"Corroboration versus Induction," *British Journal for the Philosophy of Science* **9**, 1958–59.

On simplicity, besides the work of Barker already cited, we have Nelson Goodman, "New Notes on Simplicity," *Journal of Symbolic Logic* **17**, 1952, pp. 189–91; "Axiomatic Measurement of Simplicity," *Journal of Philosophy* **52**, 1955, pp. 709–22; and "Recent Developments in the Theory of Simplicity," *Philosophy and Phenomenological Research* **19**, 1958–59, pp. 429–46; John G. Kemeny, "Two Measures of Complexity," *Journal of Philosophy* **52**, 1955, pp. 722–33, and *A Philosopher Looks at Science*, D. van Nostrand, Princeton, New Jersey, 1959; Patrick Suppes, "Nelson Goodman on the Concept of Logical Simplicity," *Philosophy of Science* **23**, 1956, pp. 153–59; Lars Svenonius, "Definability and Simplicity," *Journal of Symbolic Logic* **20**, 1955, pp. 235–50; Kyburg, "A Modest Proposal Concerning Simplicity," *Philosophical Review* **70**, 1961, pp. 390–95; R. Ackermann, "Some Remarks on Kyburg's Modest Proposal," *Philosophical Review* **71**, 1962, pp. 236–40; and Harold Jeffreys, *Scientific Inference*, Cambridge University Press, 1937, 1957.

John C. Harsanyi's attempt to make peace between probabilists and improbabilists is "Popper's Improbability Criterion for the Choice of Scientific Hypothesis," *Philosophy* **35**, 1960, pp. 332–40. Other comments on Popper's views are: D. C. Stove, review of Popper's, *Logic of Scientific Discovery*, *Australasian Journal of Philosophy* **38**, 1960, pp. 173–87; G. J. Warnock, review of Popper's, *Logic of Scientific Discovery*, *Mind* **69**, 1960, pp. 99–101; P. C. Gibbons, "On the Severity of Tests," *Australasian Journal of Philosophy* **40**, 1962, pp. 79–82. J. W. N. Watkins' comment on black ravens and white shoes comes from a long exchange of articles that will be referred to in the next chapter, as they are concerned mainly with the paradoxes of Carnapian confirmation theory; and R. H. Vincent's comment comes from "Popper on Qualitative Confirmation and Disconfirmation," *Australasian Journal of Philosophy* **40**, 1962, pp. 159–66.

Isaac Levi's compelling demonstration that Popper's corroboration concept flies in the face of ordinary conventions about utility (though it does presuppose what Popper would no doubt deny, that it makes sense at all to talk of epistemic utilities) is "Corroboration and Rules of Acceptance," *British Journal for the Philosophy of Science* **13**, 1962–63, pp. 307–13. The paper by John Kemeny and Paul Oppenheim referred to is "Degree of Factual Support," *Philosophy of Science* **19**, 1952, pp. 307–24. Rescher's formalization of the concept under discussion here is to be found in "Theory of Evidence," *Philosophy of Science* **25**, 1958, pp. 83–94; a shorter and later piece that is also relevant is his "Plausible Implication," *Analysis* **21**, 1960–61, pp. 128–35. Henry A. Finch's detailed discussion of the matter is "Confirming Power of Observations Metricized for Decisions Among Hypotheses," *Philosophy of Science* **27**, 1960, pp. 293–307, 391–404. Rudolf Carnap's corresponding formula for increase of confirmation is taken from the preface to the second edition of his *The Logical Foundations of Probability*, 2nd edition, University of Chicago Press, Chicago, 1962 (1st edition 1950). Håkan Tornebohm's article "Two Measures of Evidential Strength" appears in *Aspects of Inductive Logic* (Hintikka and Suppes, eds.), North-Holland, Amsterdam, 1966, pp. 81–95. In the same volume Jaakko Hintikka and Juhani Pietarinen ("Semantic Information and Inductive Logic" pp. 96–112) also discuss various measures of information content, and conclude that the only appropriate one is $P(H|E) - P(E)$, which is just Carnap's proposal.

13

Confirmation Theories

The inductive logic that is associated with either the logical or the subjectivistic interpretation of probability is gratifyingly simple: everything hinges on the use of Bayes' Theorem; the degree of confirmation, the probability, of a hypothesis H, relative to the total body of evidence we have at our disposal, E, is just

$$\frac{P(H \& E)}{P(E)}.$$

For the subjectivist the problem remains of evaluating these prior probabilities $P(H \& E)$ and $P(E)$; it is possible to evaluate someone's degree of belief in terms of his preferences, but as Isaac Levi has pointed out, preferences can often be just as vague and confused as degrees of belief. This is just to say that the difficulties of an inductive logic based on the subjectivistic interpretation of probability are just the difficulties of that interpretation itself. There remain, of course, a large number of practical problems—soluble ones—in the actual application of the theory; but these are primarily statistical problems rather than theoretical or philosophical problems.

The logical interpretation of the probability calculus has always been developed with the idea in mind of removing—by stipulations and conventions, if need be—the vaguenesses and confusions that surround inductive inference. But possibly as a consequence of this very clarity and precision, the logicist program for inductive logic is open to counterexamples and paradoxes.

