# 5

# Degree-of-Entailment Interpretations of Probability

A number of interpretations of probability have been offered which take probability to be a logical relation between an evidence statement (or a set of evidence statements) and a conclusion statement, in virtue of which having grounds for accepting the evidence statements constitutes partial grounds for accepting the conclusion statement, or grounds for according it a certain degree of rational belief. The first such theory to be worked out explicitly and in some symbolic detail is Keynes' theory (1921). Keynes' theory is sometimes mistakenly regarded as subjectivistic—particularly by frequency theorists—but there is a class of interpretations of probability which is subjective in a far more significant and interesting sense than Keynes' interpretation. Keynes, indeed, is at great pains to argue that the probability relation is *objective* and *logical*. Thinking so does not make something probable; it is probable only as a result of the objective logical relation it bears to its evidence.





Keynes accepts the conventional calculus of probability as holding for his probability relations, and presents it in an axiomatized form. But his probabilities are not restricted to real numbers. Probabilities are *sui generis*; although we symbolize certainty and impossibility by 1 and 0, although all other degrees of the probability relation lie between these limits, and although in many cases probabilities can be associated with real numbers, it is not by any means always the case that a probability can be associated with a real number. Although the probability of $h_1$, given evidence $e_1$, and the probability of $h_2$, given $e_2$, both lie between 0 and 1, it may not be the case that we can say either that the first probability is greater than the second, or that the second is greater than the first, or that they are equal. There are degrees of probability that are incomparable in magnitude.

For Keynes, probability is not definable. The basic probabilities must be intuited (though *what* is intuited is not subjective, but objective). Once we have an intuited stock of probabilities, however, the calculus provides us with a means of arriving at other probabilities by argument. Keynes felt that a principle of indifference was required to arrive at these basic probability intuitions. It was to be applicable only to certain special cases with a particularly simple structure. Even so, its formulation, though an improvement over earlier formulations of a principle of indifference, was not clean enough to avoid all of the classical difficulties.

Most current logical interpretations of probability are based, in one way or another, on the conception of *logical range*. This is true also of the several concepts of factual support or corroboration that play an important role in discussion of inductive logic. Carnap is the writer who has done more for and with this conception than anyone else. Like Keynes, Carnap takes probability, in many of its uses, to be a logical concept. In some contexts, he is willing to accept an empirical concept, which he calls 'probability$_2$'. But where the logical probability relation is, for Keynes, indefinable and must, at least in some cases, be intuited, Carnap regards probability$_1$, the logical concept, as at least partially definable. Carnap, in his "Replies and Expositions," develops the concept of probability$_1$ in the following way. Consider a person $X$ at a certain time $T$. It is possible to discover the degree of belief that $X$ has in a proposition $h$ at time $T$. (The problems of this determination will be discussed in the next chapter.) $Cr_X$ is the credence function of $X$; $Cr_X(h, T)$ is the function whose value (lying between 0 and 1) represents $X$'s actual degree of belief in $h$ at $T$. It is also equal to the highest betting quotient with which $X$ is willing to bet on $h$ at $T$, for small stakes. (We shall also postpone the discussion of betting quotients to the next chapter.)

Now if $X$ exhibits some degree of rationality, his degree of belief in a proposition $h$ will not depend merely on the time $T$, but on his body of knowledge at the time $T$. The function which yields the degree of belief that $X$ *would have* in $h$, were $e$ his total body of knowledge, is a *credibility function*, $Cred(h, e)$.

We now take one more step, and consider rational credibility functions.

Carnap takes the logical probability function to be a rational credibility function, but he does not assume that there can be only one such function. The technical term 'degree of confirmation' is designed as an explicatum of logical probability. This is to say that it is a term which is similar in meaning and function to 'logical probability', but free of the ambiguities and obscurities that infect 'logical probability' or 'probability$_1$'. The degree of confirmation of hypothesis *h* on evidence *e*, for example, depends entirely on the logical and semantical properties and relations of *h* and *e*. It is defined only for a particular formal language in which these logical and semantical relations can be made perfectly explicit. Carnap at one point appeared to be quite sure that only one degree of confirmation function, or *c*-function, would turn out to be rational, but he now prefers to leave that question open. Thus it may be that there are a number of different *c*-functions that may serve as equally good explicata of logical probability functions or rational credibility functions. In any event, for Carnap, the only appropriate avenue to the study of logical probability is provided by the formal development of *c*-functions.

