Explication of Inductive Probability

Patrick Maher

Abstract Inductive probability is the logical concept of probability in ordinary language. It is vague but it can be explicated by defining a clear and precise concept that can serve some of the same purposes. This paper presents a general method for doing such an explication and then a particular explication due to Carnap. Common criticisms of Carnap's inductive logic are examined; it is shown that most of them are spurious and the others are not fundamental.

1 Introduction

The word "probability" in ordinary English has two senses, which I call *inductive probability* and *physical probability*. I will begin by briefly reviewing the main features of these concepts; for a more extensive discussion see Maher (2006, 2009).

The arguments of inductive probability are two propositions and we speak of the probability *of* proposition H *given* proposition E; I will abbreviate this as $ip(H|E)$. For example, if E is that a coin is either two-headed or two-tailed and is about to be tossed, while H is that the coin will land heads, then plausibly $ip(H|E) = 1/2$. By contrast, the arguments of physical probability are an *experiment type* and an *outcome type*, and we speak of the probability of an experiment of type X having an outcome of type O; I will abbreviate this as $pp_X(O)$. For example, if X is the experiment of tossing the coin just mentioned and O is the outcome that the coin lands heads, then $pp_X(O) = 0$ or 1, depending on what is on the coin.

Let an *elementary sentence* for any function be a sentence that says the function has a specific value for specific arguments. I call a function *logical* if all true elementary sentences for it are analytic. For example, if a function f is defined by specifying its values for all possible arguments then the truth value of all elementary sentences for f follows from the definition of f and hence f is a logical function.

Clearly physical probability isn't logical, since the value of $pp_X(O)$ in my preceding example depends on what is on the coin. On the other hand, in that example $ip(H|E) =$

Patrick Maher
University of Illinois at Urbana-Champaign, 105 Gregory Hall, 810 S. Wright Street, Urbana IL 61801, USA, E-mail: patrick@maher1.net

1/2 regardless of what is on the coin. More generally, the value of $ip(H|E)$ is determined once H and E are fixed and hence inductive probability is logical.

In my coin example the inductive probability has a numeric value but many inductive probabilities don't have a numeric value. For example, the inductive probability that humans evolved in Africa, given what I know, doesn't have a numeric value. Furthermore, as Keynes (1921, 28–30) argued, some inductive probabilities are not even comparable. These facts should not be surprising, since inductive probability is a concept of ordinary language and many concepts of ordinary language are vague. However, they do make it difficult to reason rigorously about inductive probabilities, especially in complex situations. Fortunately, there is a methodology for mitigating this difficulty, namely, *explication* (Carnap 1950, 3–8).

Explication begins with a pre-existing vague concept; this is called the *explicandum*. That concept is *explicated* by identifying another concept, called the *explicatum*, that satisfies the following desiderata to a sufficient degree:

- It is clear and precise, not vague like the explicandum.
- It is similar enough to the explicandum that it can be used in place of the latter for some envisaged purposes.
- It permits the formulation of exceptionless generalizations; Carnap called this *fruitfulness*.
- It is as simple as is compatible with satisfying the preceding desiderata.

This paper will discuss how to explicate the concept of inductive probability. I begin with general considerations, then present a particular explication due to Carnap, then discuss common criticisms of Carnap.

2 General considerations

2.1 Domain of the explicatum

The propositions H and E for which $ip(H|E)$ is meaningful are enormously diverse and it isn't feasible to construct an explicatum with such a large and diverse domain. Fortunately, this is also not necessary, since the explicatum only needs to be usable for specific purposes. Therefore, the first step in explicating inductive probability is to specify a limited domain of pairs of propositions H and E for which we will explicate $ip(H|E)$.

I will assume that we aim to explicate all inductive probabilities of the form $ip(A|B.K)$, where A and B are in some algebra $\mathcal{A}$ of propositions and K is a fixed proposition.[1] Following common terminology, I will call K "background evidence," though it can be any proposition and need not be anyone's evidence. To take a very simple example, K could be that a coin will be tossed twice and land heads or tails on each toss, while $\mathcal{A}$ could be the smallest algebra that contains H_1 and H_2, where H_i is that the coin lands heads on the ith toss.

In that example K is a simple proposition but in other cases K could be a large body of information possessed by some person. If we are unable to express that information

---

[1] For any propositions A and B, I will denote the proposition that A is false by "$\sim A$," the proposition that A and B are both true by "$A.B$," and the proposition that at least one of A and B is true by "$A \vee B$." An *algebra* of propositions is a nonempty set of propositions with the property that, for every A and B that it contains, it also contains $\sim A$, $A.B$, and $A \vee B$.

in a proposition it can still be denoted "K" and treated like a proposition (Pearl 1990). At the other extreme, situations in which there is no background evidence can be handled by taking K to be an analytic proposition.

2.2 Form of the explicatum

In the method I am proposing, the explicatum for inductive probability will be a function that takes two elements of $\mathcal{A}$ as arguments and has real numbers as its values; I will call this function "p" and I will denote the value of p for arguments A and B by "$p(A|B)$." This function is to be defined in such a way that $p(A|B)$ is a good explicatum for $ip(A|B.K)$, for all A and B in $\mathcal{A}$. I don't include K in the second argument of p because it is fixed in any context.

The definition of p will consist of axioms that together specify the value of $p(A|B)$ for all A and B in $\mathcal{A}$. These values must be specified in a way that doesn't depend on contingent facts; for example, an axiom may state that $p(A|B)$ equals 1/5 but not that it equals the proportion of humans with blue eyes. By defining p in this way we ensure that it is logical and hence is, in this respect, like inductive probability.

The axioms that define p will include axioms that ensure p obeys the mathematical laws of probability. There are two reasons for this requirement. First, when inductive probabilities have numeric values they satisfy these laws, and we want $p(A|B)$ to equal $ip(A|B.K)$ when the latter has a numeric value, so we need p to satisfy the laws of probability when the corresponding inductive probabilities have numeric values. Second, a good explicatum is fruitful and simple, so it is desirable to have p satisfy the same laws even when the corresponding inductive probabilities lack numeric values.

In what follows it will be helpful to have some notation for logical relations. I will use "$A \Rightarrow B$" to mean that A logically implies B, that is, $\sim A \vee B$ is analytic. I will also use "$A \Leftrightarrow B$" to mean that A and B are logically equivalent, that is, $A \Rightarrow B$ and $B \Rightarrow A$.

The following axioms ensure that p satisfies the laws of probability; these are asserted for all A, B, C, and D in $\mathcal{A}$.[2]

Axiom 1 $p(A|B) \geq 0$.

Axiom 2 $p(A|A) = 1$.

Axiom 3 $p(A|B) + p(\sim A|B) = 1$, *provided $B.K$ is consistent*.

Axiom 4 $p(A.B|C) = p(A|C)\,p(B|A.C)$.

Axiom 5 *If $A.K \Leftrightarrow C.K$ and $B.K \Leftrightarrow D.K$ then $p(A|B) = p(C|D)$.*

One consequence of these axioms is the following:[3]

Theorem 1 *If $B.K \Rightarrow A$ then $p(A|B) = 1$.*

---

[2] Similar axiomatizations of probability have been given by von Wright (1957, 93), Carnap (1971, 38), and Roeper and Leblanc (1999, 11), though my formulation differs from all of them in some respects. Von Wright imposed the restriction to consistent evidence on Axiom 2 rather than Axiom 3, which has the result that Theorems 1 and 2 don't hold. Carnap took $p(A|C)$ to be undefined for inconsistent C. Roeper and Leblanc redundantly added Theorem 2 as an additional axiom. And none of these authors allowed for background evidence.