The oddities that are referred to as the "paradoxes of confirmation" were first noted by Janina Hosiasson-Lindenbaum in 1940; they were christened by Carl Hempel in 1945. Consider the statement 'all ravens are black'. Anyone who believes in a logic of confirmation at all would regard the observation of a black raven to be a (possibly slight) confirmation of this statement. One of the most natural properties of confirmation is that if a sentence S confirms a sentence T, and T is logically equivalent to T', then S confirms T' as well, and indeed to the same degree. Now 'all ravens are black' is logically equivalent to 'all non-black things are non-ravens', so the observation of a white shoe, which confirms the latter generalization in the same way that the observation of a black raven confirms the former, will also confirm 'all ravens are black'. Finally, the original hypothesis can be expressed: 'Everything is either a raven, or else it is not black'; and it is not hard to show that this hypothesis is confirmed by any observation of an object that is black (whether it is a raven or a cat or any other thing) as well as by any observation of an object that is not a raven (regardless of its color). Thus as confirming instances for the law that all ravens are black, we have one sensible sort of instance, a black raven, and two paradoxical sorts of instances, e.g., a white shoe, and a black cat.

J. W. N. Watkins recently offered this paradox as an argument against an inductivist view of the distinction between analytic and empirical statements ("Between Analytic and Empirical"). In answer, Hempel admitted ("Empirical Statements and Falsifiability") that these consequences of confirmation theory are "intuitively paradoxical," but insisted that they are "systematically unobjectionable." This is the tack taken by Hosiasson-Lindenbaum in her early article, and the tack Hempel himself took in his "On Studies in the Logic of Confirmation" in 1945. In his answer to Watkins, Hempel also pointed out that perfectly analogous consequences arise on the Popper–Watkins theory of falsification. A white shoe can perfectly well be regarded as the outcome of an attempt to falsify the theory that all ravens are black: when I first looked, I thought it was a raven, even though it was white, but then when I examined it, I saw that it was a shoe. The argument was carried on by Israel Scheffler, who attacked Watkins' arguments and finally pointed out that the disagreement may have stemmed from the fact that Hempel is not offering methodological prescriptions, while Popper is. Another exchange was instigated by D. C. Stove (mentioned earlier) who tried to throw some light on the argument by distinguishing between the pragmatic notion of a *test* and the logical (better, "semantical") notion of *evidence*. The "attempt" to falsify the hypothesis that all ravens are black by examining white shoes is obviously futile on anybody's theory, but "the *attempts* that might be made to falsify or instantially confirm a hypothesis are quite irrelevant to the weight of the evidence, if any, resulting from such attempts" ("Popperian Confirmation and the Paradox of the Ravens"). H. Gavin Alexander also observed that the Watkins–Popper theory of falsification was subject to the same paradoxes as the theory of confirmation, but he pointed out that if we take account of our background knowledge we can save confirmation theory by looking at the matter quantitatively; following Hosiasson-Lindenbaum, he pointed out that since we know that most things in the world aren't ravens, the observation of a white shoe is not going to confirm 'all ravens are black' as much as the observation of a black raven will.

The recent discussion was admirably summed up by J. L. Mackie, who traced the whole argument and came to much the sort of conclusion one would expect. Given a complete lack of background knowledge, Hempel is perfectly correct, and a white shoe, a black cat, and a black raven will all confirm the generalization 'all ravens are black'. Given merely some knowledge about the relative numbers of black objects and ravens and non-black objects and non-ravens, the bite may be taken out of the paradox by quantitative considerations as Hosiasson-Lindenbaum, Alexander, I. J. Good, and Patrick Suppes (the latter two from a subjectivistic point of view) all show. And if we allow unlimited background knowledge to give meaning to the notion of a *test*, the observations of black ravens and non-black non-ravens confirm the generalization to a worthwhile degree only if they are made in a genuine test.

A recent article by G. H. von Wright has done much to clarify the relation between the confirmation of instances of universally quantified generalizations and the equivalence condition which stipulates that if two formulas are logically equivalent they should be confirmed by the same instances. Von Wright introduces the concept of a *range of relevance* of a generalization; it is the set of entities which the generalization is being construed as being *about*. Thus the classical generalization about ravens may be construed as a generalization simply *about* ravens, or as a generalization *about* birds or as a generalization about all things whatever. "The generalization that all ravens are black is a different generalization, when it is about ravens and ravens only, and when it is about birds and birds only, and when it is—if it ever is—about all things in the world unrestrictedly" (p. 216). On the basis of very simple axioms for confirmation (or inductive probability) von Wright shows that since the probability that something from outside the range of the generalization will not refute the generalization is maximal (1), it will lead to no increase in degree of confirmation of the generalization. On the other hand, anything within the range of relevance of the generalization, even if it is not a raven, will contribute to its confirmation. The *natural* range of relevance of a generalization is just the set of things to which the antecedent term applies—in the case of 'all ravens are black', it is the set of ravens; in the case of 'all non-black things are non-ravens', it is the completely distinct set of non-black things. Within the natural range of relevance of 'all ravens are black', the only kinds of instances that confirm the generalization are ravens that are black.

