

Lewis, C. I., [Analysis], An Analysis of Knowledge and Evaluation, La Salle, Ill., 1946.


The Emergence of Novelty. New York, 1935.

Popper, Karl, [Forschung], Logik der Forschung. Wien, 1935.


Tolman, Edward Chase, [Behavior], "Purposeful Behavior in Animals and Men." New York, 1925.


The philosophical theory of scientific explanation first entered the twentieth century in 1948, for that was the year of publication of the earliest bona fide attempt to provide a systematic account of statistical explanation in science. 1 Although the need for some sort of inductive or statistical form of explanation had been acknowledged earlier, Hempel's essay "Deductive-Nomological vs. Statistical Explanation" (1962) contained the first sustained and detailed effort to provide a precise account of this mode of scientific explanation. Given the pervasiveness of statistics in virtually every branch of contemporary science, the late arrival of statistical explanation in philosophy of science is remarkable. Hempel's initial treatment of statistical explanation had various defects, some of which he attempted to rectify in his comprehensive essay "Aspects of Scientific Explanation" (1965a). Nevertheless, the earlier article did show unmistakably that the construction of an adequate model for statistical explanation involves many complications and subtleties that may have been largely unanticipated. Hempel never held the view—expressed by some of the more avid devotees of the D-N model—that all adequate scientific explanations must conform to the deductive-nomological pattern. The 1948 Hempel-Oppenheim paper explicitly notes the need for an inductive or statistical model of scientific explanation in order to account for some types of legitimate explanation that actually occur in the various sciences (Hempel, 1965, pp. 250–251). The task of carrying out the construction was, however, left for another occasion. Similar passing remarks regarding the need for inductive or statistical accounts were made by other authors as well, but the project was not undertaken in earnest until 1962—a striking delay of fourteen years after the 1948 essay.

One can easily form the impression that philosophers had genuine feelings of ambivalence about statistical explanation. A vivid example can be found in Carnap's Philosophical Foundations of Physics (1966), which was presented in a seminar he offered at UCLA in 1958. 2 Early in the first chapter, he says:

The general schema involved in all explanation can be expressed symbolically as follows:

1. (p) \( \rightarrow (Q \rightarrow R) \)
2. \( Q \rightarrow P \)
3. \( Q \rightarrow \alpha \)

The first statement is the universal law that applies to any object \( x \). The second statement asserts that a particular object \( x \) has the property...
P. These two statements taken together enable us to derive logically the third statement: object A has the property Q (1966, pp. 7–8, italics added).

After a single intervening paragraph, he continues:

At times, in giving an explanation, the only known laws that apply are statistical rather than universal. In such cases, we must be content with a statistical explanation (1966, p. 8, italics added).

Farther down on the same page, he assures us that "these are genuine explanations," and on the next page he points out that "in quantum theory . . . we meet with statistical laws that may not be the result of ignorance; these laws express the basic structure of these world." I must confess to a reaction of astonishment at being told that all explanations are deductive-nomological, but that some are not, because they are statistical. This lapse was removed from the subsequent paperback edition ( Carnap, 1974), which appeared under a new title.

Why did it take philosophers so long to get around to providing serious treatment of statistical explanation? It certainly was not due to any absence of statistical explanations in science. In antiquity, Lucretius (151, pp. 66–68) had based his entire cosmology upon explanations involving spontaneous swerving of atoms, and some of his explanations of more restricted phenomena can readily be interpreted as statistical. He asked, for example, why it is that Roman housewives frequently become pregnant after sexual intercourse, while Roman prostitutes to a large extent avoid doing so. Conception occurs, he explains, as a result of a collision between the male seed and a female seed. During intercourse the prostitutes wiggles their hips a great deal, but wives tend to remain passive; as everyone knows, it is much harder to hit a moving target (1951, p. 120). He said that his brothel clients in the interval period, St. Thomas Aquinas asserted:

The majority of men follow their passions, which are moviments of the sensitive appetite, in which movements of heavenly bodies can cooper-ate; but few are wise enough to resist these passions. Consequently, we are able to foetall the truth in the majority of cases, especially in a general way. But not in particular cases, for nothing prevents man resisting his passions by his free will (1947, 1 Lqs. 115, a. 3, ad Obj. 3).

Astrological explanations are, therefore, of the statistical variety. Leibniz, who like Lucretius and Aquinas was concerned about human free will, spoke of causes that incline but do not necessitate (1951, p. 515; 1965, p. 136).

When, in the latter half of the nineteenth century, the kinetic-molecular theory of gases emerged, giving rise to classical statistical mechanical explanations, became firmly entrenched in physics. In this context, it turns out, many phenomena that for all practical purposes appear amenable to strict D-N explanation—such as the melting of an ice cube placed in tepid water—must be admitted strictly speaking to be explained statistically in terms of probabilities almost indistinguishable from unity. On a smaller scale, Darwinian evolution and Mendelian genetics, provide explanations of changes in the gene pool that are statistical. In addition, social scientists approached such topics as suicide, crime, and intelligence by means of "model statistics" (Hempel 1966, p. 253). They relaxed the stringent requirements for D-N explanation in some straightforward way. It might have been felt, on the other hand, that the problems in constructing an appropriate inductive or statistical model were so formidable that one simply did not want to undertake the task. Some philosophers may unreflectively have adopted the former attitude; the latter, it turns out, is closer to the mark.

We should have suspected as much. If D-N explanations are deductive arguments, inductive or statistical explanations are, presumably, inductive arguments. This is precisely the tack Hempel took in constructing his inductive-statistical or I-S model. In proposing a D-N explanation of the fact that this penniless electrician one offers an exaplans consisting of two premises: the particular premise that this penny is composed of copper, and the universal law-statement that all pennies are composed of copper. The explanation-statement follows deductively. To provide an I-S explanation of the fact that I was tired when I arrived in Melbourne for a visit in 1978, it could be pointed out that I had been traveling by air for more than twenty-four hours (including stopovers at airports), and almost everyone who travels by air for twenty-four hours or more becomes fatigued. The explanation generalizes from the inductive premises; the event-to-be-explained is thus subsumed under a statistical generalization.

