Two

The Confirmation of Scientific Hypotheses

John Earman and Wesley C. Salmon

In Chapter 1 we considered the nature and importance of scientific explanation. If we are to be able to provide an explanation of any fact, particular or general, we must be able to establish the statements that constitute its explanans. We have seen in the Introduction that many of the statements that function as explanans cannot be established in the sense of being conclusively verified. Nevertheless, these statements can be supported or confirmed to some degree that falls short of absolute certainty. Thus, we want to learn what is involved in *confirming* the kinds of statements used in explanations, and in other scientific contexts as well.

This chapter falls into four parts. Part I (Sections 2.1–2.4) introduces the problem of confirmation and discusses some attempts to explicate the qualitative concept of support. Part II (2.5–2.6) reviews Hume's problem of induction and some attempted resolutions. Part III (2.7–2.8) develops the mathematical theory of probability and discusses various interpretations of the probability concept. Finally, Part IV (2.9–2.10) shows how the probability apparatus can be used to illuminate various issues in confirmation theory.

Parts I, II, and III can each stand alone as a basic introduction to the topic with which it deals. These three parts, taken together, provide a solid introduction to the basic issues in confirmation, induction, and probability. Part IV covers more advanced topics. Readers who prefer not to bring up Hume's problem of induction can omit Part II without loss of continuity.

Part I: Qualitative Confirmation

2.1 EMPIRICAL EVIDENCE

The physical, biological, and behavioral sciences are all empirical. This means that their assertions must ultimately face the test of observation. Some scientific statements face the observational evidence directly; for example, "All swans are white," was supported by many observations of European swans, all of which were white, but it was refuted by the observation of black swans in Australia. Other scientific statements confront the observational evidence in indirect ways; for instance, "Every proton contains three quarks," can be checked observationally only by looking at the results of exceedingly complex experiments. Innumerable cases, of course, fall between these two extremes.

Human beings are medium-sized objects; we are much larger than atoms and much smaller than galaxies. Our environment is full of other medium-sized things—for example, insects, frisbees, automobiles, and skyscrapers. These can be observed with normal unaided human senses. Other things, such as microbes, are too small to be seen directly; in these cases we can use instruments of observation—microscopes—to extend our powers of observation. Similarly, telescopes are extensions of our senses that enable us to see things that are too far away to be observed directly. Our senses of hearing and touch can also be enhanced by various kinds of instruments. Ordinary eyeglasses—in contrast to microscopes and telescopes—are not extensions of normal human senses; they are devices that provide more normal sight for those whose vision is somewhat impaired.

An observation that correctly reveals the features—such as size, shape, color, and texture—of what we are observing is called *veridical*. Observations that are not veridical are *illusory*. Among the illusory observations are hallucinations, afterimages, optical illusions, and experiences that occur in dreams. Philosophical arguments going back to antiquity show that we cannot be absolutely certain that our direct observations are veridical. It is impossible to prove conclusively, for example, that any given observation is not a dream experience. That point must be conceded. We can, however, adopt the attitude that our observations of ordinary middle-sized physical objects are reasonably reliable, and that, even though we cannot achieve certainty, we can take measures to check on the veridicality of our observations and make corrections as required (see Chapter 4 for further discussion of the topics of skepticism and antirealism).

We can make a rough and ready distinction among three kinds of entities: (i) those that can be observed directly with normal unaided human senses; (ii) those that can be observed only indirectly by using some instrument that extends the normal human senses; and (iii) those that cannot be observed either directly or indirectly, whose existence and nature can be established only by some sort of theoretical inference. We do not claim that these distinctions are precise; that will not matter for our subsequent discussion. We say much more about category (iii) and the kinds of inferences that are involved as this chapter develops.

Our scientific languages should also be noted to contain terms of two types. We have an *observational vocabulary* that contains expressions referring to entities, properties, and relations that we can observe. "Tree," "airplane," "green," "soft," and "is taller than" are familiar examples. We also have a *theoretical vocabulary* containing expressions referring to entities, properties, and relations that we cannot observe. "Microbe," "quark," "electrically charged," "ionized," and "contains more protons than" exemplify this category. The terms of the theoretical vocabulary tend to be associated with the unobservable entities of type (iii) of the preceding paragraph, but this relationship is by no means precise. The distinction between observational terms and theoretical terms—like the distinction among the three kinds of entities—is useful, but it is not altogether clear and unambiguous. One further point of terminology. Philosophers often use the expression "theoretical entity," but it would be better to avoid that term and to speak either of *theoretical terms* or *unobservable entities.*

At this point a fundamental moral concerning the nature of scientific knowledge can be drawn. It is generally conceded that scientific knowledge is not confined to what we have observed. Science provides predictions of future occurrences—such as the burnout of our sun when all of its hydrogen has been consumed in the synthesis of helium—that have not yet been observed and that may never be observed by any human. Science provides knowledge of events in the remote past—such as the extinction of the dinosaurs—before any human observers existed. Science provides knowledge of other parts of the universe—such as planets orbiting distant stars—that we are unable to observe at present. This means that much of our scientific knowledge depends upon inference as well as observation. Since, however, deductive reasoning is nonampliative (see Chapter 1, Section 1.5), observations plus deduction cannot provide knowledge of the unobserved. Some other mode of inference is required to account for the full scope of our scientific knowledge.

2.2 THE HYPOTHETICO-DEDUCTIVE METHOD

As we have seen, science contains some statements that are reports of direct observation, and others that are not. When we ask how statements of this latter type are to meet the test of experience, the answer often given is the *hypothetico-deductive (H-D) method;* indeed, the H-D method is sometimes offered as *the* method of scientific inference. We must examine its logic.

The term *hypothesis* can appropriately be applied to any statement that is intended for evaluation in terms of its consequences. The idea is to articulate some statement, particular or general, from which observational consquences can be drawn. An *observational consequence* is a statement—one that might be true or might be false—whose truth or falsity can be established by making observations. These observational consequences are then checked by observation to determine whether they are true or false. If the observational consequence turns out to be true, that is said to *confirm* the hypothesis to some degree. If it turns out to be false, that is said to *disconfirm* the hypothesis.

Let us begin by taking a look at the H-D testing of hypotheses having the form of universal generalizations. For a very simple example, consider Boyle's law of

gases, which says that, for any gas kept at a constant temperature T, the pressure P is inversely proportional to the volume V,[1] that is,

$$P \times V = \text{constant (at constant } T).$$

This implies, for example, that doubling the pressure on a gas will reduce its volume by a half. Suppose we have a sample of gas in a cylinder with a movable piston, and that the pressure of the gas is equal to the pressure exerted by the atmosphere—about 15 pounds per square inch. It occupies a certain volume, say, 1 cubic foot. We now apply an additional pressure of 1 atmosphere, making the total pressure 2 atmospheres. The volume of the gas decreases to ½ cubic foot. This constitutes a hypothetico-deductive confirmation of Boyle's law. It can be schematized as follows:

(1) At constant temperature, the pressure of a gas is inversely proportional to its volume (Boyle's law).
The initial volume of the gas is 1 cubic ft.
The initial pressure is 1 atm.
The pressure is increased to 2 atm.
The temperature remains constant.
―――――――――――――――
The volume decreases to ½ cubic ft.

Argument (1) is a valid deduction. The first premise is the *hypothesis* that is being tested, namely, Boyle's law. It should be carefully noted, however, that Boyle's law is *not* the only premise of this argument. *From the hypothesis alone it is impossible to deduce any observational prediction;* other premises are required. The following four premises state the *initial conditions* under which the test is performed. The conclusion is the *observational prediction* that is derived from the hypothesis and the initial conditions. Since the temperature, pressure, and volume can be directly measured, let us assume for the moment that we need have no serious doubts about the truth of the statements of initial conditions. The argument can be schematized as follows:

(2) *H* (test hypothesis)
I (initial conditions)
―――――――――――――――
O (observational prediction)

When the experiment is performed we observe that the observational prediction is true.

As we noted in Chapter 1, it is entirely possible for a valid deductive argument to have one or more false premises and a true conclusion; consequently, the fact that (1) has a true conclusion does not prove that its premises are true. More specifically, we cannot validly conclude that our hypothesis, Boyle's law, is true just because the observational prediction turned out to be true. In (1) the argument from premises to

[1] This relationship does not hold for temperatures and pressures close to the point at which the gas in question condenses into a liquid or solid state.

conclusion is a valid deduction but the argument *from* the conclusion *to* the premises is not. If it has any merit at all, it must be as an inductive argument.

Let us reconstruct the argument *from* the observational prediction *to* the hypothesis as follows:

(3) The initial volume of the gas is 1 cubic ft.
The initial pressure is 1 atm.
The pressure is increased to 2 atm.
The temperature remains constant.
The volume decreases to ½ cubic ft.

At constant temperature, the pressure of a gas is inversely proportional to its volume (Boyle's law).

No one would seriously suppose that (3) establishes Boyle's law conclusively, or even that it renders the law highly probable. At best, it provides a *tiny bit* of inductive support. If we want to provide solid inductive support for Boyle's law it is necessary to make repeated tests of this gas, at the same temperature, for different pressures and volumes, and to make other tests at other temperatures. In addition, other kinds of gases must be tested in a similar manner.

