JAMES H. FETZER

REICHENBACH, REFERENCE CLASSES, AND SINGLE CASE 'PROBABILITIES'*

Perhaps the most difficult problem confronted by Reichenbach's explication of physical probabilities as limiting frequencies is that of providing decision procedures for assigning singular occurrences to appropriate reference classes, i.e., *the problem of the single case.*[1] Presuming the symmetry of explanations and predictions is not taken for granted, this difficulty would appear to have two (possibly non-distinct) dimensions, namely: the problem of selecting appropriate reference classes for *predicting* singular occurrences, i.e., the problem of (single case) prediction, and the problem of selecting appropriate reference classes for *explaining* singular occurrences, i.e., the problem of (single case) explanation. If the symmetry thesis is theoretically sound, then these aspects of the problem of the single case are actually non-distinct, since any singular occurrence should be assigned to one and the same reference class for purposes of either kind; but if it is not the case that singular occurrences should be assigned to one and the same reference class for purposes of either kind, then these aspects are distinct and the symmetry thesis is not sound.[2]

The decision procedure that Reichenbach advanced to contend with the problem of the single case, i.e., the policy of assigning singular occurrences to '*the narrowest reference class for which reliable statistics can be compiled*', moreover, strongly suggests that one and the same reference class should serve for purposes of either kind. Reichenbach himself primarily focused attention on the problem of (single case) prediction, without exploring the ramifications of his resolution of the problem of the single case for the problem of (single case) explanation.[3] The theories of explanation subsequently proposed by Carl G. Hempel and by Wesley C. Salmon, however, may both be viewed as developments with considerable affinities to Reichenbach's position, which nevertheless afford distinct alternative solutions to the problem of (single case) explanation.[4] In spite of their differences, moreover, when Hempel's and Salmon's formulations are understood as incorporating the frequency criterion of statistical relevance, they appear to be saddled with theoretical difficulties whose

Synthese **34** (1977) 185–217.

resolution, in principle, requires the adoption of an alternative construction.

The purpose of this paper is to provide a systematic appraisal of Reichenbach's analysis of single case 'probabilities' *with particular concern for the frequency conceptions of statistical relevance and of statistical explanation*, especially as they may be related to the theories of explanation advanced by Hempel and Salmon. Among the conclusions supported by this investigation are the following:

(a) that the frequency criterion of relevance is theoretically inadequate in failing to distinguish between two distinct kinds of 'statistical relevance';

(b) that reliance upon this defective criterion of relevance suggests that, on frequency principles, there are no irreducibly statistical explanations; and,

(c) that these difficulties are only resolvable within the frequency framework by invoking epistemic contextual considerations.

As a result, taken together, these reflections strongly support the contention that the problem of (single case) prediction and the problem of (single case) explanation require distinct (if analogous) resolution, i.e., that the symmetry thesis is unsound; and indirectly confirm the view that the meaning of *single case* probabilities should be regarded as fundamental, where a clear distinction may be drawn between *causal* relevance and *inductive* relevance on the basis of a statistical disposition conception, i.e., that Reichenbach's limiting frequency construction should be displaced by Popper's propensity interpretation for the explication of probability as a physical magnitude.

1. REICHENBACH'S ANALYSIS OF SINGLE CASE 'PROBABILITIES'

The theoretical foundation for Reichenbach's analysis of the single case, of course, is provided by the definition of 'probability' itself: "In order to develop the frequency interpretation, we define probability as the *limit of a frequency* within an infinite sequence".[5] In other words, the probability r of the occurrence of a certain outcome attribute A within an infinite sequence of trials S is the limiting frequency with which A occurs in S, i.e.,

(I) $P(S, A) = r =_{\mathrm{df}}$ the limit of the frequency for outcome attribute A within the infinite trial sequence S equals r.