It has long been known that there are deep and striking disanalogies between inductive and deductive logic. Deductive support is transitive; strong inductive support is not. Contraposition is valid for deductive entailments; it does not hold for probabilities. Analogous entailment and the probabilities that hold in some approximate way if the probabilities involved are high enough; once we abandon strict logical entailment, and turn to probability or inductive support we must wean ourselves from the idea that the conclusions that hold in some approximate way if the probabilities involved are high enough; once we abandon strict logical entailment, and turn to probability or inductive support we must adjust our intuitions accordingly. Hempel brought out clearly in his 1962 essay, the inductive principle that permits the addition of an arbitrary term to the antecedent of an entailment does not extend into inductive logic. If A entails B, then A.C entails B, whatever C may happen to stand for. However, no matter how high the probability of A given B, there is no constraint whatever upon the probability of B given both A and C. To take an extreme case, the probability of a prime number being odd is one, but the probability that a prime number is less than any given number is zero. For those who feel uneasy about applying probability to cases of this arithmetical sort, we can readily supply empirical examples. A thirty-year-old Australian with an advanced case of lung cancer has a low
probability of surviving for five more years, even though the probability of surviving to age thirty-five for thirty-year-old Australians in general is quite high. It is this basic disanalogy between deductive and inductive (probabilistic) relations that gives rise to what Hempel called the ambiguity of inductive-statistical explanation—a phenomenon that, as he emphasized, has no counterpart in D-N explanation. His requirement of maximal specificity was designed expressly to cope with the problem of this ambiguity.

Hempel illustrates the ambiguity of I-S explanation, and the need for the requirement of maximal specificity, by means of the following example (1965, pp. 394-396). John Jones recovers quickly from a streptococcus infection, and when we ask why, we are told that he was given penicillin, and that almost all strep infections clear up quickly after penicillin is administered. The recovery is thus rendered probable relative to these explanatory facts. There are, however, certain strains of streptococcus bacteria that are resistant to penicillin. If, in addition to the above facts, we were told that the infection is of the penicillin-resistant type, then we would have to say that the prompt recovery is rendered improbable relative to the available information. It would clearly be scientifically unacceptable to ignore such relevant evidence as the penicillin-resistant character of the infection; the requirement of maximal specificity is designed to block statistical explanations that thus omit relevant facts. It says, in effect, that when the class to which the individual case is referred for explanatory purposes—such as, in this instance, the class of strep infections treated by penicillin—is chosen, we must not know how to divide it into subcases in which the probability of the fact to be explained differs from its probability in the entire class. If it has been ascertained that this particular case involved the penicillin-resistant strain, then the original explanation of the rapid recovery would violate the requirement of maximal specificity, and for that reason would be judged unsatisfactory.6

Hempel conceived of D-N explanation as valid deductive arguments satisfying certain additional conditions. Explanations that conform to his inductive-statistical or I-S model are correct inductive arguments also satisfying certain additional restrictions. Explanations of both sorts can be characterized in terms of the following four conditions:

1. The explanation is an argument with correct (deductive or inductive) logical form.
2. At least one of the premises must be a universal (or statistical) law.
3. The premises must be true, and
4. The explanation must satisfy the requirement of maximal specificity.

This fourth condition is automatically satisfied by D-N explanations by virtue of the fact that their explanatory laws are universal generalizations. If all A are B, then obviously there is no subset of A in which the probability of B is other than one. This means that all four conditions hold, with respect to explanations of the I-S variety. In general, according to Hempel (1962a, p. 10), an explanation is an argument (satisfying these four conditions) to the effect that the event-to-be-explained was to be expected by virtue of certain explanatory facts. In the case of I-S explanations, this means that the premise must lend high inductive probability to the conclusion—that is, the explanandum must be highly probable with respect to the explanans.

Explanations of the D-N and I-S varieties can therefore be schematized as follows (Hempel, 1965, pp. 336, 382):

\[
\begin{align*}
C_i, C_j, \ldots, C_f \quad \text{(particular explanatory conditions)} \\
L_1, L_2, \ldots, L_k \quad \text{(general laws)} \\
E \quad \text{(fact-to-be-explained)}
\end{align*}
\]

The single line separating the premises from the conclusion signifies that the argument is deductively valid.

\[
C_i, C_j, \ldots, C_f \quad \text{(particular explanatory conditions)} \\
L_1, L_2, \ldots, L_k \quad \text{(general laws, at least one statistical)} \\
E \quad \text{(fact-to-be-explained)}
\]

The double lines separating the premises from the conclusion signifies that the argument is inductively correct, and the number \( r \) expresses the degree of inductive probability with which the premises support the conclusion. It is presumed that \( r \) is fairly close to one.

The high-probability requirement, which seems such a natural analogue of the deductive entailment relation, leads to difficulties in two ways. First, there are arguments that fulfill all of the requirements imposed by the I-S model, but that patently do not constitute satisfactory scientific explanations. One can maintain, for example, that people who have colds will probably get over them within a fortnight if they take vitamin C, but the use of vitamin C may not explain the recovery, since almost all colds clear up within two weeks regardless. In arguing for the use of vitamin C in the prevention and treatment of colds, Linus Pauling (1970) does not base his claims upon the high probability of avoidance or quick recovery; instead, he urges that massive doses of vitamin C have a bearing upon the probability of the recovery. Conceiving the effect of these four conditions to the effect that the event-to-be-explained was to be expected by virtue of certain explanatory facts. In the case of I-S explanations, this means that the premise must lend high inductive probability to the conclusion—that is, the explanandum must be highly probable with respect to the explanans.

Explanations of the D-N and I-S varieties can therefore be schematized as follows (Hempel, 1965, pp. 336, 382):

\[
\begin{align*}
C_i, C_j, \ldots, C_f \quad \text{(particular explanatory conditions)} \\
L_1, L_2, \ldots, L_k \quad \text{(general laws)} \\
E \quad \text{(fact-to-be-explained)}
\end{align*}
\]

The single line separating the premises from the conclusion signifies that the argument is deductively valid.

\[
C_i, C_j, \ldots, C_f \quad \text{(particular explanatory conditions)} \\
L_1, L_2, \ldots, L_k \quad \text{(general laws, at least one statistical)} \\
E \quad \text{(fact-to-be-explained)}
\]

The double lines separating the premises from the conclusion signifies that the argument is inductively correct, and the number \( r \) expresses the degree of inductive probability with which the premises support the conclusion. It is presumed that \( r \) is fairly close to one.

