PROBABILITY CAPTURES THE LOGIC OF SCIENTIFIC CONFIRMATION

1. Introduction

'Confirmation' is a word in ordinary language and, like many such words, its meaning is vague and perhaps also ambiguous. So if we think of a logic of confirmation as a specification of precise rules that govern the word 'confirmation' in ordinary language, there can be no logic of confirmation. There can, however, be a different kind of logic of confirmation. This can be developed using the philosophical methodology known as *explication* (Carnap, 1950, ch. 1). In this methodology, we are given an unclear concept of ordinary language (called the *explicandum*) and our task is to find a precise concept (called the *explicatum*) that is similar to the explicandum and is theoretically fruitful and simple. Since the choice of an explicatum involves several desiderata, which different people may interpret and weight differently, there is not one "right" explication; different people may choose different explicata without either having made a mistake. Nevertheless, we can cite reasons that motivate us to choose one explicatum over another.

In this paper I will define a predicate 'C' which is intended to be an explicatum for confirmation. I will establish a variety of results about 'C' dealing with verified consequences, reasoning by analogy, universal generalizations, Nicod's condition, the ravens paradox, and projectability. We will find that these results correspond well with intuitive judgments about confirmation, thus showing that our explicatum has the desired properties of being similar to its explicandum and theoretically fruitful. In this way we will develop parts of a logic of confirmation. The predicate 'C' will be defined in terms of probability and in that sense we will conclude that probability captures the logic of scientific confirmation.

2. Explication of justified degree of belief

I will begin by explicating the concept of the degree of belief in a hypothesis H that is justified by evidence E. A little more fully, the explicandum is the degree of belief in H that we would be justified in having if E was our total evidence. We have some opinions about this but they are usually vague and different people sometimes have different opinions.

In order to explicate this concept, let us begin by choosing a formalized language, like those studied in symbolic logic; we will call this language L. We will use the letters 'D', 'E', and 'H', with or without subscripts, to denote sentences of L. Let us also stipulate that L contains the usual truth-functional connectives '$\sim$' for negation, '$\vee$' for disjunction, '.' for conjunction, and '$\supset$' for material implication.

Next we define a two-place function p which takes sentences of L as arguments and has real numbers as its values. The definition of p will specify, for each ordered pair of sentences H and E in L, a real number that is denoted $p(H|E)$. This number $p(H|E)$ is intended to be our explicatum for the degree of belief in H that is justified by evidence E; we will therefore choose a definition of p using the desiderata for such an explicatum, namely:

1. The values of p should agree with the judgments about justified degrees of belief that we have. For example, if we judge that E justifies a higher degree of belief in H_1 than in H_2 then we will want to define p in such a way that $p(H_1|E) > p(H_2|E)$. In this way we ensure that the explicatum is similar to the explicandum.

2. The function p should be theoretically fruitful, which means that its values satisfy many general principles.

3. The function p is as simple as possible.

I assume that this function p will satisfy the mathematical laws of conditional probability. We will express these laws using the following axioms. Here D, E, E', H, and H' are any sentences of L and '$=$' between sentences means that the sentences are logically equivalent.[1]

AXIOM 1. $p(H|E) \geq 0$.

AXIOM 2. $p(E|E) = 1$.

AXIOM 3. $p(H|E) + p(\sim H|E) = 1$, *provided E is consistent.*

[1] The following formulation of the axioms of probability is like that of von Wright (1957, p. 93) in several respects. In particular, I follow von Wright in taking $p(H|E)$ to be defined even when E is inconsistent. However, I differ from von Wright about which axiom is not required to hold for inconsistent E; the result is that on my axiomatization, but not von Wright's, if E is inconsistent then $p(H|E) = 1$ for all H (Proposition 2, Section 12.1). Since E entails H if E is inconsistent, the value $p(H|E) = 1$ accords with the conception of probability as a generalization of deductive logic.

AXIOM 4. $p(E.H|D) = p(E|D)p(H|E.D)$.

AXIOM 5. *If* $H' = H$ *and* $E' = E$ *then* $p(H'|E') = p(H|E)$.

3. Explication of confirmation

Now we will explicate the concept that is expressed by the word 'confirms' in statements like the following. (These are from news reports on the web.)

> 'New evidence confirms last year's indication that one type of neutrino emerging from the Sun's core does switch to another type en route to the earth.'
>
> 'New evidence confirms rapid global warming, say scientists.'
>
> 'Tree-ring evidence confirms Alaskan Inuit account of climate disaster.'

If we were to examine these examples in detail we would find that the judgment that some evidence E confirms some hypothesis H makes use of many other, previously known, pieces of evidence. Let us call this other evidence the *background evidence*. Then our explicandum may be expressed more precisely as the concept of E confirming H given background evidence D.

In looking for an explicatum for this concept, we will be guided by the idea that E confirms H given D iff (if and only if) the degree of belief in H that is justified by E and D together is higher than that justified by D alone. The corresponding statement in terms of our explicatum p is $p(H|E.D) > p(H|D)$. So let us adopt the following definition.

DEFINITION 1. $C(H, E, D)$ *iff* $p(H|E.D) > p(H|D)$.

Thus $C(H, E, D)$ will be our explicatum for the ordinary language concept that E confirms H given background evidence D.

In ordinary language we sometimes say that some evidence E_1 confirms a hypothesis H more than some other evidence E_2 does. Plausibly, such a statement is true iff the degree of belief in H that is justified by E_1 and the background evidence is higher than that justified by E_2 and the background evidence. The corresponding statement in terms of our explicatum p is $p(H|E_1.D) > p(H|E_2.D)$. So let us adopt the following definition:

DEFINITION 2. $M(H, E_1, E_2, D)$ *iff* $p(H|E_1.D) > p(H|E_2.D)$.

Thus $M(H, E_1, E_2, D)$ will be our explicatum for the ordinary language concept that E_1 confirms H more than E_2 does, given D.

In the following sections we will use these explications to derive some of the logic of scientific confirmation. Specifically, we will state and prove theorems about our explicata $C(H, E, D)$ and $M(H, E_1, E_2, D)$. These results will be seen to correspond closely to intuitive opinions about the ordinary language concept of confirmation, thus verifying that our explicata are similar to their explicanda.

4. Verified consequences

Scientists often assume that if E is a logical consequence of H then verifying that E is true will confirm H. For example, Galileo's argument that falling bodies are uniformly accelerated consisted in proving that the motion of uniformly accelerated bodies has certain properties and then experimentally verifying that the motion of falling bodies has those properties. Thus Galileo says that his hypothesis is "confirmed mainly by the consideration that experimental results are seen to agree with and exactly correspond to those properties which have been, one after another, demonstrated by us" (1638, p. 160, cf. p.167). Similarly Huygens, in the preface to his *Treatise on Light* (1690), says of that book:

> There will be seen in it demonstrations of those kinds which do not produce as great a certitude as those of geometry, and which even differ much therefrom, since, whereas the geometers prove their propositions by fixed and incontestable principles, here the principles are verified by the conclusions to be drawn from them; the nature of these things not allowing of this being done otherwise. It is always possible to attain thereby to a degree of probability which very often is scarcely less than complete proof.

Let us now examine this assumption that hypotheses are confirmed by verifying their consequences. To do so, we will first express the assumption in terms of our explicata. Although the scientists' statements of the assumption do not mention background evidence, a careful analysis of their deductions would show that when they say evidence is logically implied by a hypothesis, often this is only true given some background evidence. This suggests that the scientists' assumption might be stated in terms of our explicata thus: If E is a logical consequence of $H.D$ then $C(H, E, D)$. The following theorem says that this assumption is true with a couple of provisos.[2]

[2] Proofs of all theorems except Theorem 4 are given in Section 12.

THEOREM 1. *If E is a logical consequence of $H.D$ then $C(H,E,D)$, provided that $0 < p(H|D) < 1$ and $p(E|{\sim}H.D) < 1$.*

If the provisos do not hold then—as can be proved—$C(H,E,D)$ might be false even though E is a logical consequence of $H.D$. The provisos make sense intuitively; if H is already certainly true or false, or if the evidence is certain to obtain even if H is false, then we do not expect E to confirm H. Nevertheless, the need for these provisos is apt to be overlooked when there is no precise statement of the assumption or any attempt to prove it.