In one respect, at least, our treatment of the test of Boyle's law has been oversimplified. In carrying out the test we do not directly observe—say by feeling the container—that the initial and final temperatures of the gas are the same. Some type of thermometer is used; what we observe directly is not the temperature of the gas but the reading on the thermometer. We are therefore relying on an *auxiliary hypothesis* to the effect that the thermometer is a reliable instrument for the measurement of temperature. On the basis of an additional hypothesis of this sort we claim that we can observe the temperature indirectly. Similarly, we do not observe the pressures directly, by feeling the force against our hands; instead, we use some sort of pressure gauge. Again, we need an auxiliary hypothesis stating that the instrument is a reliable indicator.

The need for auxiliary hypotheses is not peculiar to the example we have chosen. In the vast majority of cases—if not in every case—auxiliary hypotheses are required. In biological and medical experiments, for example, microscopes of various types are employed—from the simple optical type to the tunneling scanning electron microscope, each of which requires a different set of auxiliary hypotheses. Likewise, in astronomical work telescopes—refracting and reflecting optical, infrared, radio, X-ray, as well as cameras are used. The optical theory of the telescope and the chemical theory of photographic emulsions are therefore required as auxiliary hypotheses. In sophisticated physical experiments using particle accelerators, an elaborate set of auxiliary hypotheses concerning the operation of all of the various sorts of equipment is needed. In view of this fact, schema (2) should be expanded:

(4) H (test hypothesis)
A (auxiliary hypotheses)
I (initial conditions)

O (observational prediction)

Up to this point we have considered the case in which the observational prediction turns out to be true. The question arises, what if the observational prediction happens to be false? To deal with this case we need a different example.

At the beginning of the nineteenth century a serious controversy existed about the nature of light. Two major hypotheses were in contention. According to one theory light consists of tiny particles; according to the other, light consists of waves. If the corpuscular theory is true, a circular object such as a coin or ball bearing, if brightly illuminated, will cast a uniformly dark circular shadow. The following H-D test was performed:

(5) Light consists of corpuscles that travel in straight lines.[2]
A circular object is brightly illuminated.

The object casts a uniform circular shadow.

Surprisingly, when the experiment was performed, it turned out that the shadow had a bright spot in its center. Thus, the result of the test was negative; the observational prediction was false.

Argument (5) is a valid deduction; accordingly, if its premises are true its conclusion must also be true. But the conclusion is not true. Hence, at least one of the premises must be false. Since the second premise was known to be true on the basis of direct observation, the first premise—the corpuscular hypothesis—must be false.

We have examined two examples of H-D tests of hypotheses. In the first, Boyle's law, the outcome was positive—the observational prediction was found to be true. We saw that, even assuming the truth of the other premises in argument (1), the positive outcome could, at best, lend a small bit of support to the hypothesis. In the second, the corpuscular theory of light, the outcome was negative—the observational prediction was found to be false. In that case, assuming the truth of the other premise, the hypothesis was conclusively refuted.

The negative outcome of an H-D test is often less straightforward than the example just discussed. For example, astronomers who used Newtonian mechanics to predict the orbit of the planet Uranus found that their observational predictions were incorrect. In their calculations they had, of course, taken account only of the gravitational influences of the planets that were known at the time. Instead of taking the negative result of the H-D test as a refutation of Newtonian mechanics, they postulated the existence of another planet that had not previously been observed. That planet, Neptune, was observed shortly thereafter. An auxiliary hypothesis concerning the constitution of the solar system was rejected rather than Newtonian mechanics.

It is interesting to compare the Uranus example with that of Mercury. Mercury also moves in a path that differs from the orbit calculated on the basis of Newtonian mechanics. This irregularity, however, could not be successfully explained by postulating another planet, though this strategy was tried. As it turned out, the perturbation of Mercury's orbit became one of three primary pieces of evidence supporting Einstein's general theory of relativity—the theory that has replaced Newtonian me-

[2] Except when they pass from one medium (e.g., air) to another medium (e.g., glass or water).

chanics in the twentieth century. The moral is that negative outcomes of H-D tests sometimes do, and sometimes do not, result in the refutation of the test hypothesis. Since auxiliary hypotheses are almost always present in H-D tests, we must face the possibility that an auxiliary hypothesis, rather than the test hypothesis, is responsible for the negative outcome.

2.3 PROBLEMS WITH THE HYPOTHETICO-DEDUCTIVE METHOD

The H-D method has two serious shortcomings that must be taken into account. The first of these might well be called *the problem of alternative hypotheses.* Let us reconsider the case of Boyle's law. If we represent that law graphically, it says that a plot of pressures against volumes is a smooth curve, as shown in Figure 2.1.

The result of the test, schematized in argument (1), is that we have two points (indicated by arrows) on this curve—one corresponding to a pressure of 1 atmosphere and a volume of 1 cubic foot, the other corresponding to a pressure of 2 atmospheres and a volume of ½ cubic foot. While these two points conform to the solid curve shown in the figure, they agree with infinitely many other curves as well—for example, the dashed straight line through those two points. If we perform another test, with a pressure of 3 atmospheres, we will find that it yields a volume of ⅓ cubic foot. This is incompatible with the straight line curve, but the three points we now have are still compatible with infinitely many curves, such as the dotted one, that go through these three. Obviously we can make only a finite number of tests; thus, it is clear that, no matter how many tests we make, the results will be compatible with infinitely many different curves.

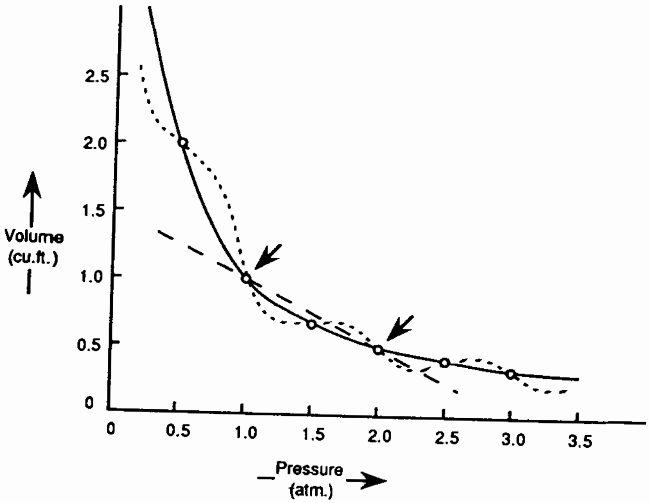

Figure 2.1

This fact poses a profound problem for the hypothetico-deductive method. *Whenever an observational result of an H-D test confirms a given hypothesis, it also confirms infinitely many other hypotheses that are incompatible with the given one.* In that case, how can we maintain that the test confirms our test hypothesis in preference to an infinite number of other possible hypotheses? This is the problem of alternative hypotheses. The answer often given is that we should prefer the *simplest* hypothesis compatible with the results of the tests. The question then becomes, what has simplicity got to do with this matter? Why are simpler hypotheses preferable to more complex ones? The H-D method, as such, does not address these questions.

The second fundamental problem for the H-D method concerns cases in which observational predictions cannot be *deduced.* The situation arises typically where statistical hypotheses are concerned. This problem may be called *the problem of statistical hypotheses.* Suppose, to return to an example cited in Chapter 1, that we want to ascertain whether massive doses of vitamin C tend to shorten the duration of colds. If this hypothesis is correct, the probability of a quick recovery is increased for people who take the drug. (As noted in Chapter 1, this is a fictitious example; the genuine question is whether vitamin C lessens the frequency of colds.) As suggested in that chapter, we can conduct a double-blind controlled experiment. However, we *cannot deduce* that the average duration of colds among people taking the drug will be smaller than the average for those in the control group. We can only conclude that, if the hypothesis is true, *it is probable* that the average duration in the experimental group will be smaller than it is in the control group. If we predict that the average duration in the experimental group will be smaller, the inference is *inductive.* The H-D method leaves no room for arguments of this sort. Because of the pervasiveness of the testing of statistical hypotheses in modern science, this limitation constitutes a severe shortcoming of the H-D method.

2.4 OTHER APPROACHES TO QUALITATIVE CONFIRMATION

The best known alternative to the H-D method is an account of qualitative confirmation developed by Carl G. Hempel (1945). The leading idea of Hempel's approach is that hypotheses are confirmed by their positive instances. Although seemingly simple and straightforward, this intuitive idea turns out to be difficult to pin down. Consider, for example, Nicod's attempt to explicate the idea for universal conditionals; for example:

$$H: (x)\,(Rx \supset Bx) \text{ (All ravens are black).}$$

(The symbol "(x)" is the so-called universal quantifier, which can be read, "for every object x"; "$\supset$" is the sign of material implication, which can be read *very roughly* "if . . . then . . . .") Although this statement is too simpleminded to qualify as a serious scientific hypothesis, the logical considerations that will be raised apply to all universal generalizations in science, no matter how sophisticated—see Section

2.11 of this chapter. According to Nicod, E Nicod-confirms such an H just in case E implies that some object is an instance of the hypothesis in the sense that it satisfies both the antecedent and the consequent, for example, E is $Ra.Ba$ (the dot means "and"; it is the symbol for conjunction). To see why this intuitive idea runs into trouble, consider a plausible constraint on qualitative confirmation.

Equivalence condition: If E confirms H and $\vdash H \equiv H'$, then E confirms H'.