In order to minimize the technical difficulties we shall consider a sample of Carnap's theory in a very simple language. Suppose the Language **L** contains two individual constants '*a*' and '*b*', two one-place predicates '*G*' and '*F*', and the usual sentential connectives including '&' for conjunction, '~' for negation, and '∨' for disjunction. We may go about constructing statements expressing possible states of the world (relative to this language) in a fairly mechanical way: does the object designated '*a*' have the property designated by '*F*'? the property designated by '*G*'? Does the object designated by '*b*' have the property designated by '*G*'? the property designated by '*F*'? The number of possible sets of answers to these questions is 16; there are 16 possible *states of the world* that can be described in this language, or, putting it another way, there are 16 *state descriptions* that can be expressed in this language. They are:

1. *Fa & Ga & Fb & Gb*
2. *Fa & Ga & Fb & –Gb*
3. *Fa & Ga & –Fb & Gb*
4. *Fa & Ga & –Fb & –Gb*
5. *Fa & –Ga & Fb & Gb*
6. *Fa & –Ga & Fb & –Gb*
7. *Fa & –Ga & –Fb & Gb*
8. *Fa & –Ga & –Fb & –Gb*
9. *–Fa & Ga & Fb & Gb*
10. *–Fa & Ga & Fb & –Gb*
11. *–Fa & Ga & –Fb & Gb*
12. *–Fa & Ga & –Fb & –Gb*
13. *–Fa & –Ga & Fb & Gb*
14. *–Fa & –Ga & Fb & –Gb*
15. *–Fa & –Ga & –Fb & Gb*
16. *–Fa & –Ga & –Fb & –Gb*



Given a certain language, then, a state description is a statement that describes a state of the world in as much detail as is possible in that language.

In addition to state descriptions, we shall want to consider structure descriptions. A structure description is a disjunction of state descriptions each of which can be obtained from the others by some permutation of the individual constants. For example, '*Fa & –Ga & –Fb & Gb*' and '*–Fa & Ga & Fb & –Gb*' can be obtained from each other by permuting '*a*' and '*b*', and thus belong to the same structure description. In the example we are considering here there are ten structure descriptions:

1. *Fa & Ga & Fb & Gb*
2. *(Fa & Ga & Fb & –Gb) ∨ (Fa & –Ga & Fb & Gb)*
3. *(Fa & Ga & –Fb & Gb) ∨ (–Fa & Ga & Fb & Gb)*
4. *(Fa & Ga & –Fb & –Gb) ∨ (–Fa & –Ga & Fb & Gb)*
5. *(Fa & –Ga & Fb & –Gb)*
6. *(Fa & –Ga & –Fb & Gb) ∨ (–Fa & Ga & Fb & –Gb)*
7. *(Fa & –Ga & –Fb & –Gb) ∨ (–Fa & –Ga & Fb & –Gb)*
8. *(–Fa & Ga & –Fb & Gb)*
9. *(–Fa & Ga & –Fb & –Gb) ∨ (–Fa & –Ga & –Fb & Gb)*
10. *–Fa & –Ga & –Fb & –Gb*

We now assign numbers to the state descriptions in such a way that the numbers assigned to all possible state descriptions add up to one. This assignment constitutes the definition of a *measure function* for the state descriptions. We extend the definition of the measure function by stipulating that the value of the measure function for the disjunction of any two logically exclusive sentences be the sum of the values of the measure function for the two sentences: thus if $S_1$ and $S_2$ are inconsistent, the measure of their disjunction is to be the sum of the measure of $S_1$ and the measure of $S_2$. Since every consistent statement in the language can be expressed as a disjunction of state descriptions, and since every two state descriptions are inconsistent, the measure of every statement in the language is determined, and will be simply the sum of the measures of the state descriptions whose disjunction it is equivalent to. In our example, the measure of the statement '*Fa & Ga & –Gb*' will be the sum of the measures of the two state descriptions '*Fa & Fa & Fb & –Gb*' and '*Fa & Ga & –Fb & –Gb*'. Every statement that we can form in the language now has a measure determined by the measure function for state descriptions with which we started.