Another assumption that is made intuitively by scientists is that a hypothesis is confirmed more strongly the more consequences are verified. This can be expressed more formally by saying that if E_1 and E_2 are different logical consequences of $H.D$ then $E_1.E_2$ confirms H more than E_1 does, given D. The corresponding statement in terms of our explicata is that if E_1 and E_2 are different logical consequences of $H.D$ then $M(H, E_1.E_2, E_1, D)$. The following theorem shows that this is correct, with a few provisos like those of Theorem 1.

THEOREM 2. *If E_2 is a logical consequence of $H.D$ then $M(H, E_1.E_2, E_1, D)$, provided that $0 < p(H|E_1.D) < 1$ and $p(E_2|{\sim}H.E_1.D) < 1$.*

This theorem holds whether or not E_1 is a logical consequence of H.

Comparing E_1 and E_2 separately, it is also intuitive that E_1 confirms H better than E_2 does if E_1 is less probable given ${\sim}H$ than E_2 is. This thought may be expressed in terms of our explicata by saying that if E_1 and E_2 are both logical consequences of $H.D$ and $p(E_1|{\sim}H.D) < p(E_2|{\sim}H.D)$ then $M(H, E_1, E_2, D)$. The next theorem shows that this is also correct, with a natural proviso.

THEOREM 3. *If E_1 and E_2 are logical consequences of $H.D$ and $p(E_1|{\sim}H.D) < p(E_2|{\sim}H.D)$ then $M(H, E_1, E_2, D)$, provided that $0 < p(H|D) < 1$.*

5. Probabilities for two properties

Often data about some sample is held to confirm predictions about other individuals. In order to deal with these kinds of cases we need to specify the language L and the function p more completely than we have done so far. This section will give those further specifications and subsequent sections will apply them to questions about confirmation from samples.

Let us stipulate that our language L contains two primitive predicates 'F' and 'G' and infinitely many individual constants 'a_1', 'a_2', ...[3] A predicate followed by an individual constant means that the individual denoted by the individual constant has the property denoted by the predicate. For example, 'Fa_3' means that a_3 has the property F.

Let us define four predicates 'Q_1' to 'Q_4' in terms of 'F' and 'G' by the following conditions. Here, and in the remainder of this paper, 'a' stands for any individual constant.

$$Q_1a = Fa.Ga;\ Q_2a = Fa.{\sim}Ga;\ Q_3a = {\sim}Fa.Ga;\ Q_4a = {\sim}Fa.{\sim}Ga.$$

A *sample* is a finite set of individuals. A *sample description* is a sentence that says, for each individual in some sample, which Q_i applies to that individual. For example Q_3a_1 is a sample description for the sample consisting solely of a_1 while $Q_1a_2.Q_4a_3$ is a sample description for the sample consisting of a_2 and a_3. We will count any logically true sentence as a sample description for the "sample" containing no individuals; this is artificial but convenient.

Let us also stipulate that L contains a sentence I which means that the properties F and G are statistically independent. Roughly speaking, this means that in a very large sample of individuals, the proportion of individuals that have both F and G is close to the proportion that have F multiplied by the proportion that have G. A more precise definition is given in Maher (2000, p. 65).

So far we have not specified the values of the function p beyond saying that they satisfy the axioms of probability. There are many questions about confirmation that cannot be answered unless we specify the values of p further. Maher (2000) states eight further axioms for a language like L. The following theorem characterizes the probability functions that satisfy these axioms. In this theorem, and in the remainder of this paper, '$\bar{F}$' is the predicate that applies to a iff '${\sim}Fa$' is true, and similarly for '$\bar{G}$'. (This theorem is a combination of Theorems 3, 5, and 6 of Maher, 2000.)

THEOREM 4. *There exist constants λ, γ_1, γ_2, γ_3, and γ_4, with $\lambda > 0$ and $0 < \gamma_i < 1$, such that if E is a sample description for a sample that does not include a, n is the sample size, and n_i is the number of individuals to which E ascribes Q_i, then*

$$p(Q_ia|E.{\sim}I) = \frac{n_i + \lambda\gamma_i}{n + \lambda}.$$

[3] The restriction to two primitive predicates is partly for simplicity but also because we do not currently have a satisfactory specification of p for the case where there are more than two predicates (Maher, 2001).

If $\gamma_F = \gamma_1 + \gamma_2$*,* $\gamma_{\bar{F}} = \gamma_3 + \gamma_4$*,* $\gamma_G = \gamma_1 + \gamma_3$*, and* $\gamma_{\bar{G}} = \gamma_2 + \gamma_4$ *then* $\gamma_1 = \gamma_F\gamma_G$*,* $\gamma_2 = \gamma_F\gamma_{\bar{G}}$*,* $\gamma_3 = \gamma_{\bar{F}}\gamma_G$*, and* $\gamma_4 = \gamma_{\bar{F}}\gamma_{\bar{G}}$*. Also, if* n_ϕ *is the number of individuals to which* E *ascribes property* ϕ *then*

$$\begin{aligned}
p(Q_1a|E.I) &= \frac{n_F + \lambda\gamma_F}{n+\lambda}\,\frac{n_G + \lambda\gamma_G}{n+\lambda} \\
p(Q_2a|E.I) &= \frac{n_F + \lambda\gamma_F}{n+\lambda}\,\frac{n_{\bar{G}} + \lambda\gamma_{\bar{G}}}{n+\lambda} \\
p(Q_3a|E.I) &= \frac{n_{\bar{F}} + \lambda\gamma_{\bar{F}}}{n+\lambda}\,\frac{n_G + \lambda\gamma_G}{n+\lambda} \\
p(Q_4a|E.I) &= \frac{n_{\bar{F}} + \lambda\gamma_{\bar{F}}}{n+\lambda}\,\frac{n_{\bar{G}} + \lambda\gamma_{\bar{G}}}{n+\lambda}.
\end{aligned}$$

I will now comment on the meaning of this theorem. In what follows, for any sentence H of L I use '$p(H)$' as an abbreviation for '$p(H|T)$', where T is any logically true sentence of L.

The γ_i represent the initial probability of an individual having Q_i, γ_F is the initial probability of an individual having F, and so on. This is stated formally by the following theorem.

THEOREM 5. *For* $i = 1, \ldots, 4$*,* $\gamma_i = p(Q_ia)$*. Also* $\gamma_F = p(Fa)$*,* $\gamma_{\bar{F}} = p(\sim Fa)$*,* $\gamma_G = p(Ga)$*, and* $\gamma_{\bar{G}} = p(\sim Ga)$*.*

As the sample size n gets larger and larger, the probability of an unobserved individual having a property moves from these initial values towards the relative frequency of the property in the observed sample (Maher, 2000, Theorem 10). The meaning of the factor λ is that it controls the rate at which these probabilities converge to the observed relative frequencies; the higher λ the slower the convergence.

The formulas given $\sim I$ and I are similar; the difference is that when we are given $\sim I$ (F and G are dependent) we use the observed frequency of the relevant Q_i and when we are given I (F and G are independent) we use the observed frequencies of F (or $\bar{F}$) and G (or $\bar{G}$) separately.

To get numerical values of p from Theorem 4 we need to assign values to γ_F, γ_G, λ, and $p(I)$. (The other values are fixed by these. For example, $\gamma_1 = \gamma_F\gamma_G$.) I will now comment on how these choices can be made.

The choice of γ_F and γ_G will depend on what the predicates 'F' and 'G' mean and may require careful deliberation. For example, if 'F' means 'raven' then, since this is a very specific property and there are vast numbers of alternative properties that seem equally likely to be exemplified a priori, γ_F should be very small, surely less than 1/1000. A reasoned choice of a precise value would require careful consideration of what exactly is meant by 'raven' and what the alternatives are.

Turning now to the choice of λ, Carnap (1980, pp. 111–119) considered the rate of learning from experience that different values of λ induce and came to the conclusion that λ should be about 1 or 2. Another consideration is this: So far as we know a priori, the statistical probability (roughly, long run relative frequency) of an individual having F might have any value from 0 to 1. In the case where $\gamma_F = 1/2$ it is natural to regard all values of this statistical probability as equally probable a priori (more precisely, to have a uniform probability distribution over the possible values), and it can be shown that this happens if and only if $\lambda = 2$. For this reason, I favor choosing $\lambda = 2$.

Finally, we need to choose a value of $p(I)$, the a priori probability that F and G are statistically independent. The alternatives I and $\sim I$ seem to me equally plausible a priori and for that reason I favor choosing $p(I) = 1/2$.