[3] All theorems are proved in Section 6 unless I refer to a proof elsewhere.

A corollary of this is:

Theorem 2 *If $K \Rightarrow \sim B$ then $p(A|B) = 1$.*

Hence if $B.K$ is inconsistent we have $p(A|B) + p(\sim A|B) = 1 + 1 = 2$; that is the reason for the proviso in Axiom 3.

Axioms 1–5 also entail the following additivity law

Theorem 3 *If $K \Rightarrow \sim(A.B)$ then $p(A \vee B|C) = p(A|C) + p(B|C)$, provided $C.K$ is consistent.*

2.3 Alternative formulations

The approach described in the two preceding subsections incorporates a number of choices that could be done differently. I will now indicate the main alternatives and my reasons for making the choices that I did.

I took the arguments of p to be propositions, that is, the kind of thing that is expressed by a declarative sentence and can be expressed by different sentences in different languages. One could instead take the arguments of p to be sentences of a formal language that can express the relevant propositions. I decided not to use the latter method because it requires attention to linguistic details that are a distraction from the main issues involved in explicating inductive probability. Also, the apparently greater concreteness and rigor involved in using sentences is mostly illusory, since we need to specify the semantics of the formal language and this is done ultimately by stating, in ordinary language, the propositions that are the meanings of the sentences of the formal language.

I treated the concept of a proposition as primitive but propositions could instead be identified with sets of possible states of affairs; the latter approach derives from Kolmogorov (1933) and is standard among mathematicians. I have not done this because it would require me to give an exposition of set theory and explain how propositions can be correlated with sets. Explanations are all the more necessary because this is not a natural way of representing propositions. Freudenthal (1974) makes further criticisms of the set representation.

I assumed that the ordinary language concept of propositions is sufficiently clear for present purposes. Some philosophers don't think so, and for them I would propose an explication along the lines sketched by Carnap (1956, 209–11). In this explication, the explicatum is neither sentences nor sets.

My way of accommodating background evidence is new, I believe, though it is merely an attempt to explicitly allow for a common Bayesian practice. An alternative approach would be to include K in $\mathcal{A}$ and have $p(A|B)$ defined only for those B in $\mathcal{A}$ that entail K; however, that is messier and isn't the way Bayesian probability models are normally formulated.

Most presentations of probability theory follow Kolmogorov (1933) in beginning with an unconditional function $p(\cdot)$. Kolmogorov's elementary axioms for this function, stated in my notation, are:

K1. $p(A) \geq 0$.
K2. If A is analytic then $p(A) = 1$.
K3. If $A.B$ is inconsistent then $p(A \vee B) = p(A) + p(B)$.

Conditional probability is then introduced by adopting as a definition:

K4. $p(A|B) = p(A.B)/p(B)$, provided $p(B) > 0$.

These axioms follow from mine, in the following sense:

Theorem 4 *If $p(A)$ is defined to be $p(A|T)$, where T is analytic, then K1–K4 all hold.*

Since all the usually-recognized elementary laws of probability follow from K1–K4, this theorem shows that those laws also follow from Axioms 1–5.

My main reason for not starting with unconditional probability is that K4 leaves $p(A|B)$ undefined when $p(B) = 0$, although $ip(A|B.K)$ can exist even when $ip(B|K) = 0$. For example, let X be tossing a coin, O that the coin lands heads, and R_r the proposition that $pp_X(O) = r$. Also let Xa be the proposition that a is a token of type X and similarly for Oa. Then, by the principle of direct inference (Maher 2009), $ip(Oa|R_r.Xa) = r$, even though $ip(R_r|Xa) = 0$ for most, if not all, r. See Hájek (2003) for further discussion of the drawbacks of taking unconditional probability as primitive, including attempts to evade the problem by using infinitesimals.

Even writers who take conditional probability as primitive often say it is undefined when the second argument is inconsistent, whereas I have taken $p(A|B)$ to be defined for all B in $\mathcal{A}$, including inconsistent B. This has no practical significance, since our evidence is always consistent, but it has some advantages in simplicity and uniformity. Also, if we think of conditional probability as a generalization of logical implication then, since $B \Rightarrow A$ for inconsistent B, we should likewise have $p(A|B) = 1$ for inconsistent B.

3 Carnap's Basic System

Axioms 1–5 imply that $p(A|B) = 1$ if $B.K \Rightarrow A$ and $p(A|B) = 0$ if $B.K \Rightarrow \sim A$ and $K \not\Rightarrow \sim B$. However, these axioms don't fix the value of $p(A|B)$ in any other case and so additional axioms must be added to complete the definition of p. Unlike Axioms 1–5, these additional axioms must depend on the content of K and the propositions in $\mathcal{A}$. I will now present an example of such additional axioms, due to Carnap.

Carnap called the explication of inductive probability "inductive logic" and he worked on it from the 1940s until his death in 1970. Most discussions of Carnap's inductive logic only talk about his early explications published between 1945 and 1952, though his later explications were much better. Here I will present one special case from Carnap's posthumous "Basic System of Inductive Logic" (1971; 1980). I won't always state things exactly the way Carnap did; in particular, I will restate Carnap's proposals in the notation I have been using.

3.1 Domain of the explicatum

Carnap (1971, 43) assumed there is a denumerable set of *individuals*, denoted a_1, a_2, ...; they could be balls in an urn, outcomes of tossing a die, birds, people, or almost anything else. It is assumed that the names "a_i" are chosen in such a way that, for $i \neq j$, it is analytic that a_i and a_j are different individuals.

Carnap (1971, 43) called a *type* of property a *modality*. Some examples of modalities are color (red, blue, ...), shape (square, cubical, ...), substance (iron, stone, wood, ...), and age in years (0, 1, 2, ...). The first three of these are qualitative and the last is quantitative. Other quantitative modalities include weight and height.

A *family of properties* is a set of properties that belong to one modality, are mutually exclusive, and jointly exhaustive. A *primitive property* is a property that isn't defined in terms of other properties in our analysis. In the explication I am presenting, Carnap (1971, 121) took the primitive properties to be the elements of a finite family of properties. These primitive properties will here be denoted $F_1, F_2, \ldots, F_k$.

Goodman (1979, 74) defined the predicate "grue" as follows: It applies to things examined before time t iff they are green, and to things not examined before t iff they are blue. For example, if t is the year 2000, then a green emerald that was examined in 1960 is grue and a green emerald that was first examined in 2001 isn't grue. Since grue is a combination of two modalities (color and time), and Carnap required the primitive properties to belong to one modality, grue cannot be one of Carnap's primitive properties.[4]

Philosophers are apt to object that we could regard a set of properties which includes grue as a modality. However, this is using "modality" in a sense other than what Carnap intended. Faced with this response, philosophers object that Carnap didn't define, with sufficient clarity and explicitness, what he meant by "modality." Carnap (1971, 74) acknowledged that "it would certainly be desirable to give further clarification for the concept of modality"; however, the explanations and examples he gave suffice to enable readers of good will to correctly identify, at least in most cases, which second-order properties are modalities in Carnap's intended sense. Note also that in any particular explication the primitive properties are specified by definition, without reference to the concept of modality, so the unclarity of the concept of modality doesn't prevent explications of inductive probability from being clear. The concept of modality is merely used by Carnap to give informal guidance on how to select primitive properties in a way that will result in a satisfactory explication of inductive probability.