The high-probability requirement, which seems such a natural analogue of the deductive entailment relation, leads to difficulties in two ways. First, there are arguments that fulfill all of the requirements imposed by the I-S model, but that patently do not constitute satisfactory scientific explanations. One can maintain, for example, that people who have colds will probably get over them within a fortnight if they take vitamin C, but the use of vitamin C may not explain the recovery, since almost all colds clear up within two weeks regardless. In arguing for the use of vitamin C in the prevention and treatment of colds, Linus Pauling (1970) does not base his claims upon the high probability of avoidance or quick recovery; instead, he urges that massive doses of vitamin C have a bearing upon the probability of the recovery. Conceiving the effect of these four conditions to the effect that the event-to-be-explained was to be expected by virtue of certain explanatory facts. In the case of I-S explanations, this means that the premise must lend high inductive probability to the conclusion—that is, the explanandum must be highly probable with respect to the explanans.

Explanations of the D-N and I-S varieties can therefore be schematized as follows (Hempel, 1965, pp. 336, 382):

\[
\begin{align*}
C_i, C_j, \ldots, C_f \quad \text{(particular explanatory conditions)} \\
L_1, L_2, \ldots, L_k \quad \text{(general laws)} \\
E \quad \text{(fact-to-be-explained)}
\end{align*}
\]
for some requirement of this latter sort. To the best of my knowledge, the advocates of the 'received view' have not, until very recently, put forth any such additional condition, nor have they come to terms with counterexamples of these types in any other way. James Fetzer’s requirement of strict maximal specificity, which rules out the use in explanations of laws that mention nominally irrelevant properties (Fetzer, 1981, pp. 125–126), seems to do the job. In fact, in (Salmon, 1979, pp. 691–692), I showed how Reichenbach’s theory of nomological statements could be used to accomplish the same end.

The second problem that arises out of the high-probability requirement is illustrated by an example furnished by Michael Scriven (1959, p. 480). If someone contracts paresis, the straightforward explanation is that he was infected with syphilis, which had progressed to paralysis. This is secondary, and latent stages without treatment with penicillin. Paresis is one form of tertiary syphilis, and it never occurs except in syphilitics. Yet far less than half of those victims of untreated syphilis ever develop paresis. Untreated latent syphilis is the explanation of paresis, but it does not provide any basis on which to say that the explanation of a man’s paralysis was to be expected by virtue of these explanations. For, for any victim of latent untreated syphilis, the odds are that he will not develop paresis. Many other examples can be found to illustrate the same point. As I understand it, mushroom poisoning may afflict only a small percentage of individuals who eat a particular type of mushroom (Smith, 1958, Introduction), but the eating of the mushroom would thus insignificantly be offered as the explanation in instances of the illness in question. The point is illustrated by remarks on certain species in a guide for mushroom hunters (Smith, 1955, pp. 34, 185).

*Helvella infusa*, “Poisonous to some, edible for most people. Not recommended.”

*Chlorophyllum molybdites*, “Poisonous to some but not to others. Those who are not made ill by it consider it a fine mushroom. The others suffer acutely.”

These examples show that high probability does not constitute a necessary condition for legitimate statistical explanations. Taking them together with the vitamin C example, we must conclude that, to a first approximation, at least—that a high probability of the explanation relative to the explanandum is neither necessary nor sufficient for correct statistical explanations, even if all of Hempel’s other conditions are fulfilled. Hempel’s model is needed, perhaps, to give some of the probabilities. A crucial feature of the explanation will be the comparison between the prior and posterior probabilities. In Hempel’s case of the streptococcus infection, for instance, it is obvious that the probability, in the entire class of people with streptococcal infections, of a quick recovery. We realize, however, that the administration of penicillin is statistically relevant to quick recovery, so we compare the probability of quick recovery among those who have received penicillin with the probability of quick recovery among those who have not received penicillin. Hempel warns, however, that there is another relevant factor, namely, the existence of the penicillin-resistant strain of bacteria. We must, therefore, take that factor into account as well. Our original model can be divided into four parts: (1) infection by non-penicillin-resistant bacteria, (2) infection by non-penicillin-resistant bacteria, (3) infection by penicillin-resistant bacteria, and (4) infection by penicillin-resistant bacteria, no penicillin given. Since the administration of penicillin is irrelevant to quick recovery in case of penicillin-resistant infections, the subclass of people infected by penicillin-resistant bacteria must be removed from (4) of the original reference class should be merged to yield (3’) infection by penicillin-resistant bacteria. If John Jones is a member of (1), we have an explanation of his quick recovery, according to the S-R approach, not because the probability is high, but, rather, because it differs significantly from the probability in the original reference class. We shall see later what must be done if John Jones happens to fall into class (3). By contrast, Hempel’s earlier high-probability requirement demands only that the posterior probability be sufficiently large—whatever that might mean—but makes no reference at all to any prior probability. According to Hempel’s abstract model, we ask, “Why is individual x a member of B?” The answer consists of an inducement argument having the following form:

\[
P(B|A) = r
\]

As we have seen, even if the first premise is a statistical law, r is high, the premises are true, and the requirement of maximal specificity has been fulfilled, our ‘explanation’ may be perhaps inadequate, due to failure of relevancy.

In (Salmon, 1970, pp. 220–221), I advocated what came to be called the statistical relevance or S-R model of scientific explanation. At that time, I thought that anything that satisfied the conditions that define that model would qualify as a legitimate scientific explanation. I no longer hold that view. It now seems to me that the relations specified in the S-R model constitute the statistical basis for a bona fide scientific explanation, but that this basis must be supplemented by certain *causal* factors in order to constitute a satisfactory scientific explanation. In chapters 5–9 I shall discuss the causal aspects of explanation. In this chapter, however, I shall confine attention to the statistical basis, as articulated in terms of the S-R model. Indeed, I have already on one or more occasions spoken of the S-R model, but, rather, of the S-R basis.