I have made these remarks about the choice of parameter values to indicate how it may be done but, except in examples, I will not assume any particular values of the parameters. What will be assumed is merely that $\lambda > 0$, $0 < \gamma_i < 1$, and $0 < p(I) < 1$.

6. Reasoning by analogy

If individual b is known to have property F then the evidence that another individual a has both F and G would normally be taken to confirm that b also has G. This is a simple example of reasoning by analogy. The following theorem shows that our explication of confirmation agrees with this. (From here on, 'a' and 'b' stand for any distinct individual constants.)

THEOREM 6. $C(Gb, Fa.Ga, Fb)$.

Let us now consider the case in which the evidence is that a has G but not F. It might at first seem that this evidence would be irrelevant to whether b has G, since a and b are not known to be alike in any way. However, it is possible for all we know that the property F is statistically irrelevant to whether something has G, in which case the fact that a has G should confirm that b has G, regardless of whether a has F. Thus I think educated intuition should agree that the evidence does confirm that b has G in this case too, though the confirmation will be weaker than in the preceding case. The following theorems show that our explication of confirmation agrees with both these judgments.

THEOREM 7. $C(Gb, \sim Fa.Ga, Fb)$.

THEOREM 8. $M(Gb, Fa.Ga, \sim Fa.Ga, Fb)$.

There are many other aspects of reasoning by analogy that could be investigated using our explicata but we will move on.

7. Universal generalizations

We are often interested in the confirmation of scientific laws that assert universal generalizations. In order to express such generalizations in L, we will stipulate that L contains an individual variable 'x' and the universal quantifier '(x)', which means 'for all x'. Then for any predicate 'ϕ' in L, the generalization that all individuals have ϕ can be expressed in L by the sentence '$(x)\phi x$'. Here the variable 'x' ranges over the infinite set of individuals a_1, a_2, ... We now have:

THEOREM 9. *If ϕ is any of the predicates*

$$F, \bar{F}, G, \bar{G}, Q_1, Q_2, Q_3, Q_4$$

and if $p(E) > 0$ then $p[(x)\phi x|E] = 0$.

I do not know whether this result extends to other universal generalizations, such as $(x)(Fx \supset Gx)$ or $(x){\sim}Q_i x$.

A corollary of Theorem 9 is:

THEOREM 10. *If ϕ is as in Theorem 9 and T is a logical truth then, for all positive integers m, ${\sim}C[(x)\phi x, \phi a_1 \ldots \phi a_m, T]$.*

On the other hand, we ordinarily suppose that if many individuals are found to have ϕ, with no exceptions, then that confirms that all individuals have ϕ. Thus our explicatum C appears to differ from its explicandum, the ordinary concept of confirmation, on this point. However, the discrepancy depends on the fact that the variables in L range over an infinite set of individuals. The following theorem shows that the discrepancy does not arise when we are concerned with generalizations about a finite set of individuals.

THEOREM 11. *If ϕ is as in Theorem 9 and T is a logical truth then, for all positive integers m and n, $C(\phi a_1 \ldots \phi a_n, \phi a_1 \ldots \phi a_m, T)$.*

We could modify our explicata to allow universal generalizations about infinitely many individuals to be confirmed by sample descriptions. However, that would add considerable complexity and the empirical generalizations that I will discuss in what follows can naturally be taken to be concerned with finite populations of individuals. Therefore, instead of modifying the explicata I will in what follows restrict attention to universal generalizations about finite populations.

8. Nicod's condition

Jean Nicod (1923) maintained that a law of the form 'All F are G' is confirmed by the evidence that some individual is both F and G. Let us call this *Nicod's condition*.

For the reason just indicated, I will take the generalization 'All F are G' to be about individuals $a_1, \ldots, a_n$, for some finite $n > 1$. I will denote this generalization by A, so we have

$$A = (Fa_1 \supset Ga_1) \ldots (Fa_n \supset Ga_n).$$

Nicod's condition, translated into our explicata, is then the claim that $C(A, Fa.Ga, D)$, where a is any one of $a_1, \ldots, a_n$ and D remains to be specified.

Nicod did not specify the background evidence D for which he thought that his condition held. We can easily see that Nicod's condition does not hold for any background evidence D. For example, we have:

THEOREM 12. $\sim C(A, Fa.Ga, Fa.Ga \supset \sim A)$.

On the other hand,

THEOREM 13. $C(A, Fa.Ga, Fa)$.

Thus Nicod's condition may or may not hold, depending on what D is. So let us now consider the case in which D is a logically true sentence, representing the situation in which there is no background evidence. Some authors, such as Hempel (1945) and Maher (1999), have maintained that Nicod's condition holds in this case. I will now examine whether this is correct according to our explicata.

In order to have a simple case to deal with, let $n = 2$, so

$$A = (Fa_1 \supset Ga_1).(Fa_2 \supset Ga_2).$$

In Section 5 I said that I favored choosing $\lambda = 2$ and $p(I) = 1/2$, while the values of γ_F and γ_G should depend on what 'F' and 'G' mean and may be very small. Suppose then that we have

$$\gamma_F = 0.001; \quad \gamma_G = 0.1; \quad \lambda = 2; \quad p(I) = 1/2.$$

In this case calculation shows that, with $a = a_1$ or a_2,

$$p(A|Fa.Ga) = .8995 < .9985 = p(A).$$

So in this case $\sim C(A, Fa.Ga, T)$, which shows that according to our explications Nicod's condition does not always hold, even when there is no background evidence.

Table I. Probabilities of $Q_i b$ in the counterexample to Nicod's condition

	G	$\bar{G}$
F	.0001	.0009
$\bar{F}$	.0999	.8991

$p(Q_i b)$

	G	$\bar{G}$
F	.2335	.1005
$\bar{F}$	.1665	.4995

$p(Q_i b|Q_1 a)$

	G	$\bar{G}$
F	.0001	.0006
$\bar{F}$	.0667	.9327

$p(Q_i b|Q_4 a)$

Since Nicod's condition has seemed very intuitive to many, this result might seem to reflect badly on our explicata. However, the failure of Nicod's condition in this example is intuitively intelligible, as the following example shows.

According to standard logic, 'All unicorns are white' is true if there are no unicorns. Given what we know, it is almost certain that there are no unicorns and hence 'All unicorns are white' is almost certainly true. But now imagine that we discover a white unicorn; this astounding discovery would make it no longer so incredible that a non-white unicorn exists and hence would disconfirm 'All unicorns are white.'

The above numerical counterexample to Nicod's condition is similar to the unicorn example; initially it is improbable that Fs exist, since $\gamma_F = 0.001$, and the discovery of an F that is G then raises the probability that there is also an F that is not G, thus disconfirming that all F are G. Table I shows $p(Q_i b)$ and $p(Q_i b|Q_1 a)$ for $i = 1$, 2, 3, and 4. This shows how the evidence $Q_1 a$ raises the probability that b is Q_2 and hence a counterexample to the generalization that all F are G. Initially that probability is 0.0009 but given $Q_1 a$ it becomes 0.1005.[4]

9. The ravens paradox

The following have all been regarded as plausible:

(i) Nicod's condition holds when there is no background evidence.

(ii) Confirmation relations are unchanged by substitution of logically equivalent sentences.

[4] Good (1968) argued that Nicod's condition could fail when there is no background evidence for the sort of reason given here. Maher (1999, sec. 4.6) showed that Good's stated premises did not entail his conclusion. The present discussion shows that Good's reasoning goes through if we use a probability function that allows for analogy effects between individuals that are known to differ in some way. The probability functions used by Maher (1999) do not allow for such analogy effects.

(iii) In the absence of background evidence, the evidence that some individual is a non-black non-raven does not confirm that all ravens are black.

However, (i)–(iii) are inconsistent. For (i) implies that a non-black non-raven confirms "all non-black things are non-ravens" and the latter is logically equivalent to "all ravens are black." Thus there has seemed to be a paradox here (Hempel, 1945).

We have already seen that (i) is false, which suffices to resolve the paradox. But let us now use our explicanda to assess (ii) and (iii).

In terms of our explicanda, what (ii) says is that if $H = H'$, $E = E'$, and $D = D'$, then $C(H, E, D)$ iff $C(H', E', D')$. It follows from Axiom 5 and Definition 1 that this is true. So we accept (ii).