An *atomic proposition* is a proposition that ascribes one of the primitive properties to one of the individuals. I will use "$F_i a_j$" to denote the atomic proposition that individual a_j has primitive property F_i. A *sample* is a finite set of individuals. A *sample proposition* is a conjunction of atomic propositions, one for each individual in some sample. For example, $F_2 a_1 . F_5 a_2$ is a sample proposition for the sample $\{a_1, a_2\}$. As a matter of formal convenience, we count the empty set as a sample and we deem an analytic proposition to be a sample proposition for the empty set.

We now fix the domain of the explicatum by taking $\mathcal{A}$ to be the algebra generated by the atomic propositions and taking K to be an analytic proposition. Note that $\mathcal{A}$ contains every sample proposition. Also, since K is analytic, no background evidence is assumed.

3.2 Definition of p

We have already partially defined p by Axioms 1–5. We will now complete the definition of p, for the domain of explication just described, by adding further axioms that were proposed by Carnap. As in Theorem 4, $p(A)$ here means $p(A|T)$, where T is analytic; similarly, $ip(A)$ will mean $ip(A|T)$. Also, E is here any sample proposition, i is any integer between 1 and k, and m and n are any positive integers.

---

[4]   Carnap (1971, 74) also had another objection to grue, which I am omitting here.

Carnap assumed that none of the F_i is infinitely precise (for example, specifying the exact wavelength of light reflected by an object). In that case, $ip(E) > 0$, for every sample proposition E. Hence Carnap (1971, 101) adopted:

Axiom 6 (Regularity) $p(E) > 0$.

The individuals are supposed to be identified in a way that carries no information about which primitive property any individual has. Therefore, permuting the individuals will not change the inductive probability of any sample proposition; for example, $ip(F_1a_3.F_2a_5) = ip(F_1a_5.F_2a_3)$. Hence Carnap (1971, 118) adopted:

Axiom 7 (Symmetry[5]) $p(E)$ *isn't changed by permuting individuals.*

A characteristic property of inductive probability is that evidence that one individual has a property raises the probability that other individuals have the same property. For example, evidence that one bird is white raises the probability that another bird is white. Hence Carnap (1971, 161) adopted:

Axiom 8 (Instantial relevance) $p(F_ia_n|E.F_ia_m) > p(F_ia_n|E)$ *provided E does not involve a_m or a_n.*

If someone is given the outcome of past tosses of a die and asked to state the probability that the die will come up six on the next toss, usually the person will take account of the number of past tosses and the number of these that have come up six, ignoring the specific results of the tosses that didn't come up six. This suggests that it would be appropriate to adopt:

Axiom 9 (λ-condition) *If a is any individual not involved in E then $p(F_ia|E)$ depends only on the number of individuals mentioned in E and the number that E says have F_i.*

But if, for example, F_1 is more similar to F_2 than to F_3, then reasoning by analogy suggests $ip(F_1a_1|F_2a_2) > ip(F_1a_1|F_3a_2)$, whereas Axiom 9 implies $p(F_1a_1|F_2a_2) = p(F_1a_1|F_3a_2)$. Carnap (1980, 84) was aware of this but considered the λ-condition to be appropriate when such differential similarity relations are insigificant.

Carnap (1980, §19) proved that Axioms 1–9 imply:

Theorem 5 ($\lambda\gamma$ theorem) *If $k > 2$ then there exist $\lambda > 0$ and $\gamma_1, \dots, \gamma_k \in (0, 1)$ such that the following holds: if E is a sample proposition for a sample of s individuals, s_i is the number of individuals to which E ascribes F_i, and a is any individual not involved in E, then*

$$p(F_ia|E) = \frac{s_i + \lambda\gamma_i}{s + \lambda}.$$

For example, if $\gamma_1 = 1/4$ and $\lambda = 2$ then

$$p(F_1a_4|F_1a_1.F_2a_2.F_3a_3) = \frac{1 + 2/4}{3 + 2} = \frac{3}{10}.$$

Extension of Theorem 5 to the case where $k = 2$ requires a further assumption (Carnap 1980, 98).

To get numeric values from Theorem 5 we must fix the values of λ and the γ_i. I'll now discuss how to to do that, starting with the γ_i.

---

[5] Following de Finetti (1937), this is also called *exchangeability.*

By setting $s = 0$ in Theorem 5, we see that $\gamma_i = p(F_i a)$; thus γ_i needs to be a good explicatum for the a priori inductive probability that something has F_i. Let the *attribute space* for the F_i be the logical space whose points are the most specific properties of the relevant modality. Carnap (1971, 43–45) noted that each F_i corresponds to a region of the attribute space and he proposed (1980, 33–34) that γ_i be set equal to the proportion of the attribute space that corresponds to F_i.

For example, suppose the F_i are colors; then the attribute space could be taken to be the unit cube whose axes represent the degree of saturation (from 0 to 1) of red, green, and blue. If F_1 is the color red, it occupies a region around the point $(1,0,0)$; if that region occupies 1/20 of the volume of the cube then we would set $\gamma_1 = 1/20$ (assuming that the object is monochromatic).

Of course, attribute spaces can be represented in different ways; for example, Carnap (1980, 8) described a representation of the space of colors different to the unit cube that I just mentioned. These different representations will, in general, lead to different values for the γ_i, but any natural representation is likely to give values for the γ_i that are sufficiently in accord with the vague a priori inductive probability that an individual has F_i. If that isn't so then we need to choose a representation that does give such values.

I now turn to λ. The formula in Theorem 5 can be rewritten as:

$$p(F_i a | E) = \left(\frac{s}{s + \lambda} \right) \frac{s_i}{s} + \left(\frac{\lambda}{s + \lambda} \right) \gamma_i.$$

This shows that $p(F_i a | E)$ is a weighted average of two factors, s_i/s and γ_i. The factor s_i/s is the relative frequency of F_i in the sample and hence is empirical, whereas γ_i is our explicatum for the a priori probability of $F_i a$, which is logical. The larger λ is, the more weight is put on the logical factor and the slower someone using p will learn from experience. In the limit as $\lambda \to \infty$, $p(F_i a | E) \to \gamma_i$ and there is no learning from experience; at the other extreme, as $\lambda \to 0$, $p(F_i a | E) \to s_i/s$. Carnap (1980, 107–19) considered the effect of different values of λ in a variety of examples and concluded that, in order for p to agree with inductive probability (to put it in my terms), λ should not be much less than 1 or much greater than 2. Since integer values are simplest, he further concluded that λ should be set equal to either 1 or 2. I think this is correct as far as it goes but we can go further, as follows.