Adopting the S-R approach, we begin with an explanatory question in a form somewhat different from that given by Hempel. Instead of asking, for instance, “Why did x get well within a fortnight?” we ask, “Why did this person with a cold get well within a fortnight?” Instead of asking,
Statistical Explanation and Its Models

"Why is x a B?" we ask. "Why is x, which is an A, also a C?" The answer—at least for preliminary purposes—is that x is also a C, where C is relevant to B within A. Thus, we have a prior probability P(B|A)—in this case, the probability that a person with a cold (A) gets well within a fortnight (B). Then we let C stand for the taking of vitamin C. We are interested in the posterior probability P(B|A,C) that a person with a cold who takes vitamin C recovers within a fortnight. If the prior and posterior probabilities are equal to one another, the taking of vitamin C can play no role in explaining why this person recovered from the cold within the specified period of time. If the posterior probability is not equal to the prior probability, then C may, under certain circumstances, furnish part or all of the desired explanation. A large part of the purpose of the present book is to investigate the way in which considerations such as those statistically relevant to a given occurrence have or lack explanatory import.

We cannot, of course, expect that every request for a scientific explanation will be phrased in this monosyllabic form. Someone might ask, for example, "Why did Mary Jones get well in no more than a fortnight’s time?" It might be clear from the context that she was suffering from a cold, so that the question could be reformulated as, "Why did this person who was suffering from a cold get well within a fortnight?" In some cases, it might be necessary to seek additional clarification from the person requesting the explanation, but presumably it will be possible to discover what explanation is being called for. This point about the form of the explanation-seeking question has fundamental importance. We can easily imagine circumstances in which an altogether different explanation is sought by means of the same initial question. Perhaps Mary had exhibited symptoms strongly suggesting that she had meningitis; in this case, the fact that it was only an ordinary cold might constitute the explanation of her quick recovery. A given why-question, construed in one way, may elicit an explanation, while otherwise construed, it asks for an explanation that cannot be given. "Why did the Married contract paroxysm?" might mean, "Why did this adult human develop paroxysms?", or, "Why did this syphilitic develop paroxysms?" On the first construal, the question has a suitable answer, which we have already discussed. On the second construal, it has no answer—for we do not know of any fact in addition to syphilis that is relevant to the occurrence of paroxysms. Some philosophers have argued against these considerations, that scientific explanation has an unavoidably pragmatic aspect (e.g., vanFraassen, 1977, 1980). It means simply that there are cases in which people ask for explanations in unclear or ambiguous terms, so that we cannot tell what explanation is being requested without further clarification, then so be it. No one would deny that we cannot be expected to supply explanations unless we know what it is we are being asked to explain. To this extent, scientific explanation surely has pragmatic or contextual components. Dealing with these considerations is, I believe, tantamount to choosing a suitable reference class with respect to which the prior probabilities are to be taken and specifying an appropriate sample space for purposes of a particular explanation. The lesson of these two items in the next section is, in chapter 4—in an extended discussion of van Fraassen’s theory—we shall return to this issue of pragmatic aspects of explanation, and we shall consider the question of whether there are any others.

THE STATISTICAL-RELEVANCE APPROACH

Let us now turn to the task of giving a detailed elaboration of the S-R basis. For purposes of initial presentation, let us construct the terms A, B, C . . . (with or without subscripts) as referring to classes, and let us construe our probabilities in some sense as relative frequencies. This does not mean that the statistical-relevance approach is tied in any crucial way to a frequency theory of probability. I am simply adopting the heuristic device of picking examples involving frequencies because they are easily grasped. Those who prefer propensities, for example, can easily make the appropriate terminological adjustments, by speaking of changes in the set of trials where I refer to reference classes and attributes. With this understanding in mind, let us consider the steps involved in constructing an S-R basis for a scientific explanation:

1. We begin by selecting an appropriate reference class A with respect to which the prior probabilities P(B|A) of the Bs are to be taken.

2. We impose an explanandum-partition upon the initial reference class A in terms of an exclusive and exhaustive set of attributes B1, . . . , Bn; this defines a sample space for purposes of the explanation under consideration. (This partition was not required in earlier presentations of the S-R model.)

3. Invoking a set of statistically relevant factors C1, . . . , Cm, we partition the initial reference class A into a set of mutually exclusive and exhaustive cells A1C1, . . . , A1Cm. The properties C1, . . . , Cm furnish the explanandum-partition.

4. We ascertain the associated probability relations:

   prior probabilities

   \[ P(B|A) = p_i \]

   for all \( i \) (1 \( \leq i \leq m \))

   posterior probabilities

   \[ P(B_i|A,C_j) = p_{ij} \]

   for all \( i \) and \( j \) (1 \( \leq i \leq m \) and \( 1 \leq j \leq s \))

5. We require that each of the cells ACj be homogeneous with respect to the explanandum-partition {B1}; that is, none of the cells in the partition can be further subdivided in any manner relevant to the occurrence of any Bi.

   (This requirement is somewhat analogous to Hempel's requirement of maximal specificity, but as we shall see, it is a much stronger condition.)

6. We ascertain the relative sizes of the cells in our explanandum-partition in terms of the following marginal probabilities:

   \[ P(C_j|A) = q_j \]

   (These probabilities were not included in earlier versions of the S-R model; the reasons for requiring them now were discussed above.)

7. We require that the explanandum-partition be a maximal homogeneous partition, that is—with an important exception to be noted later—for \( i \neq k \) we require that \( p_{ij} \neq p_{kj} \). (This requirement assures us that our partition in terms of C1, . . . , Cm does not introduce any irrelevant subdivision in the initial reference class \( A \).

8. We determine which cell ACi contains the individual \( x \) whose possession of the attribute Bi was to be explained. The probability of \( B_i \) within the cell is given in the list under 4.