Given a measure function defined for all the statements of the language, it is only a brief step to degrees of confirmation. We define the *a priori* probability of a statement to be equal to the value of the measure function for that statement (if the measure of a statement is *p*, its *a priori* probability will be *p*). Let us write '*m*' for the measure function, '*m*(*e*)' for its value for the argument

$e$. We define the degree of confirmation of a hypothesis $h$ relative to evidence $e$ as the conditional probability of $h$ on $e$, $c(h, e)$:

$$c(h, e) = \frac{m(h \,\&\, e)}{m(e)}.$$

This is to say, we define the degree of confirmation of the hypothesis $h$ on the evidence $e$ to be the ratio of the *a priori* probability of the conjunction of $e$ and $h$ to the *a priori* probability of $e$ alone. The degree of confirmation of $h$ on $e$ is just the conditional probability of $h$, given $e$. In this way Carnap's analysis has reduced the problem of defining degree of confirmation for the classes of languages with which he has been concerned to the problem of choosing a measure function for state descriptions.

There are two very natural functions defined by Carnap which have often been referred to in the literature. One is $m^\dagger$, which assigns to each state description the same number: if there are $N$ state descriptions, then the measure of each is taken to be $1/N$. In our example the measure of each state description is $\frac{1}{16}$. This is a very natural assignment of prior probabilities; one could imagine it arising from an application of a principle of indifference to state descriptions. And yet this particular measure function leads to a concept of probability$_1$ on the basis of which it is impossible to learn from experience. In the example, if $e$ is '$Fa$' and $h$ is '$Fb$', the degree of confirmation of $h$ on $e$ is just the same as the *a priori* probability of $h$—i.e., $e$ is irrelevant to $h$:

$$c(Fb, Fa) = \frac{m(Fb \,\&\, Fa)}{m(Fa)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = m(Fb).$$

In general let $e$ state that of $n_e$ individuals, $m_e$ individuals satisfy the primitive predicate $F$. Let $h$ state that of $n_h$ individuals (not mentioned in $e$), $m_h$ individuals satisfy the predicate $F$. Let $S_h$ be the number of state descriptions in which $h$ holds, $S_e$ the number in which $e$ holds, and $S_{e\&h}$ the number in which $h$ and $e$ both hold. We may calculate these numbers as follows:

The number of ways in which a specified individual may satisfy a specified primitive predicate is $2^\pi - 1$, where $\pi$ is the number of primitive one-place predicates in the language; it may satisfy any of the $Q$-predicates but that one $Q$-predicate that entails that it does not satisfy that primitive predicate. The number of ways in which $m_e$ specified individuals may satisfy $F$ is $(2^\pi - 1)^{m_e}$. Similarly the number of ways in which $n_e - m_e$ individuals may fail to satisfy $F$ is $(2^\pi - 1)^{(n_e - m_e)}$. The number of ways of specifying $m_e$ out of $n_e$ individuals is $\binom{n_e}{m_e}$. The number of ways of assigning $Q$-predicates to the remaining $N - n_e$ individuals is $(2^\pi)^{N - n_e}$. Thus the total number of state descriptions in which $e$ holds is

$$S_e = \binom{n_e}{m_e}[2^\pi - 1]^{m_e}[2^\pi - 1]^{(n_e - m_e)}(2^\pi)^{(N - n_e)}.$$



Similarly

$$S_h = \binom{n_h}{m_h}[2^\pi - 1]^{m_h}[2^\pi - 1]^{(n_h - m_h)}(2^\pi)^{(N - n_e)}$$

and

$$S_{h\&e} = \binom{n_e}{m_e}\binom{n_h}{m_h}[2^\pi - 1]^{(m_e + m_h)}[2^\pi - 1]^{(n_e - m_e + n_h - m_h)}(2^\pi)^{(N - n_e - n_h)}.$$