A theorem of de Finetti shows that we can think of the individuals a_i as tokens of some experiment type X which has an unknown physical probability of giving an outcome of type F_i.[6] If $\gamma_i = 1/2$ then the expected value of $pp_X(F_i)$ must be 1/2 and it is then natural to explicate the a priori inductive probability distribution for $pp_X(F_i)$ as uniform from 0 to 1. These assumptions imply:

Theorem 6 (Rule of succession) *If E says that in a sample of s individuals all have F_i, then*

$$p(F_i a | E) = \frac{s + 1}{s + 2}.$$

But by Theorem 5 and the assumption that $\gamma_i = 1/2$, we also have:

$$p(F_i a | E) = \frac{s + \lambda/2}{s + \lambda}.$$

---

[6] The theorem is called de Finetti's representation theorem; Jeffrey (1971, 217–21) gives an exposition of it.

These two identities imply that $\lambda = 2$.

Having thus fixed the values of λ and the γ_i, we have fixed the value of $p(A|B)$ for all A and B in $\mathcal{A}$, and hence the explication is complete.

4 Spurious criticisms of Carnap

There are many criticisms of Carnap's inductive logic that are frequently repeated by philosophers; Hájek (2007, sec. 3.2) gives a useful compendium of them. Most of these criticisms are spurious, at least with respect to Carnap's Basic System, so in this section I will point out the errors in the spurious criticisms presented by Hájek.

4.1 Arbitrariness

Hájek writes:

> Is there a correct setting of λ, or said another way, how "inductive" should the confirmation function be? The concern here is that any particular setting of λ is arbitrary in a way that compromises Carnap's claim to be offering a *logical* notion of probability.

But the choice of λ isn't arbitrary; it is designed to ensure that p is a good explicatum for inductive probability and I have argued that setting $\lambda = 2$ is best for this purpose. Furthermore, even if the choice of λ were arbitrary, p would still be logical in the sense that Carnap (1950, 30) claimed, because its values are specified by its definition in a way that doesn't depend on contingent facts.[7]

A little later Hájek expands on the objection this way:

> The whole point of the theory of logical probability is to explicate ampliative inference, although given the apparent arbitrariness in the choice of language and in the setting of λ—thus, in the choice of confirmation function—one may wonder how well it achieves this.

Here Hájek suggests that, in addition to the alleged arbitrariness in the choice of λ, there is also "arbitrariness in the choice of language." My presentation has used propositions rather than sentences of a language but, abstracting from this detail, the objection is that the choice of the domain of p is arbitrary. However, if $\mathcal{A}$ is chosen to contain the propositions whose inductive probabilities we want to explicate, as I proposed in Section 2.1, then the choice isn't arbitrary. Furthermore, even if the choice were arbitrary, that wouldn't prevent p being a good explicatum within its domain.

Hájek believes that in Carnap's inductive logic, the value of $p(H|E)$, for fixed H and E, changes when new predicates are added to the language.[8] Since the new predicates do not appear in H or E, our decision to include or exclude them from the language is

---

[7] All of the four philosophers who have written comments on this paper have objected that Hájek doesn't mean what Carnap means by "logical." So I need to remind readers that Hájek criticized "Carnap's claim" and that what claim a person made is determined by the meanings *that person* was using. I may add that, while Carnap explained what *he* meant by "logical," I have seen no account of what Hájek means by this term.

[8] Hájek states later that "by Carnap's lights, the degree of confirmation of a hypothesis depends on the language in which the hypothesis is stated" and he gives as examples "the addition of new predicates and the deletion of old ones."

irrelevant to $ip(H|E)$. Thus I think the objection Hájek intended to make is not what he said (that the choice of language is arbitrary) but rather that the value of $p(H|E)$ depends on irrelevant features of the language (or of the algebra $\mathcal{A}$). The answer to this objection is that there is no such dependence in Carnap's Basic System. In the special case that I presented, the primitive properties were required to belong to one family, so new ones can only be added by replacing existing ones. For example, we might subdivide an existing property into several more specific properties. Doing that will not change λ or the γ_i for the F_i that have not been replaced, hence it will not change $p(H|E)$ for any H and E that don't involve the new properties. We can also enrich $\mathcal{A}$ by allowing more than one family of properties; I haven't discussed how to do that but Carnap did and the proposals he made ensure that the value of $p(H|E)$, for given H and E, isn't altered by adding new families of properties (Carnap 1971, 46).

4.2 Axioms of symmetry

Hájek writes:

> Significantly, Carnap's various axioms of symmetry are hardly logical truths.

In the explication I have described there is just one "axiom of symmetry," namely Axiom 7. That axiom, like all the other axioms, is part of the definition of p, hence analytic, and in that sense a logical truth. Furthermore, if there were additional symmetry axioms, they would also be part of the definition of p and hence also logical truths.

Hájek continues:

> Moreover, Fine (1973, 202) argues that we cannot impose further symmetry constraints that are seemingly just as plausible as Carnap's, on pain of inconsistency.

There are two things wrong with this. First:

Theorem 7 *There are uncountably many probability functions that satisfy all the constraints that Fine (1973, 193) claimed are not jointly satisfiable.*

Second, one of Fine's constraints (his L6) is not something that an explicatum for inductive probability should satisfy. It implies that all γ_i have the same value, which is not desirable in general. It also implies that, when there are multiple families of properties, the explicatum is insensitive to analogies between individuals that the evidence says differ in any respect, which is never desirable.[9]

4.3 Syntactic approach

Hájek writes:

> Another Goodmanian lesson is that inductive logic must be sensitive to the meanings of predicates, strongly suggesting that a purely syntactic approach such as Carnap's is doomed.

---

[9] The problem here is essentially the one discussed in Maher (2001, sec. 3).

This criticism assumes that Carnap's inductive logic uses "a purely syntactic approach," that is, it assigns p values to pairs of expressions based on the form of the expression, without regard to what the expression means. However, Carnap's Basic System assigns p values to pairs of propositions, not expressions; hence it isn't a "syntactic approach."

Hájek's criticism seems to be that, because of its allegedly syntactic approach, Carnap's inductive logic is unable to distinguish between predicates like Goodman's "grue" and normal predicates like "green" and "blue." Stated non-linguistically, the objection would be that Carnap has no way of distinguishing properties like grue from normal properties like green and blue. But Carnap did distinguish between these properties, as we saw in Section 3.1.

4.4 No canonical language

Hájek writes:

> Finding a canonical language seems to many to be a pipe dream, at least if we want to analyze the "logical probability" of any argument of real interest— either in science, or in everyday life.

This objection appears to assume that Carnap's inductive logic requires a "canonical language," though Hájek does not explain what this is or why he thinks Carnap is committed to it. In fact, one of Carnap's central philosophical principles was that there is no uniquely correct or right language.

> Everyone is at liberty to build up his own logic, i.e. his own form of language, as he wishes. (Carnap 1937, 52)

> Let us grant to those who work in any special field of investigation the freedom to use any form of expression which seems useful to them. (Carnap 1956, 221)

I've been told that Hájek's point here is that, since Carnap allows green but not grue to be a primitive property, his explication is "language-relative," whereas a good explication shouldn't be language-relative. But in the explication I presented, p is a function of two propositions and there is no reference to any language, so there is no legitimate sense in which p is language-relative. It is true is that $p(A|B)$ depends on what the propositions A and B are; for example, if A says that an individual is green and A' says it is grue then $p(A|B)$ may differ from $p(A'|B)$, but in this respect p agrees with inductive probability, so this isn't a defect of p.[10]

I've also been told that Hájek's point here is that the explicata actually developed by Carnap have very simple domains that aren't applicable to arguments of real interest to science. This is a legitimate criticism and I will discuss it in Section 5. It is not, however, what the preceding quotation from Hájek says.