Consider in a rather rough and informal manner the way in which the foregoing pattern of explanation might be applied in a concrete situation; an example of this sort was offered by James Greeno (1971a, pp. 89–90). Suppose that Albert has committed a delinquent act—say, stealing a car, a major crime—and we ask for an explanation of that fact. We ascertain from the context that he is an American teen-ager, and so we ask, "Why did this American teen-ager commit a serious delinquent act?" The prior probabilities, which we take as our point of departure, so to speak, are the products of the various degrees of juvenile delinquency (\( B_i \) among American teen-agers (A)—that is, \( P(B_i|A) \)). We will need a suitable
explanandum-partition; Greeno suggests $B_1$ = no criminal convictions, $B_2$ = conviction for minor infractions only, $B_3$ = conviction for major offense. Our sociological theories tell us that such factors as sex, religious background, marital status of parents, type of residential community, socioeconomic status, and several others are relevant to delinquency. We therefore take the initial reference class of American teen-agers and divide it into males and females; Jewish, Protestant, Roman Catholic, no religion; parents married, parents divorced, parents separated; urban, suburban, rural place of residence; younger, middle, lower class; and so forth. Taking such considerations into account, we arrive at a large number of cells in our partition. We assign probabilities to each of these delinquency behavior to each of the cells in accordance with 4, and we ascertain the probability of a randomly selected American male belonging to each of the cells in accordance with 6. We find the cell to which Albert belongs—for example, male, from a Protestant background, parents divorced, living in a suburban area, belonging to the middle class. If we have taken into account all of the relevant factors, and if we have correctly ascertained the probabilities associated with the various cells of our partition, then we have an S-R basis for the explanation of what we have observed. We may, for example, predict that if we were not to use sex in partitioning our original reference class. By condition 5 we must employ every relevant factor; by condition 7 we must employ only relevant factors. In many concrete situations, including the present example, we know that we have not found all relevant considerations; however, as Noreata Koerge rightly emphasized (1978), this is only for which we may aim. Our philosophical analysis is the only one to capture the notion of a fully satisfactory explanation.

Nothing has been said, so far, concerning the rationale for conditions 2 and 6, which are here added to the S-R basis for the first time. We must see why these requirements are needed. Condition 2 is quite straightforward; it amounts to the requirement that the sample space for the problem at hand be specified. As we shall see when we discuss Greeno's information-theoretic approach in chapter 4, both the explanans-partition and the explanandum-partition are needed to measure the information transmitted in any explanatory scheme. This is a useful measure of the explanatory value of a theory. In Fraassen's treatment of why-questions in chapter 4, his contrast class, which is the same as our explanandum-partition, is needed in some cases to specify precisely what explanation is being sought. In dealing with the question “Why did Albert steal a car?”, we used Greeno's suggested explanandum-partition. H, however, we had used different partitions (contrast classes), other explanations might have been called forth. Suppose that the contrast class included: Albert steals a car, Bill steals a car, Charlie steals a car, and so forth. Then the answer might have involved no sociology whatever; the explanation might have been that, among the members of his gang, Albert is most adept at hot-wiring. Suppose, instead, that the contrast class had included: Albert steals a car, Bill steals a car, Albert steals a diamond ring, Albert steals a bottle of whiskey, and so forth. In that case, the answer might have been that he wanted to go joyriding.

The need for the marginal probabilities mentioned in 6 is in the following way. In many cases, such as the foregoing delinquency example, the terms $C_i$ that furnish the explanans-partition of the initial reference class are conjunctive. A given cell is determined by several factors: sex and religious background and marital status of parents and type of residential community and socioeconomic status and . . . These factors are called the probabilistic contributing causes and counteracting causes that tend, respectively, to produce or prevent delinquency. In attempting to understand one's own delinquency, it is of utmost importance to know how each factor is relevant whether positively or negatively, and how strongly—both in the population at large and in various subgroups of the population. Consider, for example, the matter of sex. It may be that within the entire class of American teen-agers (A) the probability of serious delinquency ($B_i$) is greater among males than it is among females. If so, we would want to know how much the probability among males exceeds the probability among females and by how much it exceeds the probability in the entire population. We also want to know whether either being male is always positively relevant to serious delinquency, or whether in combination with certain other factors, it may be negatively relevant or irrelevant. Given two groups of teen-agers—one consisting entirely of boys and the other entirely of girls, but alike with respect to all of the other factors—we want to know how the probabilities associated with delinquency in each of the two groups are related to one another. It might be that in each case of two cells in the explanandum-partition that differ from one another only on the basis of gender, the probability of serious delinquency in the male group is greater than the probability in the female group. It might turn out, however, that sometimes the two probabilities are equal, or that in some cases the probability is higher in the female group than it is in the corresponding male group. Relationships of all of these kinds are logically possible.

It is a rather obvious fact that each of two circumstances can individually be positively relevant to our own delinquency, but their conjunction can be negatively relevant or irrelevant. Each of two drugs can be positively relevant to good health, but taken together, the combination may be detrimental—for example, if a treatment is effective when taken in conjunction with various remedies for the common cold can greatly increase the chance of dangerously high blood pressure (Goodwin and Guze, 1979). A factor that is a contributing cause in some circumstances can be a countering cause in others. For example, problems of this sort have been discussed, sometimes under the heading of “Simpson's paradox,” by Nancy Cartwright (1983, essay 1) and Bas van Fraassen (1980, pp. 108, 148–151). In (Salmon, 1975c), I have spelled out what is required to ascertain all the statistical relevance relations that are required. We therefore need to build in a way to extract this information. This is the function of the marginal probabilities $P(B_i | A)$ of explanatory variables. As is well known, such conditional probabilities as $P(B_1 | A, D_1)$, $P(B_1 | A, E_1)$, and $P(B_1 | A, D_1, E_1)$ can be derived.11 When 2 and 6 are added to the earlier characterization of the S-R model (Salmon et al., 1977), then we have gone as far as possible in characterizing scientific explanations at the level of statistical relevance relations.

