to which the *composite hypothesis* supports each of the *simple statistical hypotheses* that it encompasses; and *likelihoodists* consider such factors too subjective to be permitted a role in a logic that countenances only objective likelihoods.[23]

---

[23] To see the point more clearly, consider an example. To keep things simple, let's suppose

Taking all of this into account, the version of the Law of Likelihood appropriate to *likelihoodists* may be stated as follows.

> Special Law of Likelihood:
> Given a pair of incompatible hypotheses h_i and h_j that imply simple statistical models regarding outcomes e^n given $(b \cdot c^n)$, the likelihoods $P[e^n \mid h_j \cdot b \cdot c^n]$ and $P[e^n \mid h_i \cdot b \cdot c^n]$ are well defined. For such likelihoods, the evidence $(c^n \cdot e^n)$ supports h_i over h_j, given b, *if and only if* $P[e^n \mid h_i \cdot b \cdot c^n] > P[e^n \mid h_j \cdot b \cdot c^n]$; the ratio of likelihoods $P[e^n \mid h_i \cdot b \cdot c^n] / P[e^n \mid h_j \cdot b \cdot c^n]$ measures the *strength of the evidence* for h_i over h_j given b.

Notice that when either version of the Law of Likelihood holds, the absolute size of a likelihood is irrelevant to the strength of the evidence. All that matters is the relative size of the likelihoods – i.e., the size of their ratio. Here is a way to see the point. Let c_1 and c_2 be the conditions for two different experiments having outcomes e_1 and e_2, respectively. Suppose that e_1 is 1000 times more likely on h_i (given $b \cdot c_1$) than is e_2 on h_i (given $b \cdot c_2$); and suppose that e_1 is also 1000 times more likely on h_j (given $b \cdot c_1$) than is e_2 on h_j (given $b \cdot c_2$) – i.e., suppose that $P_\alpha[e_1 \mid h_i \cdot b \cdot c_1] = 1000 \cdot P_\alpha[e_2 \mid h_i \cdot b \cdot c_1]$, and $P_\alpha[e_1 \mid h_j \cdot b \cdot c_1] = 1000 \cdot P_\alpha[e_2 \mid h_j \cdot b \cdot c_2]$. Which piece of evidence, $(c_1 \cdot e_1)$ or $(c_2 \cdot e_2)$, is stronger evidence with regard to the comparison of h_i to h_j? The Law of Likelihood implies both are equally strong. All that matters evidentially are the ratios of the likelihoods, and they are the same: $P_\alpha[e_1 \mid h_i \cdot b \cdot c_1] / P_\alpha[e_1 \mid h_j \cdot b \cdot c_1] = P_\alpha[e_2 \mid h_i \cdot b \cdot c_2] / P_\alpha[e_2 \mid h_j \cdot b \cdot c_2]$. Thus, the General Law of Likelihood implies the following principle.

> General Likelihood Principle:
> Suppose two different experiments or observations (or two sequences of them) c_1 and c_2 produce outcomes e_1 and e_2, respectively. Let $\{ h_1, h_2, \ldots\}$ be any set of alternative hypotheses. If there is a constant K such that for each hypothesis h_j from the set, $P_\alpha[e_1 \mid h_j \cdot b \cdot c_1] = K \cdot P_\alpha[e_2 \mid h_j \cdot b \cdot c_2]$, then the *evidential import* of $(c_1 \cdot e_1)$ for distinguishing among hypotheses in the set (given b) is precisely the same as the *evidential import* of $(c_2 \cdot e_2)$.

Similarly, the Special Law of Likelihood implies a corresponding Special Likelihood Principle that applies only to hypotheses that express simple statistical models.[24]

---

background b says that the chances of *heads* for tosses of this coin one of the 101 fractions $m/100$ for natural numbers m. Let c say that the coin is tossed in the usual random way; let e say that the coin comes up heads; and for each $r = m/100$, let $h_{[r]}$ be the *simple statistical hypothesis* asserting that the chance of heads on each toss of this coin is r. Now consider the *composite statistical hypothesis* $h_{[>.65]}$, which asserts that the chance of heads on each toss is greater than .65. From the axioms of probability theory we derive the following relationship: $P_\alpha[e \mid h_{[>.65]} \cdot b \cdot c] = \sum_{m=66}^{100} P[e \mid h_{[m/100]} \cdot b \cdot c] \cdot P_\alpha[h_{[m/100]} \mid h_{[>.65]} \cdot b \cdot c]$. The issue for the *likelihoodist* is that although the likelihoods of form $P[e \mid h_{[r]} \cdot b \cdot c]$ are completely objective, the values of the terms of form $P_\alpha[h_{[r]} \mid h_{[>.65]} \cdot b]$ are not objectively specified by the composite hypothesis $h_{[>.65]}$ (together with b). The value of the likelihood $P_\alpha[e \mid h_{[>.65]} \cdot b]$ depends essentially these non-objective factors. So it fails to possess the kind of objectivity that *likelihoodists* require.

24 The Law of Likelihood and the Likelihood Principle have been formulated in somewhat different ways by various logicians and statisticians. R.A. Fisher (1922) argued for the

Bayesians agree with likelihoodists that likelihood ratios completely characterize the impact of the evidence in support of one hypothesis over another. So they buy the Law of Likelihood and the Likelihood Principle. Indeed, for Bayesians these follow for theorems about the likelihoods. Furthermore, Bayesian confirmationists may agree that it's important to keep likelihoods separate from other factors, such as prior probabilities, in scientific reports about the evidence. However, Bayesians go further than likelihoodists in that they find a legitimate role for prior plausibility assessments to play in the full evaluation of scientific hypotheses. They propose to measure the net impact of evidence together with such plausibility assessments in terms of a probabilistic measure of confirmation, the posterior probabilities of hypotheses.

Throughout the remainder of this article we will not generally assume that likelihoods must be based on simple statistical hypotheses, as *likelihoodist* would have them. However, most of what will be said about likelihoods, including the convergence results in section 5, applied to the *likelihoodist* conception of likelihoods as well. We will, however, continue to suppose that likelihoods are *objective* in the sense that all members of the scientific community agree on their numerical values. In section 6 we will see how to relax even this supposition for contexts where it is unrealistic.

3.5 The Representation of Vague and/or Diverse Prior Probabilities

Given that a scientific community should largely agree on the values of the likelihoods, any significant disagreement regarding the posterior probabilities of hypotheses should derive from disagreements over prior probabilities. Their confirmational probabilities are *diverse*. Furthermore, individual agents may not be able to specify *precisely* how plausible they consider a hypothesis to be; so their prior probabilities for hypotheses may be vague. Both *diversity* due to disagreement among agents and *vagueness* for each individual agent can be represented by sets of confirmation functions, $\{P_\beta, P_\delta, \dots\}$, that agree on the likelihoods, but encompass a range of values for the prior plausibilities of hypotheses. *Diversity* and *vagueness* are different issues, but they may be represented in much the same way. We consider each in turn.

An individual's assessments of evidence-independent plausibilities of hypotheses will often be vague – not subject to the kind of precise quantitative treatment that a probabilistic logic of confirmation seems to require for prior probabilities. So it is sometimes objected that the kind of assessment of prior probabilities required to get the Bayesian algorithm going cannot be had in practice. Bayesian confirmation theory has a way of addressing this worry. An agent's vague assessments of prior plausibilities may be represented by a collection of probability functions, a *vagueness set*, which covers the range of plausibility values that the agent finds acceptable. Notice that if accumulating evidence drives the likelihood ratios to extremes, the range of functions in the agent's *vagueness set* will come to near agreement, near 0 or 1, on values for posterior probabilities of hypotheses. Thus, as evidence accumulates, the agent's vague initial

---

Likelihood Principle early in the 20th century, though he didn't call it that. One of the first places it is discussed under that name is in (Savage, et. al., 1962). The Law of Likelihood was first identified by that name in Hacking (1965), and has been invoked more recently by the *likelihoodist* statisticians A.F.W. Edwards (1972) and R. Royall (1997).

plausibility assessments may transform into quite sharp posterior probabilities that indicate the strong refutation or support of the various hypotheses. Intuitively this seems like quite a reasonable way for the logic to work.

The various agents in a community may widely disagree over the non-evidential plausibilities of hypotheses. Bayesian confirmation theory may represent this kind of *diversity* across the community of agents as a collection containing all functions in the agents' *vagueness sets*. Let's call such a collection a *diversity set*. So, while there may well be disagreement among agents regarding the prior plausibilities of hypotheses, and while individual agents may only have vague priors, the logic of probabilistic confirmation may readily represent this feature. Furthermore, if accumulating evidence drives the likelihood ratios to extremes, the range of functions in a *diversity set* will come to near agreement on sharp values, near 0 or 1, on the values for posterior probabilities of hypotheses. So, not only can such evidence *firm up* each agent's vague initial plausibilities, it also brings the whole community into agreement on the near refutation and on the strong support of the various alternative hypotheses.

Under what conditions might the likelihood ratios go to such extremes as evidence accumulates, effectively washing out vagueness and diversity? The *Likelihood Ratio Convergence Theorem* (discussed in detail in section 5) implies that if a true hypothesis disagrees with false alternatives on the likelihoods of possible outcomes for a long enough stream of experiments or observations, then that evidence stream will very probably produce *actual outcomes* that drive the likelihood ratios of false alternative as compared to the true hypothesis to approach 0. As this happens, almost any range of prior plausibility assessments will be driven to agreement on the posterior plausibilities for hypotheses. Thus, the accumulating evidence will very probably bring all confirmation functions in the *vagueness* and *diversity sets* for a community of agents to near agreement on posterior plausibility values – near 0 for the false competitors, and near 1 for the true hypothesis.

One more point about prior probabilities and Bayesian convergence is worth noting. Some subjectivist versions of Bayesianism seem to suggest that an agent's prior plausibility assessments for hypotheses should stay fixed once and for all, and that all plausibility updating should be brought about via the likelihoods in accord with Bayes' Theorem. Critics argue that this is unreasonable. The members of a scientific community may quite legitimately revise their prior plausibility assessments for hypotheses from time to time as they rethink plausibility arguments and bring new considerations to bear. This seems a natural part of the conceptual development of a science. It turns out that such reassessments of priors poses no difficulty for Bayesian confirmation theory. Reassessments may sometimes come about by the addition of explicit statements that supplement or modify the background information *b*. Or they may take the form of (non-Bayesian) transitions to new *vagueness sets* for individual agents and to new *diversity sets* for the community. The *logic* of Bayesian confirmation theory places no restrictions on how values for prior plausibility assessments might change. Provided that the series of reassessments of prior plausibilities doesn't push the prior of the true hypothesis ever nearer to zero, the *Likelihood Ratio Convergence Theorem* implies that the evidence will very probably bring the posterior probabilities of its empirically distinct rivals to approach 0 via decreasing likelihood ratios; and as this happens, the posterior probability of the true hypothesis will head towards 1.