---

[10]   In Carnap (1971), the explicatum is likewise a function of propositions. It is true that Carnap there represents the propositions as models of a language, but this is merely a way of specifying propositions and doesn't make the explicatum relative to the language.

4.5 Total evidence isn't well defined

Hájek writes:

> If one's credences are to be based on logical probabilities, they must be relativized to an evidence statement, e. But which is it to be? Carnap's recommendation is that e should be one's *total evidence* ... However, when we go beyond toy examples, it is not clear that this is well-defined. Suppose I have just watched a coin toss, and thus learned that the coin landed heads. Perhaps "the coin landed heads" is my total evidence? But I also learned a host of other things: as it might be, that the coin landed at a certain time, bouncing in a certain way, making a certain noise as it did so ... Call this long conjunction of facts X. I also learned a potentially infinite set of *de se* propositions: "I learned that X," "I learned that I learned that X" and so on. Perhaps, then, my total evidence is the infinite intersection of all these propositions, although this is still not obvious—and it is not something that can be represented by a sentence in one of Carnap's languages, which is finite in length.

It is true that the concept of a person's total evidence is vague, but most concepts of ordinary language are vague and that doesn't prevent them being useful. So I will take the objection to be that a person's total evidence is too vague or complex to be represented in Carnap's inductive logic.

One answer to this objection is that a person's total evidence in a given context may be *explicated* by a relatively precise proposition. Such an explication is satisfactory if it captures sufficiently well the part of the person's total evidence that is relevant to the hypotheses under consideration in that context. Hájek's sequence of *de se* propositions would normally be irrelevant and could be omitted. A second answer was mentioned in Section 2.1: we can simply denote a person's total evidence as "K," without attempting to articulate all that it contains, and explicate inductive probabilities conditional on K.

4.6 Foundationalism

Hájek continues:

> The total evidence criterion goes hand in hand with positivism and a foundationalist epistemology according to which there are such determinate, ultimate deliverances of experience. But perhaps learning does not come in the form of such "bedrock" propositions, as Jeffrey (1992) has argued—maybe it rather involves a shift in one's subjective probabilities across a partition, without any cell of the partition becoming certain.

Carnap (1936, 425; 1963, 57) denied that there are "bedrock" propositions. On the other hand, the inductive probability of any proposition given itself is 1, so if I use inductive probabilities given my evidence to guide my actions, I will act as if I am certain that my evidence is true. Carnap never explained how to reconcile these things.

The apparent contradiction can be resolved by recognizing that what we count as evidence isn't completely certain but only sufficiently certain that it can be treated as certain in the context at hand. Thus what counts as my evidence can change when the context changes. So if K is my total evidence in a particular context, then the principle

of total evidence implies that I should treat K as certain in that context but it doesn't imply that K is a "bedrock" proposition; on the contrary, there may be other contexts in which I need to consider the possibility that K is false, and K won't be evidence for me in those contexts. See Maher (1996, 158–62) for further discussion of this account of evidence.

Before moving on it may be worth noting that the requirement of total evidence has been, and continues to be, widely endorsed. Carnap (1950, 212) cited endorsements by Jacob Bernoulli, Peirce, and Keynes; more recent endorsements include the following:

> Your assignment of 1/2 to the coin landing heads superficially seems unconditional; but really it is conditional on tacit assumptions about the coin, the toss, the immediate environment, and so on. In fact, it is conditional on your total evidence. (Hájek 2003, 315)

> The point of view I maintain is based on the thesis that *it is senseless to speak of the probability of an event unless we do so in relation to the body of knowledge possessed by a given person.* (de Finetti 2008, 3)

So if there were a problem with the requirement of total evidence, it would not be a problem peculiar to Carnap.

4.7 Circularity

Hájek writes:

> By Carnap's lights, the degree of confirmation of a hypothesis depends on the language in which the hypothesis is stated and over which the confirmation function is defined. But scientific progress often brings with it a change in scientific language (for example, the addition of new predicates and the deletion of old ones), and such a change will bring with it a change in the corresponding c-values. Thus, the growth of science may overthrow any particular confirmation theory. There is something of the snake eating its own tail here, since logical probability was supposed to explicate the confirmation of scientific theories.

As I pointed out in Section 4.1, Carnap's Basic System isn't language sensitive in the way Hájek here supposes, that is, with respect to addition or deletion of predicates. However, referees tell me that Hájek's objection can be reformulated with reference to modalities. In this revised version, the objection starts from the premise that the growth of science may change what the modalities are, and infers that Carnap's inductive logic has some undesirable kind of circularity. I will now discuss this revised version of the objection.

Suppose we have defined an explicatum p for inductive probability in the way Carnap recommends. Thus, the primitive properties are the members of a set of families of properties, with the properties in each family belonging to one modality. Now suppose that the growth of science somehow makes some former non-modalities count as modalities; this doesn't require p to be changed in any way, since the definition of p still satisfies Carnap's requirements. Thus, in order for the growth of science to force a change in p, it must cause some former modalities to become non-modalities.

Humans have learned a vast amount about color over the past four centuries but those discoveries haven't prevented color from continuing to be a modality. Similarly

with the modalities of substance and weight. Is there then any example of the growth of science causing a modality to cease to be one? Perhaps someone will claim that the quantity of phlogiston in a body was once a modality and no longer is. But if "quantity of phlogiston" refers to the amount of the hypothetical substance phlogiston in a body, then what modern science tells us is that all bodies are alike in having none of it; this doesn't prevent quantity of phlogiston from being a modality, it just makes it a modality that is useless for most scientific purposes.

I haven't been able to think of any example of the growth of science causing a modality to cease to be one. But with no examples, it is hard to think clearly about what the consequences of such a development would be for Carnap's inductive logic. So then, let the critics give us an example of the growth of science causing a modality to become a non-modality; when they have done so, then it will make sense to consider the consequences of such a development for Carnap's inductive logic.

5 Legitimate criticisms of Carnap

Although most of the common criticisms of Carnap's inductive logic are spurious, there are some legitimate criticisms of it. I will discuss them in this section.

The explication that I presented in Section 3 has the property that, for any sample proposition E, $p(F_i a_1 \ldots F_i a_n | E) \to 0$ as $n \to \infty$. So, if we were to add to $\mathcal{A}$ a proposition A_i that all individuals have F_i, we would have $p(A_i|E) = 0$, for every sample proposition E. This is true of all Carnap's explications of inductive probability and many authors (including Hájek) regard it as unsatisfactory. However, there are a variety of ways of modifying Carnap's explications to avoid this result (Zabell 1997) and Carnap (1980, 145) himself had this issue on his agenda, so this defect (if it is one) is correctable, not fundamental.