Since the total number of state descriptions is $(2^\pi)^N$, and each is to be assigned the measure by $m^\dagger$, we have

$$m^\dagger(h) \cdot m^\dagger(e) = m^\dagger(h \,\&\, e)$$

and

$$c^\dagger(h, e) = m^\dagger(h).$$

In other words, the evidence that $m_e$ out of $n_e$ sampled items are $F$ is *irrelevant* to the hypothesis that $m_h$ out of $n_h$ as yet unexamined items will be $F$. But we would hardly take the logical probability of $h$ to be unaffected by $e$—we would generally suppose that the statistical makeup of the new sample would be similar to the statistical makeup of the old one.

Another natural measure function defined by Carnap is $m^*$. This measure function assigns numbers to the state descriptions in the following way: first, every *structure* description is assigned an equal measure; and then each state description belonging to a *given* structure description is assigned an equal part of the measure assigned to the structure description.

In the example above, we have for measure and confirmation functions on these two schemes, the following values:

$$\begin{aligned} m^\dagger(Fa \,\&\, Ga \,\&\, -Fb \,\&\, Gb) &= \tfrac{1}{16} \\ m^\dagger((Fa \,\&\, Ga \,\&\, -Fb \,\&\, Gb) \vee (-Fa \,\&\, Ga \,\&\, Gb \,\&\, Fb)) &= \tfrac{1}{8} \\ m^\dagger(Fa) &= \tfrac{1}{2} \\ c^\dagger(Fa, Fb) &= \tfrac{1}{2} \\ m^*(Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb) &= \tfrac{1}{10} \\ m^*(Fa \,\&\, Ga \,\&\, -Fb \,\&\, Gb) &= \tfrac{1}{20} \\ m^*(Fa) &= \tfrac{1}{2} \\ c^*(Fa, Fb) &= \tfrac{3}{5} \end{aligned}$$

Note that according to the measure function $m^\dagger$, the degree of confirmation of $Fa$, given $Fb$, is simply the same as the *a priori* probability of $Fa$, despite the fact that the same property is involved, while on the basis of the measure function $m^*$, the degree of confirmation of $Fa$ on the evidence $Fb$ is $\frac{3}{5}$, representing an increase over its *a priori* probability.

If we are dealing with a language having a large number of individual constants, even this confirmation function may lead to rather implausible values. In *The Continuum of Inductive Methods* Carnap offers a general definition of confirmation for a continuum of confirmation functions that depend on a parameter $\lambda$, which reflects the speed with which we may rationally learn from experience, neither depending too much on our *a priori* ideas, nor leaping to conclusions on insufficient evidence. The languages for which $c_\lambda$ is defined are still first-order languages (comprising individual constants, predicates, and logical machinery). The value of $\lambda$ may or may not be taken to depend on the number of predicates we have in the language. (It is possible to see here that the relation between probability and induction works both ways: not only is probability supposed to be our guide to induction, but we may look to our habits of induction to provide clues as to the particular function we should adopt as our probability function.)

In the initial development of his theory of probability$_1$ Carnap imposed severe restrictions on the languages to which his definition applied. The functions $c^\dagger$ and $c^*$, for example, are defined only for first-order functional calculi consisting of a finite or denumerable number of individual constants and a finite number of logically independent one-place predicates. These restrictions were all rapidly removed. In 1953 Kemeny published a system for languages subject only to the following mild restrictions:

1. The object language must be consistent.
2. It must not contain an axiom of infinity—i.e., it must not be provable that these are an infinite number of distinct individuals.
3. The number of types of individuals must be finite.
4. There must be a finite number of constants.
5. Each constant must be of finite order.

In Kemeny's system logical interdependence among the sentences of the language is handled in the following way: Let $m$ be a measure function defined for *all* state descriptions, including those which are logically impossible. We define the measure function $m'$ reflecting the facts of logical interdependence by the relation

$$m'(S) = \frac{m(A \,\&\, S)}{m(A)},$$

where $S$ is an arbitrary statement, and $A$ is the conjunction of the axioms or meaning postulates of the language.