(The triple bar "≡" is the symbol for material equivalence; it can be translated *very roughly* as "if and only if," which is often abbreviated "iff." The turnstile "⊢" preceding a formula means that the formula is a truth of logic.) The failure of this condition would lead to awkward situations since then confirmation would depend upon the mode of presentation of the hypothesis. Now consider

$$H' : (x)\,(\sim Bx \supset \sim Rx).$$

(The tilde "~" signifies negation; it is read simply as "not.") H' is logically equivalent to H. But

$$E\text{: } Ra.Ba$$

does not Nicod-confirm H' although it does Nicod-confirm H. Or consider

$$H''\text{: } (x)\,[(Rx.\sim Bx) \supset (Px.\sim Px)].$$

Again H'' is logically equivalent to H. But by logic alone, nothing can satisfy the consequent of H'' and if H is true nothing can satisfy the antecedent. So if H is true nothing can Nicod-confirm H''.[3]

After rejecting the Nicod account because of these and other shortcomings, Hempel's next step was to lay down what he regarded as conditions of adequacy for qualitative confirmation—that is, conditions that should be satisfied by any adequate definition of qualitative confirmation. In addition to the equivalence condition there are (among others) the following:

Entailment condition: If $E \vdash H$, the E confirms H.

(When the turnstile is preceded by a formula ("E" in "$E \vdash H$"), it means that whatever comes before the turnstile *logically entails* that which follows the turnstile—E logically entails H.)

[3] Such examples might lead one to try to build the equivalence condition into the definition of Nicod-confirmation along the following lines:

(N′) E Nicod-confirms H just in case there is an H' such that $\vdash H \equiv H'$ and such that E implies that the objects mentioned satisfy both the antecedent and consequent of H'.

But as the following example due to Hempel shows, (N′) leads to confirmation where it is not wanted in the case of multiply quantified hypotheses. Consider

H: $(x)\,(y)\,Rxy$
H': $(x)\,(y)[\sim(Rxy\,.\,Ryx) \supset (Rxy.\sim Ryx)]$
E: $Rab.\sim Rba$

E implies that the pair a, b satisfies both the antecedent and the consequent of H', and H' is logically equivalent to H. So by (N′) E Nicod-confirms H. But this is an unacceptable result since E contradicts H.

Special consequence condition: If E confirms H and $H \vdash H'$ then E confirms H'.

Consistency condition: If E confirms H and also confirms H' then H and H' are logically consistent.

As a result, he rejects

Converse consequence condition: If E confirms H and $H' \vdash H$ then E confirms H'.

For to accept the converse consequence condition along with the entailment and special consequence conditions would lead to the disastrous result that any E confirms any H. (Proof of this statement is one of the exercises at the end of this chapter.) Note that the H-D account satisfies the converse consequence condition but neither the special consequence condition nor the consistency condition.

Hempel provided a definition of confirmation that satisfies all of his adequacy conditions. The key idea of his definition is that of the *development*, $dev_I(H)$, of a hypothesis H for a set I of individuals. Intuitively, $dev_I(H)$ is what H says about a domain that contains exactly the individuals of I. Formally, universal quantifiers are replaced by conjunctions and existential quantifiers are replaced by disjunctions. For example, let $I = \{a, b\}$, and take

$$H\text{: } (x)\,Bx \text{ (Everything is beautiful)}$$

then

$$dev_I(H) = Ba.Bb.$$

Or take

$$H'\text{: } (\exists x)\,Rx \text{ (Something is rotten)}$$

then

$$dev_I(H') = Ra \vee Rb.$$

(The wedge "∨" symbolizes the *inclusive disjunction;* it means "and/or"—that is, "one, or the other, or both.") Or take

$$H''\text{: } (x)\,(\exists y)\,Lxy \text{ (Everybody loves somebody);}$$

then

$$dev_I(H'') = (Laa \vee Lab).(Lba \vee Lbb).\text{[4]}$$

Using this notion we can now state the main definitions:

[4] In formulas like H'' that have mixed quantifiers, we proceed in two steps, working from the inside out. In the first step we replace the existential quantifier by a disjunction, which yields

$$(x)\,(Lxa \vee Lxb).$$

In the next step we replace the universal quantifier with a conjunction, which yields $dev_I(H'')$.

Def. E directly-Hempel-confirms H just in case $E \vdash dev_I(H)$ for the class I of individuals menioned in E.

Def. E Hempel-confirms H just in case *E* directly confirms every member of a set of sentences K such that $K \vdash$ H.

To illustrate the difference between the two definitions, note that *Ra.Ba* does not directly-Hempel-confirm $Rb \supset Bb$ but it does Hempel-confirm it. Finally, disconfirmation can be handled in the following manner.

Def. E Hempel-disconfirms H just in case *E* confirms $\sim H$.

Despite its initial attractiveness, there are a number of disquieting features of Hempel's attempt to explicate the qualitative concept of confirmation. The discussion of these features can be grouped under two queries. First, is Hempel's definition too stringent in some respects? Second, is it too liberal in other respects? To motivate the first worry consider

$$H\text{: } (x)\ Rxy.$$

(The expression "*Rxy*" means "*x* bears relation *R* to *y*.") *H* is Hempel-confirmed by

$$E\text{: } Raa.Rab.Rbb.Rba.$$

But it is not confirmed by

$$E'\text{ : } Raa.Rab.Rbb$$

even though intuitively the latter evidence does support *H*. Or consider the compound hypothesis

$$(x)\ (\exists y)\ Rxy.(x) \sim Rxx.(x)\ (y)\ (z)\ [(Rxy.Ryz) \supset Rxz],$$

which is true, for example, if we take the quantifiers to range over the natural numbers and interpret *Rxy* to mean that *y* is greater than *x*. (Thus interpreted, the formula says that for any number whatever, there exists another number that is larger. Although this statement is true for the whole collection of natural numbers, it is obviously false for any finite set of integers.) This hypothesis cannot be Hempel-confirmed by any consistent evidence statement since its development for any finite *I* is inconsistent. Finally, if *H* is formulated in the theoretical vocabulary, then, except in very special and uninteresting cases, *H* cannot be Hempel-confirmed by evidence *E* stated entirely in the observational vocabulary. Thus, Hempel's account is silent on how statements drawn from such sciences as theoretical physics—for example, all protons contain three quarks—can be confirmed by evidence gained by observation and experiment. This silence is a high price to pay for overcoming some of the defects of the more vocal H-D account.

This last problem is the starting point for Glymour's (1980) so-called *bootstrapping* account of confirmation. Glymour sought to preserve the Hempelian idea that hypotheses are confirmed by deducing instances of them from evidence statements, but in the case of a theoretical hypothesis he allowed that the deduction of instances can proceed with the help of auxiliary hypotheses. Thus, for Glymour the

basic confirmation relation is three-place—*E* confirms *H* relative to *H'*—rather than two-place. In the main intended application we are dealing with a scientific theory *T* which consists of a network of hypotheses, from which *H* and *H'* are both drawn. If *T* is finitely axiomatizable—that is, if *T* consists of the set of logical consequences of a finite set of hypotheses, $H_1, H_2, \ldots, H_n$—we can say that *T* is bootstrap-confirmed if for each H_i there is an H_j such that *E* confirms H_i relative to H_j. These ideas are most easily illustrated for the case of hypotheses consisting of simple linear equations.

Consider a theory consisting of the following four hypotheses (and all of their deductive consequences):

$$H_1\text{: } O_1 = X$$
$$H_2\text{: } O_2 = Y + Z$$
$$H_3\text{: } O_3 = Y + X$$
$$H_4\text{: } O_4 = Z$$

The *O*s are supposed to be observable quantities while the *X*s and *Y*s are theoretical.

For purposes of a concrete example, suppose that we have samples of four different gases in separate containers. All of the containers have the same volume, and they are at the same pressure and temperature. According to Avogadro's law, then, each sample contains the same number of molecules. Observable quantities O_1–O_4 are simply the weights of the four samples:

$$O_1 = 28\text{ g}, O_2 = 44\text{ g}, O_3 = 44\text{ g}, O_4 = 28\text{ g}.$$

Our hypotheses say

- H_1: The first sample consists solely of molecular nitrogen—N_2—molecular weight 28; *X* is the weight of a mole of N_2 (28 g).
- H_2: The second sample consists of carbon dioxide—CO_2—molecular weight 44; *Y* is the weight of a mole of atomic oxygen O (16 g), *Z* is the weight of a mole of carbon monoxide CO (28 g).
- H_3: The third sample consists of nitrous oxide—N_2O—molecular weight 44; *Y* is the weight of a mole of atomic oxygen O (16 g) and *X* is the weight of a mole of molecular nitrogen (28 g).
- H_4: The fourth sample consists of carbon monoxide—CO—molecular weight 28; *Z* is the weight of a mole of CO (28 g).

(The integral values for atomic and molecular weights are not precisely correct, but they furnish a good approximation for this example.)

To show how H_1 can be bootstrap-confirmed relative to the other three hypotheses, suppose that an experiment has determined values O_1, O_2, O_3, O_4, for the observables. From the values for O_2 and O_4 we can, using H_2 and H_4, compute values for $Y + Z$ and for *Z*. Together these determine a value for *Y*. Then from the value for O_3 we can, using H_3, compute a value for $Y + X$. Then from these latter two values we get a value for *X*. Finally, we compare this computed value for *X* with the observed value for O_1. If they are equal, H_1 is confirmed. Although this simple example may seem a bit contrived, it is in principle similar to the kinds of measurements and reasoning actually used by chemists in the nineteenth century to establish molecular and atomic weights.