The several features of the new version of the S-R basis described above have been included, but their conjunction can be negatively relevant or irrelevant. Each of two drugs can be positively relevant to good health, but taken together, the combination may be detrimental—for example, if a treatment is effective when taken in conjunction with various remedies for the common cold can greatly increase the chance of dangerously high blood pressure (Goodwin and Guze, 1979). A factor that is a contributing cause in some circumstances can be a countering cause in others. For example, problems of this sort have been discussed, sometimes under the heading of “Simpson's paradox,” by Nancy Cartwright (1983, essay 1) and Bas van Fraassen (1980, pp. 108, 148–151). In (Salmon, 1975c), I have spelled out what is required to ascertain all the statistical relevance relations that are required. We therefore need to build in a way to extract this information. This is the function of the marginal probabilities $P(B_i | A)$ of explanatory variables. As is well known, such conditional probabilities as $P(B_1 | A, D_1)$, $P(B_1 | A, E_1)$, and $P(B_1 | A, D_1, E_1)$ can be derived.11 When 2 and 6 are added to the earlier characterization of the S-R model (Salmon et al., 1977), then we have gone as far as possible in characterizing scientific explanations at the level of statistical relevance relations.
A fundamental philosophical difference between our S-R basis and Hempel's I-S model lies in the interpretation of the concept of homogeneity that appears in condition 5. Our requirement of maximal specificity, which is designed to achieve a certain kind of homogeneity in the reference class, is epistemically relativized. This means, in effect, that we must not know of any way to make a relevant partition, but it certainly does not demand that no possibility of a relevant partition can exist unknown to us. As I view the S-R basis, in contrast, condition 5 demands that the cells of our explanation-partition be objectively homogeneous; for this model, homogeneity is not epistemically relativized. Since this issue of epistemic relativization versus objective homogeneity is discussed at length in chapter 3, it is sufficient for now merely to call attention to this complex problem.

Condition 7 has been the source of considerable criticism. One such objection rests on the fact that the initial reference class A to which the S-R basis is referred, may not be maximal. Replacing Kyburg's hexed salt example, mentioned above, it has been pointed out that the class of samples of table salt is not a maximal homogeneous class with respect to solubility, for there are many other chemical substances that have the same solubility—namely, unity—of dissolving when placed in water. Baking soda, potassium chloride, various sugars, and many other compounds have this property. Consequently, if we take the maximality condition seriously, it has been argued, we should not ask, "Why does this sample of table salt dissolve in water?" but rather, "Why does this sample of matter in the solid state dissolve when it comes in contact with water?" And indeed, one can argue, as Koertge has done persuasively (1975), that to follow such a policy often leads to significant scientific progress. Without underplaying his important point, I would nevertheless suggest, for purposes of elaborating the formal schema, that we take the initial reference class A as given by the explanation-seeing-why-question, and look for relevant partitions within it. A significantly different explanation, which often undeniably represents scientific progress, may result if a different why-question, embodying a broader initial reference class, is posed. If the original question is not presented in a form that unambiguously determines a reference class A, we can reasonably discuss the advantages of choosing a wider or a narrower class in the case at hand.

Another difficulty with condition 7 arises incidentally—so to speak—when different cells in the partition, A_C, and A_C, happen to have equal associated probabilities P and p for all cells B in the explanation-partition. Such a circumstance might arise if the cells are determined conjunctively by a number of relevant factors, and if the differences between the two cells cancel one another out. It might happen, for example, that the probabilities of the various days of delinquency, hospitalization, court intervention, and offense, minor offense, no offense—for an upper-class, urban, Jewish girl would be equal to those for a middle-class, rural, Protestant boy. In this case, we may proceed to relax condition 7, allowing A_C, and A_C, as separate cells, provided they differ with respect to at least two of the terms in the conjunction, so that we are faced with a fortuitous canceling of relevant factors. If, however, A_C, and A_C, differed with respect to only one condition, they would have to be merged into a single cell. Such would be the case if, for example, among upper-class, urban-dwelling, American teenagers whose religious background is atheist and whose parents are divorced, the probability of delinquent behavior was the same for both boys as for girls. Indeed, we have already encountered this situation in chapter 2, with Hempel's example of the streptococcus infection. If the infection is of the penicillin-resistant variety, the probability of recovery in a given period of time is the same whether penicillin is administered or not. In such cases, we want to say, there is no relevant difference between the two classes—not that relevant factors were canceling one another out. I bring this problem up for consideration at this point, but I shall not make a consequent modification in the formal characterization of the S-R basis, for I believe that problems of this sort are best handled in the light of causal relevance relations. Indeed, it seems advisable to postpone detailed consideration of the whole matter of regarding the cells A_C, and A_C, determined conjunctively until causation has been explicitly introduced into the discussion. As we shall see in chapter 9, (Humphreys, 1951, 1953) and (Rogers, 1981) provide useful suggestions for handling just this issue.

Perhaps the most serious objection to the S-R model of scientific explanation—as it was originally presented—is based upon the principle that mere statistical correlations explain nothing. A rapidly falling barometric reading is a sign of an imminent storm, and it is highly correlated with the onset of storms, but it certainly does not explain the occurrence of a storm. The S-R approach does, however, have a way of dealing with examples of this sort. A factor C, which is relevant to the occurrence of B in the presence of A, may be screened off the occurrence of some additional factor D; the screening-off relation is defined by equations (3) and (4), which follow. To illustrate, given a series of days (A) in some particular locale, the probability of a storm on day B is in general quite different from the probability of a storm if there has been a consecutive sharp drop in the barometric reading (D). Thus, C is in some respect independent of A within B. If, however, we take into consideration the further fact that there is an actual drop in atmospheric pressure (D) in the region, it is irrelevant whether that drop is registered on a barometer. In the presence of D and A, C becomes irrelevant to B; we say that D screens off C from B—"in symbols,

\[ P(B|A,C,D) = P(B|A,D) \]

(3)

However, C does not screen off D from B, that is,

\[ P(B|A,C,D) \neq P(B|A,C) \]

(4)
for barometers sometimes malfunction, and it is the atmospheric pressure, not the reading on the barometer per se, that is directly relevant to the occurrence of the storm. A fact is causally screened off if irrelevant, and according to the definition of the S-R basis (condition 7), it is not to be included in the explanation. The falling rain simply does not explain the storm.

Screening off is frequent enough and important enough to deserve further illustration. A study, reported in the news media a few years ago, revealed a positive correlation between drinking and heart disease, but further investigation showed that this correlation results from a correlation between coffee drinking and cigarette smoking. It turned out that cigarette smoking causes heart disease, thus rendering coffee drinking statistically (as well as causally and explanatorily) irrelevant to heart disease. Returning to our previous example for another illustration, one could reasonably suppose that there is some correlation between low socioeconomic status and parenthesiform syphilis, for there may be a higher degree of sexual promiscuity, a higher incidence of venereal disease, and a lesser likelihood of adequate medical attention if the disease occurs. But the contraction of syphilis screens off such factors as the degree of promiscuity, since syphilis is present if syphilis goes untreated. Therefore, screens off any tendency to fail to get adequate medical care. Thus when an individual has latent untreated syphilis, all other such circumstances as low socioeconomic status are screened off from the development of paresis.

