Bayesian Confirmation Theory

James Hawthorne

Forthcoming in the *Continuum Companion to Philosophy*, Steven French and Juha Saatsi, eds.

Scientific theories and hypotheses make claims that go well beyond what we can immediately observe. How can we come to know whether such claims are true? The obvious approach is to see what a hypothesis *says* about the observationally accessible parts of the world. If it gets that wrong, then it must be false; if it gets that right, then it may have some claim to being true. Any sensible attempt to construct a logic that captures how we may come to reasonably believe the falsehood or truth of scientific hypotheses must be built on this idea. Philosophers refer to such logics as *logics of confirmation* or as *confirmation theories*.

Among philosophers and logicians the most influential contemporary approach to the logic of the hypothesis confirmation is *Bayesian Confirmation Theory*. This approach employs probability functions to represent two distinct things: what a hypothesis *says* about *how likely it is* that specific evidential events will occur, and how strongly the hypothesis is confirmed or refuted by such evidence. The hypothesis-based *probabilities of evidence claims* are called *likelihoods*. When the evidence is *more likely* according to one hypothesis than according to an alternative, that increases the probabilistic *degree of confirmation* of the former hypothesis over the later. Any probabilistic confirmation theory that employs the same probability functions to represent *both* the likelihoods hypotheses confer on evidence claims *and* the degrees to which hypotheses are confirmed by these evidence claims will be a *Bayesian confirmation theory*, because a simple theorem of probability theory, called Bayes' Theorem, expresses precisely how these likelihoods contribute to the confirmational probabilities of hypotheses.

This article describes the essential features of *Bayesian confirmation theory*. Section 1 presents the probabilistic axioms for confirmation functions. Section 2 describes how this logic is applied via Bayes' Theorem to represent the evidential support of hypotheses. Section 3 draws on a *Bayesian Convergence Theorem* to show why this logic may be expected to refute false hypotheses and support true ones. Section 4 generalizes the treatment of Bayesian likelihoods described in section 2. Section 5 concludes by briefly commenting on what confirmation functions *are* conceptually.

1. The Axioms for Confirmation Functions

A confirmation function is a binary function, $P_\alpha[A|B]$, on sentences of a language capable of expressing scientific hypotheses and theories. Logicians make this idea precise by taking the language and its deductive logic to be that of predicate logic (including the identity relation) because that language is known to have the expressive power needed to represent the deductive logic of any scientific theory. Such a language possesses a *non-logical* vocabulary consisting of names (and variables) and predicate and relation terms, and a *logical* vocabulary consisting of the standard *logical terms*: '$\sim$' for "not", '$\cdot$' for "and", '$\vee$' for "inclusive or", '$\supset$' for truth-functional "if-then", '$\equiv$' for "if and only if", '$\forall$' for "all", '$\exists$' for "some", and '$=$' for the relation "is the same thing as". This language permits the expression of any scientific theory, including set theory and all the rest of mathematics employed by the sciences.

The axioms for confirmation functions are essentially *semantic rules* that constrain each *possible confirmation function* to respect the meanings of the *logical terms* (*not*, *and*, *or*, etc.), much as the axioms for truth-value assignments in the semantics for deductive logic constrain each *possible truth-value assignment* to respect the meanings of the *logical terms*. These rules don't determine which confirmation functions are *correct* (just as the semantic rules for truth-value assignments don't determine which way of assigning truth-values to sentences captures the actual truths). The *correctness* of various *measures of confirmation* may depend on additional considerations, including what the non-logical terms and sentences of the language mean.

Here are the axioms for the confirmation functions, treated as semantic rules on an object language L that's powerful enough to express any scientific theory.

> Let L be a language whose deductive logic is predicate logic with identity – where '$C \models B$' abbreviates 'C logically entails B', and '$\models B$' abbreviates 'B is a tautology'. A confirmation function is any function P_α from pairs of sentences of L to real numbers between 0 and 1 that satisfies:
>
> 1. $P_\alpha[D|E] < 1$ for some D, E;
>
> for all A, B, C,
>
> 2. if $B \models A$, then $P_\alpha[A|B] = 1$;
> 3. If $\models (B \equiv C)$, then $P_\alpha[A|B] = P_\alpha[A|C]$;
> 4. If $C \models \sim(B \cdot A)$, then either $P_\alpha[(A \vee B)|C] = P_\alpha[A|C] + P_\alpha[B|C]$ or, for every D, $P_\alpha[D|C] = 1$;
> 5. $P_\alpha[(A \cdot B)|C] = P_\alpha[A|(B \cdot C)] \ \ P_\alpha[B|C]$.

Each function satisfying these rules is a possible confirmation function. The subscript 'α' reminds us that many alternative functions $\{P_\beta, P_\gamma, \ldots\}$ obey these rules. All the usual theorems of probability theory follow from these axioms.

Some Bayesian logicians have explored the idea that, like deductive logic, a logic of confirmation might be made to depend only on the logical structures of sentences. It's now widely agreed that this project cannot be carried out in a plausible way. The logic of confirmation must also draw on the meanings of the sentences involved. Thus, one should also associate with each confirmation function P_α an assignment of meanings to non-logical terms, and thereby to sentences. This suggests two additional axioms (or rules).

> 6. If A is analytically true (given the meanings that P_α associates with the language) or an axiom of set theory or pure mathematics employed by the sciences, then $P_\alpha[A|B] = 1$ for each sentence B.

It follows that if a sentence C *analytically entails* A, then $P_\alpha[A|C] = 1$.

When a contingent sentence E is considered *certainly true* it's common practice to employ a confirmation function that assigns it probability 1 on every premise B. This saves the trouble of writing E as an explicit premise, because when $P_\beta[E|C] = 1$, $P_\beta[H|E\cdot C] = P_\beta[H|C]$. But writing $P_\beta[H|C] = r$ instead of $P_\beta[H|E\cdot C] = r$ when certain of E hides the confirmational dependence of H on E. This is a bad thing to do in a *logic* of confirmation, where the whole idea is to provide a measure of the extent to which premise sentences indicate the likely truth of conclusion sentences. It makes the logic enthymematic. In deductive logic one wouldn't write '$(E\supset A)\models A$' just because one's already certain that E. One shouldn't do this in a *logic of confirmation* either. The *logic* should represent E as maximally confirmed by every possible premise (including $\sim E$) only in cases where E is logically or analytically true or an axiom of pure mathematics. This motivates the *Axiom of Regularity*.

7. If A is neither logical nor analytically true, nor a consequence of set theory or some other piece of pure mathematics employed by the sciences, then $P_\alpha[A|\sim A] < 1$.

Taken together, axioms 6 and 7 imply that a confirmation function assigns *probability* 1 *on every possible premise* to precisely the sentences that are non-contingently true according to its associated meaning assignment.[1]

Perhaps additional axioms should further constrain confirmation functions. In particular, when hypotheses describe chance situations, a rule like David Lewis's (1980) *Principal Principle* seems appropriate. Consider a hypothesis that says systems in state Y have objective chance (or *propensity*) r to acquire an attribute X – $Ch(\mathrm{X},\mathrm{Y}) = \mathrm{r}$. When c is a system in state Y (i.e., $c\in \mathrm{Y}$), this should give rise to a *direct inference likelihood*: $P_\alpha[c\in \mathrm{X}|Ch(\mathrm{X},\mathrm{Y})=\mathrm{r} \cdot c\in \mathrm{Y}] = r$. One might add an axiom requiring that confirmation functions satisfy this principle. A general axiom of this kind would also specify precisely what sorts of information B should interfere with these likelihoods: when should $P_\alpha[c\in \mathrm{X}|Ch(\mathrm{X},\mathrm{Y})=\mathrm{r} \cdot c\in \mathrm{Y}\cdot B]$ continue to equal r, and for what B may *degree r* no longer hold? Spelling out such a *direct inference likelihood axiom* in full generality turns out to be quite difficult, so we'll not pursue it here. Nevertheless, *chancy claims* in scientific theories should often lead to objective values for likelihoods, on which all confirm functions should agree.