4. What IS Confirmational Probability Anyway?

If confirmation functions aren't some sort of normative guide to what we may legitimately believe, then they are of no use to us at all, and probabilistic confirmation theory is a pointless enterprise. A confirmation function is presumably supposed to be a kind of truth-indicating index. When things are working right, a confirmation function should eventually indicate false hypotheses by sticking them with *confirmational probability numbers* near 0, and indicate true hypotheses by assigning them *confirmation numbers* that approach 1. Given this truth-indicating feature, it makes good epistemic sense to have *degree-of-confirmation* influence *belief-strength*. But exactly how is this supposed to work. How is the *degree-of-confirmation* for a hypothesis supposed to hook up with one's *level-of-confidence* or *degree-of-belief* in its truth or falsehood? Views about the nature of confirmation functions, about what they really are, should be sensitive to this question. A view that cannot reasonably tie confirmation to appropriate belief presents us with a useless contrivance.

4.1 Against Syntactic-Structural Versions of a Logical Interpretation: Grue-Hypotheses

Bayesian logicists like Keynes and Carnap maintained that confirmation is logical in the same way that deductive logic is logical, and that it should play an analogous role in informing belief. We need more details about how this is supposed to go. But we first need to get a better handle on what these logical confirmation functions are supposed to be like.

The leading idea is that posterior probabilities for hypotheses should be determined by logical structure alone. This idea of basing probabilities on syntactic structure is plausible for *likelihoods* that deductively or statistically relate hypotheses to the evidence. So, if logical form might also be made to determine the values of prior probabilities, then the logic of confirmation would be *fully formal* in the same way that deductive logical entailment is formal – i.e., it would be based only of the logical syntactic structure of the sentences involved. Such confirmation functions would then be *logical probabilities* in the sense that they would be completely specified by the syntactic structures of the sentences of the language. A sufficiently rigorous version of this approach might even specify a uniquely best way of assigning logically appropriate priors, resulting in a single uniquely best logical confirmation function. This confirmation function would be completely objective in that it would not be influenced by anyone's subjection opinions about which of the hypotheses are more or less plausible.

Both Keynes and Carnap tried to implement this kind of approach through syntactic versions of the principle of indifference. The idea is that hypotheses that share the same syntactic structure should be assigned the same prior probability values. Carnap showed how to carry out this project in detail, but only for extremely simple formal languages. Most logicians now take the project to have failed because of a fatal flaw with the whole idea that reasonable prior probabilities can be made to depend on logical form alone. Semantic content should matter.

Goodmanesque grue-predicates provide one way to illustrate this point.[25]

Call an object *grue* at a given time *just in case* "either the time is earlier than the first second of the year 2030 and the object is green, or the time not earlier than the first second of 2030 and the object is blue". Now the statement 'All emeralds are grue (at all times)' has the same syntactic structure as 'All emeralds are green (at all times)'. So, if syntactic structure determines priors, then these two hypotheses should have the same prior probabilities. Indeed, both should have prior probabilities approaching 0. For, there are an infinite number of competitors of these two hypotheses, each sharing the same syntactic structure: consider the hypotheses 'All emeralds are grue$_n$ (at all times)', where an object is grue$_n$ at a given time *just in case* either the time is earlier than the first second of the n^{th} day after January 1, 2030, and the object is green, or the time is not earlier than the first second of the n^{th} day after January 1, 2030, and the object is blue. A purely syntactic specification of the priors should assign all of these hypotheses the same prior probability. But these are mutually exclusive hypotheses; so their prior probabilities must sum to a value no greater than 1. The only way this can happen is for 'All emeralds are green' and each of its grue$_n$ competitors to have prior probability values equal to 0, or extremely close to it. In that case, the green hypothesis can never receive a posterior probability much above 0.

One might object that the predicate 'grue' is defined in terms of 'green', and so hides the extra syntactic complexity. But from a purely formal, syntactic point of view (which is all this view is entitled to), the predicates we happen to actually employ is simply an accident of the language we happen to speak. We could have spoken the grue-language, where 'grue' is the more primitive predicate, and where the predicate 'green' is defined and hides the extra syntactic complexity. Here's how to spell out this point in detail. Suppose the grue-language also contains a predicate 'bleen' which, translated to our usual language works like this: an object is bleen at a given time *just in case* "either the time is earlier than the first second of the year 2030 and the object is blue, or the time not earlier than the first second of 2030 and the object is green". Now, it is easy to show that from the perspective of the grue-language our predicate 'green' is *defined* as follows: "an object is green at a given time *just in case* "either the time is earlier than the first second of the year 2030 and the object is grue, or the time not earlier than the first second of 2030 and the object is bleen" (and 'blue' may be similarly defined). The point is, from a purely logical perspective, there is no reason to prefer one set of primitive predicates over another. Presumably part of the mission of confirmation theory is to discover what hypotheses, couched in terms of what primitive predicates best describes the world. The syntactic-structural view attempts to avoid prejudicing the confirmatory process by assigning prior probabilities in a "logically/syntactically unbiased" way. This example shows why that can't work. If you pick a preferred set of predicates, you build in a bias. If you don't pick a preferred set, then the all of the grue-like hypotheses must be given equal footing to the green hypothesis. But then all prior probabilities must be 0, or so close to 0 that no significant amount of confirmation can occur.

And how should such syntactically specified confirmation functions inform belief? That is, even if some version of the syntactic-structural approach could be made to work, its advocates still owe us an account of how, and why, such confirmation functions should inform our belief-

---

[25] Goodman (1955) introduced predicates of the following sort as a challenge to inductive logic. However, the details of my example and the use to which I'll put it differs from Goodman's.

strengths for various hypotheses. For instance, in cases where the evidence is not yet sufficient to strongly favor a specific hypothesis over an alternative (so that the likelihood ratio is near 1, why should our *belief-strength* (or *level of confidence*) be governed solely by the *syntactic structure* of these hypotheses, rather than by their *semantic meanings* together with whatever plausibility considerations make the most sense to the scientific community? The defenders of the syntactic-structural view owe us credible reasons for conforming belief to their confirmation functions.

4.2 Against the Subjective Belief-Function Interpretation: the Problem of Old Evidence

The *subjectivist* or *personalist* Bayesian view solves the problem of how confirmation is supposed to influence belief in the most direct way possible. It says that the agent's confirmation function P_α should just *be* his belief function, Bel_α, which is a probability function that measures how confident the agent is (or should be) in the truth of various statements. Belief is, of course, dynamic. We learn new truths, including evidence claims. On the subjectivist account, upon learning new evidence e, an agent α is supposed to update his belief-strengths via Bayesian conditionalization: for all sentences S (including the hypotheses h_j), $Bel_{\alpha\text{-}new}[S] = Bel_{\alpha\text{-}old}[S \mid e]$. This is where Bayes' Theorem comes in. In particular, when sentence S is a hypothesis h_i, we have (combining Equations 10 and 11, and suppressing '*c*' and '*b*', as subjectivists usually do):

$$Bel_{\alpha\text{-}new}[h_i] \;=\; Bel_{\alpha\text{-}old}[h_i \mid e] = 1 \,/\, (1 + \textstyle\sum_{j \neq i} \dfrac{Bel_{\alpha\text{-}old}[e \mid h_j]}{Bel_{\alpha\text{-}old}[e \mid h_i]} \cdot \dfrac{Bel_{\alpha\text{-}old}[h_j]}{Bel_{\alpha\text{-}old}\,[h_i]}\;)\,,$$

where the catch-all hypothesis, if needed, is included among the hypotheses h_j.

So Bayes' Theorem governs how belief-strengths are updated on new evidence.

Formally this account works just fine. However there are reasons for thinking that *confirmation functions* must be distinct from subjectivist or personalist *degree-of-belief functions*. One such problem is the *problem of old evidence.*[26] To understand the problem we need to first consider more carefully what *belief functions* are supposed to represent.

Belief functions are supposed to provide an idealized model of belief strengths for agents. They extend the notion of ideally consistent belief to a probabilistic notion of ideally coherent belief strengths. I see no problem with this kind of idealization. It is supposed to supply a normative guide for real decision making. An agent is supposed to make decisions based on her belief-strengths about the state of the world, her belief strengths about possible consequences of her actions, and her assessment of the desirability (or *utility*) of these consequences. But the very role that *belief functions* are supposed to play in decision making makes them ill-suited to hypothesis confirmation, where the *likelihoods* are often supposed to be objective, or at least possess inter-subjectively agreed values that represent the empirical import of hypotheses. That is, for the purposes of decision making, degree-of-belief functions *should* represent the agent's belief strengths *based on everything she presently knows*. But then, the degree-of-belief

---

[26] Glymour (1980) first raised this problem. Eells (1985) extends the problem. For a more extensive version of the following treatment see (Hawthorne 2005).

likelihoods must represent how strongly the agent would believe the evidence if a hypothesis h_i were added to *everything else she presently knows*. This makes them quite different than confirmation function likelihoods, which represent what the hypothesis (together with explicit background and experimental conditions) *says* or *implies* about the evidence. In particular, *degree-of-belief* likelihoods are saddled with a version of the *problem of old evidence* – a problem not shared by confirmation function likelihoods.

Here is the problem. An evidence statement e may be well-known far in advance of the time when we first attempt to account for it with some new hypothesis or theory. For example, the rate of advance in Mercury's perihelion was known long before Einstein developed the theory of General Relativity, and figured out how the theory could account for that phenomenon. If the agent is already certain of an evidence statement e before testing a theory with it, then her *belief-function* likelihoods for e must have value 1 on every hypothesis. That is, if $Bel_{\alpha\text{-}old}$ is her *belief function* and she already knows that e, then $Bel_{\alpha\text{-}old}[e] = 1$. It then follows from the axioms of probability theory that $Bel_{\alpha\text{-}old}[e \mid h_i] = 1$ as well, regardless of what h_i says – indeed, even if h_i implies that e is quite unlikely.

And the problem goes even deeper. It not only applies to evidence that the agent *knows with certainty*. It turns out that almost anything the agent learns that can change how strongly she believes e will also influence the value of her *belief-function likelihood* for e, because a belief function $Bel_{\alpha\text{-}old}[e \mid h_i]$ represents the agent's belief strength *given everything she knows*.