Another way of achieving the same result is to construe all generalizations as essentially statistical generalizations—though of a limiting form—about a subject matter corresponding to von Wright's range of relevance. In the case of statistical generalizations, there is no contraposition: from 'nearly all A's and B's' one cannot infer (even probabilistically) 'nearly all non-B's are non-A's'. If we construe 'all ravens are black' as the limiting form of a generalization like 'nearly all ravens are black', then clearly only ravens are relevant to the generalization. If, on the other hand, we construe the same generalization as being implicitly about birds in general, then it is the limiting form of quite a different statistical generalization, namely 'nearly all birds are such that they are not nonblack ravens'. To confirm this generalization, we not only may but should examine birds of various other feathers, as well as ravens.

In all of this there is an element which does not meet the eye when one encounters a generalization running around loose in a philosopher's garden. When we consider a generalization apart from the body of knowledge of which it is a part, it is difficult to know that the *intended* range or relevance is, or what statistical hypothesis the generalization is a generalization and purification of. On the other hand, it is possible, within a specific area of inquiry, to be quite clear what the background statistical assertion concerned. In all of this there is a vast improvement over the oversimplification and artificiality that Max Black has complained of in the classical discussions of the paradoxes of confirmation.

Another paradox engendered by the logicist interpretation of probability, which also arises for the subjectivist interpretation, is the paradox of ideal evidence. This has been discussed both by Karl Popper and by R. H. Vincent ("The Paradox of Ideal Evidence"). Let a be the assertion that a particular toss of a given coin yields heads. It is clear that in the absence of any knowledge at all the *a priori* probability of a may plausibly be supposed to be $\frac{1}{2}$. Now let us subject the coin to extensive tests, and suppose that (say on the basis of a million tosses) we become very sure that the relative frequency of heads is $\frac{1}{2} \pm \epsilon$. Let the body of evidence for this assertion be E. Then the probability of a given E is also $\frac{1}{2}$. Therefore the examination of a million tosses of the coin is utterly irrelevant to a. This is clearly paradoxical. The problem here is that of finding a way to take account of the weight of the evidence, but according to Popper this cannot be done in view of the fact that "the fundamental postulate of the subjective theory [confirmation theory] is the postulate that degrees of the rationality of beliefs in the light of evidence exhibit a linear order" (p. 408).

Although this is a serious difficulty for most theories of degrees of confirmation, it is not as decisive as Popper makes it sound. These theories always contain, as an extraformal requirement, a principle of total evidence which demands that if we have information about the behavior (or structure) of the coin, we use it; this may be taken as a principle demanding that we maximize the weight of evidence; this is intended in the sense that we should use all the information we have, not in the sense that we should go on collecting information indefinitely. The question of when to stop gathering information is a pragmatic one, and one that has been considered by relatively few statisticians and almost no philosophers. It could also be maintained that there is no behavioristic way of distinguishing between the two situations Popper describes: one would, in either case, be willing to wager on the truth of a at even money. If there is no way of distinguishing them behavioristically, then there is no point in distinguishing them epistemologically.

Another reply to Popper is that although the logical (or subjectivistic) theory will lead to the same behavior with respect to a bet on a *single* toss of an untested coin as on a *single* toss of a thoroughly tested coin, the two situations are easily distinguished in the logical or subjectivistic theory when it comes to a bet on a sequence of tosses. If the coin is well tested, the probability of heads on a given toss will be relatively insensitive to what has happened on previous tosses, and the probability, say, of ten heads in a row will be very close to $(\frac{1}{2})^{10} = \frac{1}{1024}$. On the other hand, if the coin is untested, the probability of heads on a given toss may depend a lot on what has happened on earlier tosses. Given that a coin has yielded heads on each of its first nine tosses, and that that's all we know about the coin, we may be tempted (properly!) to take the probability of heads on the tenth toss to be close to 1, rather than close to $\frac{1}{2}$. Thus the probability of ten heads in a row, for the untested coin, will be much greater than $\frac{1}{1024}$. Thus we can, after all, given a body of beliefs about the behavior of a coin, distinguish between the body of beliefs based upon lack of knowledge and the body of beliefs based on ideal evidence.

This answer, however, leads directly to another difficulty already alluded to in Chapter 6 in connection with the subjectivistic interpretation of probability. Our purely *a priori* probabilities may lead us to assign very high probabilities to very powerful empirical statements; e.g., the *a priori* probability that less than 10% of the balls drawn from an urn in an infinitely long sequence of trials will be purple might be over 0.99. More specifically, this result follows from the assumption that the draws are regarded as *exchangeable* (the probability of drawing a particular sequence of purple and nonpurple balls depends only on the number of purple balls in the sequence); that the prior probability that a ball will be purple is 0.01; and that the conditional probability that the second ball will be purple, given that the first ball is purple, is 0.02. Again the logicist will point to the principle of total evidence and remind us that as soon as we collect empirical evidence, the *a priori* probability statement becomes academic and irrelevant. But one may well wonder why, in the face of such high prior probabilities, one should bother to find evidence. Particularly if one takes an epistemological view, according to which probability should serve as a guide to rational belief, something seems to have gone awry. The discussion still goes on. Further references will be found in the bibliographical notes to this chapter.