Now let F mean 'raven' and G mean 'black'. Then in terms of our explicata, what (iii) asserts is $\sim C(A, \sim Fa.\sim Ga, T)$. But using the same parameter values as in the numerical example of the preceding section, we find:

$$p(A|\sim Fa.\sim Ga) = .9994 > .9985 = p(A).$$

So in this case $C(A, \sim Fa.\sim Ga, T)$. This is contrary to the intuitions of many but when we understand the situation better it ceases to be unintuitive, as I will now show.

The box on the right in Table I shows $p(Q_i b|Q_4 a)$, which on the current interpretation is the probability of $Q_i b$ given that a is a non-black non-raven. We see that

$$p(Q_2 b|Q_4 a) = 0.0006 < 0.0009 = p(Q_2 b)$$

and so $Q_4 a$ reduces the probability that b is a counterexample to "All ravens are black." This should not be surprising. In addition, $Q_4 a$ tells us that a is not a counterexample to "All ravens are black," which a priori it might have been. So for these two reasons together, it ought not to be surprising that a non-black non-raven can confirm that all ravens are black.

Thus our response to the ravens paradox is to reject both (i) and (iii). Neither proposition holds generally according to our explicata and the reasons why they do not hold make intuitively good sense.

10. Projectability

Goodman (1979, p. 74) defined the predicate 'grue' by saying that "it applies to all things examined before t just in case they are green but to other things just in case they are blue." Goodman claimed, and it has generally been accepted, that 'grue' is not "projectable" although most discussions, including Goodman's own, do not say precisely what they mean by 'projectable'. Goodman's discussion is also entangled with his mistaken acceptance of Nicod's condition. Our explications allow us to clarify this confused situation.

One precise concept of projectability is the following:

DEFINITION 3. *Predicate ϕ is absolutely projectable iff $C(\phi b, \phi a, T)$.*

The basic predicates in L are projectable in this sense, that is:

THEOREM 14. *The predicates F, $\bar{F}$, G, and $\bar{G}$ are absolutely projectable.*

Now let us define a predicate G' as follows:

DEFINITION 4. $G'a = (Fa.Ga) \vee (\sim Fa.\sim Ga)$.

If 'F' means 'observed before t' and 'G' means 'green' then 'G'' has a meaning similar to 'grue'.

THEOREM 15. *G' is absolutely projectable.*

This is contrary to what many philosophers seem to believe but careful consideration will show that our explications here again correspond well to their explicanda. Theorem 15 corresponds to this statement of ordinary language: The justified degree of belief that an individual is grue, given no evidence except that some other individual is grue, is higher than if there was no evidence at all. If we keep in mind that we do not know of either individual whether it has been observed before t then this statement should be intuitively acceptable.

Philosophers regularly claim that if 'green' and 'grue' were both projectable then the same evidence would confirm both that an unobserved individual is green and that it is not green. It is a demonstrable fact about our explicata that the same evidence cannot confirm both a sentence and its negation, so this claim is definitely false when explicated as above. When philosophers say things like this they are perhaps assuming that we know of each individual whether or not it has been observed before t; however, the concept of absolute projectability says nothing about what is true with this background evidence. So let us now consider a different concept of projectability.

DEFINITION 5. *Predicate ϕ is projectable across predicate ψ iff $C(\phi b, \phi a.\psi a, \sim\psi b)$.*

THEOREM 16. *G is, and G' is not, projectable across F.*

We saw that if 'F' means 'observed before t' and 'G' means 'green' then 'G'' has a meaning similar to 'grue'. With those meanings, Theorem 16 fits the usual views of what is and is not projectable.

However, we could specify that 'G' means 'observed before t and green or not observed before t and not green.' Then, with 'F' still meaning 'observed before t', 'G'' would mean 'green'; in that case Theorem 16 would be just the opposite of the usual views of what is projectable. This shows that the acceptability of our explicata depends on the meanings assigned to the primitive predicates 'F' and 'G' in the language L. We get satisfactory results if the primitive predicates express ordinary concepts like 'green' and we may not get satisfactory results if some primitive predicates express gerrymandered concepts like 'grue'.

11. Conclusion

The predicate 'C' is a good explicatum for confirmation because it is similar to its explicandum and theoretically fruitful. This predicate was defined in terms of probability. In that sense, probability captures the logic of scientific confirmation.

12. Proofs

12.1. Propositions

This section states and proves some propositions that will later be used in the proofs of the theorems.

PROPOSITION 1. *If H is a logical consequence of E then $p(H|E) = 1$.*

Proof. Suppose H is a logical consequence of E. Then

$$\begin{aligned} 1 &= p(E|E), \text{ by Axiom 2} \\ &= p(H.E|E), \text{ by Axiom 5} \\ &= p(H|E)\, p(E|H.E), \text{ by Axiom 4} \\ &= p(H|E)\, p(E|E), \text{ by Axiom 5} \\ &= p(H|E), \text{ by Axiom 2.} \end{aligned}$$

PROPOSITION 2. *If E is inconsistent then $p(H|E) = 1$.*

Proof. If E is inconsistent then H is a logical consequence of E and so, by Proposition 1, $p(H|E) = 1$.

PROPOSITION 3. *If E is consistent and $E.H$ is inconsistent then $p(H|E) = 0$.*

Proof. Suppose E is consistent and $E.H$ is inconsistent. Then $\sim H$ is a logical consequence of E and so

$$\begin{aligned} p(H|E) &= 1 - p(\sim H|E), \text{ by Axiom 3} \\ &= 1 - 1, \text{ by Prop. 1} \\ &= 0. \end{aligned}$$

PROPOSITION 4. *$p(E|D) = p(E.H|D) + p(E.\sim H|D)$, provided D is consistent.*

Proof. Suppose D is consistent. If $E.D$ is consistent then:

$$\begin{aligned} p(E|D) &= p(E|D)[p(H|E.D) + p(\sim H|E.D)], \text{ by Axiom 3} \\ &= p(E.H|D) + p(E.\sim H|D), \text{ by Axiom 4.} \end{aligned}$$

If $E.D$ is inconsistent then Proposition 3 implies that $p(E|D)$, $p(E.H|D)$, and $p(E.\sim H|D)$ are all zero and so again $p(E|D) = p(E.H|D) + p(E.\sim H|D)$.

PROPOSITION 5 (Law of total probability). *If D is consistent then*

$$p(E|D) = p(E|H.D)p(H|D) + p(E|\sim H.D)p(\sim H|D).$$

Proof. Suppose D is consistent. Then

$$\begin{aligned} p(E|D) &= p(E.H|D) + p(E.\sim H|D), \text{ by Prop. 4} \\ &= p(E|H.D)p(H|D) + p(E|\sim H.D)p(\sim H|D), \text{ by Axiom 4.} \end{aligned}$$

PROPOSITION 6. *If D is consistent and $p(E|D) > 0$ then*

$$p(H|E.D) = \frac{p(E|H.D)p(H|D)}{p(E|D)}.$$

Proof. Suppose D is consistent and $p(E|D) > 0$. Then

$$\begin{aligned} p(H|E.D) &= \frac{p(E.H|D)}{p(E|D)}, \text{ by Axiom 4} \\ &= \frac{p(E|H.D)p(H|D)}{p(E|D)}, \text{ by Axiom 4.} \end{aligned}$$

PROPOSITION 7 (Bayes's theorem). *If D is consistent and $p(E|D) > 0$ then*

$$p(H|E.D) = \frac{p(E|H.D)p(H|D)}{p(E|H.D)p(H|D) + p(E|\sim H.D)p(\sim H|D)}.$$

Proof. Immediate from Propositions 5 and 6.

PROPOSITION 8. *If E is a logical consequence of $H.D$ then $p(E|D) \geq p(H|D)$.*

Proof. If D is consistent then

$$\begin{aligned} p(E|D) &= p(E|H.D)p(H|D) + p(E|{\sim}H.D)p({\sim}H|D), \text{ by Prop. 5} \\ &\geq p(E|H.D)p(H|D), \text{ by Axiom 1} \\ &= p(H|D), \text{ by Prop. 1.} \end{aligned}$$

If D is inconsistent then by Proposition 2 $p(E|D) = p(H|D) = 1$, so again $p(E|D) \geq p(H|D)$.