To see the difficulty with less-than-certain evidence, consider the following example (where I'll continue to suppress the 'b' and 'c' terms.) A physician intends use a treadmill test to find evidence about whether her patient has heart disease, h. She knows from medical studies that there is a 10% false negative rate for this test; so her belief-strength for a negative result, e, given heart disease is present, h, is $Bel_{\alpha\text{-}old}[e \mid h] = .10$. Now, her nurse is very professional and is usually unaffected by patients' test results. So, if asked, the physician would say her belief strength that her nurse will "feel devastated", d, if the test is positive (i.e. if $\sim e$) is around $Bel_{\alpha\text{-}old}[d \mid \sim e] = .05$. Let us suppose, as seems reasonable, that this belief-strength is independent of whether h is in fact true – i.e. $Bel_{\alpha\text{-}old}[d \mid \sim e \cdot h] = Bel_{\alpha\text{-}old}[d \mid \sim e]$. The nurse then says to the physician, in a completely convincing way, "he is such a nice guy – if *his* test comes out positive, I'll be devastated." The physician's new belief function likelihood for a false negative must then become $Bel_{\alpha\text{-}new}[e \mid h] = Bel_{\alpha\text{-}old}[e \mid h \cdot (\sim e \supset d)] = .69$.[27] Now, if a negative test result comes back from the lab, which likelihood is the physician supposed to use in her evaluation of the patient's prospects for having heart disease, her present personal belief-function likelihood, $Bel_{\alpha\text{-}new}[e \mid h] = .69$, or the "real" false-negative rate likelihood, $P[e \mid h] = .10$?

The main point is that even the most trivial knowledge of conditional (or disjunctive) claims involving e may completely upset the objective values for likelihoods for an agent's belief function. And an agent will almost always have some such trivial knowledge. E.g., the physician

[27] Since $Bel_{\alpha\text{-}old}[e \mid h \cdot (\sim e \supset d)] = Bel_{\alpha\text{-}old}[\sim e \supset d \mid h \cdot e] \cdot Bel_{\alpha\text{-}old}[e \mid h] / (Bel_{\alpha\text{-}old}[\sim e \supset d \mid h \cdot e] \cdot Bel_{\alpha\text{-}old}[e \mid h] + Bel_{\alpha\text{-}old}[\sim e \supset d \mid h \cdot \sim e] \cdot Bel_{\alpha\text{-}old}[\sim e \mid h]) = Bel_{\alpha\text{-}old}[e \mid h] / (Bel_{\alpha\text{-}old}[e \mid h] + Bel_{\alpha\text{-}old}[d \mid \sim e \cdot h] \cdot Bel_{\alpha\text{-}old}[\sim e \mid h]) = .1/(.1 + (.05)(.9)) = .69$.

in the previous example may also learn that if the treadmill test is negative for heart disease, *then*, (1) the patient's worried mother will throw a party, (2) the patient's insurance company won't cover additional tests, (3) it will be the thirty-seventh negative treadmill test result she has received for a patient this year,…, etc. Updating on such conditionals can force physicians' *belief functions* to deviate widely from the evidentially objective, textbook values for likelihoods.

More generally, it can be shown that the incorporation into Bel_α of almost any kind of evidence for or against the truth of a prospective evidence claim *e* – even uncertain evidence for *e*, as may come through Jeffrey updating[28] – completely undermines the objective or inter-subjectively agreed likelihoods that a belief function might have otherwise expressed.[29] This should be no surprise. The agent's belief function likelihoods reflect her *total degree-of-belief* in *e*, based on *h* together with *everything else she knows* about *e*. So the agent's present belief function may capture appropriate, public likelihoods for *e* only if *e* is completely isolated from all of the agents other beliefs. And this will rarely be the case.

One Bayesian subjectivist response to this kind of problem is that the belief functions employed in scientific inferences should often be "counterfactual belief functions", which represent what the agent *would believe* if *e* were subtracted (in some suitable way) from everything else she knows (see, e.g. Howson & Urbach, 1993). However, our example shows that merely subtracting *e* won't do. One must also subtract any conditional (or disjunctive) statements containing *e*. And one must subtract any uncertain evidence for or against *e* as well. So the counterfactual belief function idea needs a lot of working out if it is to rescue the idea that subjectivist Bayesian belief functions can provide a viable account of the likelihoods employed by the sciences.

There is important work for the *degree-of-belief* notion to do as part of our best formal account of belief and decision. But degree-of-confirmation functions, associated with objective or public likelihoods, do different work. It seems that confirmation functions should help guide changes in belief, but not as part of the belief function itself. Taking probabilistic confirmation functions to *be* degree-of-belief functions, even counterfactual ones, forces the *degree-of-belief* conception into a mold that doesn't suit it given the other work it does. Better to keep these two notions distinct, and bridge them with an account of how *degree-of-confirmation* should inform *degree-of-belief*.

4.3 How Confirmational Support should influence Belief-Strength: the *Truth-Index* Interpretation

Rather than ask what a confirmation function *is*, perhaps it's more fruitful to ask what a confirmation function is supposed to do. That is, I want to suggest a kind of *functionalist* view of the nature of confirmation functions. You might call this the *they-are-what-they-do* interpretation. But what is a confirmation function designed to do? What is it's function?

As I see it, a confirmation function is supposed to be a kind of *truth-indicating index*. And it can be expected to perform successfully in this role when things are working right. That is, when things are working right a confirmation function will eventually indicate the *falsehood* of false hypotheses by sticking them with *confirmational probability numbers* near 0, and it will come to

---

[28] See Jeffrey (1965, 1987, 1992).
[29] See (Hawthorne 2005) for more details.

indicate the *truth* of true hypotheses by assigning them *confirmation numbers* that approach 1. But what does it take for "things to work right"? Although it's not completely trivial, it's also not as daunting as one might think.

If, among the alternative hypotheses proposed to account for a given subject-matter, we are fortunate enough to think up a hypothesis that happens to in fact be true, and if we find ways to empirically test it against rivals, then all that's needed for success is persistence and not too much bad luck with how the evidence actually turns out. For, according to the Likelihood Ratio Convergence Theorem (section 5), the true hypothesis itself *says*, via its likelihoods, that a long enough (but finite) stream of observations or experiments is very likely to produce outcomes that will drive the likelihood ratios of empirically distinct false competitors to approach 0. And as this happens, the confirmation index of these competitors also approaches 0, and the confirmation index of the true hypothesis (or at least of its disjunction with empirically equivalent rivals) will approach 1.

However, one must be careful about how one reads this result. The result does not imply that whatever hypothesis has index near 1 at a given moment is likely to be the true one. The convergence theorem doesn't say that. Rather, the result suggests the pragmatic strategy of continually testing hypotheses, and taking whichever of them has an index near 1 (if there is one) as the *best current candidate* for being true. The convergence theorem implies that maintaining this strategy of continual testing is very likely to eventually promote the true hypothesis (or its disjunction with empirically indistinguishable rivals) to the status of *best current candidate*, and maintain it there. Thus, this strategy is very likely to eventually produce the truth for us. However, the theorem doesn't imply that we'll ever be in a position to justly be certain that our *best current candidate* is the true alternative. Thus, this eliminative strategy promises to work *only if* we continue to look for rivals and continue to test the best alternative candidates against them. This eliminative strategy shouldn't seem novel or surprising. It's merely a rigorously justified version of scientific common sense.

When the empirical evidence is meager or unable to distinguish between a pair of hypotheses, the confirmation index must rest on whatever our most probative non-evidential considerations may be able to tell us. We may often have good reasons besides the evidence to strongly reject some logically possible alternatives as *just too implausible*, or at least as much less plausible than some *better conceived* candidates. In fact, in determining what hypotheses to believe, we always do bring such considerations to bear, at least implicitly. For, given any hypothesis, logicians can always cook up numerous alternatives that agree with it on all the evidence thus far. Any reasonable scientist will reject (most of) these inventions immediately, because they look *ad hoc*, contrived, or "just foolish". Such reasons for rejection appeal to neither purely logical characteristics of these hypotheses, nor to evidential considerations. All such reasons ultimately rest on plausibility consideration (at least implicitly) that are not part of the evidence itself. This is not to say that such considerations are purely conceptual. They may involve some broadly empirical components as well. I call these considerations "non-evidential" because they are not directly represented by the likelihoods. Perhaps some kinds of broadly empirical considerations cannot be fully captured by statements describing observation conditions c and their evidential outcomes e, and so cannot be captured by likelihoods. On a Bayesian account of confirmation, whatever cannot be represented by the likelihoods may only be introduced via the "prior" probabilities. They are the conduit through which considerations not expressed by likelihoods may be brought to bear in the net evaluation of scientific hypotheses.

All this suggests that the normative connection between confirmation and belief should go something like this:

> *The Belief-Confirmation Alignment Condition*:
> Each agent should bring her belief-strengths for hypotheses into alignment with their degrees-of-confirmation on all of the relevant evidence she is aware of – where the confirmation function she employs draws on prior probabilities that represent her best estimates of the comparative plausibilities of alternative hypotheses based on all relevant non-evidential considerations of which she is aware. That is, if P_α is her confirmation function (as just described), and she is certain of background and auxiliaries b and evidence $c^n \cdot e^n$, and this is the totality of her evidence that is relevant to h_i, then her belief strength Bel_α should be (or become) $Bel_\alpha[h_i] = P_\alpha[h_i \mid b \cdot c^n \cdot e^n]$.
>
> Furthermore, if (as is often the case) she has partial or uncertain evidence that's relevant to h_i, then her belief strength should be the weighted sum of the degrees-of-confirmation of the hypothesis on *each possible evidence sequence* $c^n \cdot e^n$ (among those that represent the ways that her uncertain evidence could turn out to be true), weighted by her belief-strengths for each of those possible evidence sequences (and similarly for possible alternative auxiliaries b, if uncertain), as follows:[30]
>
> $$Bel_\alpha[h_i] = \sum_{\{b \cdot c^n \cdot e^n\}} P_\alpha[h_i \mid b \cdot c^n \cdot e^n] \cdot Bel_\alpha[b \cdot c^n \cdot e^n].$$

The *Alignment Condition* may be difficult for real agents to follow precisely. It should be a normative guide for real agents – much as Bayesian decision theory is supposed to be a normative guide. The *Alignment Condition* merely recommends that a real agent's confidence in scientific hypotheses should conform to the level indicated by her confirmation function, moderated by how confident she is in the truth of the evidence claims. It shouldn't be overly difficult for real agents to approximately align belief to confirmation in this way. Furthermore, supposing (as argued earlier) that probabilistic confirmation functions should not *just be* belief functions, the *Alignment Condition* shows how probabilistic confirmation can plausibly be made to mesh with the usual Bayesian account of belief and decision. What recommends the *Alignment Condition* as a norm is the fact that, if the agent comes to strongly believe true evidence statements, then *alignment* takes advantage of *Likelihood Ratio Convergence* to very probably bring the agent to strongly doubt false hypotheses and strongly believe true ones.[31] What better recommendation for the formation of belief-strengths about scientific hypotheses could one reasonably expect to have?