Carnap's most recent work has been concerned with first-order languages containing a finite number of families of predicates. The confirmation functions depend on two parameters, $\lambda$ and $\eta$. The special case of a language containing only one-place primitive predicates has been formulated axiomatically in "Replies and Expositions." The primitive predicates are classified



into families such that to each individual exactly one member of each family applies. The axioms of this new system are as follows:

○ **A–1**

*If e and e′ are logically equivalent, then* $c(h, e) = c(h, e')$.

○ **A–2**

*If h and h′ are logically equivalent, then* $c(h, e) = c(h', e)$.

○ **A–3**

$c(h \,\&\, j, e) = c(h, e) \times c(j, e \,\&\, h)$.

○ **A–4**

*If* $e \,\&\, h \,\&\, j$ *is logically false, then* $c(h \vee j, e) = c(h, e) + c(j, e)$.

○ **A–5**

*If t is a tautology, and e is not logically false,* $c(t, e) = 1$.

○ **A–6**

*In a language with a finite number of models,* $c(h, e) = 1$ *only if h is a logical consequence of e.*

○ **A–7**

*The value of* $c(h, e)$ *remains unchanged under any finite permutation of individuals.*

○ **A–8**

*The value of* $c(h, e)$ *remains unchanged under any permutation of the predicates of any family.*

○ **A–9**

*The value of* $c(h, e)$ *remains unchanged under any permutation of families with the same number of predicates.*

○ **A–10**

*The value of* $c(h, e)$ *remains unchanged if the domain of individuals in language is enlarged, provided no quantifiers occur in e or h.*

### ○ A–11

*The value of $c(h, e)$ remains unchanged if further families of predicates are added to the language.*

The following axioms concern a single family of $k$ predicates, $P_1, \ldots P_k$. Let $e_s$ be a statement specifying that of $s$ individuals, $s_1$ specified individuals have the property $P_1$, $s_2$ the property $P_2, \ldots$, and $s_k$ the property $P_k$; $h_j$ is the statement specifying that individual $a_{s+1}$ has the property $P_j$; $h'_j$ is the statement specifying that the individual $a_{s+2}$ has the property $P_j$.

### ○ A–12

(*a*) $c(h_j, e_s \,\&\, h'_j) \geq c(h_j, e_s)$.
(*b*) $c(h_j, e_s \,\&\, h'_j) \neq c(h_j, e_s)$.

### ○ A–13

*Consider a sequence $e_1, e_2, \ldots e_s, \ldots$ in which each $e_i$ is an individual distribution for i individuals, as before, and in which $e_{s+1}$ logically implies $e_s$. Then*

$$\lim_{s\to\infty}\left[c(h_j, e_s) - \frac{s_j}{s}\right] = 0.$$

### ○ A–14

*Let i, j, l be three distinct numbers among $1, \ldots k$. Let $e'_s$ be like $e_s$ except that one of the s individuals described that has the property $P_j$, according to $e_s$, has the property $P_l$ according to $e'_s$, then $c(h_j, e_s) = c(h_j, e'_s)$.*

The foregoing axioms yield the following theorem for a family with more than two predicates:

### ○ THEOREM

$$c(h_j, e_s) = \frac{s_j + \lambda/k}{s + \lambda},$$

*where $\lambda$ is a finite positive real number characterizing the function c.*

To say that $\lambda$ is a number characterizing the function $c$ is to say that to choose $\lambda$ is to fix the value of $c$ for every pair of sentences in the language. The final axiom of the system presented in "Replies and Expositions" extends this theorem to the case $k = 2$.



### ○ A–15

*For $k = 2$ and fixed s, $c(h_j, e_s)$ is a linear function of $s_j$.*

Insofar as the $c$-function is applied to predicates belonging to only one family, it is completely specified by the choice of a value of $\lambda$. If $\lambda = k$ where $k$ is the number of predicates in a family, we have $c^*$. If $\lambda = 0$, we have a rule which identifies the probability of $h_j$ with the relative frequency of $P_j$ among the sample of $s$ individuals. This rule violates A-6. If we let $\lambda$ approach infinity, for $k = 2$, we obtain in the limit the function $c^\dagger$; this function violates A-12.