The next paradox arises, like the classical one of the white shoes and the black ravens, from a conflict between our intuitive assessments of probability and our lightning calculations of degrees of confirmation. R. H. Vincent uses it to attempt to show that no plausible theory of confirmation can accept the multiplication axiom ("A Note on Some Quantitative Theories of Confirmation"). Let q consist of a hypothesis (such as Newton's laws) together with boundary conditions sufficient to entail p, where p is the statement, 'A freely falling body near the earth will fall 144 feet in 3 seconds'. Since q entails p, the conjunction of q and p is equivalent to p, and (writing $c(x, y)$ for the degree of confirmation of x given y)

$$c(p \,\&\, q, r) = c(q, r)$$

According to the multiplication axiom,

$$c(p \,\&\, q, r) = c(p, r) \times c(q, p \,\&\, r)$$

or

$$c(p, r) = \frac{c(q, r)}{c(q, p \,\&\, r)}$$

Since p cannot alone contribute much to the confirmation of q, $c(q, p \,\&\, r) \approx c(q, r)$ and $c(p, r) \approx 1$. In other words, under the circumstances outlined, it is practically certain, in the absence of Newton's laws, that a freely falling body near the earth will fall 144 feet in 3 seconds!

A numerical example shows that the conclusion is not valid, and that its paradoxical air results from thinking loosely about numbers. I borrow the example from Vincent, making it only slightly more specific. He says that if q is 'all balls in urn U are red' and p is 'the first three balls drawn are red', then by the above argument p will be practically certain relative to our background knowledge. But assuming the equiprobability of structure descriptions (because the cards would be stacked in our favor if we used the hypothesis of equiprobable state descriptions), and assuming that we have five balls in the urn, we have

$$\begin{aligned} c(p, r) &= \tfrac{3}{10} \qquad \text{(drawing with replacement)} \\ c(p, q \,\&\, r) &= 1 \\ c(q, p \,\&\, r) &= \tfrac{5}{9} \\ c(q, r) &= \tfrac{1}{6} \end{aligned}$$

so that

$$c(p, r) = \tfrac{3}{10} = \frac{c(q, r)}{c(q, p \,\&\, r)} = \frac{\frac{1}{6}}{\frac{5}{9}}.$$

To say that if the number of balls were larger so that $c(q, p \,\&\, r)$ were closer to $c(q, r)$ we would find more of a paradox is false: $c(q, r)$ and $c(q \,\&\, r)$ will both decrease, but their ratio need not at all approach unity; indeed on the supposition of equiprobable structure descriptions it will *decrease*.

Here is a last paradox that is regarded by Popper as an out-and-out contradiction in Carnap's system. The paradox is that

> "There are cases in which x is strongly supported by z and y is strongly undermined by z, while at the same time x is confirmed by z to a lesser degree than is y" (p. 390).

For example, if x is the statement that a six will turn up on the next throw of a die, and y is the statement that some number other than six will turn up, and z is the statement that an even number has turned up, then z increases the probability of x and decreases that of y, while at the same time confirming x less than it confirms y:

$$\begin{aligned} c(x) &= \tfrac{1}{6} \qquad & c(x, z) &= \tfrac{1}{3} \\ c(y) &= \tfrac{5}{6} \qquad & c(y, z) &= \tfrac{2}{3} \end{aligned}$$

while

$$c(x, z) < c(y, z)$$

Popper claims that this is a self-contradictory state of affairs, for he claims that it is always self-contradictory to say that "x has the property P . . . and y has not the property P and y has the property P in a higher degree than x" (p. 391).

This is a valid objection to Carnap's way of talking, for Carnap himself claimed that the statement 'a is warm and b is not warm and b is warmer than a' was self-contradictory, and the statement about confirmation sounds the same. In fact, however, it is not the same, and the similarity stems from a confusion in the *informal* characterization of the classificatory concept of confirmation. This confusion is cleared up in the preface to the new edition of the *Logical Foundation of Probability*, where Carnap distinguishes two distinct triples of concepts, neither of which allows Popper's paradox.

Firmness	Increase in Firmness
1 h is firm on e.	h is made firmer by i.
2 h on e is firmer than h' on e'.	h is made firmer by i than h' is by i'.
3 The degree of firmness of h on e is u.	The increase in firmness of h by i is u.

Each of these concepts admits of a simple and obvious explication in terms of confirmation functions. The paradox observed by Popper only arose through Carnap's mistaken selection of (1) under *increase in firmness* and (2) and (3) under *firmness* as a triple of concepts analogous to 'warm' 'warmer', and 'of such and such a temperature'. It is not the case, of course, that x can be firm on z, and y not firm on z, and y firmer on z than is x on z. One of the most serious-sounding objections to the basic idea of confirmation theory thus turns out to be (as Bar-Hillel has observed) a matter of terminological confusion.

We have already encountered Nelson Goodman's famous predicate 'grue' in Chapter 10 on estimation. Such predicates imply the existence of a serious difficulty with confirmation theory as it is ordinarily approached. I have not included this difficulty among the "paradoxes" of confirmation theory, both because it seems to be more in the nature of a straightforward problem to be overcome by straightforward methods, than something as cute as a paradox, and because it is a problem which affects all inductive theories equally, as well as theories of scientific inference which call themselves noninductive. Something is grue, recall, if it is green up until the year 2000, and blue thereafter. Goodman's original paradox amounted simply to the observation that whatever evidence we have which justifies our belief that all emeralds are green, or which gives to that generalization a high degree of confirmation, is equally a justification for believing that all emeralds are grue, or gives an equally high degree of confirmation to the generalization 'all emeralds are grue'. As observed in Chapter 10, speaking of estimation won't help us. The universality of the generalization is irrelevant: we face the same problem if we consider only the statistical assertions 'almost all emeralds are green' and 'almost all emeralds are grue'. Since they are contraries, and since whatever is evidence for one is also evidence for another, neither can have a probability of more than a half. And since we can multiply predicates *ad libitum* (gred = green now and red later; grellow, grentian griolet, etc), we can reduce the degree of confirmation of the generalization 'all emeralds are green' to as low a level as we choose.