That's the axiomatic basis for the logic of confirmation functions. However, this characterization leaves two important questions untouched: (1) what, conceptually, *is* a *confirmational probability function*?; and (2) why should we consider a confirmation function to be a good way of measuring evidential support? These issues can only be adequately addressed after we see how the logic applies evidence to the confirmation of hypotheses. However, the subjectivist reading of Bayesian confirmation functions has become so prominent in the literature that I will say something about it before proceeding.

---

[1] $P_\alpha[A|C] = 1$ for all C just when $P_\alpha[A|\sim A] = 1$. So $P_\alpha[A|C] = 1$ for all C *implies* A is either logical or analytically true, or a consequence of set theory or some other piece of pure mathematics employed by the sciences. Axiom 6 yields the converse implication.

The subjectivist *interpretation* takes $P_\alpha[A|B]$ to express the degree of belief (or confidence) an agent α would have in A were she to become certain that B but possess no other information that's relevant to A. Although this kind of belief-strength interpretation may be appropriate for probabilities in decision theory, it faces very significant difficulties as a way of understanding confirmation functions. I'll discuss some of these difficulties later, but please forgo this reading for now, since it can be very misleading. Instead think of a confirmation function as a kind of *truth-indicating index*. Later I'll bolstered this idea with an account of how evidence can bring confirmation functions to point towards the truth-values of hypotheses. Because of this feature, confirmation functions should influence one's belief-strengths regarding the truth of hypotheses, although they are not themselves *measures of belief-strength*.

2. The Bayesian Logic of Evidential Support

Let's see how the logic of confirmation functions represents evidential support for scientific hypotheses. Let $<H_1,H_2,\ldots,H_m,\ldots>$ be an exhaustive list of alternative hypotheses about some specific subject. The list may contain a simple pair of alternatives – e.g., <Joe is infected by HIV, Joe is not infected by HIV> – or it may be a long list of competitors (e.g., of alternative ailments that may be responsible for the patient's symptoms). The competitors may make claims about some single event (e.g. about what disease(s) afflicts Joe), or they may be grand theories (e.g., about what laws of nature govern the universe). The list of alternative hypotheses or theories may, in principle, be infinitely long. The idea of testing infinitely many alternatives may seem extraordinary, but nothing about the logic itself forbids it. The alternative hypotheses need not be entertained all at once. They may be constructed and assessed over millennia.

Practically speaking, in order for the list of competing hypotheses to be exhaustive it may need to contain a *catch-all hypothesis* H_K that says none of the other hypotheses is true (e.g., "the patient has an unrecognized disease"). When only a finite number m of explicit alternative hypotheses is under consideration, the catch-all alternative H_K will be equivalent to the sentence that denies each explicit alternative. Thus, if the list of alternatives, including the catch-all, is $<H_1,H_2,\ldots,H_m,H_K>$, then H_K is equivalent to $(\sim H_1\cdot\ldots\cdot\sim H_m)$.

The evidence employed to test hypotheses consists of experiments or observations that each hypothesis *says something about*. On the older *hypothetico-deductive* account of confirmation each hypothesis H_i *speaks* about observations by *deductively entailing* evidence claims. However, hypotheses cannot usually accomplish this on their own. They usually draw on statements C that describe the conditions under which evidential outcome E occurs. In addition, hypotheses often rely on background knowledge and auxiliary hypotheses B (e.g. about how measuring devices function) to connect them via experimental circumstances C to evidential outcomes E.[2] So the deductive logical relationship through which a hypothesis *speaks about* evidence takes the form: $H_i\cdot B\cdot C\models E$. If the observation condition and evidential outcome $(C\cdot E)$

[2] B may itself contain hypotheses that are subject to confirmation via the same kind of treatment described for hypotheses H_i and H_j below (though their confirmation may be relative to some simpler auxiliaries B^*; perhaps even tautological B^*).

occurs, this *may* provide *good evidence* for H_i, provided B holds up. On the other hand, if C holds but E is observed to be false, then deductive logic alone gives us $B\cdot C\cdot{\sim}E \models {\sim}H_i$, and hypothesis H_i is *falsified* by $B\cdot C\cdot{\sim}E$. Thus a hypothesis H_i usually fails to entail evidential claims on its own, but only *speak about evidence deductively* with the assistance of background and auxiliary claims together with descriptions of the experimental or observational circumstances. Similarly, when hypotheses *speak about evidence probabilistically* via likelihoods, conditions C and background B play a comparable enabling role.

In Bayesian confirmation theory the degree to which a hypothesis H_i is confirmed on evidence $C\cdot E$, relative to background B, is represented by the *posterior probability* of H_i, $P_\alpha[H_i|B\cdot C\cdot E]$. Bayes' Theorem shows how this *posterior probability* depends on two kinds of probabilistic factors. It depends on the *prior probability* of H_i, $P_\alpha[H_i|B]$, and on the *likelihood of evidential outcome*, E, according to H_i together with B and C, $P_\alpha[E|H_i\cdot B\cdot C]$. Let's consider the nature of each. Then we'll see how they come together in the logic of hypothesis evaluation.

Likelihoods. Likelihoods express what hypotheses *say* about observationally accessible parts of the world. If a hypothesis together with auxiliaries and observation conditions deductively entails an evidence claim, axiom 2 guarantees that every confirmation function P_α assigns the likelihood value 1 – i.e., if $H_i\cdot B\cdot C \models E$, then $P_\alpha[E|H_i\cdot B\cdot C] = 1$. Similarly, if $H_i\cdot B\cdot C \models {\sim}E$ and $H_i\cdot B\cdot C$ is contingent, the axioms yield $P_\alpha[E|H_i\cdot B\cdot C] = 0$. However, quite often the hypothesis H_i will only imply the evidence to some probabilistic degree. For instance, H_i may itself be an explicitly statistical or chance hypothesis, or H_i may be a non-statistical hypothesis that's probabilistically related to the evidence by statistical auxiliary hypotheses that reside within background B. In either case the likelihoods may be the kind of *direct inference likelihoods* described near the end of section 1. That is, when $H_i\cdot B \models Ch(\mathrm{X},\mathrm{Y})=\mathrm{r}$ (the chances of acquiring X for systems having Y is r) and C is of form $c\in\mathrm{Y}$ and E is of form $c\in\mathrm{X}$, we should have $P_\alpha[E|H\cdot B\cdot C] = r$ (provided $H_i\cdot B$ doesn't also entail some relevant defeater of this *direct inference*).[3] Such likelihoods should be completely objective in that all confirmation functions should agree on their values, just as all confirmation functions agree on likelihoods when evidence is logically entailed. Functions P_α that satisfy the confirmation function axioms but get such *direct inference likelihoods* wrong should be discarded as illegitimate.