Another legitimate criticism is that the explicata developed by Carnap have very simple domains and, as a result, aren't applicable to most situations of real interest. For example, Carnap never developed explications for domains involving relations or continuous magnitudes, or for situations with rich background evidence, though these are all common in science and everyday life. While this is true, it is merely a fact about the explications that Carnap actually developed; it doesn't show that the methodology of explicating inductive probability is similarly limited. On the contrary, Bayesian statisticians have developed probability models, which I would interpret as explications of inductive probability, for a wide variety of realistic domains; there are many examples in Gelman et al. (2003), Congdon (2007), and elsewhere. Furthermore, an explication of inductive probability for an artificially simple domain isn't necessarily useless, since it may help to clarify fundamental questions about the properties of confirmation and resolve philosophical paradoxes, as I have shown elsewhere.[11]

I conclude that explication of inductive probability is a valuable methodology for reasoning about inductive probability and that the particular explication of Carnap's that I have presented is a creditable simple example of such an explication. Contrary to what most philosophers today believe, Carnap's inductive logic isn't fundamentally misconceived.

---

[11] See Maher (2004). Today I would replace the term "justified degree of belief" used in that paper with "inductive probability."

6 Proofs

6.1 Proof of Theorem 1

$$
\begin{aligned}
p(A|B) &= p(A|B)p(B|B) && \text{by Axiom 2} \\
&= p(A|B)p(B|A.B) && \text{by Axiom 5 and } B.K \Rightarrow A \\
&= p(A.B|B) && \text{by Axiom 4} \\
&= p(B|B) && \text{by Axiom 5 and } B.K \Rightarrow A \\
&= 1 && \text{by Axiom 2.}
\end{aligned}
$$

6.2 Proof of Theorem 2

If $K \Rightarrow \sim B$ then $B.K$ is inconsistent, so trivially $B.K \Rightarrow A$, and hence $p(A|B) = 1$ by Theorem 1.

6.3 Lemmas used in the proof of Theorem 3

Lemma 1 *If $C.K \Rightarrow A$ then $p(\sim A|C) = 0$, provided $C.K$ is consistent.*

Proof:

$$
\begin{aligned}
p(\sim A|C) &= 1 - p(A|C) && \text{by Axiom 3} \\
&= 1 - 1 && \text{by Theorem 1 and } C.K \Rightarrow A \\
&= 0.
\end{aligned}
$$

Lemma 2 *$p(A|C) = p(A.B|C) + p(A.\sim B|C)$, provided $C.K$ is consistent.*

Proof: If $A.C.K$ is consistent then

$$
\begin{aligned}
p(A|C) &= p(A|C)[p(B|A.C) + p(\sim B|A.C)] && \text{by Axiom 3} \\
&= p(A.B|C) + p(A.\sim B|C) && \text{by Axiom 4.}
\end{aligned}
$$

If $A.C.K$ is inconsistent then $C.K \Rightarrow \sim A$ so, by Lemma 1, all quantities in Lemma 2 are zero.

6.4 Proof of Theorem 3

$$
\begin{aligned}
p(A \vee B|C) &= p[(A \vee B).A|C] + p[(A \vee B).\sim A|C] && \text{by Lemma 2} \\
&= p(A|C) + p(B.\sim A|C) && \text{by Axiom 5} \\
&= p(A|C) + p(B|C) && \text{by Axiom 5 and } K \Rightarrow \sim(A.B).
\end{aligned}
$$

6.5 Proof of Theorem 4

K1 follows from Axiom 1, K2 from Axiom 2, K3 from Theorem 3, and K4 from Axiom 4.

6.6 Proof of Theorem 6

Let E be $Oa_1 \ldots Oa_n$, let K be $Xa_1 \ldots Xa_{n+1}$, and let R_r be the proposition that $pp_X(O) = r$. Then we have:

$$
\begin{aligned}
ip(E|R_r.K) &= ip(Oa_1 \ldots Oa_n|R_r.K) \\
&= ip(Oa_1|R_r.K).ip(Oa_2|R_r.K.Oa_1) \ldots ip(Oa_n|R_r.K.Oa_1 \ldots Oa_{n-1}) \\
&= r^n, \quad \text{by Maher (2009, Theorem 4 and Definition 1).} \tag{1}
\end{aligned}
$$

For any proposition A, let

$$
ip'(R_r|A) = \lim_{\delta \to 0^+} \frac{ip(r \le pp_X(O) \le r + \delta|A)}{\delta}.
$$

The assumption that the a priori inductive probability distribution of $pp_X(O)$ is uniform on $[0, 1]$ implies:

$$
ip'(R_r|K) = 1 \text{ for all } r \in [0, 1). \tag{2}
$$

Applying a generalized form of Bayes's theorem[12] we have, for all $r \in [0, 1)$:

$$
\begin{aligned}
ip'(R_r|E.K) &= \frac{ip(E|R_r.K) \, ip'(R_r|K)}{\int_0^1 ip(E|R_s.K) \, ip'(R_s|K) \, ds} \\
&= \frac{ip(E|R_r.K)}{\int_0^1 ip(E|R_s.K) \, ds}, \quad \text{by (2)} \\
&= \frac{r^n}{\int_0^1 s^n \, ds}, \quad \text{by (1)} \\
&= (n+1)r^n. \tag{3}
\end{aligned}
$$

Applying a generalized form of the law of total probability, we now have:

$$
\begin{aligned}
ip(Oa_{n+1}|E.K) &= \int_0^1 ip(Oa_{n+1}|E.R_r.K) \, ip'(R_r|E.K) \, dr \\
&= \int_0^1 ip(Oa_{n+1}|E.R_r.K)(n+1)r^n \, dr, \quad \text{by (3)} \\
&= \int_0^1 (n+1)r^{n+1} \, dr, \quad \text{by Maher (2009, Theorem 4 and Definition 1)} \\
&= \frac{n+1}{n+2}.
\end{aligned}
$$

---

[12]   The generalization consists in replacing summation with integration. Derivations of this generalized version can be found in many textbooks of probability and statistics. Similarly for the generalized law of total probability used below.

6.7 Proof of Theorem 7

Lemma 5 of Fine (1973, 193) asserts that five constraints, called L1, L2, L3, L6, and L7 by Fine, are not jointly satisfiable. I will show that, on the contrary, there are uncountably many functions that satisfy those constraints. To facilitate comparison with Fine's text, I will mostly use Fine's notation in this proof.

Let a *family of properties* be a finite set of properties that are pairwise exclusive and jointly exhaustive.[13] Let $\mathcal{P}^1, \ldots, \mathcal{P}^n$ be logically independent families of properties with $\mathcal{P}^i = \{P_1^i, \ldots, P_{k_i}^i\}$. Let $P_{l_1 \ldots l_n}^{1 \ldots n}$ be the property of having all of $P_{l_1}^1, \ldots, P_{l_n}^n$ and let

$$\mathcal{P}^{1 \ldots n} = \{P_{l_1 \ldots l_n}^{1 \ldots n} : \ 1 \leq l_j \leq k_j, \ j = 1, \ldots, n\}.$$

Thus $\mathcal{P}^{1 \ldots n}$ is a family of properties formed by combining $\mathcal{P}^1, \ldots, \mathcal{P}^n$.

Let a finite set of individuals be given and, for any property ϕ, let ϕa be the proposition that individual a has ϕ. A proposition of the form $P_l^i a$ will be called an *atomic proposition*. Let a *sample proposition* with respect of family of properties $\mathcal{P}$ be a proposition that ascribes a property from $\mathcal{P}$ to each member of some sample. Logically true propositions will be regarded as sample propositions with respect to any family of properties, the sample in this case being the empty set.