The same kind of thing happens to probability statements made on the basis of statistical hypotheses. Everyone who would agree to assigning a probability to heads on the next toss of this coin at all would, if they knew it was a coin which yielded heads half the time, assign the value $\frac{1}{2}$ to that probability. But the next toss is also a member of the set of all tosses that either yield heads or belong to the intersection of the unit class of the next toss, and the set of tosses yielding tails:

$$\vee \cup (\{\text{next toss}\} \cap T),$$

and the proportion of heads in this class is arbitrarily close to 1; why not say the probability of heads is 1, even though the coin is well tested? As noted earlier, the problems we are led to in defining randomness seem insuperable

if we take any definable class to be a legitimate reference class. It seems that one cannot consider all definable classes on an equally legitimate footing in inductive logic.

Furthermore we can raise an analogous problem in terms of such purely statistical predicates as 'exhibits a relative frequency of A's lying between the limits 0.3 and 0.4'. Consider, for example:

> Practically all samples of 1000 B's exhibit a proportion of A's lying outside the limits $r - \epsilon$ and $r + \epsilon$, where r is the measure of A's among B's.
>
> Practically all samples of 1000 B's exhibit a proportion of A's lying between $r - \delta_1$ and $r + \delta_2$.
>
> Practically all samples of 1000 B's exhibit a proportion of A's lying between $r - \gamma_1$ and $r + \gamma_2$.

We may choose $\epsilon, \delta_1, \delta_2, \gamma_1, \gamma_2$ in such a way that all three of these statements are true, and can be used in conjunction with a statement reflecting the observation of a particular sample of 1000 B's, to arrive at high probabilities or acceptance of the three statements:

> The measure of A's among B's does not lie between $k - \epsilon$ and $k + \epsilon$.
>
> The measure of A's among B's lies between $k - \delta_1$ and $k + \delta_2$.
>
> The measure of A's among B's lies between $k - \gamma_1$ and $k + \gamma_2$.

where these statements cannot be jointly true.

For example, if we are drawing balls from an urn (with replacement) and find that 600 of them have been black, it is more probable than not that between 0.00 and 0.61 of the balls drawn from the urn will turn out to be black, and it is more probable than not that between 0.59 and 1.00 of the balls drawn from the urn will turn out to be black, and more probable than not that it will *not* be the case that between 0.59 and 0.61 of the balls will turn out to be black.

Nearly all of the more or less formal theories of scientific inference that I have discussed so far have at one point or another run into the 'grue' problem or some variant of it. The problem has many ramifications that haven't been explored yet. So far, the peculiar predicates have resisted any general characterization; no solution even to the simplest cases has gained anything like universal acceptance. Goodman offered a relative solution in terms of "entrenchment"—the degree to which a predicate is entrenched in our language. But might it not be that we could be speaking a different language? For example a language in which 'grue' and 'bleen' (understood as referring to objects blue before 2000 and green thereafter) are primitive color words. Barker and Achinstein attempt to show that we could not;

Goodman, in "Positionality and Pictures," rejects their attempt to establish a logical asymmetry between 'grue' and 'green', and so does Ullian, in another article. Salmon's solution is that we should project only ostensively definable predicates, and that a predicate can only be ostensively definable if it applies to things that *look alike*. Salmon's solution is like Goodman's in presupposing a principle of induction of the very sort it is supposed to elucidate. Salmon presupposes that things that look alike will continue to look alike, while Goodman admits that entrenchment in the past must be assumed a good guide to usefulness in the future. Even granted all this, we must accept the conventional meaning of 'looks alike'; I find it easy to imagine creatures (grue-blees) to whom all grue things look alike and who would be horribly surprised if in the year 2000 all grue things suddenly turned bleen! Salmon's solution, like Goodman's, applies only to ostensive predicates, and thus not to predicates like 'has a mass of 5 until the year 2000 and a mass of $\frac{1}{5}$ thereafter' or like the statistical predicates mentioned above.

Both Hugues Leblanc and Carl Hempel ("Inductive Inconsistencies") have dealt recently with the problem created by Goodman's peculiar predicates. Leblanc carefully discusses the alleged inconsistencies into which one is led by the usual rules of inductive extrapolation when one allows Goodman's peculiar predicates, and shows that we are led to no *formal* inconsistencies. Hempel points out that the problem is not one of being able to determine which statements are lawlike and which are not (as Goodman claimed), because the same problem arises in connection with the extrapolation of numerical relations as quantitative laws: there are an infinite number of laws of the form $y = f_1(x)$, $y = f_2(x)$, ... which are mutually inconsistent, which are equally supported by the evidence, and which are equally lawlike. He also has no solution to offer, though both he and Leblanc point out that in formal theories of confirmation no *inconsistencies* are generated by the peculiar predicates. They nevertheless have unwelcome consequences.