All of the $c$-functions considered by Carnap in his published works have the property that the prior probability of a factual universal generalization in a language concerning an infinite number of individuals is 0, and as a consequence no amount of evidence in favor of such a generalization can change that zero probability: $c(l, e) = 0$, for any $c$-function belonging to the $\lambda$-continuum, for any universally quantified factual generalization $l$, and for any finite quantifier-free sentence $e$.

Some writers have found this counterintuitive, and sufficient reason for rejecting Carnap's $c$-functions as adequate explicata for logical probability. Among them, Jaakko Hintikka has offered an alternative confirmation function in which it is possible for laws to have nonzero probabilities. For the sake of simplicity, we shall present Hintikka's approach only in relation to a language containing $k$ logically independent monadic predicates. As before, we define $2^k$ $Q$-predicates. To attribute a $Q$-predicate to an individual is to say all that can be said about it in this language. We next define constituents; a constituent is a general sentence asserting that such and such $Q$-predicates are exemplified in the world, and that *only* such $Q$-predicates are exemplified. For example, if $Q_1$, $Q_2$, and $Q_3$ are $Q$-predicates,

$$(\exists x)Q_1(x) \,\&\, (\exists x)Q_2(x) \,\&\, (\exists x)Q_3(x) \,\&\, (x)(Q_1(x) \vee Q_2(x) \vee Q_3(x))$$

is a constituent. There are $(2)^{2^k}$ subsets of the $Q$-predicates, but since one of these is the empty subset, and at least one $Q$-predicate must be exemplified, there are only $(2)^{2^k} - 1$ consistent constituents that can be formed. Hintikka's new assignment of *a priori* probabilities, which leads to the result that universal generalizations can have nonzero probabilities, consists in first assigning probabilities to the constituents. The probabilities thus assigned are independent of the number of individuals to which the language is to be applied. In "On a Combined System of Inductive Logic," Hintikka proposed assigning to each constituent the same probability, $1/[(2)^{2^k} - 1]$; in "A Two-dimensional Continuum of Inductive Methods," he proposes the assignment $[(\alpha + w - 1)!]/[(w - 1)!]$, where $w$ is the number of components in the constituent, and $\alpha$ is a free parameter, like Carnap's $\lambda$. In both of the papers mentioned, Hintikka then proceeds to divide the probability of a constituent evenly among the structure descriptions that make it true, and then to divide the probability of each structure description among the state descriptions that make it true, thus proceeding at the lower levels in precisely the way Carnap proceeds in developing $c^*$.

There are two kinds of questions that may be raised about such systems: one concerns the justification for the assumption that probability$_1$ should obey the usual probability axioms; the other concerns the question, *which* confirmation function? On the first question, definite answers have been forthcoming, in two ways. First, a number of writers (Shimony, Lehman, Kemeny) have shown that *if* probability$_1$ is to be taken as a guide to action—e.g., in making bets or in choosing between uncertain alternatives—*then* it must obey the usual probability axioms. (The first arguments establishing the relation between behavior under circumstances of uncertainty and the probability axioms were set forth by F. P. Ramsey; his work will be referred to in the next chapter.) This kind of defense of the usual probability axioms is characteristic of the subjectivistic theorists, but is also welcomed by those who accept a logical interpretation of probability. Abner Shimony, for example, has presented careful arguments in favor of the axioms of the probability calculus, but so far from being a subjectivist, he hopes to find a version of the principle of indifference which can be certified on *a priori* grounds.

In a very different way, the physicist R. T. Cox has offered arguments in favor of the axioms of the conventional probability calculus. His method is functional analysis. He begins by assuming that the probability is a functional relationship whose domain is pairs of propositions (evidence; hypothesis) and whose range is included in the set of real numbers in the closed interval (0, 1). He shows that a very few, very innocuous and natural, conditions suffice to ensure that the probability function will satisfy the axioms of the probability calculus. The arguments employed by Cox are both beautiful and sophisticated.