Not all scientific likelihoods are warranted deductively or by explicitly stated chance claims. Nevertheless, the likelihoods that relate hypotheses to evidence in scientific contexts will often have widely recognized objective or intersubjectively agreed values. For, likelihoods represent the empirical content of hypotheses – what hypotheses *say* about the observationally accessible parts of the world. So the empirical objectivity of a science relies on a high degree of agreement among scientists on their values.

Consider what a science would be like if scientists disagreed widely about the values of likelihoods for important hypotheses? Whereas expert α takes H_1 to say E is much more likely

---

[3] See David Lewis's (1980) argument for the objectivity of likelihoods based on chance statements. His *Principal Principle* is a direct inference principle governing such likelihoods. Lewis maintains that objective chance is a purely theoretical concept, and that the *Principal Principle* captures "all we know about chance."

than does H_2 ($P_\alpha[E|H_1{\cdot}B{\cdot}C] >> P_\alpha[E|H_2{\cdot}B{\cdot}C]$), her colleague β sees it just the opposite way ($P_\beta[E|H_1{\cdot}B{\cdot}C] << P_\beta[E|H_2{\cdot}B{\cdot}C]$). Thus, whereas α considers $C{\cdot}E$ (given B) to be powerful evidence for H_1 over H_2, β takes the very same evidence to forcefully support H_2 over H_1. If this kind of disagreement occurs often or for important hypotheses in a scientific discipline, the empirical objectivity of that discipline would be a shambles. Each scientist understands the *empirical import* of these hypotheses so differently that each H_j *as understood by* α is an empirically different hypothesis than H_j *as understood by* β. Thus, the empirical objectivity of the sciences requires that experts understand significant hypotheses in similar enough ways that the values of their likelihoods are closely aligned.

For now let's suppose that each hypothesis H_j in the list of alternatives has *precise*, objective or intersubjectively agreed values for its likelihoods (relative to appropriate background and auxiliaries).[4] We'll mark this agreement by dropping the subscript 'α', 'β', etc., from expressions that represent likelihoods, because all confirmation functions under consideration agree on them. Nevertheless, there are perfectly legitimate scientific contexts where *precise agreement* on the values of likelihoods isn't realistic; so later, in section 4, we'll see how the present supposition of *precise agreement* may be relaxed. But for now the main ideas will be more easily explained if we focus on cases where all confirmation functions precisely agree on the values of likelihoods.

Scientific hypotheses are usually tested by a stream of evidence: $C_1{\cdot}E_1$, $C_2{\cdot}E_2$, …, $C_n{\cdot}E_n$. Let's use the expression 'C^n' to represent the conjunction ($C_1{\cdot}C_2{\cdot}\ldots{\cdot}C_n$) of descriptions of the first n observation conditions, and use 'E^n' to represent the conjunction ($E_1{\cdot}E_2{\cdot}\ldots{\cdot}E_n$) of descriptions of their outcomes. So a likelihood for a stream of n observations and their outcomes will take the form '$P[E^n|H_i{\cdot}B{\cdot}C^n] = r$'. Furthermore, the evidence should be representable as probabilistically independent components relative to a given hypothesis $H_i{\cdot}B$:

$$P[E^n|H_i{\cdot}B{\cdot}C^n] = P[E_1|H_i{\cdot}B{\cdot}C_1]\ P[E_2|H_i{\cdot}B{\cdot}C_2]...P[E_n|H_i{\cdot}B{\cdot}C_n].$$[5]

Prior and Posterior Probabilities. The degree to which a hypothesis is confirmed on the evidence, $P_\alpha[H_i|B{\cdot}C^n{\cdot}E^n]$, is called the *posterior probability* of the hypothesis – its probabilistic degree of confirmation *posterior* to taking account of the evidence. Bayes' Theorem will show that posterior confirmation depends on two kinds of factors: *likelihoods*, $P[E^n|H_i{\cdot}B{\cdot}C^n]$, and *prior probabilities* $P_\alpha[H_i|B]$. Prior probabilities represent the degree to which a hypothesis H_i is supported by non-evidential plausibility considerations, *prior* to taking the evidence into account. The notion of *prior*ity for *prior* probabilities isn't temporal – it might make better sense to call them *non-evidential probabilities*. Though *non-evidential*, the plausibility considerations that inform values for *priors* may not be purely *a priori*. They may include both conceptual and broadly empirical considerations not captured by the *likelihoods*.

Because plausibility assessments are usually less objective than likelihoods, critics sometimes

---

[4] The only exception is the catch-all hypothesis H_K, which seldom yields objective likelihoods.
[5] If the evidence were not parsible into independent parts in this way, then hypothesis $H_i{\cdot}B$ would always have to consult a large number of past evidential results, ($C^n{\cdot}E^n$), in order to *say* how likely the various outcomes E of the next experiment C are – since $P[E|H_i{\cdot}B{\cdot}C{\cdot}(C^n{\cdot}E^n)]$ would differ from $P[E|H_i{\cdot}B{\cdot}C]$.

brand *priors* as *merely subjective*, and take their role in the evaluation of hypotheses to be highly problematic. But plausibility assessments often play a crucial role in the sciences, especially when evidence is insufficient to distinguish among some alternative hypotheses. Furthermore, the epithet "*merely subjective*" is unwarranted. Plausibility assessments are often backed by extensive arguments that draw on forceful conceptual and empirical considerations not captured by likelihoods. That's the epistemic role of the thought experiment, for example.

Indeed, we often have good reasons besides the evidence to strongly reject some *logically possible* alternatives as *just too implausible*, or as at least as *much less plausible* than better conceived candidates. In evaluating hypotheses we often 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 available thus far. Any reasonable scientist will reject most such inventions immediately, because they look *ad hoc*, contrived, or plain foolish. Such reasons for rejection appeal to neither purely logical characteristics of these hypotheses, nor to evidence. All such reasons are "mere" plausibility assessments, not part of the evidential likelihoods.

Prior plausibilities are "subjective" in the sense that scientists may disagree on the relative merits of plausibility arguments, so disagree on the values for priors. Furthermore, the plausibility of a hypothesis is usually somewhat vague or imprecise. So it's reasonable to represent priors by an interval of values, a *plausibility range*, rather than by specific numbers.[6] We'll see more about how that works a bit later. The main point is that plausibility assessments in the sciences are far from *mere subjective whims*. They play an important role in the epistemology of the sciences. So it's a virtue of Bayesian confirmation theory that it provides a place for such assessments to figure into the logic of hypothesis evaluation.

Forms of Bayes' Theorem. Let's now examine several forms of Bayes' Theorem, each derivable from our axioms. Here is the simplest:

(6) Simple Form of Bayes' Theorem:

$$P_\alpha[H_i|B\cdot C^n\cdot E^n] = \frac{P[E^n|H_i\cdot B\cdot C^n]\ \ P_\alpha[H_i|B]\ \ P_\alpha[C^n|H_i\cdot B]}{P_\alpha[E^n|B\cdot C^n]\ \ P_\alpha[C^n|B]}$$

$$= \frac{P[E^n|H_i\cdot B\cdot C^n]\ \ P_\alpha[H_i|B]}{P_\alpha[E^n|B\cdot C^n]} \qquad \text{provided } P_\alpha[C^n|H_i\cdot B] = P_\alpha[C^n|B]$$

Here the posterior probability of a hypothesis is seen to depend on the *likelihood* it assigns the evidence, its *prior probability*, and the *simple probability of the evidence*, $P_\alpha[E^n|B\cdot C^n]$. If an outcome E_k occurs with likelihood $P[E_k|H_i\cdot B\cdot C_k] = 0$, then the cumulative likelihood $P[E^n|H_i\cdot B\cdot C^n] = 0$ as well. As a result the posterior degree of confirmation of H_i crashes to 0; H_i is *falsified* by the evidence.