12.2. Proof of Theorem 1

Suppose that E is a logical consequence of $H.D$, $0 < p(H|D) < 1$, and $p(E|{\sim}H.D) < 1$. By Proposition 2, D is consistent. So by Axiom 3, $p({\sim}H|D) > 0$. Since $p(E|{\sim}H.D) < 1$ it then follows that

$$p(E|{\sim}H.D)p({\sim}H|D) < p({\sim}H|D). \tag{1}$$

Since $p(H|D) > 0$, it follows from Proposition 8 that $p(E|D) > 0$. So

$$\begin{aligned} p(H|E.D) &= \frac{p(E|H.D)p(H|D)}{p(E|H.D)p(H|D) + p(E|{\sim}H.D)p({\sim}H|D)}, \text{ by Prop. 7} \\ &= \frac{p(H|D)}{p(H|D) + p(E|{\sim}H.D)p({\sim}H|D)}, \text{ by Prop. 1} \\ &> \frac{p(H|D)}{p(H|D) + p({\sim}H|D)}, \text{ by (1) and } p(H|D) > 0 \\ &= p(H|D), \text{ by Axiom 3.} \end{aligned}$$

So, by Definition 1, $C(H, E, D)$.

12.3. Proof of Theorem 2

Suppose that E_2 is a logical consequence of $H.D$. It follows that E_2 is a logical consequence of $H.E_1.D$. Suppose further that $0 < p(H|E_1.D) < 1$ and $p(E_2|{\sim}H.E_1.D) < 1$. It then follows from Theorem 1 that $C(H, E_2, E_1.D)$. By Definition 1, this means that $p(H|E_1.E_2.D) > p(H|E_1.D)$. By Definition 2, it follows that $M(H, E_1.E_2, E_1, D)$.

12.4. Proof of Theorem 3

Suppose that E_1 and E_2 are logical consequences of $H.D$, $p(E_1|{\sim}H.D) < p(E_2|{\sim}H.D)$, and $0 < p(H|D) < 1$. By Proposition 2, D is consistent. So by Axiom 3, $p({\sim}H|D) > 0$. Since $p(E_1|{\sim}H.D) < p(E_2|{\sim}H.D)$, it then follows that

$$p(E_1|{\sim}H.D)p({\sim}H|D) < p(E_2|{\sim}H.D)p({\sim}H|D). \qquad (2)$$

Since $p(H|D) > 0$, it follows from Proposition 8 that $p(E_1|D) > 0$ and $p(E_2|D) > 0$. So

$$\begin{aligned}
p(H|E_1.D) &= \frac{p(E_1|H.D)p(H|D)}{p(E_1|H.D)p(H|D) + p(E_1|{\sim}H.D)p({\sim}H|D)}, \text{ by Prop. 7} \\
&= \frac{p(H|D)}{p(H|D) + p(E_1|{\sim}H.D)p({\sim}H|D)}, \text{ by Prop. 1} \\
&> \frac{p(H|D)}{p(H|D) + p(E_2|{\sim}H.D)p({\sim}H|D)}, \text{ by (2) and } p(H|D) > 0 \\
&= \frac{p(E_2|H.D)p(H|D)}{p(E_2|H.D)p(H|D) + p(E_2|{\sim}H.D)p({\sim}H|D)}, \text{ by Prop. 1} \\
&= p(H|E_2.D), \text{ by Prop. 7.}
\end{aligned}$$

So by Definition 2, $M(H, E_1, E_2, D)$.

12.5. Proof of Theorem 5

$$\begin{aligned}
p(Q_1a) &= p(Q_1a|I)p(I) + p(Q_1a|{\sim}I)p({\sim}I), \text{ by Prop. 5} \\
&= \gamma_1 p(I) + \gamma_1 p({\sim}I), \text{ by Theorem 4} \\
&= \gamma_1, \text{ by Axiom 3.}
\end{aligned}$$

Similarly, $\gamma_2 = p(Q_2a)$, $\gamma_3 = p(Q_3a)$, and $\gamma_4 = p(Q_4a)$.

$$\begin{aligned}
p(Fa) &= p(Q_1a) + p(Q_2a), \text{ by Prop. 4} \\
&= \gamma_1 + \gamma_2, \text{ as just shown} \\
&= \gamma_F(\gamma_G + \gamma_{\bar{G}}), \text{ by Theorem 4} \\
&= \gamma_F(\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4), \text{ by definition of } \gamma_G \text{ and } \gamma_{\bar{G}} \\
&= \gamma_F[p(Q_1a) + p(Q_2a) + p(Q_3a) + p(Q_4a)], \text{ as just shown} \\
&= \gamma_F[p(Fa) + p({\sim}Fa)], \text{ by Prop. 4} \\
&= \gamma_F, \text{ by Axiom 3.}
\end{aligned}$$

Similarly, $\gamma_{\bar{F}} = p({\sim}Fa)$, $\gamma_G = p(Ga)$, and $\gamma_{\bar{G}} = p({\sim}Ga)$.

12.6. Propositions used in the proof of Theorems 6–8

This section states and proves some propositions that will be used in proving the theorems in Section 6.

PROPOSITION 9. *$p(I|Q_i a) = p(I)$ and $p(\sim I|Q_i a) = p(\sim I)$ for $i = 1, \ldots, 4$.*

Proof.

$$\begin{aligned} p(I) &= \frac{\gamma_1 p(I)}{\gamma_1}, \text{ trivially} \\ &= \frac{p(Q_1 a|I)p(I)}{\gamma_1}, \text{ by Theorem 4} \\ &= \frac{p(Q_1 a|I)p(I)}{p(Q_1 a)}, \text{ by Theorem 5} \\ &= p(I|Q_1 a), \text{ by Prop. 6.} \end{aligned}$$

Similarly, $p(I|Q_i a) = p(I)$ for $i = 2$, 3, and 4. It follows from Axiom 3 that $p(\sim I|Q_i a) = p(\sim I)$.

PROPOSITION 10. *$\gamma_F + \gamma_{\bar{F}} = \gamma_G + \gamma_{\bar{G}} = 1$.*

Proof. By Theorem 5, $\gamma_F = p(Fa)$ and $\gamma_{\bar{F}} = p(\sim Fa)$. By Axiom 3, the sum of these is 1. Similarly, $\gamma_G + \gamma_{\bar{G}} = 1$.

PROPOSITION 11. *If ϕ is F, $\bar{F}$, G, or $\bar{G}$ then $0 < \gamma_\phi < 1$.*

Proof. By Theorem 4, $\gamma_F = \gamma_1 + \gamma_2 > 0$. Also

$$\begin{aligned} \gamma_F &= 1 - \gamma_{\bar{F}}, \text{ by Prop. 10} \\ &< 1, \text{ since } \gamma_{\bar{F}} > 0. \end{aligned}$$

The argument for $\bar{F}$, G, and $\bar{G}$ is similar.

PROPOSITION 12. *If ϕ is F, $\bar{F}$, G, or $\bar{G}$ then $\gamma_\phi < (1+\lambda\gamma_\phi)/(1+\lambda)$.*

Proof. By Proposition 11, $\gamma_\phi < 1$. Adding $\lambda\gamma_\phi$ to both sides and then dividing both sides by $1 + \lambda$ gives the proposition.

PROPOSITION 13. *$\gamma_G < (1 + \lambda\gamma_1)/(1 + \lambda\gamma_F)$.*

Proof. Substitute G for ϕ and $\lambda\gamma_F$ for λ in Proposition 12.

PROPOSITION 14. *$p(Ga|Fa) = \gamma_G$.*

Proof.

$$\begin{aligned}\gamma_F\gamma_G &= \gamma_1, \text{ by Theorem 4}\\ &= p(Q_1a), \text{ by Theorem 5}\\ &= p(Fa)p(Ga|Fa), \text{ by Axiom 4}\\ &= \gamma_F p(Ga|Fa), \text{ by Theorem 5.}\end{aligned}$$

Dividing both sides by γ_F gives the proposition.