It has been suggested—e.g., by Stephen Barker—that this is a problem peculiar to the view that enumerative induction is basic to scientific inference. But it is clearly just as much a matter of concern to those who claim that scientific inference is a matter of selecting the simplest from a class of acceptable hypotheses, for, if we spoke in a grue–bleen language, the hypothesis that all emeralds are grue would be much simpler (it would involve no reference to time, for example) than the hypothesis that they are green—i.e., grue until the year 2000, and bleen thereafter. The problem of finding some way of distinguishing between sensible predicates like 'blue' and 'green' and the outlandish ones suggested by Goodman, Barker, and others is thus surely one of the most important problems to come out of recent discussions of inductive logic. It is also one of the most pressing, since there are now only some thirty years left in which to solve it.

Furthermore the existence of these problems has not discouraged people from following the path of confirmation theory. For a thoroughgoing subjectivist like Richard Jeffrey, the Goodmanesque predicates present no more than an idle academic amusement; no one ever assigned a positive probability to the hypothesis that all emeralds are grue, so that no matter how many we observe to be grue, we will still not be compelled to find the queer hypothesis probable. From this point of view, admirably adumbrated in Jeffrey's *Logic of Preference*, the whole point of inductive logic is to explicate the notion of *rational behavior*. The data are people's beliefs (probabilities) and people's values (utilities), and the upshot is an explication of the rationale which underlies or ought to underlie the choice of a course of action under circumstances of uncertainty. There is no question by itself of believing or disbelieving hypotheses, much less any questions of *accepting* hypotheses. Indeed, on Jeffrey's refinement of the subjectivistic position, one no longer has even to suppose that *evidence* statements come to be accepted; it suffices to consider new subjectivistic probability functions appropriate to changed circumstances.

Carnap's new system of inductive logic circumvents the difficulties introduced by peculiar predicates by the simple expedient of not being designed for a language rich enough to allow the definition of these predicates. But even in a more powerful language, the way of avoiding those difficulties would be clear: one would assign exceedingly small (or zero) *a priori* probabilities to state descriptions which exhibited radical discontinuities—i.e., in which every object up to a_n had a certain property P_i, and every subsequent object had a different property P_j.

Even in Carnap's new system, however, we are faced with the problem that universal generalizations (in a language referring to an infinite number of individuals) always have the *a priori* probability 0, and hence also the *a posteriori* probability 0. Carnap conceives of inductive logic in the manner of Jeffrey, as a guide to decision under uncertainty, and therefore can get along with instance confirmation—i.e., the probability that the next instance will satisfy the generalization. Instead of demanding that the generalization 'all ravens are black' be itself highly probable, on this view it is sufficient that the instantial statement 'the next raven will be black' is highly probable.

This brings us to a very live issue that affects both the subjectivistic theories of induction, such as Jeffrey's, and certain of the logicist theories, like Carnap's. The issue is the problem of whether or not inductive conclusions are ever accepted. Jeffrey and Carnap say "no"; other subjectivists and logicists say "yes". Those who regard inductive logic as essentially the hypothetico-deductive method tend to answer "yes". The more carefully one attends to the arguments between proponents of the hypothetico-deductive method and inductivists (falsification theorists and confirmation theorists), the more important does this problem appear. Popper regards Carnap as going beyond the pale of true Humean skepticism in saying that hypotheses can ever be "probable"; Popper will *accept* a hypothesis (tentatively, of course; until further notice), but will never make the assertion that it is probable. Ironically, Carnap frowns upon Popper for nearly the same reason: it would on his view be altogether unwarranted, and flying in the face of sound Humean skeptical arguments, ever to *accept* a hypothesis, tentatively or in any other way. All we can hope for is to be able to assign some degree of probability to the hypothesis.

What is at issue here is precisely the question of whether we should speak of accepting hypotheses; it is the question of whether inductive logic is to have a rule of acceptance. Carnap is one of the few philosophers to take a very definite stand on this issue and to recognize it as a fundamentally important one. He says that the great majority of contemporary writers make "one basic mistake": they regard the result of inductive reasoning as the acceptance of a new proposition ("The Aim of Inductive Logic"). On Carnap's view, the result of an induction is the assignment of a degree of confirmation to a new proposition. Now this is all, he would argue, that we need or can possibly want in the way of inductive conclusions. On the basis of these degrees of confirmation, we can define mathematical expectations and make decisions among courses of action.

Aside from general philosophical considerations, the most serious argument against the rule of acceptance is the lottery paradox. Consider a fair, 1,000,000-ticket lottery, with one prize. On almost any view of probability, the probability that a given ticket (say ticket No. 1) will win is 0.000001, and the probability that it will not win is 0.999999. Surely if a sheer probability is ever sufficient to warrant the acceptance of a hypothesis, this is a case. It is hard to think of grounds on which to base a distinction between this case and the cases of thoroughly acceptable statistical hypotheses, The same argument, however, goes through for ticket No. 7, ticket No. 156, etc. In fact, for any i between 1 and 1,000,000 inclusive, the argument goes through and we should rationally be entitled—indeed, obligated—to accept the statement 'ticket i will not win'. A commonly accepted principle of acceptability is that if S and T are acceptable statements, then their conjunction is also acceptable. But this means that in the lottery case, since each statement of the form 'ticket i will not win' is acceptable, so is the conjunction of 1,000,000 of these statements, which is equivalent to the statement than *no* ticket will win the lottery, which contradicts the statement which we initially took to be acceptable that one ticket would win the lottery. A similar paradox can be formulated with statistical hypotheses in the place of lottery tickets. One response to this paradox has been to reject the possibility of a rule of acceptance for inductive logic.