Reichenbach himself assumes no properties other than the limiting frequency of A within S as necessary conditions for the existence of probabilities and thereby obtains an interpretation of the broadest possible generality.[6] It is important to note, however, that Reichenbach envisions *finite* sequences as also possessing 'limits' in the following sense:

> Notice that a limit exists even when only a finite number of elements x_i belong to S; the value of the frequency for the last element is then regarded as the limit. This trivial case is included in the interpretation and does not create any difficulty.[7]

One justification for the inclusion of such 'limits', moreover, is that when the sequence S contains only a finite number of members, those members may be counted repetitiously an endless number of times to generate trivial limiting frequencies.[8]

As the basis for a theoretical reconciliation of these concepts, therefore, let us assume that a sequence S is *infinite* if and only if S contains at least one member and the description of its reference class does not impose any upper bound to the number of members of that class on syntactical or semantical grounds alone. Although 'limits' may be properties of finite sequences under this interpretation, they are not supposed to be properties of single individual trials *per se* and may only be predicated of single individual trials *as a manner of speaking*:

> I regard the statement about the probability of the single case, not as having a meaning of its own, but as representing an elliptical mode of speech. In order to acquire meaning, the statement must be translated into a statement about a frequency in a sequence of repeated occurrences. The statement concerning the probability of the single case thus is given a *fictitious meaning*, constructed by a *transfer of meaning from the general to the particular case.*[9]

Strictly speaking, therefore, probabilities are only properties of single trials as members of reference classes collectively; but Reichenbach nevertheless countenances referring to 'probabilities' as properties of such trials distributively "not for cognitive reasons, but because it serves the purpose of action to deal with such statements as meaningful".[10] Indeed, in order to distinguish the meaning of 'probability' with respect to the occurrence of singular trials and of infinite trial sequences, Reichenbach introduces a different term, i.e., 'weight', for application to the single case.

In spite of this difference in meaning, the numerical value of the weight to be assigned to attribute A as the outcome of a singular trial T, in principle, is determined by the limiting frequency with which A occurs within a trial sequence of kind K, where $T \in K$. The existence of a single case weight for the occurrence of attribute A thus requires (i) the existence of an infinite sequence of trials of kind K (ii) with a limiting frequency for A equal to r, where (iii) it is not the case that there exists some other infinite sequence of kind K such that the limiting frequency for A is not equal to r; hence, '$P(K, A) = r$' is true if and only if there exists an infinite sequence of trials of kind K such that the limiting frequency for A is equal to r and, for all trial sequences S, if S is an infinite sequence of kind K, then the limiting frequency for outcomes of kind A is equal to r.[11] In effect, therefore, every single individual trial T that happens to be a trial of kind K with respect to the occurrence of attribute A must be regarded as a member of a unique trial sequence K consisting of all and only those single trials that are trials of kind K with respect to the occurrence of outcome attribute A; otherwise, violation of the uniqueness condition would generate explicit contradictions of the form, '$P(K, A) \neq P(K, A)$'.[12]

Any single individual trial, however, may be exhaustively described if and only if *every* property of that trial is explicitly specifiable including, therefore, the spatial and the temporal relations of that instantiation of properties relative to every other. Let us assume that any single individual trial T is a property of some object or collection of objects x such that, for each such individual trial, 'Tx' is true if and only if '$F^1x \cdot F^2x \cdot \ldots$', is true, where '$F^1$', '$F^2$', ..., and so on are predicates designating distinct properties of that single individual trial. Then the single individual trial T is subject to exhaustive description, in principle, if and only if there exists some number m such that F^1 through F^m exhausts every property of that single individual trial; otherwise, the single individual trial T is not exhaustively describable, even in principle, since there is no end to the number of properties that would have to be specified in order to provide an exhaustive description of that trial T. The last time I turned the ignition to start my car, for example, was a single individual trial involving a 1970 Audi 100LS 4-door red sedan, with a Kentucky license plate, a recorded mileage of 88 358.4 miles, that had been purchased in 1973 for $2600.00 and driven to California during the Summer of the same year, and so on, which was parked in the right-most section of a three-car wooden garage to the rear of a two-story building at 159 Woodland Avenue in Lexington, on a somewhat misty morning at approximately 10:30 A.M. on 12 March 1976, one half-hour after I had drunk a cup of coffee, and so forth. As it happened, the car would not start.