PROPOSITION 15.

$$p(Gb|Fa.Ga.Fb) = \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1}{1+\lambda\gamma_F}p(\sim I).$$

Proof.

$$\begin{aligned}p(Q_1b|Q_1a) &= p(Q_1b|Q_1a.I)p(I|Q_1a) + p(Q_1b|Q_1a.\sim I)p(\sim I|Q_1a),\\ &\quad \text{by Prop. 5}\\ &= p(Q_1b|Q_1a.I)p(I) + p(Q_1b|Q_1a.\sim I)p(\sim I), \text{ by Prop. 9}\\ &= \frac{1+\lambda\gamma_F}{1+\lambda}\frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1}{1+\lambda}p(\sim I), \text{ by Thm. 4.} \quad (3)\end{aligned}$$

Similarly,

$$p(Q_2b|Q_1a) = \frac{1+\lambda\gamma_F}{1+\lambda}\frac{\lambda\gamma_{\bar{G}}}{1+\lambda}p(I) + \frac{\lambda\gamma_2}{1+\lambda}p(\sim I). \quad (4)$$

$$\begin{aligned}p(Fb|Q_1a) &= p(Q_1b|Q_1a) + p(Q_2b|Q_1a), \text{ by Prop. 4}\\ &= \frac{1+\lambda\gamma_F}{1+\lambda}\frac{1+\lambda\gamma_G+\lambda\gamma_{\bar{G}}}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1+\lambda\gamma_2}{1+\lambda}p(\sim I),\\ &\quad \text{by (3) and (4)}\\ &= \frac{1+\lambda\gamma_F}{1+\lambda}p(I) + \frac{1+\lambda\gamma_F}{1+\lambda}p(\sim I), \text{ by Prop. 10 and def. of } \gamma_F\\ &= \frac{1+\lambda\gamma_F}{1+\lambda}, \text{ by Axiom 3.} \quad (5)\end{aligned}$$

By Axioms 4 and 5,

$$p(Q_1b|Q_1a) = p(Fb|Q_1a)p(Gb|Fa.Ga.Fb).$$

Substituting (3) and (5) in this gives:

$$\frac{1+\lambda\gamma_F}{1+\lambda}\frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1}{1+\lambda}p(\sim I) = \frac{1+\lambda\gamma_F}{1+\lambda}p(Gb|Fa.Ga.Fb).$$

Dividing both sides by $(1+\lambda\gamma_F)/(1+\lambda)$ gives the proposition.

PROPOSITION 16.

$$p(Gb|{\sim}Fa.Ga.Fb) = \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \gamma_G p({\sim}I).$$

Proof. Similar to Proposition 15.

12.7. Proof of Theorem 6

$$\begin{aligned} p(Gb|Fa.Ga.Fb) &= \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1}{1+\lambda\gamma_F}p({\sim}I), \text{ by Prop. 15} \\ &> \gamma_G p(I) + \gamma_G p({\sim}I), \text{ by Props. 12 and 13} \\ &= \gamma_G, \text{ by Axiom 3} \\ &= p(Gb|Fb), \text{ by Prop. 14.} \end{aligned}$$

So by Definition 1, $C(Gb, Fa.Ga, Fb)$.

12.8. Proof of Theorem 7

$$\begin{aligned} p(Gb|{\sim}Fa.Ga.Fb) &= \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \gamma_G p({\sim}I), \text{ by Prop. 16} \\ &> \gamma_G p(I) + \gamma_G p({\sim}I), \text{ by Prop. 12 and } p(I) > 0 \\ &= \gamma_G, \text{ by Axiom 3} \\ &= p(Gb|Fb), \text{ by Prop. 14.} \end{aligned}$$

So by Definition 1, $C(Gb, {\sim}Fa.Ga, Fb)$.

12.9. Proof of Theorem 8

$$\begin{aligned} p(Gb|Fa.Ga.Fb) &= \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \frac{1+\lambda\gamma_1}{1+\lambda\gamma_F}p({\sim}I), \text{ by Prop. 15} \\ &> \frac{1+\lambda\gamma_G}{1+\lambda}p(I) + \gamma_G p({\sim}I), \text{ by Prop. 13 and } p(I) < 1 \\ &= p(Gb|{\sim}Fa.Ga.Fb), \text{ by Prop. 16.} \end{aligned}$$

So by Definition 2, $M(Gb, Fa.Ga, {\sim}Fa.Ga, Fb)$.

12.10. Propositions used in the proof of Theorem 9

PROPOSITION 17. *If $E_1, \ldots, E_n$ are pairwise inconsistent and D is consistent then*

$$p(E_1 \vee \cdots \vee E_n|D) = p(E_1|D) + \cdots + p(E_n|D).$$

Proof. If $n = 1$ then the proposition is trivially true. Now suppose that the proposition holds for $n = k$ and let $E_1, \ldots, E_{k+1}$ be pairwise inconsistent propositions. Then

$$\begin{aligned} p(E_1 \vee \cdots \vee E_{k+1}|D) &= p((E_1 \vee \cdots \vee E_{k+1}).{\sim}E_{k+1}|D) + \\ &\quad\; p((E_1 \vee \cdots \vee E_{k+1}).E_{k+1}|D), \text{ by Prop. 4} \\ &= p(E_1 \vee \cdots \vee E_k|D) + p(E_{k+1}|D), \text{ by Axiom 5} \\ &= p(E_1|D) + \cdots + p(E_{k+1}|D), \text{ by assumption.} \end{aligned}$$

Thus the proposition holds for $n = k+1$. So by mathematical induction the proposition holds for all positive integers n.

PROPOSITION 18. *If $\lambda > 0$ and $0 < \gamma < 1$ then*

$$\prod_{i=0}^{\infty} \frac{i + \lambda\gamma}{i + \lambda} = 0.$$

Proof. Let $\bar{\gamma} = 1 - \gamma$. Then for all $i \geq 0$,

$$\frac{i + \lambda\gamma}{i + \lambda} = 1 - \frac{\lambda\bar{\gamma}}{i + \lambda}. \tag{6}$$

Also $0 < \lambda\bar{\gamma}/(i + \lambda) < 1$ for all $i \geq 0$. Now

$$\int_{x=0}^{\infty} \frac{\lambda\bar{\gamma}}{x + \lambda}\, dx = \lambda\bar{\gamma}\, [\ln(x + \lambda)]_{x=0}^{\infty} = \infty.$$

So by the integral test for convergence of infinite series (Flatto, 1976, Theorem 5.10),

$$\sum_{i=0}^{\infty} \frac{\lambda\bar{\gamma}}{i + \lambda} = \infty.$$

Hence, by (6) and Theorem 5.32(2) of Flatto (1976),

$$\prod_{i=0}^{\infty} \frac{i + \lambda\gamma}{i + \lambda} = 0.$$

PROPOSITION 19. *If ϕ is any of the predicates F, $\bar{F}$, G, or $\bar{G}$ then, for all positive integers n,*

$$p(\phi a_1 \ldots \phi a_n) = \prod_{i=0}^{n-1} \frac{i + \lambda\gamma_\phi}{i + \lambda}.$$

Proof. Let E be any sample description for $a_1, \ldots, a_n$ that ascribes either Q_1 or Q_2 to each of $a_1, \ldots, a_n$. Let n_i be the number of individuals to which E ascribes Q_i. Then

$$
\begin{aligned}
& p(Fa_{n+1}|Fa_1 \ldots Fa_n.I) \\
&= \sum_E p(Fa_{n+1}|E.I)p(E|Fa_1 \ldots Fa_n.I), \text{ by Prop. 5} \\
&= \sum_E \left[p(Q_1a_{n+1}|E.I) + p(Q_2a_{n+1}|E.I)\right] p(E|Fa_1 \ldots Fa_n.I), \\
&\quad \text{by Prop. 4} \\
&= \sum_E \frac{n+\lambda\gamma_F}{n+\lambda}\left[\frac{n_G+\lambda\gamma_G}{n+\lambda} + \frac{n_{\bar{G}}+\lambda\gamma_{\bar{G}}}{n+\lambda}\right] p(E|Fa_1 \ldots Fa_n.I), \\
&\quad \text{by Theorem 4} \\
&= \sum_E \frac{n+\lambda\gamma_F}{n+\lambda} p(E|Fa_1 \ldots Fa_n.I), \text{ by Prop. 10 and } n_G + n_{\bar{G}} = n \\
&= \frac{n+\lambda\gamma_F}{n+\lambda} \sum_E p(E|Fa_1 \ldots Fa_n.I), \text{ since } n \text{ is the same for all } E \\
&= \frac{n+\lambda\gamma_F}{n+\lambda} p\left(\bigvee_E E|Fa_1 \ldots Fa_n.I\right), \text{ by Prop. 17} \\
&= \frac{n+\lambda\gamma_F}{n+\lambda}, \text{ by Prop. 1.} \qquad (7)
\end{aligned}
$$