---

[6] I.e., vague prior plausibilities may be represented by a set of confirmation functions that jointly cover the *plausibility ranges* for hypotheses.

This version of Bayes' Theorem includes the terms $P_\alpha[C^n|H_i{\cdot}B]$ and $P_\alpha[C^n|B]$, which express how likely it is that the conditions for the experiments or observations actually hold. These factors are often suppressed in presentations of Bayes' Theorem, perhaps by hiding conditions C^n in the background B. However, that approach makes B continually change as new evidence is accumulated. So it's preferable to make these factors explicit, and deal with them directly. That's easy because in realistic cases the ratio $(P_\alpha[C^n|H_i{\cdot}B]/P_\alpha[C^n|B])$ should be 1, or nearly 1, because the truth of the hypothesis should not be relevant to whether the observation conditions hold.

The *simple probability of the evidence* represents a weighted average of likelihoods across all the alternative hypotheses: $P_\alpha[E^n|B{\cdot}C^n] = \sum_j P_\alpha[E^n|H_j{\cdot}B{\cdot}C^n]\ P_\alpha[H_j|B{\cdot}C^n] = \sum_j P_\alpha[E^n|H_j{\cdot}B{\cdot}C^n]\ P_\alpha[H_j|B]$ when $P_\alpha[C^n|H_i{\cdot}B] = P_\alpha[C^n|B]$. This factor is hard to assess if one isn't aware of all hypotheses worthy of consideration. So, in most cases another form of Bayes' Theorem is more useful – a form that compares one pair of hypotheses at a time.

(7) Ratio Form of Bayes' Theorem:

$$\frac{P_\alpha[H_j|B{\cdot}C^n{\cdot}E^n]}{P_\alpha[H_i|B{\cdot}C^n{\cdot}E^n]} = \frac{P[E^n|H_j{\cdot}B{\cdot}C^n]}{P[E^n|H_i{\cdot}B{\cdot}C^n]}\ \frac{P_\alpha[H_j|B]}{P_\alpha[H_i|B]}\ \frac{P_\alpha[C^n|H_j{\cdot}B]}{P_\alpha[C^n|H_i{\cdot}B]}$$

$$= \frac{P[E^n|H_j{\cdot}B{\cdot}C^n]}{P[E^n|H_i{\cdot}B{\cdot}C^n]}\ \frac{P_\alpha[H_j|B]}{P_\alpha[H_i|B]} \quad \text{for } P_\alpha[C^n|H_j{\cdot}B] = P_\alpha[C^n|H_i{\cdot}B].$$

The second line follows when neither hypothesis makes the occurrence of the observation conditions more likely than the other: $P_\alpha[C^n|H_j{\cdot}B] = P_\alpha[C^n|H_i{\cdot}B]$. This should hold for most real applications, so let's suppose it holds throughout the remaining discussion.[7]

This form of Bayes' Theorem is the most useful for many scientific applications, where few alternative hypotheses are considered. It shows that *likelihood ratios* carry the full import of the evidence. Evidence influences the evaluation of hypotheses in no other way. Although this version has not received much attention in the philosophical literature, it's so central to a *realistic Bayesian Confirmation Theory* that I'll discuss it in detail.

Notice that the ratio form of the theorem easily accommodates situations where we don't have precise values for prior probabilities. For one thing, it only depends on our ability to assess *how much more or less plausible* alternative H_j is than H_i – the ratio $P_\alpha[H_j|B]/P_\alpha[H_i|B]$. Such *relative plausibilities* are much easier to judge than are specific numerical values for individual hypotheses. This results in assessments of *ratios of posterior confirmational probabilities* – e.g. $P_\alpha[H_j|B{\cdot}C{\cdot}E]/P_\alpha[H_i|B{\cdot}C{\cdot}E] = 1/10$ says "on the evidence, H_i is a ten times more plausible than H_j". Although such posterior ratios don't supply values for the individual posterior probabilities,

---

[7] This supposition also avoids inessential complexity. Nothing I'll say below changes much when ratios $P_\alpha[C^n|H_j{\cdot}B]/P_\alpha[C^n|H_i{\cdot}B]$ don't get exceptionally far from 1. If they did, the experimental conditions themselves would count as significant evidence.

they place an important constraint on the posterior confirmation of H_j, since logically $P_\alpha[H_j|B\cdot C\cdot E] \le P_\alpha[H_j|B\cdot C\cdot E]/P_\alpha[H_i|B\cdot C\cdot E]$.

Furthermore, this form of Bayes' Theorem tolerates a good deal of vagueness or imprecision in assessments of the ratios of prior plausibilities. In practice one need only assess bounds for the prior plausibility ratios to achieve meaningful results. Given a prior ratio in a specific interval, $q \le P_\alpha[H_j|B]/P_\alpha[H_i|B] \le r$, a likelihood ratio $P[E^n|H_j\cdot B\cdot C^n]/P[E^n|H_i\cdot B\cdot C^n] = s_n$ produces a posterior confirmation ratio in the interval $s_n\cdot q \le P_\alpha[H_j|B\cdot C^n\cdot E^n]/P_\alpha[H_i|B\cdot C^n\cdot E^n] \le s_n\cdot r$. As the likelihood ratio value s_n approaches 0, the interval value for the posterior ratio gets smaller, and its upper bound $s_n\cdot r$ approaches 0; so the absolute degree of confirmation of H_j, $P_\alpha[H_j|B\cdot C^n\cdot E^n]$, also must approach 0. This is really useful because it can be shown that when $H_i\cdot B\cdot C^n$ is true and H_j is empirically distinct from H_i, the values of likelihood ratios $P[E^n|H_j\cdot B\cdot C^n]/P[E^n|H_i\cdot B\cdot C^n]$ will very likely approach 0 as the amount of evidence increases. (I'll discuss this *Likelihood Ratio Convergence* result below.) When that happens, the upper bound on the posterior probability ratio also approaches 0, driving the posterior probability of H_j to approach 0, effectively refuting H_j. Thus, false competitors of a true hypothesis are eliminated.