If we want the bootstrap procedure to constitute a test in the sense that it carries with it the potential for falsification, then we should also require that there are possible values for the observables such that, using these values and the very same bootstrap calculations that led to a confirmatory instance, values for the theoretical quantities are produced that contradict the hypothesis in question. This requirement is met in the present example.

In Glymour's original formalization of the bootstrapping idea, macho bootstrapping was allowed; that is, in deducing instances of H, it was allowed that H itself could be used as an auxiliary assumption. To illustrate, consider again the earlier example of the perfect gas law P(ressure) $\times$ V(olume) $= K \times T$(emperature), and suppose P, V, T to be observable quantities while the gas constant K is theoretical. We proceed to bootstrap-test this law relative to itself by measuring the observables on two different occasions and then comparing the values k_1 and k_2 for K deduced from the law itself and the two sets of observation values p_1, v_1, t_1 and p_2, v_2, t_2. However, macho bootstrapping can lead to unwanted results, and in any case it may be unnecessary since, for instance, in the gas law example it is possible to analyze the logic of the test without using the very hypothesis being tested as an auxiliary assumption in the bootstrap calculation (see Edidin 1983 and van Fraassen 1983). These and other questions about bootstrap testing are currently under discussion in the philosophy journals. (The original account of bootstrapping, Glymour 1980, is open to various counterexamples discussed in Christensen 1983; see also Glymour 1983.)

Let us now return to Hempel's account of confirmation to ask whether it is too liberal. Two reasons for giving a positive answer are contained in the following paradoxes.

Paradox of the ravens. Consider again the hypothesis that all ravens are black: $(x)\ (Rx \supset Bx)$. Which of the following evidence statements Hempel-confirm the ravens hypothesis?

E_1: $Ra_1 . Ba_1$
E_2: $\sim Ra_2$
E_3: Ba_3
E_4: $\sim Ra_4 . \sim Ba_4$
E_5: $\sim Ra_5 . Ba_5$
E_6: $Ra_6 . \sim Ba_6$

The answer is that E_1–E_5 all confirm the hypothesis. Only the evidence E_6 that refutes the hypothesis fails to confirm it. The indoor ornithology of some of these Hempel-confirmation relations—the confirmation of the ravens hypothesis, say, by the evidence that an individual is a piece of white chalk—has seemed to many to be too easy to be true.

Goodman's paradox. If anything seems safe in this area it is that the evidence $Ra.Ba$ that a is a black raven confirms the ravens hypothesis $(x)\ (Rx \supset Bx)$. But on Hempel's approach nothing rides on the interpretation of the predicates Rx and Bx. Thus, Hempel confirmation would still obtain if we interpreted Bx to mean that x is blite, where "blite" is so defined that an object is blite if it is examined on or before December 31, 2000, and is black or else is examined afterwards and found to be white. Thus, by the special consequence condition, the evidence that a is a black raven confirms the prediction that if b is a raven examined after 2000, it will be white, which is counterintuitive to say the least.

Part II: Hume's Problem of Induction

2.5 THE PROBLEM OF JUSTIFYING INDUCTION

Puzzles of the sort just mentioned—involving blite ravens and grue emeralds (an object is grue if it is examined on or before December 31, 2000 and is green, or it is examined thereafter and is blue)—were presented in Nelson Goodman (1955) under the rubric of *the new riddle of induction*. Goodman sought the basis of our apparent willingness to generalize inductively with respect to such predicates as "black," "white," "green," and "blue," but not with respect to "blite" and "grue." To mark this distinction he spoke of *projectible predicates* and *unprojectible predicates*, and he supposed that there are predicates of each of these types. The problem is to find grounds for deciding which are which.

There is, however, a difficulty that is both historically and logically prior. In his *Treatise of Human Nature* ([1739–1740] 1978) and his *Enquiry Concerning Human Understanding* (1748) David Hume called into serious question the thesis that we have any logical or rational basis for any inductive generalizations—that is, for considering any predicate to be projectible.

Hume divided all reasoning into two types, reasoning concerning *relations of ideas* and reasoning concerning *matters of fact and existence*. All of the deductive arguments of pure mathematics and logic fall into the first category; they are unproblematic. In modern terminology we say that they are necessarily truth-preserving because they are nonampliative (see Chapter 1, Section 1.5). If the premises of any such argument are true its conclusion must also be true because the conclusion says nothing that was not said, at least implicitly, by the premises.

Not all scientific reasoning belongs to the first category. Whenever we make inferences from observed facts to the unobserved we are clearly reasoning ampliatively—that is, the content of the conclusion goes beyond the content of the premises. When we predict future occurrences, when we retrodict past occurrences, when we make inferences about what is happening elsewhere, and when we establish generalizations that apply to all times and places we are engaged in reasoning concerning matters of fact and existence. In connection with reasoning of the second type Hume directly poses the question: What is the foundation of our inferences from the observed to the unobserved? He readily concludes that such reasoning is based upon relations of cause and effect. When we see lightning nearby (cause) we infer that the sound of thunder (effect) will ensue. When we see human footprints in the sand state description 1, all by itself, describes a particular structure, namely, all threeentities have property F. Similarly, state description 8 describes the structure in which no object has that property.

Having identified the structure descriptions, Carnap proceeds to assign equal weights to them (each gets 1/4); he then assigns equal weights to the state descriptions within each structure description. The resulting system of weights is shown above. These weights are then used as a measure of the ranges of statements;[11] this system of measures is called m^*. A confirmation function c^* is defined as follows:[12]

$$c^*(H|E) = m^*(H.E)/m^*(E).$$

To see how it works, let us reconsider the hypothesis Fc in the light of different bits of evidence. First, the range of Fc consists of state description 1, which has weight 1/4, and 3, 4, and 7, each of which has weight 1/12. The sum of all of them is 1/2; that is, the probability of our hypothesis before we have any evidence. Now, we find that Fa; its measure is 1/2. The range of $Fa.Fc$ is state descriptions 1 and 3, whose weights are, respectively, 1/4 and 1/12, for a total of 1/3. We can now calculate the degree of confirmation of our hypothesis on this evidence:

$$c^*(H|E) = m^*(E.H)/m^*(E) = 1/3 \div 1/2 = 2/3.$$

Carrying out the same sort of calculation for evidence $Fa.Fb$ we find that our hypothesis has degree of confirmation 3/4. If, however, our first bit of evidence had been $\sim Fa$, the degree of confirmation of our hypothesis would have been 1/3. If our second bit of evidence had been $\sim Fb$, that would have reduced its degree of confirmation to 1/4. The confirmation function c^* seems to do the right sorts of things. When the evidence is what we normally consider to be positive, the degree of confirmation goes up. When the evidence is what we usually take to be negative, the degree of confirmation goes down. Clearly, c^* allows for learning from experience.

A serious philosophical problem arises, however. Once we start playing the game of assigning weights to state descriptions, we face a huge plethora of possibilities. In setting up the machinery of state descriptions and weights, Carnap demands only that the weights for all of the state descriptions add up to 1, and that each state description have a weight greater than 0. These conditions are sufficient to guarantee an admissible interpretation of the probability calculus. Carnap recognized the obvious fact that infinitely many confirmation functions satisfying this basic requirement are possible. The question is how to make an appropriate choice. It can easily be shown that choosing a confirmation function is precisely the same as assigning prior probabilities to all of the hypotheses that can be stated in the given language.

Consider the following possibility for a measure function:

[11] The measure of the range of any statement H can be identified with the prior probability of that statement in the absence of any background knowledge K. It is an *a priori* prior probability.

[12] Wittgenstein's measure function assigns the weight ⅛ to each state description; the confirmation function based upon it is designated $c^\dagger$.

TABLE 2.3

State Description	Weight	Structure Description	Weight
1. $Fa.Fb.Fc$	1/20	All F	1/20
2. $Fa.Fb.\sim Fc$	3/20		
3. $Fa.\sim Fb.Fc$	3/20	2 F, 1 $\sim F$	9/20
4. $\sim Fa.Fb.Fc$	3/20		
5. $Fa.\sim Fb.\sim Fc$	3/20		
6. $\sim Fa.Fb.\sim Fc$	3/20	1 F, 2 $\sim F$	9/20
7. $\sim Fa.\sim Fb.Fc$	3/20		
8. $\sim Fa.\sim Fb.\sim Fc$	1/20	No F	1/20

(The idea of a confirmation function of this type was given in Burks 1953; the philosophical issues are further discussed in Burks 1977, Chapter 3.) This method of weighting, which may be designated $m^\diamond$, yields a confirmation function $C^\diamond$, which is a sort of counterinductive method. Whereas m^* places higher weights on the first and last state descriptions, which are state descriptions for universes with a great deal of uniformity (either every object has the property, or none has it), $m^\diamond$ places lower weights on descriptions of uniform universes. Like c^*, $c^\diamond$ allows for "learning from experience," but it is a funny kind of anti-inductive "learning." Before we reject $m^\diamond$ out of hand, however, we should ask ourselves if we have any a priori guarantee that our universe is uniform. Can we select a suitable confirmation function without being totally arbitrary about it? This is the basic problem with the logical interpretation of probability.

Part IV: Confirmation and Probability

2.9 THE BAYESIAN ANALYSIS OF CONFIRMATION

We now turn to the task of illustrating how the probabilistic apparatus developed above can be used to illuminate various issues concerning the confirmation of scientific statements. Bayes's theorem (Rule 9) will appear again and again in these illustrations, justifying the appellation of Bayesian confirmation theory.