$$
\begin{aligned}
& p(Fa_{n+1}|Fa_1 \ldots Fa_n.{\sim}I) \\
&= \sum_E p(Fa_{n+1}|E.{\sim}I)p(E|Fa_1 \ldots Fa_n.{\sim}I), \text{ by Prop. 5} \\
&= \sum_E \left[p(Q_1a_{n+1}|E.{\sim}I) + p(Q_2a_{n+1}|E.{\sim}I)\right] p(E|Fa_1 \ldots Fa_n.{\sim}I), \\
&\quad \text{by Prop. 4} \\
&= \sum_E \left[\frac{n_1+\lambda\gamma_1}{n+\lambda} + \frac{n_2+\lambda\gamma_2}{n+\lambda}\right] p(E|Fa_1 \ldots Fa_n.{\sim}I), \text{ by Theorem 4} \\
&= \sum_E \frac{n+\lambda\gamma_F}{n+\lambda} p(E|Fa_1 \ldots Fa_n.{\sim}I), \text{ since } n_1 + n_2 = n \\
&= \frac{n+\lambda\gamma_F}{n+\lambda}, \text{ using Prop. 17 as above.} \qquad (8)
\end{aligned}
$$

$$
\begin{aligned}
& p(Fa_{n+1}|Fa_1 \ldots Fa_n) \\
&= p(Fa_{n+1}|Fa_1 \ldots Fa_n.I)p(I|Fa_1 \ldots Fa_n) + \\
&\quad p(Fa_{n+1}|Fa_1 \ldots Fa_n.{\sim}I)p({\sim}I|Fa_1 \ldots Fa_n), \text{ by Prop. 5}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{n+\lambda\gamma_F}{n+\lambda}[p(I|Fa_1\dots Fa_n) + p(\sim I|Fa_1\dots Fa_n)], \text{ by (7) and (8)} \\
&= \frac{n+\lambda\gamma_F}{n+\lambda}, \text{ by Axiom 3.} \qquad (9)
\end{aligned}
$$

$$
\begin{aligned}
p(Fa_1\dots Fa_n) &= p(Fa_1)\prod_{i=1}^{n-1} p(Fa_{i+1}|Fa_1\dots Fa_i), \text{ by Axiom 4} \\
&= \prod_{i=0}^{n-1}\frac{i+\lambda\gamma_F}{i+\lambda}, \text{ by Theorem 5 and (9).}
\end{aligned}
$$

So by mathematical induction, the proposition holds for $\phi = F$. Parallel reasoning shows that it also holds for $\phi = \bar{F}$, $\phi = G$, and $\phi = \bar{G}$.

12.11. Proof of Theorem 9

Case (i): ϕ is F, $\bar{F}$, G, or $\bar{G}$. Let $\varepsilon > 0$. By Theorem 4, $\lambda > 0$ and $0 < \gamma_\phi < 1$, so by Proposition 18 there exists an integer N such that

$$\prod_{i=0}^{N-1}\frac{i+\lambda\gamma_\phi}{i+\lambda} < \varepsilon.$$

Now

$$
\begin{aligned}
p[(x)\phi x] &\le p(\phi a_1\dots\phi a_N), \text{ by Prop. 8} \\
&= \prod_{i=0}^{N-1}\frac{i+\lambda\gamma_\phi}{i+\lambda}, \text{ by Prop. 19} \\
&< \varepsilon, \text{ by choice of } N.
\end{aligned}
$$

Hence $p[(x)\phi x] = 0$. Also

$$
\begin{aligned}
p[(x)\phi x|E] &= \frac{p[(x)\phi x.E]}{p(E)}, \text{ by Axiom 4 and } p(E) > 0 \\
&\le \frac{p[(x)\phi x]}{p(E)}, \text{ by Prop. 8} \\
&= 0, \text{ since } p[(x)\phi x] = 0.
\end{aligned}
$$

Case (ii): ϕ is Q_1, Q_2, Q_3, or Q_4.

$$
\begin{aligned}
p[(x)Q_1x|E] &\le p[(x)Fx|E], \text{ by Prop. 8} \\
&= 0, \text{ from case (i).}
\end{aligned}
$$

Similar reasoning shows that the result also holds if ϕ is Q_2, Q_3, or Q_4.

12.12. Proof of Theorem 10

Using Proposition 19 and Axiom 5, $p(\phi a_1 \ldots \phi a_m.T) > 0$. Also $p(T) > 0$. So by Theorem 9,

$$p[(x)\phi x|\phi a_1 \ldots \phi a_m.T] = p[(x)\phi x|T] = 0.$$

So by Definition 1, $\sim C[(x)\phi x, \phi a_1 \ldots \phi a_m, T]$.

12.13. Proof of Theorem 11

If $m \geq n$ then

$$\begin{aligned} p(\phi a_1 \ldots \phi a_n|\phi a_1 \ldots \phi a_m) &= 1, \text{ by Prop. 1} \\ &> p(\phi a_1 \ldots \phi a_n), \text{ by Prop. 19.} \end{aligned}$$

By Proposition 19, $0 < p(\phi a_1 \ldots \phi a_m) < 1$. So if $m < n$ then

$$\begin{aligned} p(\phi a_1 \ldots \phi a_n|\phi a_1 \ldots \phi a_m) &= \frac{p(\phi a_1 \ldots \phi a_n)}{p(\phi a_1 \ldots \phi a_m)}, \text{ by Axiom 4} \\ &> p(\phi a_1 \ldots \phi a_n). \end{aligned}$$

Thus $p(\phi a_1 \ldots \phi a_n|\phi a_1 \ldots \phi a_m) > p(\phi a_1 \ldots \phi a_n)$ for all m and n. So by Definition 1, $C(\phi a_1 \ldots \phi a_n, \phi a_1 \ldots \phi a_m, T)$.

12.14. Proof of Theorem 12

Let $D = Fa.Ga \supset \sim A$. Then $Fa.Ga.D$ is consistent and $Fa.Ga.D.A$ is inconsistent, so by Proposition 3, $p(A, Fa.Ga.D) = 0$. By Axiom 1, $p(A, D) \geq 0$ and so, by Definition 1, $\sim C(A, Fa.Ga, D)$.

12.15. Proof of Theorem 13

$$\begin{aligned} p(A|Fa.Ga.Fa) &= p(A|Fa.Ga), \text{ by Axiom 5} \\ &> p(A|Fa.Ga)p(Ga|Fa), \\ &\quad \text{since } p(Ga|Fa) < 1 \text{ by Props. 11 and 14} \\ &= p(A|Fa.Ga)p(Ga|Fa) + p(A|Fa.{\sim}Ga)p({\sim}Ga), \\ &\quad \text{since } p(A|Fa.{\sim}Ga) = 0 \text{ by Prop. 3} \\ &= p(A|Fa), \text{ by Prop. 5.} \end{aligned}$$

So by Definition 1, $C(A, Fa.Ga, Fa)$.

12.16. Proof of Theorem 14

Let ϕ be F, $\bar{F}$, G, or $\bar{G}$. Then

$$\begin{aligned} p(\phi b|\phi a) &= \frac{p(\phi a.\phi b)}{p(\phi a)}, \text{ by Axiom 4} \\ &= \frac{1+\lambda\gamma_\phi}{1+\lambda}, \text{ by Prop. 19} \\ &> \gamma_\phi, \text{ by Prop. 12} \\ &= p(\phi b), \text{ by Theorem 5.} \end{aligned}$$

So by Definition 1, $C(\phi b, \phi a, T)$. Hence by Definition 3, ϕ is absolutely projectable.