Relative to each hypothesis, evidential events should be probabilistically independent of one another (or at least parsible into independent clusters). So the likelihood ratio for the total evidence decomposes into a product of likelihood ratios for each observation:

$$\frac{P[E^n|H_j\cdot B\cdot C^n]}{P[E^n|H_i\cdot B\cdot C^n]} = \frac{P[E_1|H_j\cdot B\cdot C_1]}{P[E_1|H_i\cdot B\cdot C_1]} \cdots \frac{P[E_{n-1}|H_j\cdot B\cdot C_{n-1}]}{P[E_{n-1}|H_i\cdot B\cdot C_{n-1}]} \frac{P[E_n|H_j\cdot B\cdot C_n]}{P[E_n|H_i\cdot B\cdot C_n]}$$

It follows from (7) that a previous *confirmation ratio* (based on the previous evidence) is updated on new evidence via multiplication by the likelihood ratio for the new evidence:

(8) Ratio Bayesian Updating Formula:

$$\frac{P_\alpha[H_j|B\cdot C^n\cdot E^n]}{P_\alpha[H_i|B\cdot C^n\cdot E^n]} = \frac{P[E_n|H_j\cdot B\cdot C_n]}{P[E_n|H_i\cdot B\cdot C_n]} \frac{P_\alpha[H_j|B\cdot C^{n-1}\cdot E^{n-1}]}{P_\alpha[H_i|B\cdot C^{n-1}\cdot E^{n-1}]}$$

$$= \frac{P[E_n|H_j\cdot B\cdot C_n]}{P[E_n|H_i\cdot B\cdot C_n]} \cdot\ldots\cdot \frac{P[E_1|H_j\cdot B\cdot C_1]}{P[E_1|H_i\cdot B\cdot C_1]} \frac{P_\alpha[H_j|B]}{P_\alpha[H_i|B]}$$

The second line of (8) shows how the contribution of any individual piece of evidence may be reassessed (even tossed out) if it comes into doubt. Similarly, prior probability ratios (or intervals for them) may be reassessed and changed to reflect additional plausibility considerations.[8]

[8] Technically, changing the value of the (interval covering the) prior plausibility ratio means switching to a different confirmation function (or different set of functions with priors that span the new interval).

From a Bayesian perspective, when scientists report on their experimental findings in research journals, they should indicate the impact of the evidence on various hypotheses by reporting the values of the *likelihood ratios*, $P[E|H_j \cdot B \cdot C]/P[E|H_i \cdot B \cdot C]$, for the evidence $C \cdot E$ obtained from their research.[9] Although they may say little or nothing about the (ratios of) prior plausibilities, some conception of the plausibility of the hypotheses must be in play, at least implicitly, because if no one in the relevant scientific community takes hypothesis H_j at all seriously (i.e. if the relevant scientific community takes $P_\alpha[H_j|B]$ to be almost 0 to begin with), then no one will care about an experimental study that "finds strong new evidence against H_j" by establishing some result ($C \cdot E$) that makes the likelihood ratio $P[E|H_j \cdot B \cdot C]/P[E|H_i \cdot B \cdot C]$ extremely small. No respectable scientific journal would bother to publish such results. If prior plausibility played no role, such results would deserve as much consideration as any.

Bayesian confirmation is a version of *eliminative induction*. Suppose H_i is a 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 $P[E^n|H_j \cdot B \cdot C^n]/P[E^n|H_i \cdot B \cdot C^n]$ toward 0. Equation (7) shows that each alternative H_j is effectively refuted because $P_\alpha[H_j|B \cdot C \cdot E] \leq P_\alpha[H_j|B \cdot C \cdot E]/P_\alpha[H_i|B \cdot C \cdot E]$ approaches 0. As this happens the posterior probability of H_i must approach 1, as the next two forms of Bayes' Theorem show.

The *odds against A* given *B* is, by definition, $\Omega_\alpha[{\sim}A|B] = P_\alpha[{\sim}A|B]/P_\alpha[A|B]$. If we sum the ratio versions of Bayes' Theorem over all alternatives to hypothesis H_i (including the catch-all H_K, if needed), we get an Odds Form of Bayes' Theorem:

(9) Odds Form of Bayes' Theorem

$$\Omega_\alpha[{\sim}H_i|B \cdot C^n \cdot E^n] = \sum_{j \neq i} \frac{P_\alpha[H_j|B \cdot C^n \cdot E^n]}{P_\alpha[H_i|B \cdot C^n \cdot E^n]} + \frac{P_\alpha[H_K|B \cdot C^n \cdot E^n]}{P_\alpha[H_i|B \cdot C^n \cdot E^n]}$$

$$= \sum_{j \neq i} \frac{P[E^n|H_j \cdot B \cdot C^n]}{P[E^n|H_i \cdot B \cdot C^n]} \; \frac{P_\alpha[H_j|B]}{P_\alpha[H_i|B]} + \frac{P_\alpha[E^n|H_K \cdot B \cdot C^n]}{P[E^n|H_i \cdot B \cdot C^n]} \; \frac{P_\alpha[H_K|B]}{P_\alpha[H_i|B]}$$

If the catch-all alternative isn't needed, just drop the expression after the '+' sign. We represent the term for the catch-all hypothesis separately because the *likelihood* of evidence relative to it will not generally enjoy the kind of objectivity possessed by likelihoods for *specific* hypotheses. We indicate this by leaving the subscript 'α' on catch-all likelihoods.

Although the catch-all hypothesis lacks objective likelihoods, the influence of the whole catch-all term should diminish as additional specific hypotheses become articulated. When a new hypothesis H_{m+1} is made explicit, the old catch-all H_K is replaced by a new one, H_{K^*}, of form $({\sim}H_1 \cdot \ldots \cdot {\sim}H_m \cdot {\sim}H_{m+1})$. The prior probability for the new catch-all hypothesis is peeled off the prior of the old catch-all: $P_\alpha[H_{K^*}|B] = P_\alpha[H_K|B] - P_\alpha[H_{m+1}|B]$. So the influence of the catch-all term

[9] Here Bayesian Confirmation Theory agrees with the view about how statistical hypotheses should be tested called *Likelihoodism*. See (Edwards, 1972) and (Royall, 1997).

should diminish towards 0 as new alternative hypotheses are developed.[10] Thus, if increasing evidence drives the likelihood ratios that test H_i against each rival towards 0, and if the influence of the catch-all term also approaches 0, then the posterior odds against H_i must approach 0. As $\Omega_\alpha[{\sim}H_i|B{\cdot}C^n{\cdot}E^n]$ 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:

(10) Bayes' Theorem: From Posterior Odds to Posterior Probability

$$P_\alpha[H_i|B{\cdot}C^n{\cdot}E^n] \;=\; 1/(1+\Omega_\alpha[{\sim}H_i|B{\cdot}C^n{\cdot}E^n]).$$

For scientific contexts where not all significant alternative hypotheses can be surveyed, the formulas for posterior odds and posterior probabilities provided by equations (9) and (10) are only of conceptual interest. They tell us about the nature of the logic, but may not permit us to compute actual posterior probabilities of hypotheses that remain unrefuted by likelihood ratios. In practice the best we can usually do in such contexts is compare pairs of hypotheses, and find evidence enough to drive one of each pair to near extinction via extreme likelihood ratios. Thus, Bayesian confirmation is fundamentally a variety of eliminative induction, where the hypothesis that remains unrefuted is our best candidate for the truth.

If we are fortunate enough to develop the true alternative, then each of its evidentially distinct rivals may be laid low by evidence via the likelihood ratios. As that happens the true hypothesis will climb to the top of the list of alternatives and remain there – its posterior plausibility $P_\alpha[H_i|B{\cdot}C^n{\cdot}E^n]$ will become many times larger than the posterior plausibilities of alternatives. In principle it's posterior probability heads towards 1, but in practice we merely recognize such a hypothesis as *very strongly confirmed* – superior to all alternatives considered thus far. Thus, this Bayesian logic is a formal representation of how the theoretical sciences actually operate.