---

[30] If the agent is *certain* of some particular bit of evidence $c_k \cdot e_k$ in the evidence stream, her belief function will assign belief-strength 0 to each possible evidence sequence $c^n \cdot e^n$ that fails to contain $c_k \cdot e_k$ – i.e., $Bel_\alpha[b \cdot c^n \cdot e^n] = 0$ for all such $c^n \cdot e^n$.

[31] See (Hawthorne 2005) for more about the alignment of belief with confirmational support.

5. The Likelihood Ratio Convergence Theorem

The *Likelihood Ratio Convergence Theorem* shows that when h_i is true and h_j is empirically distinct from h_i, it's *very likely* that a sequence of outcomes e^n will occur that yields a sequence of likelihood ratios $P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n]$ that approach 0 as the evidence accumulates (i.e., as n increases). The theorem places an explicit lower bound on the *rate of probable convergence*. That is, it puts bound that approaches 1 *on the likelihood that*, when h_i is true, some stream of outcomes will occur that yields a likelihood ratio within any chosen small distance of 0. When that happens, it counts heavily against the truth of alternative h_j.

This convergence theorem draws only on likelihoods. Neither the statement of the theorem nor its proof employs prior probabilities of any kind. Likelihoodists and Bayesian confirmationists agree that when the ratios $P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n]$ approach 0 for increasing n, the evidence goes strongly against h_j as compared to h_i. So even likelihoodists, who eschew the use of prior probabilities, may embrace this result.

For Bayesians, the *Likelihood Ratio Convergence Theorem* has the additional implication that the posterior probabilities of empirically distinct false competitors of a true hypothesis are very likely to converge to 0. That's because, whenever the ratios $P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n]$ approach 0 for increasing n, the Ratio Form of Bayes' Theorem, Equation 9, says that the posterior probability of h_j will also approach 0. The values of prior probabilities only accelerate or retard this process of convergence. This also implies that all confirmation functions in a collection that constitutes a *vagueness set* (that represent the range of vagueness in an agent's assessments of the prior plausibilities of hypotheses) will very likely come to agreement, all coming to agree that the posterior probability of false alternatives approach 0.[32] And as the posterior probabilities of false competitors approach 0, the posterior probability of the true hypothesis heads towards 1.

The *Likelihood Ratio Convergence Theorem* avoids or overcomes the usual objections raised against Bayesian convergence results:

- The theorem does not employ *second-order probabilities* – it doesn't relay on assessing the probability of a probability. The theorem only concerns the probability of particular disjunctive sentences representing possible sequences of outcomes.

- The theorem does not rely on countable additivity (to which some commentators have objected with regard to other convergence results).

- The theorem does not require evidence to consist of sequences of outcomes that, according to the hypotheses, are *identically distributed* (like repeated tosses of a die). The version of the theorem I'll present does, however, suppose that the evidential outcomes in the sequence of experiments or observations are probabilistically independent given each hypothesis – or at least can be grouped into probabilistically independent clusters.

---

[32] And the same goes for *diversity sets*, which represent the range of plausibility assessments among members of a scientific community.

However, there is a variant of this theorem (not presented here) that applies even when probabilistic independence fails. Nevertheless, I will argue that the sort of probabilistic independence the present version of the theorem draws on is almost always present in scientific contexts.

- The rate of likely convergence of the likelihood ratios is explicitly calculable from the likelihoods specified by individual hypotheses. (The theorem is a kind of *Weak Law of Large Numbers*.) So the theorem doesn't warrant the often heard objection that Bayesian convergence results only apply to the infinite long run (when we'll all be long dead).

- The values of the prior probabilities need not be permanently "locked in" for the theorem to apply. The probable convergence occurs even if agents reassess the non-evidential plausibilities of hypotheses from time to time, and assign new priors accordingly.

This last point needs some explanation. It is sometimes objected that Bayesian convergence results only work when prior probabilities are held fixed – that the theorems fall through if an agent is permitted to change her evidence-independent assessments of prior plausibilities from time to time. Critics point out that real agents may quite legitimately change their assessments of the evidence-independent plausibilities of hypotheses, perhaps due to newly developed plausibility arguments, or due to the reassessment of old ones. A Bayesian confirmation theory has to represent such reassessments as non-Bayesian shifts from one confirmation function (or from one *vagueness* or *diversity set* of confirmation functions) to another. But, critics object, Bayesian convergence theorems always assume that the only dynamic element in the confirmational process is due to the addition of new evidence, which brings to bear the associated likelihoods, and results in the updating of posterior probabilities via Bayes' Theorem. So, it looks like Bayesian confirmation is severely handicapped as an account of scientific hypothesis evaluation.

However, the *Likelihood Ratio Convergence Theorem* is not subject to this objection. It applies even if agents revise their evidence-independent priors from time to time. For, the theorem itself only involves the values of likelihoods. Thus, provided that the reassessments of prior plausibilities don't push the prior plausibility of the true hypothesis down towards 0 *too rapidly*, the theorem continues to show that posterior probabilities of the empirically distinct false competitors of a true hypothesis will *very probably* approach 0 as evidence increases.[33]

---

[33] That is, for each confirmation function P_α, the posterior $P_\alpha[h_j \mid b\cdot c^n\cdot e^n]$ must go to 0 if the ratio $P_\alpha[h_j \mid b\cdot c^n\cdot e^n] / P_\alpha[h_i \mid b\cdot c^n\cdot e^n]$ goes to 0; and that will occur if the likelihood ratios $\mathrm{P}[e^n \mid h_j\cdot b\cdot c^n] / P[e^n \mid h_i\cdot b\cdot c^n]$ approach 0 and the prior $P_\alpha[h_i \mid b]$ is greater than 0. The Likelihood Ratio Convergence Theorem will show that when $h_i\cdot b$ is true, it is very likely that the evidence will indeed be such as to drive the likelihood ratios as near to 0 as you please (given a long enough evidence stream). As that happens, the only way a Bayesian agent can avoid having his confirmation function yield *posterior probabilities* for h_i that approach 0 (as n gets large) is to continually switch among confirmation functions (moving from P_α to P_β to ... to P_γ to …) in a way that revises the pre-evidential prior probability of h_i downward towards 0. And even then, he can only avoid having the posterior probability for alternative h_j approach 0 for each *current confirmation function* by continually switching to new confirmation functions at a rate that keeps

I raise these point in advance so that the reader can be on the look-out, to see that the theorem really does avoid these challenges. The version of the theorem I'll present has its roots in L.J. Savage's (1954) convergence theorem, but generalizes that result considerably. In particular, Savage's theorem is subject to several of the above mentioned objection, while the present theorem overcomes them. We now turn to the details.

5.1 The Space of Possible Outcomes of Experimental and Observational Conditions

To spell out the details of the *Likelihood Ratio Convergence Theorem* we'll need a few additional notational conventions and definitions. Here they are.

For a sequence of n experiments or observations c^n, consider the set of those possible sequences of outcomes that would result in likelihood ratios for h_j over h_i that are less than some chosen small number $\varepsilon > 0$. This set is represented by the following expression:

$$\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\}$$

One may choose any small value of ε that seems interesting, and then form the corresponding set. Placing the disjunction symbol '$\vee$' in front of this expression yields an expression

$$\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\},$$

that represents the disjunction of all outcome sequences in this set. So, the expression '$\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\}$' just represents a particular sentence that says, in effect, "one of those sequences of outcomes from the first n experiments or observations will occur that makes the likelihood ratio for h_j over h_i less than ε."

The *Likelihood Ratio Convergence Theorem* says, for any specific ε you choose, that the likelihood of a disjunctive sentence of this sort, given that '$h_i \cdot b \cdot c^n$' is true, i.e.,

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n],$$

must have a value of at least $1-(\psi/n)$, for some explicitly calculable term ψ. And clearly this

---

the new priors for h_i diminishing towards 0 at least as quickly as the likelihood ratios diminish towards 0 (with increasing n). To see this, suppose, *to the contrary*, that $P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n]$ approaches 0 *faster* than sequence $P_\gamma[h_i \mid b]$, for changing P_γ and increasing n – i.e. *approaches 0 faster* in the sense that $(P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n]) \,/\, P_\gamma[h_i \mid b]$ goes to 0, for changing P_γ and increasing n. Then, we'd have $(P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n]) \,/\, P_\gamma[h_i \mid b] > (P[e^n \mid h_j \cdot b \cdot c^n] \,/\, P[e^n \mid h_i \cdot b \cdot c^n]) \cdot (P_\gamma[h_j \mid b] \,/\, P_\gamma[h_i \mid b]) = P_\gamma[h_j \mid b \cdot c^n \cdot e^n] \,/\, P_\gamma[h_i \mid b \cdot c^n \cdot e^n]$. So, $P_\gamma[h_j \mid b \cdot c^n \cdot e^n] \,/\, P_\gamma[h_i \mid b \cdot c^n \cdot e^n]$ must still go to 0, for changing P_γ and increasing n; and thus, so must $P_\gamma[h_j \mid b \cdot c^n \cdot e^n]$.

For a nice presentation of the most prominent Bayesian convergence results and a discussion of their weaknesses see (Earman, 1992, Ch. 6). Earman was not aware of the convergence theorems under consideration here.

lower bound, $1-(\psi/n)$, will approach 0 as n increases. Thus, the true hypothesis h_i implies that as the amount of evidence, n, increases, it is highly likely (as close to 1 as you please) that one of the outcome sequences e^n will occur that yields a likelihood ratio $P[e^n \mid h_j{\cdot}b{\cdot}c^n] / P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ less than ε, for any value of ε you may choose. As this happens, the posterior probability of h_i's false competitor, h_j, must approach 0, as required by the Ratio Form of Bayes' Theorem.