12.17. Proof of Theorem 15

$$\begin{aligned} p(G'a.G'b|I) &= p(Q_1a.Q_1b|I) + p(Q_1a.Q_4b|I) + p(Q_4a.Q_1b|I) + \\ &\quad\; p(Q_4a.Q_4b|I), \text{ by Prop. 4} \\ &= \gamma_1 \frac{1+\lambda\gamma_F}{1+\lambda}\frac{1+\lambda\gamma_G}{1+\lambda} + 2\gamma_1\gamma_4\frac{\lambda^2}{(1+\lambda)^2} + \\ &\quad\; \gamma_4\frac{1+\lambda\gamma_{\bar{F}}}{1+\lambda}\frac{1+\lambda\gamma_{\bar{G}}}{1+\lambda}, \text{ by Axiom 4 and Theorem 4} \\ &= \gamma_1\left(\gamma_F + \frac{\gamma_{\bar{F}}}{1+\lambda}\right)\left(\gamma_G + \frac{\gamma_{\bar{G}}}{1+\lambda}\right) + 2\gamma_1\gamma_4\frac{\lambda^2}{(1+\lambda)^2} + \\ &\quad\; \gamma_4\left(\gamma_{\bar{F}} + \frac{\gamma_F}{1+\lambda}\right)\left(\gamma_{\bar{G}} + \frac{\gamma_G}{1+\lambda}\right) \\ &= (\gamma_1+\gamma_4)^2 + \frac{1}{1+\lambda}\left(\gamma_1\gamma_2 + \gamma_1\gamma_3 + \gamma_2\gamma_4 + \gamma_3\gamma_4 - 4\gamma_1\gamma_4\frac{\lambda}{1+\lambda}\right) \\ &> (\gamma_1+\gamma_4)^2 + \frac{1}{1+\lambda}\left(\gamma_1\gamma_2 + \gamma_1\gamma_3 + \gamma_2\gamma_4 + \gamma_3\gamma_4 - 4\gamma_1\gamma_4\right) \\ &= (\gamma_1+\gamma_4)^2 + \frac{1}{1+\lambda}[\gamma_G\gamma_{\bar{G}}(\gamma_F-\gamma_{\bar{F}})^2 + \gamma_F\gamma_{\bar{F}}(\gamma_G-\gamma_{\bar{G}})^2] \\ &\geq (\gamma_1+\gamma_4)^2. \end{aligned} \tag{10}$$

$$\begin{aligned} p(G'a.G'b|{\sim}I) &= p(Q_1a.Q_1b|{\sim}I) + p(Q_1a.Q_4b|{\sim}I) + p(Q_4a.Q_1b|{\sim}I) + \\ &\quad\; p(Q_4a.Q_4b|{\sim}I), \text{ by Prop. 4} \\ &= \gamma_1\frac{1+\lambda\gamma_1}{1+\lambda} + 2\gamma_1\gamma_4\frac{\lambda}{1+\lambda} + \gamma_4\frac{1+\lambda\gamma_4}{1+\lambda}, \\ &\quad\; \text{by Axiom 4 and Theorem 4} \\ &= \gamma_1\left(\gamma_1 + \frac{1-\gamma_1}{1+\lambda}\right) + 2\gamma_1\gamma_4\left(1 - \frac{1}{1+\lambda}\right) + \end{aligned}$$

$$\begin{aligned}
&\gamma_4\left(\gamma_4+\frac{1-\gamma_4}{1+\lambda}\right)\\
&= (\gamma_1+\gamma_4)^2+\frac{(\gamma_1+\gamma_4)(\gamma_2+\gamma_3)}{1+\lambda}\\
&> (\gamma_1+\gamma_4)^2, \text{ by Theorem 4.} \qquad (11)
\end{aligned}$$

$$\begin{aligned}
p(G'b|G'a) &= \frac{p(G'a.G'b)}{p(G'a)}, \text{ by Axiom 4}\\
&= \frac{p(G'a.G'b|I)p(I)+p(G'a.G'b|{\sim}I)p({\sim}I)}{p(G'a)}, \text{ by Prop. 5}\\
&> \frac{(\gamma_1+\gamma_4)^2p(I)+(\gamma_1+\gamma_4)^2p({\sim}I)}{p(G'a)}, \text{ by (10) and (11)}\\
&= \frac{(\gamma_1+\gamma_4)^2}{p(G'a)}, \text{ by Axiom 3}\\
&= p(G'b), \text{ since } p(G'a)=p(G'b)=\gamma_1+\gamma_4.
\end{aligned}$$

So by Definition 1, $C(G'b, G'a, T)$. Hence by Definition 3, G' is absolutely projectable.

12.18. Proof of Theorem 16

Interchanging F and $\sim F$ in Theorem 7 gives $C(Gb, Fa.Ga, {\sim}Fb)$. The proof of this is the same, *mutatis mutandis*, as the proof of Theorem 7. So by Definition 5, G is projectable across F.

$$\begin{aligned}
p(G'b|Fa.G'a.{\sim}Fb) &= \frac{p(G'b.{\sim}Fb|Fa.G'a)}{p({\sim}Fb|Fa.G'a)}, \text{ by Axiom 4}\\
&= \frac{p({\sim}Gb.{\sim}Fb|Fa.Ga)}{p({\sim}Fb|Fa.Ga)}, \text{ by Def. 4 and Axiom 5}\\
&= p({\sim}Gb|Fa.Ga.{\sim}Fb), \text{ by Axiom 4}\\
&= 1-p(Gb|Fa.Ga.{\sim}Fb), \text{ by Axiom 3}\\
&< 1-p(Gb|{\sim}Fb), \text{ since } G \text{ is projectable across } F\\
&= p({\sim}Gb|{\sim}Fb), \text{ by Axiom 3}\\
&= \frac{p({\sim}Gb.{\sim}Fb)}{p({\sim}Fb)}, \text{ by Axiom 4}\\
&= \frac{p(G'b.{\sim}Fb)}{p({\sim}Fb)}, \text{ by Def. 4 and Axiom 5}\\
&= p(G'b|{\sim}Fb), \text{ by Axiom 4.}
\end{aligned}$$

So by Definition 1, ${\sim}C(G'b, Fa.G'a, {\sim}Fb)$. Hence by Definition 5, G' is not projectable across F.

References

Carnap, R. (1950). *Logical Foundations of Probability.* Chicago: University of Chicago Press. Second edition 1962.

Carnap, R. (1980). A Basic System of Inductive Logic, Part II. In: R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability* Vol. 2 (pp. 7–155). Berkeley: University of California Press.

Flatto, L. (1976). *Advanced Calculus.* Baltimore: Williams & Wilkins.

Galileo Galilei (1638). *Discorsi e Dimostrazioni Matematiche intorno à due nuove scienze.* Leiden: Elsevier. (References are to the English translation, *Dialogues Concerning Two New Sciences*, by H. Crew and A. de Salvio, New York: Dover 1954.)

Good, I. J. (1968). The White Shoe *qua* Herring is Pink. *British Journal for the Philosophy of Science*, 19, 156–157.

Goodman, N. (1979). *Fact, Fiction, and Forecast* (3rd edition). Indianapolis: Hackett.

Hempel, C. G. (1945) Studies in the Logic of Confirmation. *Mind* **54**. (Page references are to the reprint in Hempel, 1965.)

Hempel, C. G. (1965). *Aspects of Scientific Explanation.* New York: The Free Press.

Huygens, C. (1690). *Traite de la Lumiere.* Leiden: Pierre vander Aa. (English translation, *Treatise on Light*, by S.P. Thompson, New York: Dover, 1962.)

Maher, P. (1999). Inductive Logic and the Ravens Paradox. *Philosophy of Science*, 66, 50–70.

Maher, P. (2000). Probabilities for Two Properties. *Erkenntnis*, 52, 63–91.

Maher, P. (2001). Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny. *Erkenntnis*, 55, 183–216.

Nicod, J. (1923). *Le Problème Logique de l'Induction.* Paris: Alcan. (English translation in Nicod, 1970.)

Nicod, J. (1970). *Geometry and Induction.* Berkely and Los Angeles: University of California Press. (English translation of works originally published in French in 1923 and 1924.)

von Wright, G. H. (1957). *The Logical Problem of Induction* (second edition). Oxford: Basil Blackwell.

Further reading

Corfield, D. and Williamson, J. (Eds). (2001). *Foundations of Bayesianism.* Dordrecht: Kluwer.

Earman, J. (1992). *Bayes or Bust?* Cambridge, Mass.: MIT Press.

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

Howson, C. and Urbach, P. (1993). *Scientific Reasoning: The Bayesian Approach* (second edition). Chicago: Open Court.

Jaynes, E.T. (2003). *Probability Theory: The Logic of Science.* Cambridge: Cambridge University Press.

Maher, P. (1990). How Prediction Enhances Confirmation. In J.M. Dunn and A. Gupta (Eds). *Truth or Consequences: Essays in Honor of Nuel Belnap* (pp. 327–343). Dordrecht: Kluwer.

Maher, P. (1996). Subjective and Objective Confirmation. *Philosophy of Science*, 63, 149–173.

PATRICK MAHER