3. The Likelihood Ratio Convergence Theorem

When $H_i{\cdot}B$ is true, the series of likelihood ratios $P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n]$ will very probably favor H_i over empirically distinct alternatives H_j by heading towards 0 for as the evidence accumulates (as n increases). A *Bayesian Convergence Theorem* establishes this fact. Before stating the theorem I'll first explain some notation.

For observation sequence C^n, consider each of the *possible outcomes sequences* E^n. Some would result in likelihood ratios for H_j over H_i that are less than ε, for some chosen small increment ε > 0 (e.g. you might choose ε = 1/1000). For specific ε, the set of all possible such outcome sequences is expressed by '$\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}$'. This will be some particular finite set of sentences. Now, consider the disjunction of all sentences in that set; the resulting disjunctive sentence asserts that one of the outcome sequences, described by one of the sentences in the set, is true. We indicate this disjunctive sentence by placing the "or" symbol '∨' in front of the expression for the set: $\vee\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}$. *How likely is it*, if $H_i{\cdot}B{\cdot}C^n$ is true, that this disjunctive sentence will be true? – i.e., *how likely is it*, if $H_i{\cdot}B{\cdot}C^n$ is true, that "one

---

[10] $P_\alpha[H_K|B] = \mathrm{P}_\alpha[{\sim}H_1{\cdot}\ldots{\cdot}{\sim}H_m{\cdot}(H_{m+1}\vee{\sim}H_{m+1})|B] = P_\alpha[{\sim}H_1{\cdot}\ldots{\cdot}{\sim}H_m{\cdot}{\sim}H_{m+1}|B]+P_\alpha[H_{m+1}|B]$.

of the outcome sequences E^n will occur that makes $P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon$"? The *Likelihood Ratio Convergence Theorem* answers this question by providing a lower bound on *how likely this is*, and that lower bound approaches 1 as n increases. The *Theorem* expresses this in terms of a likelihood: $P[\vee\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}|H_i{\cdot}B{\cdot}C^n]$.

The full statement of the theorem comes in two parts. The first part encompasses cases where H_j says some outcome is *impossible* that H_i counts as *possible*; the second part encompasses evidential sequences where such extreme disagreement doesn't happen.

Likelihood Ratio Convergence Theorem:[11]

1. Suppose a subsequence C^m of the whole evidence stream consists of observations where for each one, C_k, there is some possible outcome E_k *deemed possible by* $H_i{\cdot}B$ to at least some small degree $\delta > 0$ but *deemed impossible by* $H_j{\cdot}B$ – i.e. for each C_k there is a possible E_k such that $P[E_k|H_i{\cdot}B{\cdot}C_k] \geq \delta > 0$ but $P[E_k|H_j{\cdot}B{\cdot}C_k] = 0$. Then,

   $$P[\vee\{E^m : P[E^m|H_j{\cdot}B{\cdot}C^m] = 0\}|H_i{\cdot}B{\cdot}C^m] \;\geq\; 1-(1-\delta)^m$$, which approaches 1 for large m.

2. Suppose the evidence stream C^n consists of observations where for each one, C_k, each possible outcome E_k *deemed possible* by $H_i{\cdot}B$ is also *deemed possible* by $H_j{\cdot}B$ – i.e. for each E_k, if $P[E_k|H_i{\cdot}B{\cdot}C_k] > 0$, then $P[E_k|H_j{\cdot}B{\cdot}C_k] > 0$. And further suppose that for each E_k such that $P[E_k|H_i{\cdot}B{\cdot}C_k] > 0$, $P[E_k|H_j{\cdot}B{\cdot}C_k] \geq \gamma{\cdot}P[E_k|H_i{\cdot}B{\cdot}C_k]$, for some small positive $\gamma \leq 1/3$ – i.e. there is some positive $\gamma \leq 1/3$ such that no such possible E_k disfavors the competitor H_j so much as to make $P[E_k|H_j{\cdot}B{\cdot}C_k]/P[E_k|H_i{\cdot}B{\cdot}C_k] < \gamma$. Then for any small positive $\varepsilon < 1$ you might choose (but large enough that for the number of observations n being contemplated, the value of $(1/n)\sum_{k=1}^{n} \mathrm{EQI}[C_k|H_i/H_j|B] > -(\log \varepsilon)/n$),

   $$P[\vee\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}|H_i{\cdot}B{\cdot}C^n] \;>$$

   $$1 \;-\; \frac{1}{n}\,\frac{(\log \gamma)^2}{((1/n)\sum_{k=1}^{n} \mathrm{EQI}[C_k|H_i/H_j|B] + (\log \varepsilon)/n)^2}$$

   which approaches 1 for large n, provided $(1/n)\sum_{k=1}^{n} \mathrm{EQI}[C_k|H_i/H_j|B]$ has a positive lower bound – i.e., provided the sequence of observation C^n has an *average expected quality of information* (*average* EQI) for empirically distinct H_j, given H_i, that *doesn't get arbitrarily near* 0 as the evidence sequence increases.[12] (The base of the log doesn't matter, but let's take it to be 2; then for $\varepsilon = 1/2^k$, $\log \varepsilon = -k$; and for $\gamma = 1/2^u$, $(\log \gamma)^2 = u^2$.)

The term on the right-hand side of the inequality is a *worst case lower bound*. The actual value of $P[\vee\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}|H_i{\cdot}B{\cdot}C^n]$ will in all but the most extreme cases be much

---

[11] For proof see (Hawthorne 2009, supplements 4-7).

[12] This provision can fail only if new observations C_n *can only produce* ever weaker evidence (whose likelihood ratio values *have to be* ever closer to 1 *for all possible outcomes of* C_n) as more evidence is obtained (as n is increased).

larger than this bound. That is, given two specific hypotheses H_i and H_j (and their associated likelihoods for possible outcomes), one can actually compute the precise value of $P[\vee\{E^n : P[E^n|H_j{\cdot}B{\cdot}C^n]/P[E^n|H_i{\cdot}B{\cdot}C^n] < \varepsilon\}|H_i{\cdot}B{\cdot}C^n]$. For most hypotheses and types of possible evidence the value of this likelihood is much larger than the lower bound given by this *worst case theorem*.

The term $(1/n)\sum_{k=1}^{n}$ EQI$[C_k|H_i/H_j|B]$ is an information theoretic measure of how good, *on average*, *the range of all possible outcomes* of C_k *are* at distinguishing between H_i and H_j, if $H_i{\cdot}B$ is true. The formula for each EQI$[C_k|H_i/H_j|B]$ is

$$\text{EQI}[C_k|H_i/H_j|B] = \sum_{\{O_{kx}:C_k\}} \log(P[O_{kx}|H_i{\cdot}B{\cdot}C_k]/P[O_{kx}|H_j{\cdot}B{\cdot}C_k])\ P[O_{kx}|H_i{\cdot}B{\cdot}C_k]$$

where the sum ranges over all the possible outcomes O_{kx} of observation C_k that H_i takes to be possible (i.e. for which $P[O_{kx}|H_i{\cdot}B{\cdot}C_k] > 0$).[13]

Thus, the Likelihood Ratio Convergence Theorem establishes that if H_i (together with $B{\cdot}C^n$) is true, as the sequence of observations C^n increases, it becomes *highly likely* (as near 1 as you like) that its outcomes will provide likelihood ratios as close to 0 as you wish.[14] This theorem is not subject to the usual criticisms of Bayesian convergence results. The theorem (and its proof) does not rely on prior probabilities in any way. It doesn't suppose that the evidence is "identically distributed" – it applies to any pair of empirically distinct hypotheses. It's a "weak law of large numbers" result that gives explicit lower bounds on the rate of convergence – so there's no need to wait for the infinite long run. It's a "convergence to truth" result (not merely "convergence to agreement"). It doesn't depend on countable additivity.[15]

Furthermore, because this theorem doesn't depend on prior probabilities, it's not undermined by if they are reassessed and changed as new conceptual and broadly empirical considerations are introduced. 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 (via equation (7)) 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 (via equations (9) and (10)).[16]

---

[13] EQI$[C_k|H_i/H_j|B]$ is the *expected value* of the *log*s *of the likelihood ratios*. Each EQI$[C_k|H_i/H_j|B]$ is greater than 0 if some O_{kx} has $P[O_{kx}|H_j{\cdot}B{\cdot}C_k] \neq P[O_{kx}|H_i{\cdot}B{\cdot}C_k]$; otherwise EQI$[C_k|H_i/H_j|B] = 0$.