Let $\mathcal{A}$ be the algebra of propositions generated by the atomic propositions and let $C(H|E)$ be defined for all $H \in \mathcal{A}$ and consistent $E \in \mathcal{A}$ by the following axioms. Here H and H' are any propositions in $\mathcal{A}$, E and E' are any consistent propositions in $\mathcal{A}$, and λ is any positive real constant.

A1. If E is a sample proposition with respect to $\mathcal{P}^{1 \ldots n}$ for a sample of s individuals, $s_{l_1 \ldots l_n}^{1 \ldots n}$ is the number of individuals to which E ascribes $P_{l_1 \ldots l_n}^{1 \ldots n}$, and a is any individual not involved in E, then

$$C(P_{l_1 \ldots l_n}^{1 \ldots n} a | E) = \frac{s_{l_1 \ldots l_n}^{1 \ldots n} + \lambda / k_1 \ldots k_n}{s + \lambda}.$$

A2. If $H \Leftrightarrow H'$ and $E \Leftrightarrow E'$ then $C(H|E) = C(H'|E')$.
A3. If $E \Rightarrow \sim(H.H')$ then $C(H \vee H'|E) = C(H|E) + C(H'|E)$.
A4. $C(H.E'|E) = C(H|E.E') \, C(E'|E)$.

Let a *state* be a proposition of the form $\phi_1 a_1 \ldots \phi_\nu a_\nu$, where each ϕ_i is a property in $\mathcal{P}^{1 \ldots n}$ and $a_1, \ldots, a_\nu$ are all the individuals. Letting $C(H)$ be an abbreviation for $C(H|T)$, where T is any logically true proposition, repeated application of A4 gives:

$$C(\phi_1 a_1 \ldots \phi_\nu a_\nu) = C(\phi_1 a_1) C(\phi_2 a_2 | \phi_1 a_1) \ldots C(\phi_\nu a_\nu | \phi_1 a_1 \ldots \phi_{\nu-1} a_{\nu-1}).$$

The value of each term on the right hand side is given by A1, hence the axioms fix the value of $C(S)$ for every state S. Every consistent $H \in \mathcal{A}$ is equivalent to a disjunction of states so, letting "S" be a variable ranging over states, we have by A2 and A3:

$$C(H) = \sum_{S \Rightarrow H} C(S), \quad \text{for all consistent } H \in \mathcal{A}. \tag{4}$$

---

[13]   This is a looser definition than the one in Section 2.1, since it doesn't require the properties to belong to one modality. It agrees with Fine's definition but not with that of Carnap (1971, 43) and will only be used in the present proof.

If H is inconsistent then

$$\begin{aligned}
C(H) &= C(H \vee \sim H) - C(\sim H), \quad \text{by A3} \\
&= C(\sim H) - C(\sim H), \quad \text{by A2} \\
&= 0.
\end{aligned}$$

Combining this with (4), we have:

$$C(H) = \sum_{S \Rightarrow H} C(S), \quad \text{for all } H \in \mathcal{A}. \tag{5}$$

By A1, $C(S) > 0$ for all states S and hence $C(E) > 0$ for all consistent $E \in \mathcal{A}$. Therefore, by A4, we have

$$C(H|E) = C(H.E)/C(E), \quad \text{for all consistent } E \in \mathcal{A}. \tag{6}$$

Since the values of $C(H.E)$ and $C(E)$ are given by (5), this shows that A1–A4 fix the value of $C(H|E)$ for all $H \in \mathcal{A}$ and consistent $E \in \mathcal{A}$. I will now show that C satisfies the constraints that Fine claimed could not be jointly satisfied.

L1

Theorem 2 of Fine (1973, 189) states that L1 is equivalent to a conjunction of five conditions. Three of these are identical to A2, A3, and A4; the other two are the following (asserted for $H, H' \in \mathcal{A}$ and consistent $E, E', E'' \in \mathcal{A}$):

(i) $0 \leq C(H|E) < \infty$.
(ii) If $E \Rightarrow H$ and $E' \nRightarrow H'$ then $C(H|E) > C(H'|E')$.

I will now show that both these conditions are satisfied.
   Proof of (i):

$$\begin{aligned}
C(H|E) &= \frac{C(H.E)}{C(E)}, \quad \text{by (6)} \\
&= \frac{\sum_{S \Rightarrow H.E} C(S)}{\sum_{S \Rightarrow E} C(S)}, \quad \text{by (5)}.
\end{aligned}$$

There are at least as many terms in the denominator as in the numerator and, by A1, each term is positive. Hence $0 \leq C(H|E) \leq 1$, which entails (i).
   Proof of (ii): Assume $E \Rightarrow H$ and $E' \nRightarrow H'$. Then:

$$\begin{aligned}
C(H|E) &= \frac{C(H.E)}{C(E)}, \quad \text{by (6)} \\
&= \frac{C(E)}{C(E)}, \quad \text{by A2 and } E \Rightarrow H \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
C(H'|E') &= \frac{C(H'.E')}{C(E')}, \quad \text{by (6)} \\
&= \frac{\sum_{S \Rightarrow H'.E'} C(S)}{\sum_{S \Rightarrow E'} C(S)}, \quad \text{by (5)}.
\end{aligned}$$

Since $E' \nRightarrow H'$, the terms in the numerator are a proper subset of those in the denominator and so, since all terms are positive, $C(H'|E') < 1$. Hence $C(H|E) > C(H'|E')$.

L2

L2 says $C(H|E)$ is invariant under any permutation of the individuals. It is satisfied because A1–A4 treat all individuals alike.

L3

L3 says that $C(H|E)$ is invariant under augmentation of the set of individuals.[14] This is satisfied because none of A1–A4 refers to the total number of individuals.

L6

L6 says that $C(H|E)$ is invariant under any permutation of properties in $\mathcal{P}^{1\ldots n}$. We have seen that
$$C(H|E) = \frac{\sum_{S \Rightarrow H.E} C(S)}{\sum_{S \Rightarrow E} C(S)}.$$

Permuting the properties in $\mathcal{P}^{1\ldots n}$ will not change the number of states S that entail $H.E$ or E and, by A1 and A4, it will not change the value of $C(S)$ for any S. Therefore, $C(H|E)$ will not change and L6 is satisfied.

L7

L7 says that $C(H|E)$ is invariant under augmentation of the set of families of properties. So let $\mathcal{P}^{n+1} = \{P_1^{n+1}, \ldots, P_{k_{n+1}}^{n+1}\}$ be a family of properties that is logically independent of $\mathcal{P}^{1\ldots n}$. Let $\mathcal{A}'$ be the algebra generated by propositions of the form $P_{l_i}^i a$, where $1 \leq i \leq n+1$ and $1 \leq l_i \leq k_i$. Let $C'(H|E)$ be defined for all $H \in \mathcal{A}'$ and consistent $E \in \mathcal{A}'$ by the following axioms.