Given that we have decided to accept a probability$_1$ function satisfying the usual probability calculus, we must face the question, *which* probability$_1$ function? It was Carnap's original belief that there was one and only one probability function that could plausibly be accepted as yielding the appropriate meaning of probability$_1$. Since that time, he has begun to have his doubts. The continuum of inductive methods proposed a continuum of probability functions, any one of which might with some degree of plausibility, be regarded as giving the meaning of probability$_1$. In his most recent work, he intends to provide a two-dimensional continuum of inductive methods (*The Basic Systems of Inductive Logic*, forthcoming). The extent to which that two-dimensional continuum of inductive methods can be narrowed down, he now regards as an open question. Hintikka and others are exploring the results of adopting various different ways of choosing $c$-functions.

We shall return to this question of the arbitrariness of probability$_1$ functions in Part Two when we consider the inductive logics that can be based on these functions. For the moment, we merely observe that *given* a measure on the statements of a language, which is additive and has a maximum value of 1,



we can in a natural and obvious way derive a probability$_1$ function that satisfies the laws of the probability calculus. On the other hand, there is no general agreement about which of the possible measures that satisfy these conditions should be adopted.

In the following chapter we shall see that this element of arbitrariness may be regarded as a virtue rather than a shortcoming. The subjectivistic interpretation of probability is offered on the basis that there is an ineradicable element of arbitrariness in the assessment of evidence and that an adequate formalization of the probability relation must reflect that arbitrariness.

## EXERCISES

1. Consider a language $\mathbf{L}_2^2$ containing two individual constants ('$a$', '$b$') and two logically independent one-place predicates ('$P$', '$Q$'). Compute $c^\dagger(Pa, t)$ and $c^*(Pa, t)$, where $t$ represents a tautology, such as '$Pa \vee \sim Pa$'. Show that it does not matter what tautology of $\mathbf{L}$ you choose. Compute $c^\dagger(Pa, Pb)$ and $c^*(Pa, Pb)$. In a finite language, a universal generalization, '$(x)\,\phi\,(x)$' may be represented as a conjunction, '$\phi(a)\ \&\ \phi(b)\ \&\cdots\phi(n)$', where $a, b, \ldots n$ are the individuals referred to in that language. Compute (for $\mathbf{L}_2^2$), $c^\dagger((x)(P(x) \supset Q(x)))$, $P(a)\ \&\ Q(a)$, and $c^*((x)P(x) \supset Q(x))$, $P(a)\ \&\ Q(a))$.
2. Repeat the calculation of problem 1 for a language $\mathbf{L}_3^2$ containing three individual constants ('$a$', '$b$', '$c$') and the same two predicates.
3. Repeat the calculations of problem 1 for a language $\mathbf{L}_2^3$ containing three predicates, '$P$', '$Q$', '$R$', and two individual constants.
4. What is the effect on $c^\dagger$, on $c^*$, on confirmation functions of this sort, of the addition to the language of new individual constants? What is the effect of the addition of new predicates?
5. Show that Axiom A-6 is equivalent (given the other axioms) to the following convention from Carnap's *Logical Foundations of Probability*:
   For any state description $S$ in a language concerning only a finite number of individuals, $c(S, t) > 0$.
6. Explain in words, and give a justifying argument for each of Carnap's axioms A47, A–8, and A–9.
7. What is the point of Axiom A–10? Of Axiom A–11?
8. Explain Axiom A–12 in words.
9. How does $\lambda = 0$ violate A–6?
10. How does $\lambda = \infty$ conflict with A–12?
11. Let $\mathbf{L}$ be a language concerning an infinite number of individuals, and $l$ a universal factual generalization in that language. Let $e$ be any finite statement not containing quantifiers. Show that if $c$ satisfies Carnap's axioms, $c(l, e) = 0$.
12. In Hintikka's system, show that a law can have a nonzero probability.
13. What are the constituents of the language $\mathbf{L}_2^3$ of problem 2? What is the effect of adding more individual constants to the language?

14. If '$(x)(Px \supset Qx)$' has been observed to hold for five individuals, what is the degree of confirmation of each of the constituents of problem 13? (Assign equal weight to each constituent.)