Various ways are available to connect the probabilistic concept of confirmation back to the qualitative concept, but perhaps the most widely followed route utilizes an incremental notion of confirmation: E confirms H relative to the background knowledge K just in case the addition of E to K raises the probability of H, that is, $Pr(H|E.K) >) Pr(H|K)$.[13] Hempel's study of instance confirmation in terms of a

[13] Sometimes, when we say that a hypothesis has been confirmed, we mean that it has been rendered highly probable by the evidence. This is a *high probability* or *absolute* concept of confirmation, and it should be carefully distinguished from the *incremental* concept now under discussion (see Carnap 1962, Salmon 1973, and Salmon 1975). Salmon (1973) is the most elementary discussion.

two-place relation can be taken to be directed at the special case where K contains no information. Alternatively, we can suppose that K has been absorbed into the probability function in the sense that $Pr(K) = 1$,[14] in which case the condition for incremental confirmation reduces to $Pr(H|E) > Pr(H)$. (The unconditional probability $Pr(H)$ can be understood as the conditional probability $Pr(H|T)$, where T is a vacuous statement, for example, a tautology. The axioms of Section 2.7 apply only to conditional probabilities.)

It is easy to see that on the incremental version of confirmation, Hempel's consistency condition is violated as is

Conjunction condition: If E confirms H and also H' then E confirms $H.H'$.

It takes a bit more work to construct a counterexample to the special consequence condition. (This example is taken from Carnap 1950 and Salmon 1975, the latter of which contains a detailed discussion of Hempel's adequacy conditions in the light of the incremental notion of confirmation.) Towards this end take the background knowledge to contain the following information. Ten players participate in a chess tournament in Pittsburgh; some are locals, some are from out of town; some are juniors, some are seniors; and some are men (M), some are women (W). Their distribution is given by

TABLE 2.4

	Locals	Out-of-towners
Juniors	*M, W, W*	*M, M*
Seniors	*M, M*	*W, W, W*

And finally, each player initially has an equal chance of winning. Now consider the hypotheses H: an out-of-towner wins, and H': a senior wins, and the evidence E: a woman wins. We find that

$$Pr(H|E) = 3/5 > Pr(H) = 1/2$$

so E confirms H. But

$$Pr(H \vee H'|E) = 3/5 < (\ Pr(H \vee H') = 7/10.$$

So E does not confirm $H \vee H'$; in fact E confirms $\sim(H \vee H')$ and so disconfirms $H \vee H'$ even though $H \vee H'$ is a consequence of H.

The upshot is that on the incremental conception of confirmation, Hempel's adequacy conditions and, hence, his definition of qualitative confirmation, are inadequate. However, his adequacy conditions fare better on the high probability conception of confirmation according to which E confirms H relative to K just in case $Pr(H|E.K) > r$, where r is some number greater than 0.5. But this notion of

[14] As would be the case if learning from experience is modeled as change of probability function through conditionalization; that is, when K is learned, Pr_{old} is placed by $Pr_{new}(\ ) = Pr_{old}(\ |K)$. From this point of view, Bayes's theorem (Rule 9) describes how probability changes when a new fact is learned.

confirmation cannot be what Hempel has in mind; for he wants to say that the observation of a single black raven (E) confirms the hypothesis that all ravens are black (H), although for typical K, $Pr(H|E.K)$ will surely not be as great as 0.5. Thus, in what follows we continue to work with the incremental concept.

The probabilistic approach to confirmation coupled with a simple application of Bayes's theorem also serves to reveal a kernel of truth in the H-D method. Suppose that the following conditions hold:

$$\text{(i) } H, K \vdash E; \text{ (ii) } 1 > Pr(H|K) > 0; \text{ and (iii) } 1 > Pr(E|K) > 0.$$

Condition (i) is the basic H-D condition. Conditions (ii) and (iii) say that neither H nor E is known on the basis of the background information K to be almost surely false or almost surely true. Then on the incremental conception it follows, as the H-D methodology would have it, that E confirms H on the basis of K. By Bayes's theorem

$$Pr(H|E.K) = \frac{Pr(H|K)}{Pr(E|K)}$$

since by (i),

$$Pr(E|H.K) = 1.$$

It then follows from (ii) and (iii) that

$$Pr(H|E.K) > Pr(H|K).$$

Notice also that the smaller $Pr(E|K)$ is, the greater the incremental confirmation afforded by E. This helps to ground the intuition that "surprising" evidence gives better confirmational value. However, this observation is really double-edged as will be seen in Section 2.10.

The Bayesian analysis also affords a means of handling a disquieting feature of the H-D method, sometimes called the problem of irrelevant conjunction. If the H-D condition (i) holds for H, then it also holds for $H.X$ where X is anything you like, including conjuncts to which E is intuitively irrelevant. In one sense the problem is mirrored in the Bayesian approach, for assuming that $1 > Pr(H.X|K) > 0$, it follows that E incrementally confirms $H.X$. But since the special consequence condition does not hold in the Bayesian approach, we cannot infer that E confirms the consequence X of $H.X$. Moreover, under the H-D condition (i), the incremental confirmation of a hypothesis is directly proportional to its prior probability. Since $Pr(H|K) \geq Pr(H.X|K)$, with strict inequality holding in typical cases, the incremental confirmation for H will be greater than for $H.X$.

Bayesian methods are flexible enough to overcome various of the shortcomings of Hempel's account. Nothing, for example, prevents the explication of confirmation in terms of a Pr-function which allows observational evidence to boost the probability of theoretical hypotheses. In addition the Bayesian approach illuminates the paradoxes of the ravens and Goodman's paradox.

In the case of the ravens paradox we may grant that the evidence that the individual a is a piece of white chalk can confirm the hypothesis that "All ravens are black" since, to put it crudely, this evidence exhausts part of the content of the hypothesis. Nevertheless, as Suppes (1966) has noted, if we are interested in subjecting the hypothesis to a sharp test, it may be preferable to do outdoor ornithology and sample from the class of ravens rather than sampling from the class of nonblack things. Let a denote a randomly chosen object and let

$$Pr(Ra.Ba) = p_1, \quad Pr(Ra.{\sim}Ba) = p_2$$
$$Pr({\sim}Ra.Ba) = p_3, \quad Pr({\sim}Ra.{\sim}Ba) = p_4.$$

Then

$$Pr({\sim}Ba|Ra) = p_2 \div (p_1 + p_2)$$
$$Pr(Ra|{\sim}Ba) = p_2 \div (p_2 + p_4)$$

Thus, $Pr({\sim}Ba|Ra) > Pr(Ra|{\sim}Ba)$ just in case $p_4 > p_1$. In our world it certainly seems true that $p_4 > p_1$. Thus, Suppes concludes that sampling ravens is more likely to produce a counterinstance to the ravens hypothesis than is sampling the class of nonblack things.

There are two problems here. The first is that it is not clear how the last statement follows since a was supposed to be an object drawn at random from the universe at large. With that understanding, how does it follow that $Pr({\sim}Ba|Ra)$ is the probability that an object drawn at random from the class of ravens is nonblack? Second, it is the anti-inductivists such as Popper (see item 4 in Section 2.8 above and 2.10 below) who are concerned with attempts to falsify hypotheses. It would seem that the Bayesian should concentrate on strategies that enhance absolute and incremental probabilities. An approach due to Gaifman (1979) and Horwich (1982) combines both of these points.

Let us make it part of the background information K that a is an object drawn at random from the class of ravens while b is an object drawn at random from the class of nonblack things. Then an application of Bayes's theorem shows that

$$Pr(H|Ra.Ba.K) > Pr(H|{\sim}Rb.{\sim}Bb.K)$$

just in case

$$1 > Pr({\sim}Rb|K) > Pr(Ba|K).$$

To explore the meaning of the latter inequality, use the principle of total probability to find that

$$Pr(Ba|K) = Pr(Ba|H.K) \cdot Pr(H|K) + Pr(Ba|{\sim}H.K) \cdot Pr({\sim}H|K)$$
$$= Pr(H|K) + Pr(Ba|{\sim}H.K) \cdot Pr({\sim}H|K)$$

and that

$$Pr({\sim}Rb|K) = Pr(H|K) + Pr({\sim}Rb|{\sim}H.K) \cdot Pr({\sim}H|K).$$

So the inequality in question holds just in case

$$1 > Pr({\sim}Rb|{\sim}H.K) > Pr(Ba|{\sim}H.K),$$

or

$$Pr({\sim}Ba|{\sim}H.K) > Pr(Rb|{\sim}H.K) > 0,$$

which is presumably true in our universe. For supposing that some ravens are nonblack, a random sample from the class of ravens is more apt to produce such a bird than is a random sample from the class of nonblack things since the class of nonblack things is much larger than the class of ravens. Thus, under the assumption of the stated sampling procedures, the evidence $Ra.Ba$ does raise the probability of the ravens hypothesis more than the evidence ${\sim}Rb.{\sim}Bb$ does. The reason for this is precisely the differential propensities of the two sampling procedures to produce counterexamples, as Suppes originally suggested.