The term ψ in the theorem depends on a measure of the empirical distinctness of the two hypotheses involved on the proposed sequence of experiments and observations. To specify this measure we need to contemplate not only the actual outcomes, but the collection of alternative possible outcomes of each experiment or observation. So, consider some sequence of experimental or observation conditions described by sentences $c_1,c_2,\ldots,c_n$. Corresponding to each condition c_k there will be some range of possible alternative outcomes; let $O_k = \{o_{k1},o_{k2},\ldots,o_{kw}\}$ be a set of statements describing the alternative possible outcomes for condition c_k. (The number of alternative outcomes will usually differ for distinct experiments $c_1,\ldots,c_n$; so, the value of w depends on c_k). For each hypothesis h_j, the alternative outcomes of c_k in O_k are mutually exclusive and exhaustive – that is, we have:

$$P[o_{ku}{\cdot}o_{kv} \mid h_j{\cdot}b{\cdot}c_k] = 0 \text{ and } \sum_{u=1}^{w} P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] = 1.$$

Expressions like 'e_k' represent possible outcomes of c_k – i.e., 'e_k' ranges over the members of O_k. As before, 'c^n' denotes the conjunction of the first n test conditions, $(c_1{\cdot}c_2{\cdot}\ldots{\cdot}c_n)$, and '$e^n$' represents possible sequences of corresponding outcomes, $(e_1{\cdot}e_2{\cdot}\ldots{\cdot}e_n)$. We'll take '$E^n$' to represent the set of all possible outcome sequences for c^n. So, for each hypothesis h_j (including h_i), we have $\sum_{e^n \in E^n} P[e^n \mid h_j{\cdot}b{\cdot}c^n] = 1$. There are no substantive assumptions in any of this – only notational conventions.

5.2 About Probabilistic Independence

In almost all scientific contexts the outcomes in a series of experiments or observations are *probabilistically independent* of one another relative to each hypothesis under consideration. We may divide the kind of independence involved into two types.

Definition: Independent Evidence Conditions:

(1) A sequence of outcomes e^k is *condition-independent* of a condition for an additional experiment or observation c_{k+1}, given $h{\cdot}b$ and its own conditions c^k, *if and only if* $P[e^k \mid h{\cdot}b{\cdot}c^k{\cdot}c_{k+1}] = P[e^k \mid h{\cdot}b{\cdot}c^k]$.

(2) An individual outcome e_k is *result-independent* of a sequence of other observations and their outcomes $(c^{k-1}{\cdot}e^{k-1})$, given $h{\cdot}b$ and its own condition c_k, *if and only if* $P[e_k \mid h{\cdot}b{\cdot}c_k{\cdot}(c^{k-1}{\cdot}e^{k-1})] = P[e_k \mid h{\cdot}b{\cdot}c_k]$.

When these two conditions hold, the likelihood for a sequence of experiments or observations may be decomposed into the product of the likelihoods for individual experiments or observations. To see how the two *independence conditions* affect the decomposition, first consider the following formula, which holds even if neither *independence condition* is satisfied:

(12) $P[e^n \mid h_j \cdot b \cdot c^n] = \prod_{k=1}^{n} P[e_k \mid h_j \cdot b \cdot c^n \cdot e^{k-1}]$.

When *condition-independence* holds, the likelihood of the whole evidence stream parses into a product of likelihoods that *probabilistically depend* on only past observation conditions and their outcomes. They do not depend on the conditions for other experiments whose outcomes are not yet specified. Here is the formula:

(13) $P[e^n \mid h_j \cdot b \cdot c^n] = \prod_{k=1}^{n} P[e_k \mid h_j \cdot b \cdot c_k \cdot (c^{k-1} \cdot e^{k-1})]$.

Finally, whenever both independence conditions are satisfied we obtain the following relationship between the likelihood of the evidence stream and the likelihoods of individual experiments or observations:[34]

(14) $P[e^n \mid h_j \cdot b \cdot c^n] = \prod_{k=1}^{n} P[e_k \mid h_j \cdot b \cdot c_k]$..

In almost all scientific contexts both clauses of the *Independent Evidence Condition* will be satisfied. To see this, let us consider each independence condition more carefully.

Condition-independence says that the mere addition of a new observation condition c_{k+1}, *without specifying one of its outcomes*, does not alter the likelihood of the outcomes e^k of other experiments c^k. To appreciate the significance of this condition, imagine how the world would be if it were violated. Suppose hypothesis h_j is some statistical theory, say, a quantum theory of superconductivity. The conditions expressed in c^k describe a number of experimental setups, perhaps conducted in numerous labs throughout the world, that test a variety of aspects of the theory (e.g., experiments that test electrical conductivity in different materials at a range of temperatures). An outcome sequence e^k describes the results of these experiments. The violation of *condition-independence* would mean that merely adding to $h_j \cdot b \cdot c^k$ a statement c_{k+1} describing how an additional experiment is constructed, but with no mention of its outcome, changes how likely the evidence sequence e^k is: i.e., $P[e^k \mid h \cdot b \cdot c^k \cdot c_{k+1}] \neq P[e^k \mid h \cdot b \cdot c^k]$. What $(h_j \cdot b)$ *says*, via likelihoods, about the outcomes e^k of experiments c^k differs as a result of merely supplying a description of another experimental arrangement, c_{k+1}. *Condition-independence*, when it holds, rules out such strange effects.

Result-independence says that the description of previous test conditions *together with their outcomes* is irrelevant to the likelihoods of outcomes for additional experiments. If this condition were widely violated, then in order to specify the most informed likelihoods for a given hypothesis one would need to include information about volumes of past observations and their outcomes. What a hypothesis says about future cases would depend on how past cases have gone. Such *dependence* had better not happen on a large scale. Otherwise, the hypothesis would be fairly useless, since its empirical import in each specific case would depend on taking into account volumes of past observational and experimental results. However, even if such dependencies occur, provided they are not too pervasive, *result-independence* can be

[34] For derivations of equations (13) and (14) see
(Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement3.html

accommodated rather easily by packaging each collection of *result-dependent* data together, treating it like a single extended experiment or observation. The *result-independence condition* will then be satisfied by letting each term 'c_k' in the statement of the independence condition represent a conjunction of test conditions for a collection of *result-dependent* tests, and by letting each term 'e_k' (and each term 'o_{ku}') stand for a conjunction of the corresponding *result-dependent* outcomes. Thus, by packaging *result-dependent* data together in this way, the *result-independence* condition is satisfied by those (conjunctive) statements that describe the separate, *result-independent* chunks.[35]

The version of the *Likelihood Ratio Convergence Theorem* I'll present depends on the usual axioms of probability theory together with the *Independent Evidence Conditions*. It depends on no other assumptions (except those explicitly stated in the antecedent of the theorem itself). Thus, from this point on, let's suppose that the following two assumptions holds.

> *Independent Evidence Assumptions*: For each hypothesis h and background b under consideration, let's assume that the experiments and observations can be packaged into condition statements, $c_1,\ldots, c_k, c_{k+1},\ldots$, and possible outcomes in a way that satisfies the following independence conditions:
>
> (1) Each sequence of possible outcomes e^k of a sequence of conditions c^k is *condition-independent* of additional conditions c_{k+1} – i.e., $P[e^k \mid h{\cdot}b{\cdot}c^k{\cdot}c_{k+1}] = P[e^k \mid h{\cdot}b{\cdot}c^k]$.
>
> (2) Each possible outcome e_k of condition c_k is *result-independent* of sequences of other observations and possible outcomes $(c^{k-1}{\cdot}e^{k-1})$—i.e., $P[e_k \mid h{\cdot}b{\cdot}c_k{\cdot}(c^{k-1}{\cdot}e^{k-1})] = P[e_k \mid h{\cdot}b{\cdot}c_k]$.

We now have all that is needed to begin to state the *Likelihood Ratio Convergence Theorem*. The convergence theorem comes in two parts. The first part applies to only those experiments or observations that have possible outcomes, according to h_i, that alternative h_j says are impossible. The second part of the theorem applies to all other experiments or observations.

5.3 Likelihood Ratio Convergence under Conditions where Falsifying Outcomes are Possible

The first part of the *Likelihood Ratio Convergence Theorem* applies whenever some of the

---

[35] In scientific contexts the most prominent kind of case where data may fail to be *result-independent* is where some quantity of past data helps tie down the numerical value of a parameter not completely specified by the hypothesis at issue, and where the value of this parameter influences the likelihoods for outcomes of lots of other experiments. Such hypotheses (with their free parameters) are effectively *disjunctions* of more specific hypotheses, where each distinct disjunct is a distinct version of the original hypothesis that has a specific value for the parameter filled in. Evidence that "fills in the value" for the parameter just amounts to evidence that refutes (via likelihood ratios) those more specific, filled-in hypotheses that possess incorrect parameter values. For any specific, filled-in hypotheses, the evidence that bears on whether it has the correct parameter value will be independent of other evidence that relies on the parameter value. So, relative to each of these more specific hypotheses, *result-independence* holds.

experiments or observations in sequence c^n have possible outcomes with non-0 likelihoods on hypothesis h_i, but 0 likelihoods on alternative h_j. Such outcomes are highly desirable. If they occur, the likelihood ratio comparing h_j to h_i will be 0, and h_j will be *falsified*. A *crucial experiment* is a special case of this, the case where, for at least one possible outcome o_{ku}, $P[o_{ku} \mid h_i \cdot b \cdot c_k] = 1$ and $P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0$. In the more general case h_i together with b says that one of the outcomes of c_k is at least minimally probable, whereas h_j says that outcome is impossible: $P[o_{ku} \mid h_i \cdot b \cdot c_k] > 0$ and $P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0$.

Likelihood Ratio Convergence Theorem 1: *The Falsification Theorem*:[36]
Suppose c^m, a subsequence of the whole evidence sequence c^n, consists of experiments or observations with the following property: there are outcomes o_{ku} of each c_k in c^m *deemed impossible* by $h_j \cdot b$ but *deemed possible* by $h_i \cdot b$ to at least some small degree δ. That is, suppose there is a $\delta > 0$ such that for each c_k in c^m, $P[\vee\{o_{ku} : P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0\} \mid h_i \cdot b \cdot c_k] \geq \delta$.
Then,

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] = 0\} \mid h_i \cdot b \cdot c^n] = P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] = 0\} \mid h_i \cdot b \cdot c^n]$$

$$\geq 1-(1-\delta)^m$$

which approaches 1 for large m.

In other words, suppose h_i says observation c_k has at least a small likelihood of producing one of the outcomes o_{ku} that h_j says is impossible – i.e., $P[\vee\{o_{ku}: P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0\} \mid h_i \cdot b \cdot c_k] \geq \delta > 0$. And suppose that some number m of experiments or observations are of this kind. If the number of such observations is large enough, and h_i (together with $b \cdot c^n$) is true, then it is highly likely that one of the outcomes held to be impossible by h_j will occur, and the likelihood ratio of h_j over h_i will then become 0. Bayes' Theorem then goes on to imply that when this happens, h_j is absolutely refuted – its posterior probability becomes 0.