A1′. If E is a sample proposition with respect to $\mathcal{P}^{1\ldots n+1}$ for a sample of s individuals, $s_{l_1\ldots l_{n+1}}^{1\ldots n+1}$ is the number of individuals to which E ascribes $P_{l_1\ldots l_{n+1}}^{1\ldots n+1}$, and a is any individual not involved in E, then

$$C'(P_{l_1\ldots l_{n+1}}^{1\ldots n+1} a|E) = \frac{s_{l_1\ldots l_{n+1}}^{1\ldots n+1} + \lambda/k_1 \ldots k_{n+1}}{s + \lambda}.$$

A2′. If $H \Leftrightarrow H'$ and $E \Leftrightarrow E'$ then $C'(H|E) = C'(H'|E')$.
A3′. If $E \Rightarrow \sim(H.H')$ then $C'(H \vee H'|E) = C'(H|E) + C'(H'|E)$.
A4′. $C'(H.E'|E) = C'(H|E.E')\, C'(E'|E)$.

These axioms fix the value of $C'(H|E)$ for all $H \in \mathcal{A}'$ and consistent $E \in \mathcal{A}'$; the proof is exactly analogous to the proof that A1–A4 fix the value of $C(H|E)$ for all $H \in \mathcal{A}$ and consistent $E \in \mathcal{A}$. I will now show that C' agrees with C on $\mathcal{A}$.

Let E be a sample proposition with respect to $\mathcal{P}^{1\ldots n}$ for a sample of s individuals and let E' be any sample proposition with respect to $\mathcal{P}^{1\ldots n+1}$ that involves the same individuals as E and is such that $E' \Rightarrow E$. Then for any individual a not involved in E:

---

[14] I omit Fine's qualification that H and E not contain universal or existential quantifiers because $\mathcal{A}$ contains only truth-functional combinations of atomic propositions.

$$C'(P_{l_1\ldots l_n}^{1\ldots n}a|E) = \sum_{E'} C'(P_{l_1\ldots l_n}^{1\ldots n}a|E')C'(E'|E), \quad \text{by A2}'\text{–A4}'$$

$$= \sum_{E'}\sum_{l_{n+1}} C'(P_{l_1\ldots l_{n+1}}^{1\ldots n+1}a|E')C'(E'|E), \quad \text{by A2}' \text{ and A3}'$$

$$= \sum_{E'}\sum_{l_{n+1}} \frac{s_{l_1\ldots l_{n+1}}^{1\ldots n+1} + \lambda/k_1\ldots k_{n+1}}{s+\lambda} C'(E'|E), \quad \text{by A1}'$$

$$= \frac{s_{l_1\ldots l_n}^{1\ldots n} + \lambda/k_1\ldots k_n}{s+\lambda} \sum_{E'} C'(E'|E)$$

$$= \frac{s_{l_1\ldots l_n}^{1\ldots n} + \lambda/k_1\ldots k_n}{s+\lambda}, \quad \text{by A2}'\text{–A4}'.$$

Hence C' satisfies the proposition that results from substituting "C'" for "C" in A1. The same is obviously true for A2–A4. Hence $C'(H|E) = C(H|E)$ for all $H \in \mathcal{A}$ and consistent $E \in \mathcal{A}$, so L7 is satisfied.

This completes the proof that C satisfies all the constraints in Fine's Lemma 5. Since λ can be any positive real number, and each choice of λ gives a different C, it follows that there are uncountably many functions that satisfy Fine's constraints.

References

Carnap, Rudolf. 1936. Testability and meaning, parts i–iii. *Philosophy of Science* 3:419–471. Reprinted by the Graduate Philosophy Club, Yale University, 1950.

———. 1937. *The Logical Syntax of Language*. Routledge and Kegan Paul. Translated by Amethe Smeaton.

———. 1945. On inductive logic. *Philosophy of Science* 12:72–97.

———. 1950. *Logical Foundations of Probability*. University of Chicago Press. 2nd ed. 1962.

———. 1952. *The Continuum of Inductive Methods*. University of Chicago Press.

———. 1956. *Meaning and Necessity*. University of Chicago Press, 2nd ed.

———. 1963. Intellectual autobiography. In *The Philosophy of Rudolf Carnap*, ed. Paul Arthur Schilpp, Open Court, 1–84.

———. 1971. A basic system of inductive logic, part I. In Carnap and Jeffrey (1971), 33–165.

———. 1980. A basic system of inductive logic, part II. In *Studies in Inductive Logic and Probability*, ed. Richard C. Jeffrey, University of California Press, vol. 2, 7–155.

Carnap, Rudolf and Jeffrey, Richard, eds. 1971. *Studies in Inductive Logic and Probability*, vol. 1. Berkeley: University of California Press.

Congdon, Peter. 2007. *Bayesian Statistical Modelling*. Wiley, 2nd ed.

de Finetti, Bruno. 1937. La prevision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* 7:1–68. English translation in Kyburg and Smokler (1980).

———. 2008. *Philosophical Lectures on Probability*. Springer.

Fine, Terrence L. 1973. *Theories of Probability*. Academic Press.

Freudenthal, Hans. 1974. The crux of course design in probability. *Educational Studies in Mathematics* 5:261–277. Errata in vol. 6 (1975), p. 125.

Gelman, Andrew, Carlin, John B., Stern, Hal S., and Rubin, Donald B. 2003. *Bayesian Data Analysis*. Chapman & Hall, 2nd ed.

Goodman, Nelson. 1979. *Fact, Fiction, and Forecast*. Hackett, 3rd ed.

Hájek, Alan. 2003. What conditional probability could not be. *Synthese* 137:273–323.

———. 2007. Interpretations of probability. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, http://plato.stanford.edu/archives/fall2007/-entries/probability-interpret/.

Jeffrey, Richard. 1971. Probability measures and integrals. In Carnap and Jeffrey (1971), 167–223.

———. 1992. *Probability and the Art of Judgment*. Cambridge University Press.

Keynes, John Maynard. 1921. *A Treatise on Probability*. Macmillan. Reprinted with corrections 1948.

Kolmogorov, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. English translation: *Foundations of Probability*, trans. Nathan Morrison, Chelsea 1956.

Kyburg, Henry E., Jr. and Smokler, Howard E., eds. 1980. *Studies in Subjective Probability*. Krieger, 2nd ed.

Maher, Patrick. 1996. Subjective and objective confirmation. *Philosophy of Science* 63:149–173.

———. 2001. Probabilities for multiple properties: The models of Hesse and Carnap and Kemeny. *Erkenntnis* 55:183–216.

———. 2004. Probability captures the logic of scientific confirmation. In *Contemporary Debates in Philosophy of Science*, ed. Christopher R. Hitchcock, Blackwell, 69–93.

———. 2006. The concept of inductive probability. *Erkenntnis* 65:185–206.

———. 2009. Physical probability. In *Logic, Methodology and Philosophy of Science: Proceedings of the Thirteenth International Congress*, eds. Clark Glymour, Wei Wang, and Dag Westerståhl, Kings College Publications. To appear.

Pearl, Judea. 1990. Jeffrey's rule, passage of experience, and *Neo*-bayesianism. In *Knowledge Representation and Defeasible Reasoning*, eds. Henry E. Kyburg, Jr., Ronald P. Loui, and Greg N. Carlson, Kluwer, 245–265.

Roeper, P. and Leblanc, H. 1999. *Probability Theory and Probability Logic*. University of Toronto Press.

von Wright, Georg Henrik. 1957. *The Logical Problem of Induction*. Blackwell, 2nd ed.

Zabell, S. L. 1997. Confirming universal generalizations. *Erkenntnis* 45:267–283.