The Bayesian analysis also casts light on the problems of induction, old and new, Humean and Goodmanian. Russell (1948) formulated two categories of induction by enumeration:

> Induction by simple enumeration is the following principle: "Given a number n of α's which have been found to be β's, and no α which has been found to be not a β, then the two statements: (a) 'the next α will be a β,' (b) 'all α's are β's,' both have a probability which increases as n increases, and approaches certainty as a limit as n approaches infinity."
>
> I shall call (a) "particular induction" and (b) "general induction." (1948, 401)

Between Russell's "particular induction" and his "general induction" we can interpolate another type, as the following definitions show (note that Russell's "α" and "β" refer to properties, not to individual things):

Def. Relative to K, the predicate "P" is *weakly projectible* over the sequence of individuals $a_1, a_2, \ldots$ just in case[15]

$$\lim_{n \to \infty} Pr(Pa_{n+1}|Pa_1. \ldots .Pa_n.K) = 1.$$

Def. Relative to K, "P" is *strongly projectible* over $a_1, a_2, \ldots$ just in case

$$\lim_{n,\, m \to \infty} Pr(Pa_{n+1}. \ldots .Pa_{n+m}| Pa_1. \ldots .Pa_n.K) = 1.$$

(The notation $\lim_{m,\, n \to \infty}$ indicates the limit as m and n both tend to infinity in any manner you like.) A sufficient condition for both weak and strong probability is that the general hypothesis H: $(i)Pa_i$ receives a nonzero prior probability. To see that it is sufficient for weak projectibility, we follow Jeffreys's (1957) proof. By Bayes's theorem

$$Pr(H|Pa_1. \ldots .Pa_{n+1}.K) = \frac{Pr(Pa_1. \ldots .Pa_{n+1}|H.K) \cdot Pr(H|K)}{Pr(Pa_1. \ldots .Pa_{n+1}|K)}$$

$$= \frac{Pr(H|K)}{Pr(Pa_1|K) \cdot Pr(Pa_2|Pa_1.K) \cdot \ldots \cdot Pr(Pa_{n+1}|Pa_1. \ldots .Pa_n.K)}$$

15 Equation $\lim_{n \to \infty} x_n = L$ means that, for any real number $\epsilon > 0$, there is an integer $N > 0$ such that, for all $n > N$, $|x_n - L| < \epsilon$.

Unless $Pr(Pa_{n+1}|Pa_1. \ldots .Pa_n.K)$ goes to 1 as $n \to \infty$, the denominator on the right-hand side of the second equality will eventually become less than $Pr(H|K)$, contradicting the truth of probability that the left-hand side is no greater than 1.

The posit that

$$\text{(P)} \quad Pr([(i)Pa_i|K] > 0$$

is not necessary for weak projectibility. Carnap's systems of inductive logic (see item 6 in Section 2.8 above) are relevant examples since in these systems (P) fails in a universe with an infinite number of individuals although weak projectibility can hold in these systems.[16] But if we impose the requirement of countable additivity

$$\text{(CA)} \quad \lim_{n \to \infty} Pr(Pa_i. \ldots .Pa_n|K) = Pr[(i)\, Pa_i|K]$$

then (P) is necessary as well as sufficient for strong projectibility.

Also assuming (CA), (P) is sufficient to generate a version of Russell's "general induction," namely

$$\text{(G)} \quad \lim_{n \to \infty} Pr[(i)Pa_i|Pa_1. \ldots .Pa_n.K] = 1.$$

(Russell 1948 lays down a number of empirical postulates he thought were necessary for induction to work. From the present point of view these postulates can be interpreted as being directed to the question of which universal hypotheses should be given nonzero priors.)

Humean skeptics who regiment their beliefs according to the axioms of probability cannot remain skeptical about the next instance or the universal generalization in the face of ever-increasing positive instances (and no negative instances) unless they assign a zero prior to the universal generalization. But

$$Pr[(i)Pa_i|K] = 0$$

implies that

$$Pr[(\exists i) \sim Pa_i|K] = 1,$$

which says that there is certainty that a counterinstance exists, which does not seem like a very skeptical attitude.

[16] A nonzero prior for the general hypothesis is a necessary condition for strong projectibility but not for weak projectibility. The point can be illustrated by using de Finetti's representation theorem, which says that if P is exchangeable over $a_1, a_2, \ldots$ (which means roughly that the probability does not depend on the order) then:

$$Pr(Pa_1.Pa_2. \ldots .Pa_n \mid K) = \int_0^1 \theta^n \, d\mu(\theta)$$

where $\mu(\theta)$ is a uniquely determined measure on the unit interval $0 \leq \theta \leq 1$. For the uniform measure $d\mu(\theta) = d(\theta)$ we have

$$Pr\,(Pa_{n+1}|Pa_1. \ldots . Pa_n. K) = n + 1/n + 2$$

and

$$Pr\,(Pa_{n+1}. \ldots . Pa_{n+m}|Pa_1. \ldots . Pa_n.K) = m + 1/n + m + 1.$$

Note also that the above results on instance induction hold whether "P" is a normal or a Goodmanized predicate—for example, they hold just as well for P^*a_i which is defined as

$$[(i \leq 2000).Pa_i] \vee [(i > 2000).\sim Pa_i)],$$

where Pa_i means that a_i is purple. But this fact just goes to show how weak the results are; in particular, they hold only in the limit as $n \to \infty$ and they give no information about how rapidly the limit is approached.

Another way to bring out the weakness is to note that (P) does not guarantee even a weak form of Hume projectibility.

Def. Relative to K, "P" is *weakly Hume projectible* over the doubly infinite sequence $\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots$ just in case for any n,
$\lim_{k \to \infty} Pr(Pa_n|Pa_{n-1}. \ldots .Pa_{n-k} . K) = 1.$

(To illustrate the difference between the Humean and non-Humean versions of projectibility, let Pa_n mean that the sun rises on day n. The non-Humean form of projectibility requires that if you see the sun rise on day 1, on day 2, and so on, then for any $\varepsilon > 0$ there will come a day N when your probability that the sun will rise on day $N + 1$ will be at least $1 - \varepsilon$. By contrast, Hume projectibility requires that if you saw the sun rise yesterday, the day before yesterday, and so on into the past, then eventually your confidence that the sun will rise tomorrow approaches certainty.)

If (P) were sufficient for Hume projectibility we could assign nonzero priors to both $(i)Pa_i$ and $(i)P^*a_i$, with the result that as the past instances accumulate, the probabilities for Pa_{2001} and for P^*a_{2001} both approach 1, which is a contradiction.

A sufficient condition for Hume projectibility is exchangeability.

Def. Relative to K, "P" is *exchangeable* for Pr over the a_is just in case for any n and m

$$Pr(\pm Pa_n. \ldots . \pm Pa_{n+m}|K) = Pr(\pm Pa_{n'}. \ldots . \pm Pa_{n'+m'}|K)$$

where $\pm$ indicates that either P or its negation may be chosen and $[a_{i'}]$ is any permutation of the a_is in which all but a finite number are left fixed. Should we then use a Pr-function for which the predicate "purple" is exchangeable rather than the Goodmanized version of "purple"? Bayesianism per se does not give the answer anymore than it gives the answer to who will win the presidential election in the year 2000. But it does permit us to identify the assumptions needed to guarantee the validity of one form or another of induction.

Having touted the virtues of the Bayesian approach to confirmation, it is now only fair to acknowledge that it is subject to some serious challenges. If it can rise to these challenges, it becomes all the more attractive.

2.10 CHALLENGES TO BAYESIANISM

1. Nonzero priors. Popper (1959) claims that "in an infinite universe . . . the probability of any (non-tautological) universal law will be zero." If Popper were right and universal generalizations could not be probabilified, then Bayesianism would be worthless as applied to theories of the advanced sciences, and we would presumably have to resort to Popper's method of corroboration (see item 4 in Section 2.8 above).

To establish Popper's main negative claim it would suffice to show that the prior probability of a universal generalization must be zero. Consider again $H: (i)Pa_i$. Since for any n

$$H \vdash Pa_1 . Pa_2 . \ldots . Pa_n,$$
$$Pr(H|K) \leq \lim_{n \to \infty} Pr(Pa_1 . \ldots . Pa_n|K).$$

Now suppose that

(I) For all n, $Pr(Pa_1 . \ldots . Pa_n|K) = Pr(Pa_1|K) \cdot \ldots \cdot Pr(Pa_n|K)$

and that

(E) For any m and n, $Pr(Pa_m|K) = Pr(Pa_n|K)$.

Then except for the uninteresting case that $Pr(Pa_n|K) = 1$ for each n, it follows that

$$\lim_{n \to \infty} Pr(Pa_1 . \ldots . Pa_n|K) = 0$$

and thus that $Pr(H|K) = 0$.

Popper's argument can be attacked in various places. Condition (E) is a form of exchangeability, and we have seen above that it cannot be expected to hold for all predicates. But Popper can respond that if (E) does fail then so will various forms of inductivism (e.g., Hume projectibility). The main place the inductivist will attack is at the assumption (I) of the independence of instances. Popper's response is that the rejection of (I) amounts to the postulation of something like a causal connection between instances. But this a red herring since the inductivist can postulate a probabilistic dependence among instances without presupposing that the instances are cemented together by some sort of causal glue.

In another attempt to show that probabilistic methods are ensnared in inconsistencies, Popper cites Jeffreys's proof sketched above that a non-zero prior for $(i)Pa_i$ guarantees that

$$\lim_{n \to \infty} Pr(Pa_{n+1}|Pa_1 . \ldots . Pa_n . K) = 1.$$

But, Popper urges, what is sauce for the goose is sauce for the gander. For we can do the same for a Goodmanized P^*, and from the limit statements we can conclude that for some $r > 0.5$ there is a sufficiently large N such that for any $N' > N$, the probabilities for $P_{a_{N'}}$ and for $P^*_{a_{N'}}$ are both greater than r, which is a contradiction for appropriately chosen P^*. But the reasoning here is fallacious and there is in fact no contradiction lurking in Jeffreys's limit theorem since the convergence is not supposed to be uniform over different predicates—indeed, Popper's reasoning shows that it cannot be.