The Falsification Theorem is very commonsensical. First, notice that when there is a *crucial experiment* in the evidence stream, the theorem is completely obvious. That is, suppose for the specific experiment c_k (in evidence stream c^n) there are two incompatible possible outcomes o_{kv} and o_{ku} such that $P[o_{kv} \mid h_j \cdot b \cdot c_k] = 1$ and $P[o_{ku} \mid h_i \cdot b \cdot c_k] = 1$. Then, clearly, $P[\vee\{o_{ku}: P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0\} \mid h_i \cdot b \cdot c_k] = 1$, since o_{ku} is "one of the o_{ku} such that $P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0$". So where there is a crucial experiment available, the theorem applies with $m = 1$ and $\delta = 1$.

The theorem is equally commonsensical when there is no crucial experiment. To see what it says in such cases, consider an example. Let h_i be some theory that implies a specific rate of proton decay, but a rate so low that there is only a very small probability that any particular proton will decay in a given year. Consider an alternative theory h_j that implies that protons *never* decay. If h_i is true, then for a persistent enough sequence of observations (i.e., if proper detectors can be built and billions of protons kept under observation for long enough), eventually a proton decay will almost surely be detected. When this happens, the likelihood ratio becomes 0. Thus, the

---

[36] For a proof see (Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement4.html

posterior probability of h_j becomes 0.

It may be instructive to plug some specific numbers into the formula given by the Falsification Theorem, to see what the convergence rate might look like. For example, the theorem tells us that if we compare any pair of hypotheses h_i and h_j on an evidence stream c^n that contains at least $m = 19$ observations or experiments, each having $\delta \geq .10$ for the likelihood of yielding a *falsifying outcome*, then the likelihood (on $h_i{\cdot}b{\cdot}c^n$) of obtaining an outcome sequence e^n that will yield a likelihood-ratio $P[e^n \mid h_j{\cdot}b{\cdot}c^n] \,/\, P[e^n \mid h_i{\cdot}b{\cdot}c^n] = 0$ must be least $1-(1-.1)^{19} = .865$.

A comment about the *need for*, and *usefulness of* such convergence theorems is in order, now that we've seen one. Given some specific pair of scientific hypotheses h_i and h_j, one may always directly compute the likelihood, given ($h_i{\cdot}b{\cdot}c^n$), that any specific sequence of experiments or observations c^n will result in one of the specific sequences of outcomes that yields low likelihood ratios. So, given a specific pair of hypotheses and a proposed sequence of experiments, we don't need a general *Convergence Theorem* to tell us the likelihood of obtaining refuting evidence. The specific hypotheses h_i and h_j tell us this *themselves*. Indeed, they tell us the likelihood of obtaining each specific outcome stream, including those that refute the competitor or produce a very small likelihood ratio for it. Furthermore, after we've actually performed an experiment and recorded its outcome, all that matters is the actual ratio of likelihoods for that outcome. Convergence theorems become moot.

The point of Likelihood Ratio Convergence Theorem (both the Falsification Theorem and the other part of the theorem that's still to come) is to assure us *in advance of the consideration of any specific pair of hypotheses* that if the possible evidence streams that test hypotheses have certain characteristics that reflect the empirical distinctness of the hypotheses, then it is highly likely that one of the sequences of outcomes will occur that results in a very small likelihood ratio. These theorems provide relatively loose, finite lower bounds on how quickly such convergence is likely to be. Thus, convergence theorems may assure us in advance of our using the logic of confirmation to test specific hypotheses, that this logic is likely to do what we want it to do – i.e., to result in the refutation of empirically distinct false alternatives to the true hypothesis, and to generate a high degree of positive confirmation for the true hypothesis.

5.4 Likelihood Ratio Convergence under Conditions where No Falsifying Outcomes are Possible

The Falsification Theorem shows what happens when the evidence stream includes possible outcomes that may *falsify* the alternative hypothesis. But what if no *possibly falsifying* outcomes are present? That is, what if hypothesis h_j only specifies various non-zero likelihoods for possible outcomes? Or what if h_j does specify 0 likelihoods for some outcomes, but only for those that h_i says are impossible? Such evidence streams are undoubtedly much more common in practice than those containing possibly falsifying outcomes. To cover evidence streams of this kind we first need to identify a useful way to measure the degree to which hypotheses are empirically distinguishable by such evidence.

Consider some particular sequence of outcomes e^n that results from observations c^n. The likelihood ratio $P[e^n \mid h_j{\cdot}b{\cdot}c^n] \,/\, P[e^n \mid h_i{\cdot}b{\cdot}c^n]$ measures the extent to which that outcome sequence distinguishes between h_i and h_j. But, as a measure of the power of evidence to distinguish among hypotheses, likelihood ratios themselves provide a rather lopsided scale, a scale that ranges from 0 to infinity with the midpoint, the point where e^n doesn't distinguish at all between h_i and h_j, at 1. So, rather than using raw likelihood ratios to measure the ability of e^n to distinguish between hypotheses, it proves more useful to employ a symmetric measure. The logarithm of the likelihood ratio provides just such a measure.

Definition: QI – *the Quality of the Information.*
For each experiment or observation c_k, define *the quality of the information* provided by possible outcome o_{ku} for distinguishing h_j from h_i, given b, as follows (where we take the log to be base 2):

$$\text{QI}[o_{ku} \mid h_i/h_j \mid b{\cdot}c_k] = \log[P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] \,/\, P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k]].$$

Similarly, define $\text{QI}[e^n \mid h_i/h_j \mid b{\cdot}c^n] = \log[P[e^n \mid h_i{\cdot}b{\cdot}c^n] \,/\, P[e^n \mid h_j{\cdot}b{\cdot}c^n]]$.

We measure the *Quality of the Information* an outcome would yield in distinguishing between two hypotheses as the base-2 logarithm of the likelihood ratio. This is clearly a measure of the outcome's evidential strength at distinguishing between the two hypotheses. By this measure, two hypotheses, h_i and h_j, assign the same likelihood value to a given outcome o_{ku} *just in case* $\text{QI}[o_{ku} \mid h_i/h_j \mid b{\cdot}c_k] = 0$. And whenever $P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] / P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] = 2^r$, $\text{QI}[o_{ku} \mid h_i/h_j \mid b{\cdot}c_k] = r$, for any real number r. So, QI measures information on a logarithmic scale that is symmetric about the natural no-information midpoint, 0. Positive information ($r > 0$) favors h_i over h_j and negative information ($r < 0$) favors h_j over h_i.

Given the *Independent Evidence Assumptions* it is easy to see that relative to each hypothesis (with background), $h_i{\cdot}b$ and $h_j{\cdot}b$, the QI for a sequence of outcomes is just the sum of the QIs of the individual outcomes in the sequence:

$$(15) \quad \text{QI}[e^n \mid h_i/h_j \mid b{\cdot}c^n] \;=\; \textstyle\sum_{k=1}^{n} \;\text{QI}[e_k \mid h_i/h_j \mid b{\cdot}c_k].$$

QI only measures the amount by which each specific outcome counts for or against the two hypotheses. But what we want to know is something about how the experiment or observation as a whole tends to produce distinguishing outcomes. The *expected value* of QI turns out to be very helpful in this regard. The *expected value* of a quantity is gotten by first multiplying each of its *possible values* by its probability of occurring, and then summing up these products. Thus, the *expected value* of QI is given by the following formula:

Definition: EQI – *the Expected Quality of the Information.*
Let's call h_j *outcome-compatible* with h_i on evidence stream c^k *just when* for each possible outcome sequence e^k of c^k, if $P[e^k \mid h_i{\cdot}b{\cdot}c^k] > 0$, then $P[e^k \mid h_j{\cdot}b{\cdot}c^k] > 0$. We also adopt the convention that if $P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] = 0$, then the term $\text{QI}[o_{ku} \mid h_i/h_j \mid b{\cdot}c_k] \cdot P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] = 0$, since the outcome o_{ku} has 0 probability of occurring given $h_i{\cdot}b{\cdot}c_k$.

For h_j *outcome-compatible* with h_i on c^k, define

$$\text{EQI}[c_k \mid h_i/h_j \mid b] = \textstyle\sum_u \text{QI}[o_{ku} \mid h_i/h_j \mid b{\cdot}c_k] \cdot P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k].$$

Also, define $\text{EQI}[c^n \mid h_i/h_j \mid b] = \sum_{e^n \in \text{E}^n} \text{QI}[e^n \mid h_i/h_j \mid b] \cdot P[e^n \mid h_i{\cdot}b{\cdot}c^n]$.

Notice that when h_j is not *outcome-compatible* with h_i on evidence stream c^m, *Likelihood Ratio Convergence Theorem 1*, the *Falsification Theorem* given in the previous section applies.

The EQI of an experiment or observation is the *Expected Quality of its Information* for distinguishing h_i from h_j when h_i is true. It is a measure of the expected evidential strength of the possible outcomes of an experiment or observation at distinguishing between the hypotheses. Whereas QI measures the ability of each particular outcome or sequence of outcomes to empirically distinguish hypotheses, EQI measures the tendency of experiments or observations to produce distinguishing outcomes. EQI tracks empirical distinctness in a very precise way, as we'll see in a moment.

The EQI for a sequence of observations c^n turns out to be just the sum of the EQIs of the individual observations c_k in the sequence:[37]

(16) $\text{EQI}[c^n \mid h_i/h_j \mid b] \;=\; \sum_{k=1}^{n} \text{EQI}[c_k \mid h_i/h_j \mid b]$

This suggests that it may be useful to average the values of the $\text{EQI}[c_k \mid h_i/h_j \mid b]$ over the number of observations n. We then obtain a measure of the *average expected quality of the information* from the experiments and observations that make up c^n.

Definition: <u>EQI</u> *– The Average Expected Quality of Information.*
The average expected quality of information, <u>EQI</u>, from c^n for distinguishing h_j from h_i, given $h_i{\cdot}b$, is defined as:

$$\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] \;=\; \text{EQI}[c^n \mid h_i/h_j \mid b] \,/\, n\,.$$

This definition together with equation (16) yields the following:

(17) $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] \;=\; (1/n) \cdot \sum_{k=1}^{n} \text{EQI}[c_k \mid h_i/h_j \mid b]$

It turns out that the value of $\text{EQI}[c_k \mid h_i/h_j \mid b]$ cannot be less than 0; and it will be greater than 0 just in case h_i is *empirically distinct* from h_j on at least one outcome o_{ku} – i.e., just in case for at least one o_{ku}, $P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] \neq P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k]$. The same goes for the average, $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b]$.