As the foregoing examples show, there are situations in which one circumstance or occurrence is correlated with another because of an indirect causal relationship, so that the correlation is spurious. In such cases, it often happens that the more proximate causal factors screen off those that are more remote. Thus 'mere correlation' factors are left in explanatory contexts with correlations that are intuitively recognized to have explanatory force.

In Statistical Explanation and Statistical Relevance, where the S-R model of statistical explanation was first explicitly named and articulated, I held out some hope (but did not try to defend the thesis) that all of the causal factors that play a role in scientific explanation could be explicated in terms of statistical relevance relations with the screening-off relation playing a crucial role.

I shall explain in chapter 6, I no longer believe this is possible. A large part of the material in the present book is devoted to an attempt to analyze the nature of the causal relations that enter into scientific explanation, and the manner in which they function in explanatory contexts. After characterizing the S-R model, I wrote:

One might ask on what grounds we can claim to have characterized explanation. The answer is this. When an explanation (as herein explicated) has been provided, we know exactly how to regard any A with respect to the property B. We know which causal factors to hold on, which to let go, and at what odds. We know precisely what degree of expectation is rational. We know how to face uncertainty about an A's being a B in the most reasonable, practical, and effective way. We know every factor that is relevant to an A having the property B. We know exactly what weight should be attached to the prediction that this A will be a B. We know all of the regularities (universal and statistical) that are relevant to our original question. What more could one ask of an explanation? (Salmon et al., 1971, p. 78)

The answer, of course, is that we need to know something about the causal relationships as well.

In acknowledging this egregious shortcoming of the S-R model of scientific explanation, I am not abandoning it completely. The attempt, rather, is to supplement it in suitable ways. While recognizing its incompleteness, I still think it constitutes a sound basis upon which to erect a more adequate account. And at a fundamental level, I still think it provides important insights into the nature of scientific explanation.

In the introduction to Statistical Explanation and Statistical Relevance, I offered the following succinct comparison between Hempel's S-I model and the S-R model:

S-I model: an explanation is an argument that renders the explanandum highly probable.
S-R model: an explanation is an assembly of facts statistically relevant to the explanandum, regardless of the degree of probability that results.

It was Richard Jeffrey (1969) who first explicitly formulated the thesis that (at least some) counter-intuitive and apparently non-argumentative facts are not arguments; it is beautifully expressed in his brief paper, "Statistical Explanation vs. Statistical Inference," which was reprinted in Statistical Explanation and Statistical Relevance. In (Salmon, 1965, pp. 145-146), I had urged the following: that positive relevance rather than high probability is the desideratum in statistical explanation. In (Salmon, 1970), I expressed the view, which many philosophers found weird and counter-intuitive (e.g., L. J. Cohen, 1975), that statistical explanations may even embody negative relevance—that is, an explanation of an event may, in some cases, show that the event to be explained is less probable than we had initially realized. I still do not regard that thesis as absurd. In an illuminating discussion of the explanatory force of negative and positively relevant factors, Paul Humphreys (1981) has introduced some felicitous terminology for dealing with such cases, and he has pointed to an important constraint. Consider a simple example. Smith is stricken with a heart attack, and the doctor says, "Despite the fact that Smith exercised regularly and had given up smoking several years ago, he contracted heart disease because he was seriously overweight." The "because" clause mentions those factors that are positively relevant and the "despite" clause cites those that are negatively relevant. Humphreys refers to them as contributing causes and countervailing causes. If we wish to discuss causal explanation in later chapters, we will want to say that a complete explanation of an event must make mention of the causal factors that tend to prevent its occurrence as well as those that tend to bring it about. Thus it is not inappropriate for the S-R basis to include the factors that are negatively relevant to the explanandum-event. As Humphreys points out, however, we would hardly consider as appropriate a putative explanation that had only negatively relevant facts, since it is a plausible claim that the explanandum is true, regardless of the degree of probability that results.

Despite the fact that Jones never smoked, exercised regularly, was not overweight, and did not have elevated levels of serum cholesterol, he had a heart attack," would hardly be considered an acceptable explanation of his fatal illness.

Before concluding this chapter on models of statistical explanation, we should take a brief look at the probabilistic (D-N-P) model of scientific explanation offered by Peter Railton (1978). By employing well-established statistical laws, such as that covering the spontaneous radioactive decay of unstable model, it is possible to deduce the fact that a decay on a particular isotope has a certain probability of occurring within a given time interval. For an atom of Carbon 14 (which is used in radiocarbon dating in determining, for example), the probability of a decay in 5,730 years is 1/2. The explanation of the probability of the decay-event conforms to Hempel's S-I model. Such an explanation does not, however, explain the actual occurrence of a decay, for, given the probability of such an event—however high or low—the event in question may or may not occur. A probabilistic fact does not qualify as an argument to the effect that the event-to-be-explained was to be expected with deductive certainty, given the explanation. Rather, it is, of course, clearly aware of the fact. He goes on to point out, nevertheless, that if we simply attach an "addendum" to the deductive argument stating that the event-to-be-explained did, in fact, occur in the case, we can claim to have a probabilistic account—wholly not a deductive or inductive argument—of the occurrence of the event. In this

Statistical Explanation and Its Models 181

Statistical Explanation and Its Models
respects, Raiton is in rather close agreement with Jeffrey (1969) that some explanations are not arguments. He also agrees with Jeffrey in emphasizing the importance of exhibiting the physical mechanisms that lead up to the occurrence that is to be explained. Raiton’s theory of explanation that of Jeffrey—has some deep affinities to the S-R model. In including a reference to physical mechanisms as an essential part of the D-N-P model, however, Raiton goes beyond the view that statistical relevance relations, in and of themselves, have explanatory import. His theory of scientific explanation is more appropriately characterized as causal or mechanistic. It is closely related to the two-tiered causal-statistical account that I am attempting to elaborate as an improvement upon the S-R model. Although with Kyburg’s help, I have offered what seem to be damaging counterexamples to the D-N model—for instance, the one about the man who explains his own avoidance of pregnancy on the basis of his control over the act of castration by his wife’s birth control pills (Salmon et al., 1971, p. 34)—the major emphasis has been upon statistical explanation, and that continues to be the case in what follows. Aside from the fact that, in some cases, contemporary science obviously provides many statistical explanations and many types of phenomena, and that any philosophical theory of scientific explanation has only lately come forth, there is a further reason for focusing upon statistical explanation. As I maintained in chapter 1, we can identify three distinct approaches to scientific explanation that do not seem to differ from one another in any important way as long as we confine our attention to contexts in which all of the explanatory laws are universal generalizations. I shall argue in chapter 4, however, that these three general conceptualizations of scientific explanation can be seen to differ radically from another when we move on to situations in which statistical explanations are in principle the only ones we can achieve. Close consideration of statistical explanations with sufficient attention to their causal ingredients, provides important insight into the underlying philosophical questions relating to our scientific understanding of the world.