[14] A short evidence sequence may suffice if the *average expected quality of information* is large.

[15] For a nice presentation of the most prominent Bayesian convergence results and a discussion of their weaknesses see (Earman, 1992, Ch. 6). Earman doesn't consider the Likelihood Ratio Convergence Theorem, which was first published in (Hawthorne, 1993).

[16] This claim depends on H_i being empirically distinct from each alternative – i.e., that the C_k have possible outcomes E_k such that $P[E_k|H_i{\cdot}B{\cdot}C_k] \neq P[E_k|H_j{\cdot}B{\cdot}C_k]$. If the true hypothesis has empirically equivalent rivals, then convergence implies posterior probability of *their disjunction* goes to 1. Among the equivalent rivals, $P_\alpha[H_j|B{\cdot}C^n{\cdot}E^n]/P_\alpha[H_i|B{\cdot}C^n{\cdot}E^n] = P_\alpha[H_j|B]/P_\alpha[H_i|B]$. So the true hypothesis can obtain a posterior probability near 1 (after evidence drives the posteriors of empirically distinct rivals near 0) *just in case* plausibility considerations result its prior plausibility being much higher than the sum of those of its empirically equivalent rivals.

4. When Likelihoods are not Precise

For some important contexts it's unreasonable to expect likelihoods to possess precise, agreed values, but the evidence remains capable of sorting among hypotheses in a reasonably objective way. Here's how that works.[17]

Consider the following *continental drift hypothesis*: the land masses of Africa and South America were once joined, then split and have drifted apart over the eons. Let's compare it to an alternative *contraction hypothesis*: 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 closely related as they are? One may not be able to determine anything like precise numerical values for such likelihoods. But experts may readily agree that each of these observations is much more likely on the *drift hypothesis* than on the *contraction hypothesis*, and jointly constitute very strong evidence in favor of *drift* over *contraction*. On a Bayesian analysis this is due to the fact that even though these likelihoods do not have precise values, it's obvious to experts that the *ratio of the likelihoods* is pretty extreme, strongly favoring *drift* over *contraction* (according to the Ratio Form of Bayes' Theorem), unless *contraction* is taken to be much more plausible than the *drift* on other grounds.[18]

I argued earlier that disagreement on likelihoods among members of a scientific community would be disastrous to the scientific enterprise were it to result in desperate assessments of which hypotheses are favored by evidence. However, precise values for likelihoods are not crucial to the way evidence sorts among hypotheses. Rather, *ratios of likelihoods* do all the heavy lifting. So when two confirmation functions P_α and P_β disagree on the values of likelihoods, they'll agree well enough on the refutation and support for hypotheses *if* they yield *directionally agreeing* likelihood ratios.

Directional Agreement Condition for Likelihoods Ratios:
The likelihood ratios for a pair of confirmation functions P_α and P_β *directionally agree* on the possible outcomes of observations relevant to a pair of hypotheses *just in case* for each possible outcome E_k of the conditions C_k in the evidence stream,
$P_\alpha[E_k|H_j{\cdot}B{\cdot}C_k]/P_\alpha[E_k|H_i{\cdot}B{\cdot}C_k] < 1$ *just when* $P_\beta[E_k|H_j{\cdot}B{\cdot}C_k]/P_\beta[E_k|H_i{\cdot}B{\cdot}C_k] < 1$, and
$P_\alpha[E_k|H_j{\cdot}B{\cdot}C_k]/P_\alpha[E_k|H_i{\cdot}B{\cdot}C_k] > 1$ *just when* $P_\beta[E_k|H_j{\cdot}B{\cdot}C_k]/P_\beta[E_k|H_i{\cdot}B{\cdot}C_k] > 1$, and

[17] Technically an imprecise likelihood is represented by a set of confirmation functions with likelihood values that span the interval of the imprecision.

[18] Historically, geologists largely dismissed the evidence described above until the 1960s. The strength of this evidence didn't suffice to overcome non-evidential (though *broadly empirical*) considerations that made the *drift hypothesis* seem much less plausible than the traditional *contraction* view. Chiefly, there appeared to be no plausible mechanisms that could move the continents through the ocean floor. Such objections were overcome when a plausible enough "convection mechanism" was articulated and evidence favoring it over the "embedded in bedrock" model was acquired.

each of these likelihood ratios is either close to 1 for both functions or for neither.

When this condition holds, the evidence supports H_i over H_j according to P_α just when it does so for P_β. Furthermore, although the *rate* at which the likelihood ratios increase or decrease as evidence accumulates may differ for these confirmation functions, the total impact of the cumulative evidence will affect the refutation and support of hypotheses in the same way. Indeed, the *Likelihood Ratio Convergence Theorem* still applies. The proof of the theorem doesn't depend on likelihoods being objective. It applies to each confirmation function P_α individually. Thus, when a family of confirmation functions satisfy the *Directional Agreement Condition* and enough empirically distinguishing observations are forthcoming, each will *very probably* yield likelihood ratios for empirically distinct false competitors of a true hypothesis that become extremely small. *Directional Agreement* guarantees that if such convergence towards 0 happens for one of the agreeing confirmation functions, it must happen for them all. As that happens, the posterior confirmation of the true hypothesis must rise to become highly confirmed according to each confirmation function in the family.

5. What *is* Confirmational Probability?

Now that we understand how confirmational probabilities may come to indicate the falsehood or truth of hypotheses, perhaps it's not so important to try to interpret them. But, let's briefly consider some prominent views about how to understand what these functions *are*.

5.1 The Syntactic Logical Interpretation

Early on some Bayesian logicists attempted to explicate confirmation functions that depend only on the *syntactic structures* of sentence, in the same way that deductive logical entailment depends only on syntactic structures.[19] Most logicians now take this project to be fatally flawed. On this view hypotheses with the same syntactic structure should have the same prior probability values. But given any hypothesis, logicians can easily construct an infinite number of alternatives with the same syntactic structure. Most such alternatives would be quickly rejected by scientists as *ad hoc* and ridiculously implausible, but such assessments are based on semantic content, not logical structure. So semantic content should matter. Moreover, how are we supposed to implement this syntactic approach in real cases? Are we to compare the syntactic structures of the various interpretations of quantum theory to see which has the higher prior probability? The defenders of the syntactic-structural view owe us credible reasons to base non-evidential plausibility assessments on syntactic structure alone.