Of course, none of this helps with the difficult questions of which hypotheses should be assigned nonzero priors and how large the priors should be. The example from item 5 in Section 2.8 above suggests that the latter question can be ignored to some extent since the accumulation of evidence tends to swamp differences in priors and force merger of posterior opinion. Some powerful results from advanced probability theory show that such merger takes place in a very general setting (on this matter see Gaifman and Snir 1982).

2. Probabilification vs. inductive support. Popper and Miller (1983) have argued that even if it is conceded that universal hypotheses may have nonzero priors and thus can be probabilified further and further by the accumulation of positive evidence, the increase in probability cannot be equated with genuine inductive support. This contention is based on the application of two lemmas from the probability calculus:

Lemma 1. $Pr(\sim H|E.K) \times Pr(\sim E|K) = Pr(H \vee \sim E|K) - Pr(H \vee \sim E|E.K)$.

Lemma 1 leads easily to

Lemma 2. If $Pr(H|E.K) < 1$ and $Pr(E|K) < 1$ then

$$Pr(H \vee \sim E|E.K) < Pr(H \vee \sim E|K).$$

Let us apply Lemma 2 to the case discussed above where Bayesianism was used to show that under certain conditions the H-D method does lead to incremental confirmation. Recall that we assumed that

$$H, K \vdash E;\ 1 > Pr(E|K) > 0;\ \text{and}\ 1 > Pr(H|K) > 0$$

and then showed that

$$Pr(H|E.K) > Pr(H|K),$$

which the inductivists want to interpret as saying that E inductively supports H on the basis of K. Against this interpretation, Popper and Miller note that H is logically equivalent to $(H \vee E).(H \vee \sim E)$. The first conjunct is deductively implied by E, leading Popper and Miller to identify the second conjunct as the part of H that goes beyond the evidence. But by Lemma 2 this part is *countersupported* by E, except in the uninteresting case that $E.K$ makes H probabilistically certain.

Jeffrey (1984) has objected to the identification of $H \vee \sim E$ as the part of H that goes beyond the evidence. To see the basis of his objection, take the case where

$$H: (i)Pa_i \text{ and } E: Pa_1 . \ldots . Pa_n.$$

Intuitively, the part of H that goes beyond this evidence is $(i)\,[\,(i > n) \ldots Pa_i]$ and not the Popper-Miller $(i)Pa_1 \vee \sim(Pa_1 . \ldots . Pa_n)$.

Gillies (1986) restated the Popper-Miller argument using a measure of inductive support based on the incremental model of confirmation: (leaving aside K) the support given by E to H is $S(H, E) = Pr(H|E) - Pr(H)$. We can then show that

Lemma 3. $S(H, E) = S(H \vee E, E) + S(H \vee \sim E, E)$.

Gillies suggested that $S(H \vee EE,)$ be identified as the deductive support given H by E and $S(H \vee {\sim}E, E)$ as the inductive support. And as we have already seen, in the interesting cases the latter is negative. Dunn and Hellman (1986) responded by dualizing. Hypothesis H is logically equivalent to $(H.E) \vee (H.{\sim}E)$ and $S(H, E) = S(H.E, E) + S(H.{\sim}E, E)$. Identify the second component as the deductive countersupport. Since this is negative, any positive support must be contributed by the first component which is a measure of the nondeductive support.

3. The problem of old evidence. In the Bayesian identification of the valid kernel of the H-D method we assumed that $Pr(E|K) < 1$, that is, there was some surprise to the evidence E. But this is often not the case in important historical examples. When Einstein proposed his general theory of relativity (H) at the close of 1915 the anomalous advance of the perihelion of Mercury (E) was old news, that is, $Pr(E|K) = 1$. Thus, $Pr(H|E.K) = Pr(H|K)$, and so on the incremental conception of confirmation, Mercury's perihelion does not confirm Einstein's theory, a result that flies in the face of the fact that the resolution of the perihelion problem was widely regarded as one of the major triumphs of general relativity. Of course, one could seek to explain the triumph in nonconfirmational terms, but that would be a desperate move.

Garber (1983) and Jeffrey (1983) have suggested that Bayesianism be given a more human face. Actual Bayesian agents are not logically omniscient, and Einstein for all his genius was no exception. When he proposed his general theory he did not initially know that it did in fact resolve the perihelion anomaly, and he had to go through an elaborate derivation to show that it did indeed entail the missing 43″ of arc per century. Actual flesh and blood scientists learn not only empirical facts but logicomathematical facts as well, and if we take the new evidence to consist in such facts we can hope to preserve the incremental model of confirmation. To illustrate, let us make the following assumptions about Einstein's degrees of belief in 1915:

(a) $Pr(H|K) > 0$ (Einstein assigned a nonzero prior to his general theory.)

(b) $Pr(E|K) = 1$ (The perihelion advance was old evidence.)

(c) $Pr(H \vdash E|K) < 1$ (Einstein was not logically omniscient and did not invent his theory so as to guarantee that it entailed the 43″.)

(d) $Pr[(H \vdash E) \vee (H \vdash {\sim}E)|K] = 1$ (Einstein knew that his theory entailed a definite result for the perihhelion motion.)

(e) $Pr[H.(H \vdash {\sim}E)|K] = Pr[H.(H \vdash {\sim}E).{\sim}E|K]$ (Constraint on interpreting $\vdash$ as logical implication.)

From (a)–(e) it can be shown that $Pr[H|(H \vdash E).K]. > Pr(H|K)$. So learning that his theory entailed the happy result served to increase Einstein's confidence in the theory.

Although the Garber-Jeffrey approach does have the virtue of making Bayesian agents more human and, therefore, more realistic, it avoids the question of whether the perihelion phenomena did in fact confirm the general theory of relativity in favor of focusing on Einstein's personal psychology. Nor is it adequate to dismiss this

concern with the remark that the personalist form of Bayesianism is concerned precisely with psychology of particular agents, for even if we are concerned principally with Einstein himself, the above calculations seem to miss the mark. We now believe that for Einstein in 1915 the perihelion phenomena provided a strong confirmation of his general theory. And contrary to what the Garber-Jeffrey approach would suggest, we would not change our minds if historians of science discovered a manuscript showing that as Einstein was writing down his field equations he saw in a flash of mathematical insight that $H \vdash E$ or alternatively that he consciously constructed his field equations so as to guarantee that they entailed E. "Did E confirm H for Einstein?" and "Did learning that $H \vdash E$ increase Einstein's confidence in H?" are two distinct questions with possibly different answers. (In addition, the fact that agents are allowed to assign $Pr\ (H \vdash E|K) < 1$ means that the Dutch book justification for the probability axioms has to be abandoned. This is anathema for orthodox Bayesian personalists who identify with the betting quotient definition of probability.)

A different approach to the problem of old evidence is to apply the incremental model of confirmation to the counterfactual degrees of belief that would have obtained had E not been known. Readers are invited to explore the prospects and problems of this approach for themselves. (For further discussion of the problem of old evidence, see Howson 1985, Eells 1985, and van Fraassen 1988.)

2.11 CONCLUSION

The topic of this chapter has been *the logic of science*. We have been trying to characterize and understand the patterns of inference that are considered legitimate in establishing scientific results—in particular, in providing support for the hypotheses that become part of the corpus of one science or another. We began by examining some extremely simple and basic modes of reasoning—the hypothetico-deductive method, instance confirmation, and induction by enumeration. Certainly (pace Popper) all of them are frequently employed in actual scientific work.

We find—both in contemporary science and in the history of science—that scientists do advance hypotheses from which (with the aid of initial conditions and auxiliary hypotheses) they deduce observational predictions. The test of Einstein's theory of relativity in terms of the bending of starlight passing close to the sun during a total solar eclipse is an oft-cited example. Others were given in this chapter. Whether the example is as complex as general relativity or as simple as Boyle's law, the logical problems are the same. Although the H-D method contains a valid kernel—as shown by Bayes's rule—it must be considered a serious oversimplification of what actually is involved in scientific confirmation. Indeed, Bayes's rule itself seems to offer a schema far more adequate than the H-D method. But—as we have seen—it, too, is open to serious objections (such as the problem of old evidence).

When we looked at Hempel's theory of instance confirmation, we discussed an example that has been widely cited in the philosophical literature—namely, the generalization "All ravens are black." If this is a scientific generalization, it is certainly at a low level, but it is not scientifically irrelevant. More complex examples raise the same logical problems. At present, practicing scientists are concerned with—and excited by—such generalizations as, "All substances having the chemical structure given by the formula $YBa_2Cu_3O_7$ are superconductors at 70 kelvins." As if indoor ornithology weren't bad enough, we see, by Hempel's analysis, that we can confirm this latter-day generalization by observing black crows. It seems that observations by birdwatchers can confirm hypotheses of solid state physics. (We realize that bird-lovers would disapprove of the kind of test that would need to be performed to establish that a raven is not a superconductor at 70°K.) We have also noted, however, the extreme limitations of the kind of evidence that can be gathered in any such fashion.