### NOTES

1. Ilkka Niiniluoto (1981, p. 444) suggests that “Peirce should be regarded as the true founder of the theory of inductive-probabilistic explanation” (italics added). To this statement, “The statistical syllogism may be conveniently termed the explanatory syllogism” (Peirce, 1932, 2:716). I am inclined to disagree, for one isolated and unmediated statement in a novel can hardly be considered even the beginnings of any genuine theory.  

2. As Carnap reports in the preface, the semin complex proceeding was presented and described by his wife. Martin Gardner edited—“it would probably be more accurate to say ‘wrote up’”—the proceedings, enclosed them to Carnap, who rewrote them extensively. There is little doubt that Carnap saw and approved the passages I have quoted.

3. Hume’s expressions: “A woman makes concep more difficult by offering no resistance and accepting Venus with a wriggling body. She diverts the frown from the straight course of the ploughshare and makes the seed fall wide of the plot. These tricks are employed by prostitutes for their own ends, so that they may be convinced for the present and be laid up by marriage” (1951, p. 170, italics added).  

4. See (Wessels, 1982), for an illuminating discussion of the history of the statistical interpretation of quantum mechanics.

5. These are spelled out in detail in (Salmon, 1965a). See (Salmon, 1967, pp. 190–191) for a discussion of the ‘almost-deduction’ conception of inductive inference.

6. We shall see in chapter 3 that the requirement of statistical specificity, and the calculation by Hempel in his (1965) and revised in his (1968), does not actually do the job. Nevertheless, this way is clearly the right one.

7. It should be mentioned in passing that Hempel (1965, pp. 380–381) offers still another model of scientific explanation that he characterizes as deductive-statistical (D-S). In an explanation of this type, a statistical regularity is explained by deducing it from other statistical laws. There is no real need, however, to treat such explanations as a distinct type, for they fall under the D-N schema, just given, provided we allow that at least some may be statistical. In the present context, we are concerned only with statistical explanations of nondeductive sorts.

Many examples are presented and analyzed in (Salmon et al., 1971, pp. 33–40). Nancy Cartwright (1983, pp. 26–27) errs when she attributes to Hempel the requirement that a statistical regularity be necessary for the probability of the explanation; this, which I first advanced in (Salmon, 1965), was never advocated by Hempel. Shortly thereafter (1983, 28–29), she provides a elaboration of the relationships among the views of Hempel, Suppes, and me.

9. Hempel’s most recent discussion of statistical explanation, he appears to maintain the astonishing view that although such examples have psychologically misleading features, they do qualify as logically satisfactory explanations (1977, pp. 107–111).

10. I am extremely sympathetic to the thesis, expounded in (Humphreys, 1985), that probable explanations—all those appearing in the S-R basis—are important tools in the construction of scientific explanations, but that they do not constitute any part of a scientific explanation per se. This thesis allows him to relax considerably the kinds of maximal specificity or homogeneity requirements that must be satisfied by statistical or probabilistic explanations. One factor that is statistically relevant may be causally irrelevant because, for example, it does not contribute any causal factors contributing to causal or casual explanation, but it is an attractive property of scientific explanation, for factors having small statistical relevance often seem to be so.

Humphreys’ approach does not show, however, that such relevance relations can be omitted from the S-R basis; on the contrary, the S-R basis must include such factors in order that we may ascertain whether the event is a causal explanation or not. I shall return to Humphreys’ concept of alaetory explanation in chapter 9.

11. Suppose, for example, that we wish to conclude P(B | A) > P(B | ¬A) from the evidence D = C1 V . . . V Cn, the cells Cj being mutually exclusive. This can be done as follows. We are given P(Cj | A) and P(Cj | ¬A).

By the multiplication theorem,

\[ P(D | B, A) = P(D | A) \cdot P(B | A) \cdot P(B | D, A) \]

Assuming P(D | A) ≠ 0, we have,

\[ P(B | A, D) = P(D | B, A) / P(D | A) \]

By the addition theorem

\[ P(D | A) = \sum_j P(C_j | A) \cdot P(D | C_j | A) \]

By the multiplication theorem

\[ P(D | B, A) = \sum_j P(C_j | A) \cdot P(B | A, C_j) \]

Substitution in (*) yields the desired relation:

\[ P(B | A, D) = \sum_j P(C_j | A) \cdot P(B | A, C_j) / P(D | A) \]

12. Cartwright (1985, p. 27) asserts that on Hempel’s account, "every scientific explanation is an objective, person-independent matter," and she applauds him for holding that view. I find it difficult to reconcile her characterization with Hempel’s repeated emphasis on observation (prior to 1977) of the doctrine of epistemic relativization of inductive-statistical explanations. Among other things, it implies that my way of dealing with problems concerning the proper formulation of the explanation-seeking why-questions—that is, problems concerning the choice of an appropriate inferential reference class—"makes explanation a subjective matter" (ibid., p. 29). "What explains what," she continues, "depends on the laws and facts true in our world, and cannot be adjusted by shifting our interest or our focus" (ibid.). This criticism seems to me to be mistaken. Clarification of the question is often required to determine what it is that can be explained, and this may have pragmatic dimensions. However, once the explanation has been unambiguously specified, on my view, the proper formulation of the appropriate explanans is fully objective. I am in complete agreement with Cartwright concerning the desirability of such objectification, however, my extensive concern with objective homogeneity is based directly upon the desire to eliminate from the theory of statistical explanation such subjective features as epistemic relativization.