Theorem: Nonnegativity of EQI.[38]
$\text{EQI}[c_k \mid h_i/h_j \mid b] \geq 0$; and, $\text{EQI}[c_k \mid h_i/h_j \mid b] > 0$ *if and only if* for at least one of its possible outcomes o_{ku}, $P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] \neq P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k]$.
Also, $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] \geq 0$; and $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] > 0$ *if and only if* at least one experiment or

[37] For a derivation see (Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement5.html
[38] For proof see (Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement6.html

observation c_k has at least one possible outcome o_{ku} such that $P[o_{ku} \mid h_i \cdot b \cdot c_k] \neq P[o_{ku} \mid h_j \cdot b \cdot c_k]$.

In fact it can be shown that increasing the *fineness* of the partition of the outcome space $O_k = \{o_{k1},\ldots,o_{kv},\ldots,o_{kw}\}$ by breaking it up into more distinct outcomes (if it can be so divided) always results in a larger value for EQI, provided that at least some of the additional outcomes have distinct likelihood ratio values.[39] Thus, EQI tracks empirical distinctness in a very precise way. The importance of the *Non-negativity of* EQI result for the *Likelihood Ratio Convergence Theorem* will become apparent in a moment.

We are now in a position to state the second part of the *Likelihood Ratio Convergence Theorem*. It applies to all evidence streams that do not contain *possibly falsifying outcomes* for h_j when h_i holds – i.e., it applies to all evidence streams for which h_j is *outcome-compatible* with h_i on each c_k in the stream.

Likelihood Ratio Convergence Theorem 2: *The Non-Falsifying Refutation Theorem.*[40]
Let $\gamma > 0$ be any number smaller than $1/e^2$ ($\approx .135$; where this 'e' is the base of the natural logarithm). And suppose that for each possible outcome o_{ku} of each observation condition c_k in c^n, either $P[o_{ku} \mid h_i \cdot b \cdot c_k] = 0$ or $P[o_{ku} \mid h_j \cdot b \cdot c_k] / P[o_{ku} \mid h_i \cdot b \cdot c_k] \geq \gamma$. Choose any positive $\varepsilon < 1$, as near to 0 as you like, but large enough that (for the number of observations n being contemplated) the value of $\text{EQI}[c^n \mid h_i/h_j \mid b] > -(\log \varepsilon)/n$. Then

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n] \; > \; 1 \; - \; (1/n) \cdot \frac{(\log \gamma)^2}{(\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] + (\log \varepsilon)/n\,)^2}$$

which approaches 1 for large n when $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b]$ has a positive lower bound – i.e., when the sequence of observation c^n has an average expected quality of information (for empirically distinguishing h_j from h_i) that doesn't diminish towards 0 as the evidence sequence increases.

This theorem provides a very reasonable sufficient condition for the *likely refutation* of false alternatives via exceedingly small likelihood ratios. The condition under which this happens draws only on a characterization of the degree to which the hypotheses involved are empirically distinct from each other. The theorem says that when these conditions of empirical distinctness are met, hypothesis h_i (together with $b \cdot c^n$) provides a likelihood that is *at least* within $(1/n) \cdot (\log \gamma)^2 / (\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] + (\log \varepsilon)/n)^2$ of 1 that some outcome sequence e^n will occur that yields a likelihood ratio smaller than chosen ε. It turns out that in almost every case the actual likelihood of obtaining such evidence will be much closer to 1 than this factor indicates. Thus, this theorem provides a rather loose lower bound on the likelihood of obtaining small likelihood ratios. It shows that the larger the value of EQI for an evidence stream, the more likely it is that the stream will produce a sequence of outcomes that yield very small likelihood ratios. But even if EQI

---

[39] See (Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement6.html
[40] For a proof see (Hawthorne 2004) http://plato.stanford.edu/entries/logic-inductive/supplement7.html

remains quite small, a long enough stream, n, will almost surely do the trick.[41]

Notice that the antecedent condition of the theorem, that "either $P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] = 0$ or $P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] / P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] \geq \gamma$, for some $\gamma > 0$ but less than $1/e^2$ ($\approx .135$)", does not favor hypothesis h_i in any way. This condition only rules out the possibility that some outcomes might furnish *extremely strong* evidence *against* h_j relative to h_i. This condition is only needed because our measure of the evidential distinguishability of pairs of hypotheses, QI, blows up when the likelihood ratio $P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] / P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k]$ is extremely small. Furthermore, this condition is really no restriction at all on the application of the theorem to possible experiments or observations. If c_k has some possible outcome description o_{ku} that would make $P[o_{ku} \mid h_j{\cdot}b{\cdot}c_k] / P[o_{ku} \mid h_i{\cdot}b{\cdot}c_k] < \gamma$ (for a given small γ of interest), one may disjunctively lump o_{ku} together with some other outcome description o_{kv} for c_k. Then, the antecedent condition of the theorem will be satisfied, but with the sentence '$(o_{ku} \vee o_{kv})$' treated as a single outcome in the formula for EQI. It can be proved that the only effect of such "disjunctive lumping" is to make EQI a bit smaller than it would otherwise be. If such a "too refuting" outcome o_{ku} actually occurs when the evidence is collected, so much the better. We merely failed to take this possibility into account in computing our lower bound on the likelihood that refutation via likelihood ratios will occur.

The point of the two *Convergence Theorems* explored in this section is to assure us, in advance of the consideration of any specific pair of hypotheses, that if the possible evidence streams that test them have certain characteristics which reflect their evidential distinguishability, it is highly likely that outcomes yielding small likelihood ratios will result. These theorems provide finite lower bounds on how quickly convergence is likely to occur, bounds that show one need not wait for convergence through some infinitely long run. Indeed, for any evidence sequence in which the probability distributions are at all well behaved, the *actual likelihood* of obtaining outcomes that yield small likelihood ratio values will inevitably be *much higher* than the lower bounds given by Theorems 1 and 2.

In sum, according to Theorems 1 and 2, each hypothesis h_i *says*, via likelihoods, the following: "given enough observations, *I* am very likely to dominate my empirically distinct rivals in a contest of likelihood ratios." Even a sequence of observations with an extremely low *average expected quality of information* is very likely to do the job, provided that the sequence is long enough. Presumably, in *saying this*, the true hypothesis speaks truthfully, and its false competitors lie. Thus (by Equation 9), as evidence accumulates, the degree of confirmation for false hypotheses will very probably approach 0, which will indicate that they are probably false; and as this happens, (by Equations 10 and 11) the degree of confirmation of the true hypothesis

---

[41] It should now be clear why the boundedness of EQI above 0 is important. Convergence Theorem 2 applies only when $\underline{\text{EQI}}[c^n \mid h_i/h_j \mid b] > -(\log \varepsilon)/n$. But this requirement is not a strong assumption. For, the *Nonnegativity of* EQI *Theorem* shows that the empirical distinctness of two hypotheses on a single possible outcome suffices to make the average EQI positive for the whole sequence of experiments. So, given any small fraction $\varepsilon > 0$, the value of $-(\log \varepsilon)/n$ (which is greater than 0) will eventually become smaller than EQI, provided that the degree to which the hypotheses are empirical distinct for the various observations c_k does not on average degrade too much as the length n of the evidence stream increases. This seems a reasonable condition on the empirical distinctness of hypotheses.

will approach 1, indicating its probable truth.

6. When the Likelihoods are Vague and/or Diverse

Up to this point we've been supposing that likelihoods possess objective or agreed numerical values. Although this supposition is often satisfied in scientific contexts, there are important settings where it is unrealistic, where individuals are pretty vague about the numerical values of likelihoods for evidence that nevertheless seems to weigh strongly against one hypothesis and in support of another. So let's see how the supposition of precise, agreed values for likelihoods may be relaxed in a reasonable way.

Let's first consider an example where the likelihoods are vague for an important scientific evidence claim. Consider, the geological *drift hypothesis*, that the land masses of Africa and South America were once joined, then split and have drifted apart over the eons. Also consider the alternative *contractionist hypothesis*, that the continents have fixed positions acquired when the earth first formed, cooled and contracted into its present configuration. On each of these hypotheses, how likely is it that: (1) the shape of the east coast of South America should match the shape of the west coast of Africa as closely as it in fact does; (2) the geology of the two coasts should match up so well; (3) the plant and animal species on these distant continents should be as similar as they are. One may not be able to determine anything like precise numerical values for such likelihoods. But experts readily agree that each of these observations is *much more likely* on the *drift hypothesis* than on the *contractionist hypothesis*. Jointly these observations constitute very strong evidence in favor of the *drift hypothesis* over its *contraction* alternative. They do so because experts in the scientific community are in widespread agreement that the *ratio of the likelihoods* strongly favors *drift* over *contraction*.[42]

---

[42] Historically the case for continental drift is somewhat more complicated. Geologists tended to largely dismiss the evidence referred to above until the 1960s. Although this evidence seems to be quite strong, it was found unconvincing because it was not sufficiently strong to overcome certain non-evidential plausibility considerations that made the *drift hypothesis* seem extremely implausible – much less plausible that the more traditional *contraction* view. The chief problem was that there appeared to be no plausible mechanism by which *drift* might occur. It was argued, quite reasonably, that no known force or mechanism could push or pull the continents apart, and that the less dense continental material cannot possibly push through the denser material that makes up the ocean floor. These objections were eventually overcome when a plausible mechanism was articulated – i.e. that the continental crust floats atop molten material and moves apart as convection currents in the molten material carry it along. The case was pretty well clinched when evidence for this mechanism was found in the form of "spreading zones" containing alternating strips of magnetized material at regular distances from mid-ocean ridges. The magnetic alignments of materials in these strips corresponds closely to the magnetic alignments found in magnetic materials in dateable sedimentary layers at other locations on the earth. These magnetic alignments indicate time periods when the direction of earth's magnetic field has reversed. And this gave geologists a way of measuring the rate at which the sea floor might spread and the continents move apart. Although geologists may not be able to determine anything like precise values for the likelihoods of any of this evidence on each of the alternative

Recall now the reasons given earlier for why agreement or near agreement on values for likelihoods is highly desirable. To the extent that members of a scientific community disagree on the likelihoods, they disagree about the empirical content of their hypotheses, about what each hypothesis *says* about how the world is likely to be. And this can result in disagreement about which hypotheses are refuted or favored by a given stream of evidence. Similarly, to the extent that the values of likelihoods are vague for an individual agent, he or she may be unable to determine which of several hypotheses is refuted or favored by a given body of evidence.