## BIBLIOGRAPHICAL NOTES FOR CHAPTER 5

The first explicit treatment of probability as a logical relation between evidence and conclusion is John Maynard Keynes' *Treatise on Probability*, Macmillan, London and New York, 1921. The system of Harold Jeffreys, a down-to-earth and practical adaptation of Keynes' ideas, is to be found in his *Scientific Inference* (second edition), Cambridge University Press, New York, 1957. B. O. Koopman's work on systems of probability which, following Keynes' suggestion, take degrees of probability to be partially ordered rather than simply ordered, and thus to form a lattice, is to be found in three articles: "The Axioms and Algebra of Intuitive Probability," *Annals of Mathematics* **41**, 1940, pp. 269–92; "Intuitive Probabilities and Sequences," *Annals of Mathematics* **42**, 1941, pp. 169–87; and "The Bases of Probability," *Bulletin of the American Mathematical Society* **46**, 1940, pp. 763–74. The last article is reprinted in Kyburg and Smokler, *Studies in Subjective Probability*, John Wiley and Sons, New York, 1964.

Rudolf Carnap's *magnum opus* on probability, to date, is *The Logical Foundations of Probability* (second edition), University of Chicago Press, Chicago, 1962 (first edition, 1950). Two works have appeared extending the system of *Logical Foundations*, and a third is promised soon. The two that have appeared are, *The Continuum of Inductive Methods*, University of Chicago Press, Chicago, 1952, and Rudolf Carnap and Wolfgang Stegmüller, *Induktive Logik und Wahrscheinlichkeit*, Springer, Vienna, 1959; the third in his projected *Basic System of Inductive Logic*. The axioms appearing in this chapter come from "Replies and Expositions," in *The Philosophy of Rudolf Carnap* (P. A. Schilpp, ed.), Open Court, La Salle, Illinois, 1963, pp. 859–1013.

Developments that have followed the lines suggested by Carnap's work include John G. Kemeny, "A Logical Measure Function," *Journal of Symbolic Logic* **18**, 1953, pp. 289–308, in which Carnap's techniques are extended to a large class of object languages, and Hilary Putnam, "A Definition of Degree of Confirmation for Very Rich Languages," *Philosophy of Science* **23**, 1956, pp. 8–62, which extends the technique still further. A further development of Carnap's original system is provided by Richard Martin, "A Formalization of Inductive Logic," *Journal of Symbolic Logic* **23**, 1958, pp. 251–56. A clear exposition of the Carnapian point of view will be found in John Kemeny's "Carnap's Theory of Probability and Induction," in *The Philosophy of Rudolf Carnap* (P. A. Schilpp, ed.), Open Court, La Salle, Illinois, 1963, pp. 711–38.

Jaakko Hintikka's work is fairly recent. The basic logical framework is provided by "Distributive Normal Forms in First-Order Logic," in *Formal Systems and Recursive Functions* (Crossley and Dummett, eds.), North-Holland, Amsterdam, 1965, pp. 47–90. Developments of the system are to be found in "Towards a Theory of Inductive Generalization," in *Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science*, North-Holland, Amsterdam, 1965, pp. 274–88; "On a Combined System of Inductive Logic," in *Studia Logico-Mathematica et Philosophica in Honorem Rolf Nevalinna*, *Acta Philosophica Fennica* **18**, 1965, pp. 21–30; and "A Two-Dimensional Continuum of Inductive Methods," in *Aspects of Inductive Logic* (Hintikka and Suppes, eds.), North-Holland, Amsterdam, 1966, pp. 113–32.



Arguments justifying the usual probability axioms have been presented (independently and almost simultaneously) by Abner Shimony, "Coherence and the Axioms of Confirmation," *Journal of Symbolic Logic* **20**, 1955, pp. 644–60; John G. Kemeny, "Fair Bets and Inductive Probabilities," *Journal of Symbolic Logic* **20**, 1955, pp. 263–73; and R. Sherman Lehman, "On Confirmation and Rational Betting," *Journal of Symbolic Logic* **20**, 1955, pp. 251–62. The rather different argument of Richard T. Cox is to be found in his book *The Algebra of Probable Inference*, Johns Hopkins University Press, Baltimore, 1961.