Popper, Black, Day, and others, however, have argued that scientists do accept hypotheses, if only provisionally. Hempel has pointed out that there are difficulties involved in supposing that they do, and Levi has shown how even on Popper's own scheme the relation between corroboration and acceptance is complicated. The issue is still wide open. Some arguments in favor of a rule of detachment are these:

1. People do in fact accept statements that are probable enough; they regard them as practically certain. I know that it is only *probable* (in one sense) that I shall find paper in my desk drawer, but in another sense I am *certain* that there is paper there, because I just now put it there. It is unrealistic, a violation of good sense, to demand that all such statements always be regarded as only probabilities.
2. Among the statements that are "practically certain" are those which report sense experience, and which writers like Carnap must regard as absolutely certain. He must regard such statements as certain in order to have statements relative to which he can compute the probabilities of other statements. By incorporating a rule of detachment in our inductive logic, we can avoid this necessity, and we can admit that observation statements are only overwhelmingly probable—probable enough to be included directly in our rational corpora, but not so probable as to be incorrigible.
3. As remarked earlier, most of the scientific inferences that one encounters in practice are deductive in form; one argues that since one sample of X melted at about t°C, all samples of X will melt at about t°C; one argues that potash is necessary for plant growth, because plants on those plots deficient in potash were stunted; etc. All such arguments can be reconstructed in a straightforward enough manner, *provided* that we can use as premises statements for which the evidence is partial and inconclusive, strictly speaking, though it should certainly be strong enough to yield practical certainty.

Whether or not these arguments are persuasive, they suggest that the search for a plausible rule of acceptance for inductive conclusions is not wholly misconceived. In the following chapter we shall consider three systems of inductive logic which do allow for a rule of acceptance.

EXERCISES

1. Devise a numerical example to support the claim that the paradox of confirmation created by the white shoe is a quantitative illusion.
2. If the observation of a bullfinch is taken to support the generalization that all ravens are black, what is the range of relevance of that generalization likely being taken to be?
3. In general how would you be able to tell what the range of relevance of a generalization was?
4. What other ways, besides eschewing a rule of inductive acceptance, might one use to avoid the lottery paradox?
5. Show that if the acceptance probability is taken to be $1/n$ no lottery with less than n tickets will lead to a paradox—not even a biased one.

6. Provide counterarguments to each of the three numbered arguments with which the chapter ends.
7. Give a detailed illustration of how the grue–bleen problem may affect a hypothetico-deductive view of induction.
8. Discuss various ways of handling the grue–bleen problem proposed in the relevant articles mentioned in the bibliography.

BIBLIOGRAPHICAL NOTES FOR CHAPTER 13

Isaac Levi's criticism of the subjectivistic interpretation of probability will be found in his book, *Gambling With Truth*, Alfred A. Knopf, New York, 1967, which will be discussed at length in the next chapter, and in his review article on Jeffrey's *Logic of Decision*, "Probability Kinematics," *British Journal for the Philosophy of Science* **18**, 1967, pp. 205–206.

Papers concerning the paradoxes of confirmation are: Janina Hosiasson-Lindenbaum, "On Confirmation," *Journal of Symbolic Logic* **5**, 1940, pp. 133–48, and Carl Hempel, "Studies in the Logic of Confirmation," *Mind* **54**, 1945, pp. 97–121; these are the two basic ones. There was a flurry (a blizzard?) of papers on this topic in the late 1950's: Hempel, "Empirical Statements and Falsifiability," *Philosophy* **33**, 1958, pp. 342–48; "Inductive Inconsistencies," *Synthese* **12**, 1960, pp. 439–69, and "Deductive-Nomological vs. Statistical Explanation," *Minnesota Studies in the Philosophy of Science*, Vol. III (Feigl, ed.), 1962, pp. 98–169; J. W. N. Watkins, "Between Analytic and Empirical," *Philosophy* **32**, 1957, pp. 112–31; "A Rejoinder to Professor Hempel's Reply," *Philosophy* **33**, 1958, pp. 349–55; "Mr. Stove's Blunders," *Australasian Journal of Philosophy* **37**, 1959, pp. 240–41; "A Reply to Mr. Stove's Reply," *Australasian Journal of Philosophy* **38**, 1960, pp. 54–58; "Confirmation without Background Knowledge," *British Journal for the Philosophy of Science* **10**, 1959–60, pp. 318–20, and "Professor Scheffler's Note," *Philosophical Studies* **12**, 1961, pp. 16–19; Israel Scheffler, "A Note on Confirmation," *Philosophical Studies* **11**, 1960, pp. 21–23, and "A Rejoinder on Confirmation," *Philosophical Studies* **12**, 1961, pp. 19–20; D. C. Stove, "Popperian Confirmation and the Paradox of the Ravens," *Australasian Journal of Philosophy* **37**, 1959, pp. 149–55; "A Reply to Mr. Watkins," *Australasian Journal of Philosophy* **38**, 1960, pp. 51–54; H. Gavin Alexander, "The Paradoxes of Confirmation," *British Journal for the Philosophy of Science* **9**, 1958–59, pp. 227–33, and "The Paradoxes of Confirmation—A Reply to Dr. Agassi," *British Journal for the Philosophy of Science* **10**, 1959–60, pp. 229–34; J. L. Mackie, "The Paradox of Confirmation," *British Journal for the Philosophy of Science* **13**, 1962–63, pp. 265–77; and I. J. Good, "The Paradox of Confirmation," *British Journal for the Philosophy of Science* **11**, 1960–61, pp. 145–49, "The Paradox of Confirmation (II)," *British Journal for the Philosophy of Science* **12**, 1961–62, and "The White Shoe is a Red Herring" *British Journal for the Philosophy of Science* **17**, 1966–67, p. 322.