5.2 Subjective Degrees of Belief

Think of α, β, γ, etc., as logically ideal agents, each having his or her own *degrees of belief* in various propositions. We may represent each agent's belief-strengths in terms of a belief-strength function, P_α, P_β, P_γ, etc, defined on statement. Taking unconditional probability as basic, read '$P_\alpha[A] = r$' as saying "the strength of α's belief that A is r"; and read $P_\alpha[A|B] = r$ as saying "the

---

[19] Keynes (1921) and Carnap's (1950) are the most widely known proponents of this idea.

strength of α's belief that *A*, were she to become newly certain of *B* (and nothing more than *B*) would be *r*." This is the widely subscribed *Subjectivist* or *Personalist* interpretation of confirmational probabilities.[20] Subjectivists intend this to be the same notion of probability employed in Bayesian decision theory.

A major difficulty for this view are versions of the *problem of old evidence*.[21] They show that *belief-function likelihoods* cannot maintain the kind of objective values that *confirmation-function likelihoods* should have. This problem is much worse than usually realized.

Suppose '*E*' says "the coin lands *heads* on the next toss", *H* says "the coin is fair" and *C* says "the coin is tossed in the usual way on the next toss", so that the confirmational likelihood should be $P[E|H \cdot C] = 1/2$. However, if the agent is already certain that *E*, then her belief function likelihood should be $P[E|H_j \cdot C] = 1$ for every hypothesis H_j, which undermines the role of the likelihood in testing any hypothesis. Furthermore, even when the agent isn't certain of *E*, but becomes certain of trivial disjunctions involving *E* – e.g., "either the outcome of the next toss will be *heads*, or Jim won't like the outcome of the next toss", $(E \vee F)$ – it can be shown that belief function likelihoods become radically altered from their objective values.

The problem is that an agent's *belief-function likelihood* has to represent her belief strength in the evidence statement when the hypothesis is added to *everything else* the agent *already holds*. But other beliefs and partial beliefs (even hunches) the agent has will almost always severely interfere with the objective values the likelihoods should have for confirmational purposes. Thus, a Bayesian account of *confirmation* and *belief* will require confirmation functions that are distinct from belief functions, and some account of how *degrees-of-confirmation* are supposed to inform an agent's *degrees-of-belief*.

One additional point: The subjectivists' ideal agents are logically omniscient. They assign belief-strength 1 to all logical truths. How is this unobtainable norm for confirmation/belief functions supposed to be relevant to the functioning of real people? However, if we don't reduce confirmation functions to ideal agents' belief functions, this "logical omniscience problem" has no hold over confirmation functions. Real people *use the logic* of confirmation functions in the same sort of way they might use any logic to inform their real (non-ideal) belief strengths.

5.3 Another *Logical* View

Rather than ask what confirmation functions *are*, it's more fruitful to ask what they *do*. Under appropriate circumstances they're *truth-indicating indices*. If, among the alternative hypotheses proposed to account for a given subject-matter we are fortunate enough to think of a hypothesis that happens to be true, and if we find enough ways to empirically test it against rivals, then all that's needed for confirmational success is persistent testing and not too much bad luck with how the evidence actually turns out. For, according to the Likelihood Ratio Convergence Theorem,

---

[20] Ramsey (1926), de Finetti (1937), Savage (1954) are well-known proponents of this view. Howson and Urbach (1993) provide a comprehensive treatment.

[21] Glymour (1980) first raised this problem. For details of the versions mentioned here see (Hawthorne 2005).

the true hypothesis itself *says*, via its likelihoods, that a long enough (but finite) stream of observations is very likely to produce outcomes that will drive the likelihood ratios of empirically distinct false competitors to approach 0. As this happens, the confirmation index of these competitors, as measured by their posterior probabilities, also approaches 0, and the confirmation index of the true hypothesis (or its disjunction with empirically equivalent rivals) will approach 1.

This result does not imply that whatever hypothesis has index near 1 at a given moment is likely to be the true alternative. Rather, it suggests the pragmatic strategy of continually testing hypotheses, and taking whichever of them has an index nearest to 1 (if there is one) as the *best current candidate* for being true. The convergence theorem implies that maintaining this strategy and continually 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 it will remain there. So if we align our belief strengths for hypotheses with their approximate confirmation indices, eventually we should (very probably) come to strongly believe the true hypothesis. But this eliminative strategy only promises to work if we continue to look for rivals and continue to test the best alternative candidates against them. This 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 rely on whatever our most probative non-evidential considerations tell us. We often have good reasons besides the observable evidence to strongly discount some logically possible alternatives as *just too implausible*, or at least as significantly less plausible than some *better conceived* rivals. We always bring some such considerations to bear, at least implicitly. It is a virtue of Bayesian confirmation theory that it provides a place for such assessments to figure into the logic of hypothesis evaluation.

6. Future Directions for Research

Although there has already been a lot of good philosophical work on the implications of Bayesian Confirmation Theory for our understanding of the epistemology of the sciences, a number of important topics and issues remain unsettled, and open to additional investigation. Here is a brief list of some of these topics: (1) does evidential diversity/variety better confirm hypotheses (why)? (2) why Occam's razor? – are simpler hypotheses more likely (why)? (3) the Duhem problem – given that hypotheses usually only "speak to the evidence" via additional auxiliary hypotheses, precisely how are we to determine (in a Bayesian context) whether negative evidence impugns the target hypothesis or an auxiliary hypothesis? (4) what more can be said about the objective evaluation of prior probabilities? – e.g. why are *ad hoc* hypothesis supposed to be less plausible? (5) what does a Bayesian approach have to say about such traditional *paradoxes of confirmation* as "the ravens paradox", "the grue problem", "the apparent confirmation of irrelevant conjuncts" (a.k.a. "the tacking problem"). Many of these topics, and more, are discussed in (Horwich, 1982), (Earman, 1992), and (Howson and Urbach, both 1993 and revised 2005). These treatments are far from definitive, and new, more probative results continue to appear, primarily an journal articles.[22] Much more generally, perhaps the most important open philosophical issue it this: how does, or should, Bayesian Confirmation Theory fit into an larger epistemology of the sciences, and how should this scientific epistemology mesh with much broader philosophical conceptions of the epistemology of everyday life?

References

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

De Finetti, B., 1937, "La Prévision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincaré*, 7, 1-68.

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

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

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

Hawthorne, J., 1993, "Bayesian Induction *Is* Eliminative Induction", *Philosophical Topics* 21, 99-138.

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

–––, 2008, "Inductive Logic", *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/entries/logic-inductive/>.

Hawthorne, J., Fitelson, B., 2004, "Re-solving Irrelevant Conjunction with Probabilistic Independence", *Philosophy of Science* 71, 505-514.

Horwich, P., 1982, *Probability and Evidence*, Cambridge: Cambridge University Press.

Howson, C., Urbach, P., 1993, *Scientific Reasoning: The Bayesian Approach*, 2nd edition, Chicago: Open Court (revised 3rd edition, 2005).

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

Lewis, D., 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.

Ramsey, F. P., 1926, "Truth and Probability", in *Foundations of Mathematics and other Essays*, R. B. Braithwaite (ed.), Routledge & P. Kegan,1931, 156-198.

[22] See (Fitelson and Hawthorne, 2004) for an example of this kind of work.

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).