Notice, however, that the values of individual likelihoods are not really crucial to the way evidence impacts hypotheses. Rather (as Equations 9-11 show), it is *ratios of likelihoods* that do the heavy lifting. So, even if two confirmation functions P_α and P_β disagree on the values of likelihoods, they may, nevertheless, largely agree on the refutation or support that accrues to various rival hypotheses when the following condition is satisfied:

Directional Agreement Condition:
The likelihood ratios due to each of a pair of confirmation functions P_α and P_β will be said to *agree in direction* (with respect to the possible outcomes of experiments or observations relevant to a pair of hypotheses) *just in case* each of the following conditions hold:

- for each possible outcome e_k of the experiments and observations c_k in the evidence stream, $P_\alpha[e_k \mid h_j{\cdot}b{\cdot}c_k] / P_\alpha[e_k \mid h_i{\cdot}b{\cdot}c_k] < 1$ just in case $P_\beta[e_k \mid h_j{\cdot}b{\cdot}c_k] / P_\beta[e_k \mid h_i{\cdot}b{\cdot}c_k] < 1$, and $P_\alpha[e_k \mid h_j{\cdot}b{\cdot}c_k] / P_\alpha[e_k \mid h_i{\cdot}b{\cdot}c_k] > 1$ just in case $P_\beta[e_k \mid h_j{\cdot}b{\cdot}c_k] / P_\beta[e_k \mid h_i{\cdot}b{\cdot}c_k] > 1$.
- each of these likelihood ratios is either close to 1 for neither confirmation function or for both functions.

When this condition holds, the evidence will support h_i over h_j according to P_α just in case it does so for P_β as well, although the strength of support may differ. In addition, although the *rate* at which the likelihood ratios increase or decrease as evidence accumulates may differ for these functions, the total impact of the cumulative evidence will ultimately affect the refutation and support of hypotheses in much the same way for each.

Thus, when *likelihoods* are vague or diverse, we may take the approach we employed for *vague* and *diverse* prior plausibility assessments. We may represent the vagueness in an agent's assessments of both prior plausibilities and likelihoods in terms of a *vagueness set* – a set of confirmation functions that covers the range of values that are acceptable to the agent. Similarly, we may extend the *diversity sets* for communities of agents to include confirmation functions for both the range of *likelihoods* and the range of prior plausibilities (from individual *vagueness sets*) that represent the considered views of the members of the relevant scientific community.

The *Likelihood Ratio Convergence Theorem* can still do its work in this context, provided the

---

hypotheses, the evidence is universally agreed to be *much* more likely on the *drift* hypothesis than on the *contractionist* alternative. And, with the emergence of the possibility of a plausible mechanism, the *drift* hypothesis no longer seems so overwhelmingly implausible *prior* to the evidence, either. Thus, the *value of a likelihood ratio* may be objective or public enough, even when precise values for likelihoods cannot be assessed.

Directional Agreement Condition is satisfied by all confirmation functions in these extended *vagueness* and *diversity sets*. The proof of the theorem doesn't depend on supposing that likelihoods are objective or have intersubjectively agreed values. It applies to each confirmation function P_α individually. The only difficulty that comes from applying the theorem to a range of confirmation functions that disagree on the likelihoods is that the specific outcome sequences that strongly favors h_i according to P_α may instead strongly favor h_j according to P_β. However, when the *Directional Agreement Condition* holds for these confirmation functions, that cannot happen. *Directional Agreement* means that the empirical import of hypotheses is represented by P_α and P_β as similar enough that each evidence sequence must favor the same hypotheses for them both. Thus, when the *Directional Agreement Condition* holds, if enough empirically distinguishing experiments or observations are forthcoming, all support functions in an *extended vagueness* or *diversity* set will very probably come to agree that the likelihood ratios for empirically distinct false competitors of a true hypothesis are extremely small. As that happens, the community comes to agree on the refutation of these competitors, and the true hypothesis rises to the top of the heap.[43]

What if the true hypothesis has empirically equivalent rivals? Then their posterior probabilities must rise as well. The Likelihood Ratio Convergence Theorem only assures us that the disjunction of the true hypothesis with its empirically equivalent rivals will be driven to 1 (as evidence lays low the empirically distinct rivals). The true hypothesis may itself approach 1 only if either it has no empirically equivalent rivals, or if whatever equivalent rivals it has are laid low as well by non-evidential plausibility considerations.

References

Carnap, R., 1950, *Logical Foundations of Probability*, Chicago: University of Chicago Press.

–––, 1952, *The Continuum of Inductive Methods*, Chicago: University of Chicago Press.

–––, 1971, "A Basic System of Inductive Logic, Part I", in R. Carnap and R. C. Jeffrey (eds.), *Studies in Inductive Logic and Probability, Vol. 1*, Berkeley: University of California Press.

–––, 1980, "A Basic System of Inductive Logic, Part II", in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability, Vol. 2*, Berkeley: University of California Press.

De Finetti, B., 1937. "La Prévision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincaré*, 7, 1-68; translated as "Foresight. Its Logical Laws, Its Subjective

[43] Even if there are a few directionally controversial likelihood ratios, where P_α says the ratio is somewhat greater than 1, while and P_β assigns a value somewhat less than 1, these may not greatly effect the trend of P_α and P_β towards agreement on the refutation and support of hypotheses *provided that* the controversial ratios are not so extreme as to overwhelm the stream of the other evidence for which the likelihood ratios do directionally agree.

Sources”, in H. E. Kyburg, Jr. and H. E. Smokler (eds.), 1980, *Studies in Subjective Probability*, Robert E. Krieger Publishing Company.

Duhem, P., 1906, *La theorie physique. Son objet et sa structure*, Paris: Chevalier et Riviere; translated by P.P. Wiener, 1954, *The Aim and Structure of Physical Theory*, Princeton, NJ: Princeton University Press.

Earman, J. 1992, *Bayes or Bust?*, Cambridge, MA: MIT Press.

Edwards, A.W.F. 1972, *Likelihood*, Cambridge: Cambridge University Press.

Eells, E., 1985, “Problems of Old Evidence”, *Pacific Philosophical Quarterly*, 66, 283-302.

Fisher, R.A., 1922, “On the Mathematical Foundations of Theoretical Statistics”, Philosophical Transactions of the Royal Society, series A, 309-368.

Fitelson, B., 1999, “The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity”, *Philosophy of Science* 66, S362-S378.

Goodman, N., 1955, "The New Riddle of Induction," Chapter 3 of *Fact, Fiction, and Forecast*, Cambridge, Mass.: Harvard University Press.

Glymour, C., 1980, *Theory and Evidence*, Princeton: Princeton University Press.

Hacking, I., 1965, *Logic of Statistical Inference*, Cambridge: Cambridge University Press.

Hájek, A., 2003a, “What Conditional Probability Could Not Be”, *Synthese* 137, 273–323.

—–, 2003b, “Interpretations of the Probability Calculus”, in the *Stanford Encyclopedia of Philosophy*, (Summer 2003 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2003/entries/probability-interpret/.

Hawthorne, J., 2004, "Inductive Logic", *The Stanford Encyclopedia of Philosophy (Winter 2006 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/win2006/entries/logic-inductive/.

—–, 2005, “Degree-of-Belief and Degree-of-Support: Why Bayesians Need Both Notions”, *Mind* 114, 277-320.

Howson, C., Urbach, P., 1993, *Scientific Reasoning: The Bayesian Approach*, 2nd edition, Chicago: Open Court.

Jaynes, E. T., 1968, “Prior Probabilities”, *I.E.E.E. Transactions on Systems Science and Cybernetics*, SSC-4, 227-241.

Jeffrey, R. C., 1965, *The Logic of Decision,* McGraw Hill; 1983, *2nd edition*, Chicago:

University of Chicago Press.

–––, 1987, “Alias Smith and Jones: The Testimony of the Senses”, *Erkenntnis* 26, 391-399.

–––, 1992, *Probability and the Art of Judgment*, New York: Cambridge University Press.

Jeffreys, H., 1939. *Theory of Probability*, Oxford: Oxford University Press.

Joyce, J., 1998, “A Nonpragmatic Vindication of Probabilism”, *Philosophy of Science*, 65 (4) 575-603.

–––, 1999, *The Foundations of Causal Decision Theory*, New York: Cambridge University Press.

Keynes, J. M., 1921, *A Treatise on Probability*, Macmillan and Co.

Kyburg, H., 1977, “Randomness and the Right Reference Class”, *Journal of Philosophy*, 74: 501-520.

Levi, I., 1967, *Gambling with Truth*, New York: Knopf.

–––, 1977, “Direct Inference,” *Journal of Philosophy*, 74, 5-29.

–––, 1978, “Confirmational Conditionalization”, *Journal of Philosophy*, 75, 730-737.

–––, 1980, *The Enterprise of Knowledge*, Cambridge, Mass.: MIT Press.

Lewis, David, 1980, “A Subjectivist's Guide to Objective Chance”, in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability, Vol. 2*, Berkeley: University of California Press, 263-293.

Maher, P., 1996. “Subjective and Objective Confirmation”, *Philosophy of Science* 63, 149-174.

–––, 2001, “Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny”, Erkenntnis 55, 183–216.

–––, 2006, “The Concept of Inductive Probability”, *Erkenntnis* 65, 185-206.

Quine, W.V. 1953, “Two Dogmas of Empiricism”, in *From a Logical Point of View*, New York: Harper Torchbooks.

Ramsey, F. P., 1926, “Truth and Probability”, in *Foundations of Mathematics and other Essays*, R. B. Braithwaite (ed.), Routledge & P. Kegan,1931, 156-198; reprinted in H. E. Kyburg, Jr. and H. E. Smokler (eds.), *Studies in Subjective Probability*, 2nd ed., R. E. Krieger Publishing Company, 1980, 23-52; reprinted in D. H. Mellor (ed.), *Philosophical Papers*, Cambridge: University Press, Cambridge, 1990.

Rosenkrantz, R.D., 1981, *Foundations and Applications of Inductive Probability*, Atascadero, CA: Ridgeview Publishing.

Royall, R. 1997, *Statistical Evidence: A Likelihood Paradigm*, New York: Chapman & Hall/CRC.

Savage, L. J., 1954, *The Foundations of Statistics*, John Wiley (2nd ed., New York: Dover 1972).

Savage, L. J., et al.1962. *The Foundations of Statistical Inference*, London: Methuen.

Skyrms, B., 1984, *Pragmatics and Empiricism*, New Haven: Yale University Press.

Williamson, J., 2006, "Philosophies of Probability: Objective Bayesianism and its Challenges", in Andrew Irvine (ed.), *Handbook of the Philosophy of Mathematics*, *Volume 4 of the Handbook of the Philosophy of Science*, Elsevier.