Patrick Suppes "A Bayesian Approach to the Paradoxes of Confirmation" is to be found in *Aspects of Inductive Logic* (Suppes and Hintikka, eds.), North-Holland Publishing Co., Amsterdam, 1966, pp. 198–207. The same volume contains Max Black's careful review of the philosophical questions raised by the paradoxes, in "Notes on the Paradoxes of Confirmation," pp. 175–97 and G. H. von Wright's very persuasive proposal involving the range of relevance of a generalization in "The Paradoxes of Confirmation," pp. 208–19.

The paradox of ideal evidence is discussed in Karl Popper, *The Logic of Scientific Discovery*, Hutchinson, London, 1959, and R. H. Vincent, "The Paradox of Ideal Evidence," *Philosophical Review* **71**, 1962, pp. 497–503. The other paradox discussed by Vincent is presented in "A Note on Some Quantitative Theories of Confirmation," *Philosophical Studies* **12**, 1961, pp. 91–92.

For Carnap's Response to Popper's alleged contradiction, see the new preface to his *Logical Foundations of Probability* (2nd edition), University of Chicago Press, Chicago, 1962; Bar-Hillel's remarks are to be found in "Comments on 'Degree of Confirmation' by Professor K. R. Popper," *British Journal for the Philosophy of Science* **6**, 1955–56, pp. 155–57.

Nelson Goodman's grue–bleen problem was introduced in *Fact, Fiction, and Forecast*, Harvard University Press, Cambridge, 1955. Discussion of the problem will be found (among other places) in Goodman, "Positionality and Pictures," *Philosophical Review* **69**, 1960, pp. 523–25; S. Barker and P. Achinstein, "On the New Riddle of Induction," *Philosophical Review* **69**, 1960, pp. 511–22; J. Ullian, "More on 'Grue' and Grue," *Philosophical Review* **70**, 1961, pp. 386–89; W. Salmon, "On Vindicating Induction," in *Induction* (Kyburg and Nagel, eds.), Wesleyan University Press, Middletown, Connecticut, 1963, pp. 27–41; Hugues Leblanc, "A Revised Version of Goodman's Paradox on Confirmation," *Philosophical Studies* **14**, 1963, pp. 49–51; Donald Davidson, "Emeroses by Other Names," *Journal of Philosophy* **63**, 1966, pp. 778–80; J. Vickers, "Characteristics of Projectible Predicates," *Journal of Philosophy* **64**, 1967, pp. 280–86; Howard Smokler, "Goodman's Paradox and the Problem of Rules of Acceptance," *American Philosophical Quarterly* **3**, 1966, pp. 1–6; Richard Jeffrey, "Goodman's Query," *Journal of Philosophy* **63**, 1966.

Richard Jeffrey's splendidly clear exposition of the subjectivistic view of induction and inductive behavior is *The Logic of Decision*, McGraw-Hill, New York, 1966. Carnap's most recent system, *Basic System of Inductive Logic*, will appear soon, it is hoped. The quotation in the text comes from his nontechnical article "The Aims of Inductive Logic," in Nagel, Suppes, and Tarski (eds.) *Logic, Methodology, and Philosophy of Science*, Stanford University Press, Stanford, California, 1962, pp. 303–18.

The lottery paradox was first presented in my *Probability and the Logic of Rational Belief*, Wesleyan University Press, Middletown, Connecticut, 1961, where, however, the moral I drew from it was that deductive closure among accepted statements should be abandoned, rather than that an acceptance rule was impossible. The problem was noticed at the same time by Carl Hempel, "Deductive-Nomological vs. Statistical Explanation," in *Minnesota Studies in the Philosophy of Science*, Vol. III (Feigl and Maxwell, eds.), University of Minnesota Press, Minneapolis, 1962, pp. 98–169. An improved version is stated in Levi, *Gambling with Truth*. The problem is also discussed in Fred Schick's "Consistency and Rationality," *Journal of Philosophy* **60**, 1963, pp. 5–19, and in my response, "A Further Note on Rationality and Consistency," *Journal of Philosophy* **60**, 1963, pp. 463–65, as well as, from a less formal point of view, by Robert C. Sleigh, "A Note on Some Epistemic Principles of Chisholm and Martin," *Journal of Philosophy* **61**, 1964, pp. 216–18, and Keith Lehrer, "Knowledge and Probability," *Journal of Philosophy* **61**, 1964, pp. 368–72.