Although induction by enumeration is used to establish universal generalizations, its most conspicuous use in contemporary science is connected with statistical generalizations. An early example is found in Rutherford's counting of the frequencies with which alpha particles bombarding a gold foil were scattered backward (more or less in the direction from which they came). The counting of instances led to a statistical hypothesis attributing stable frequencies to such events. A more recent example—employing highly sophisticated experiments—involves the detection of neutrinos emitted by the sun. Physicists are puzzled by the fact that they are detecting a much smaller frequency than current theory predicts. (Obviously probabilities of the type characterized as frequencies are involved in examples of the sort mentioned here.) In each of these cases an inductive extrapolation is drawn from observed frequencies. In our examination of induction by enumeration, however, we have found that it is plagued by Hume's old riddle and Goodman's new one.

One development of overwhelming importance in twentieth-century philosophy of science has been the widespread questioning of whether there is any such thing as a *logic* of science. Thomas Kuhn's influential work, *The Structure of Scientific Revolutions* (1962, 1970), asserted that the choice of scientific theories (or hypotheses) involves factors that go beyond observation and logic—including judgement, persuasion, and various psychological and sociological influences. There is, however, a strong possibility that, when he wrote about going beyond the bounds of observation and logic, the kind of logic he had in mind was the highly inadequate H-D schema, (see Salmon 1989 for an extended discussion of this question, and for an analysis of Kuhn's views in the light of Bayes's rule). The issues raised by the Kuhnian approach to philosophy of science are discussed at length in Chapter 4 of this book.

Among the problems we have discussed there are—obviously—many to which we do not have adequate solutions. Profound philosophical difficulties remain. But the deep and extensive work done by twentieth-century philosophers of science in these areas has cast a good deal of light on the nature of the problems. It is an area in which important research is currently going on and in which significant new results are to be expected.

DISCUSSION QUESTIONS

1. Select a science with which you are familiar and find a case in which a hypothesis or theory is taken to be confirmed by some item of evidence. Try to characterize the relationship between the

evidence and hypothesis or theory confirmed in terms of the schemas discussed here. If none of them is applicable, can you find a new schema that is?

2. If the prior probability of every universal hypothesis is zero how would you have to rate the probability of the statement that unicorns (at least one) exist? Explain your answer.
3. Show that accepting the combination of the entailment condition, the special consequence condition, and the converse consequence condition (see Section 2.4) entails that any E confirms any H.
4. Consider a population that consists of all of the adult population of some particular district. We want to test the hypothesis that all voters are literate,

$$(x)(Vx \supset Lx),$$

which is, of course, equivalent to

$$(x)(\sim Vx \supset \sim Lx).$$

Suppose that approximately 75 percent of the population are literate voters, approximately 15 percent are literate nonvoters, approximately 5 percent are illiterate nonvoters, and approximately 5 percent are illiterate voters—but this does not preclude the possibility that no voters are illiterate. Would it be best to sample the class of voters or the class of illiterate people? Explain your answer. (This example is given in Suppes 1966, 201.)

5. Goodman's examples challenge the idea that hypotheses are confirmed by their instances. Goodman holds that the distinction between those hypotheses that are and those that are not projectable on the basis of their instances is to be drawn in terms of *entrenchment*. Predicates become entrenched as antecedents or consequents by playing those roles in universal conditionals that are actually projected. Call a hypothesis *admissible* just in case it has some positive instances, no negative instances, and is not exhausted. Say that H *overrrides* H' just in case H and H' conflict, H is admissible and is better entrenched than H' (i.e., has a better entrenched antecedent and equally well entrenched consequent or vice versa), and H is not in conflict with some still better entrenched admissible hypothesis. Critically discuss the idea that H is projectable on the basis of its positive instances just in case it is admissible but not overridden.
6. Show that

$$H: (x)\ (\exists\ y)\ Rxy.(x) \sim Rxx.(x)\ (y)\ (z)\ [(Rxy.Ryz) \supset Rxz]$$

cannot be Hempel-confirmed by any consistent E.

7. It is often assumed in philosophy of science that if one is going to represent numerically the degree to which evidence E supports hypothesis H with respect to background B, then the numbers so produced — $P(H|E.B)$ — must obey the probability calculus. What are the prospects of alternative calculi? (Hint: Consider each of the axioms in turn and ask under what circumstances each axiom could be violated in the context of a confirmation theory. What alternative axiom might you choose?)
8. If Bayes's rule is taken as a schema for confirmation of scientific hypotheses, it is necessary to decide on an interpretation of probability that is suitable for that context. It is especially crucial to think about how the prior probabilities are to be understood. Discuss this problem in the light of the admissible interpretations offered in this chapter.
9. William Tell gave his young cousin Wesley a two-week intensive archery course. At its completion, William tested Wes's skill by asking him to shoot arrows at a round target, ten feet in radius with a centered bull's-eye, five feet in radius.

"You have learned *no* control at all," scolded William after the test. "Of those arrows that hit the target, five are within five feet of dead center and five more between five and ten feet from dead center." "Not so," replied Wes, who had been distracted from archery practice by his newfound love of geometry. "That five out of ten arrows on the target hit the bull's-eye shows I *do* have control. The bullseye is only one quarter the total area of the target."

Adjudicate this dispute in the light of the issues raised in the chapter. Note that an alternative form of Bayes's rule which applies when one considers the relative confirmation accrued by two hypotheses H_1 and H_2 by evidence E with respect to background B is:

$$\frac{Pr(H_1|E.B)}{Pr(H_2|E.B)} = \frac{Pr(E|H_1.B)}{Pr(E|H_2.B)} \cdot \frac{Pr(H_1|B)}{Pr(H_2|B)}$$

10. Let $\{H_1, H_2, \ldots, H_n\}$ be a set of competing hypotheses. Say that E selectively Hempel-confirms some H_j just in case it Hempel-confirms H_j but fails to confirm the alternative Hs. Use this notion of selective confirmation to discuss the relative confirmatory powers of black ravens versus nonblack nonravens for alternative hypotheses about the color of ravens.
11. Prove Lemmas 1, 2, and 3 of Section 2.10.
12. Discuss the prospects of resolving the problem of old evidence by using counterfactual degrees of belief, that is, the degrees of belief that would have obtained had the evidence E not been known.
13. Work out the details of the following example, which was mentioned in Section 2.8. There is a square piece of metal in a closed box. You cannot see it. But you are told that its area is somewhere between 1 square inch and 4 square inches. Show how the use of the principle of indifference can lead to conflicting probability values.
14. Suppose there is a chest with two drawers. In each drawer are two coins; one drawer contains two gold coins, the other contains one gold coin and one silver coin. A coin will be drawn from one of these drawers. Suppose, further, that you know (without appealing to the principle of indifference) that each drawer has an equal chance of being chosen for the draw, and that, within each drawer, each coin has an equal chance of being chosen. When the coin is drawn it turns out to be gold. What is the probability that the other coin in the same drawer is gold? Explain how you arrived at your answer.
15. Discuss the problem of ascertaining limits of relative frequencies on the basis of observed frequencies in initial sections of sequences of events. This topic is especially suitable for those who have studied David Hume's problem regarding the justification of inductive inference in Part II of this chapter.
16. When scientists are considering new hypotheses they often appeal to plausibility arguments. As a possible justification for this procedure, it has been suggested that plausibility arguments are attempts at establishing prior probabilities. Discuss this suggestion, using concrete illustrations from the history of science or contemporary science.
17. Analyze the bootstrap confirmation of the perfect gas law in such a way that no "macho" bootstrapping is used, that is, the gas law itself is not used as an auxiliary to deduce instances of itself.

SUGGESTED READINGS

GLYMOUR, CLARK (1980), *Theory and Evidence*. Princeton: Princeton University Press. This book, which is rather technical in parts, contains the original presentation of bootstrap confirmation.

GOODMAN, NELSON (1955), *Fact, Fiction, and Forecast*. 1st ed. Cambridge, MA: Harvard University Press. This twentieth-century classic is now in its 4th edition. Chapter 3 contains Goodman's dissolution of "the old riddle of induction" and presents his "new riddle of induction"—the grue-bleen paradox. Chapter 4 gives Goodman's solution of the new riddle in terms of projectibility.

HEMPEL, CARL G. (1945), "Studies in the Logic of Confirmation," *Mind 54:* 1–26, 97–121. Reprinted in Hempel (1965, see Bibliography), with a 1964 Postscript added. This classic essay contains Hempel's analysis of the Nicod criterion of confirmation, and it presents Hempel's famous paradox of the ravens, along with his analysis of it.

——— (1966), *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall. Chapters 2–4 provide an extremely elementary and readable introduction to the concept of scientific confirmation.

HUME, DAVID (1748), *An Enquiry Concerning Human Understanding*. Many editions available. This is *the* philosophical classic on the problem of induction, and it is highly readable. Sections 4–7 deal with induction, causality, probability, necessary connection, and the uniformity of nature.

POPPER, KARL R. (1972), "Conjectural Knowledge: My Solution of the Problem of Induction," in *Objective Knowledge*. Oxford: Clarendon Press, pp. 1–31. This is an introductory presentation of his deductivist point of view.

SALMON, WESLEY C. (1967), *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh. This book provides an introductory, but moderately thorough, survey of many of the issues connected with confirmation, probability, and induction.

STRAWSON, P. F. (1952), *Introduction to Logical Theory*. London: Methuen. Chapter 9 presents the ordinary language dissolution of the problem of induction.