The significance of examples such as this, I surmise, has two distinctive aspects. On the one hand, of course, it indicates the enormous difficulty, in principle, of providing an exhaustive description of any such individual trial; indeed, it strongly suggests *a density principle for single trial descriptions*, i.e., that for any description '$F^1x \cdot \ldots \cdot F^mx$' of any singular trial T that occurs during the course of the world's history, there exists some further description '$F^1x \cdot \ldots \cdot F^nx$' such that the set of predicates {'F^1', ..., 'F^m'} is a proper subset of the set of predicates {'F^1', ..., 'F^n'}.[13] On the other hand, it is at least equally important to notice that, among all of the properties that have thus far been specified, only some *but not all* would be viewed as contributing factors, i.e., as relevant variables, with respect to the outcome attribute of starting or not, as the case happened to be; in other words, even if this single individual trial is not exhaustively describable, it does not follow that the *causally relevant* properties of this trial arrangement T are themselves not exhaustively describable; for the properties whose presence or absence contributed to bringing about my car's failure to start would include, perhaps, accumulated moisture in the distributor, but would not include, presumably, my wearing a flannel overshirt at the time. It is at least logically possible, therefore, that among the infinity of properties that happened to attend this single individual trial, only a finite proper subset would exhaust all those exerting any influence upon that outcome attribute on that particular occasion.

The theoretical problem in general, therefore, may be described as follows, namely: for any single individual trial T, (i) to determine the kind of trial K that T happens to be with respect to the occurrence of outcome attribute A; and, (ii) to ascertain the limiting frequency r with which attribute A occurs within the infinite sequence of trials of kind K, if such a sequence and such a limit both happen to exist. The theoretical 'problem of the single case', therefore, is precisely that of determining the kind of trial K that any single individual trial T happens to be with respect to the occurrence of an outcome attribute of kind A, i.e., the problem of the selection of a unique reference sequence K for the assignment of a unique individual trial T. Reichenbach's solution to this problem is therefore enormously important to his analysis of the single case:

We then proceed by considering *the narrowest class for which reliable statistics can be compiled.* If we are confronted by two overlapping classes, we shall choose their common class. Thus, if a man is 21 years old and has tuberculosis, we shall regard the class of persons of 21 who have tuberculosis [with respect to his life expectancy]. Classes that are known to be irrelevant for the statistical result may be disregarded. A class C is irrelevant with respect to the reference class K and the attribute class A if the transition to the common class $K \cdot C$ does not change the probability, that is, if $P(K \cdot C, A) = P(K, A)$. For instance, the class of persons having the same initials is irrelevant for the life expectation of a person.[14]

A property C is *statistically relevant* to the occurrence of an attribute A with respect to a reference class K, let us assume, if and only if:

(II) $P(K \cdot C, A) \neq P(K, A)$;

that is, the limiting frequency for the outcome attribute A within the infinite sequence of $K \cdot C$ trials differs from the limiting frequency for that same attribute within the infinite sequence of K trials itself.[15]

Since a unique individual trial T is supposed to be assigned to the narrowest reference class K 'for which reliable statistics can be compiled', it seems clear that for Reichenbach, at least, a decision of this kind depends upon the state of knowledge $\mathscr{K}$ of an individual or collection of individuals z at a specific time t, i.e., the set of statements $\{\mathscr{K}\}$ accepted or believed by z at t, no matter whether true or not, as follows:

Whereas the probability of a single case is thus made dependent on our state of knowledge, this consequence does not hold for a probability referred to classes .... The probability of death for men 21 years old concerns a frequency that holds for events of nature and has nothing to do with our knowledge about them, nor is it changed by the fact that the death probability is higher in the narrower class of tuberculous men of the same age. The dependence of a single-case probability on our state of knowledge originates from the impossibility of giving this concept an independent interpretation; there exist only substitutes for it, given by class probabilities, and the choice of the substitute depends on our state of knowledge.[16]

Reichenbach allows that statistical knowledge concerning a reference class K may be fragmentary and incomplete, where the problem is one of "balancing the importance of the prediction against the reliability available". Nevertheless, as a general policy, Reichenbach proposes treating an individual trial T as a member of successively narrower and narrower reference classes $K^1, K^2, \ldots$ and so on, where each class is specified by taking into account successively more and more statistically relevant properties F^1, F^2, and so forth, where $K^1 \supset K^2$, $K^2 \supset K^3$, and so on and $P(K^1, A) \neq P(K^2, A)$, $P(K^2, A) \neq P(K^3, A)$, and so forth.

Reichenbach observes that, strictly speaking, the choice of a reference *class* is not identical with the choice of a reference *sequence*, since the members of that class may be ordered in different ways, which may sometimes differ in probability.[17] The intriguing question, however, is precisely how narrow the appropriate class in principle should be if our knowledge of the world were complete:

According to general experience, the probability will approach a limit when the single case is enclosed in narrower and narrower classes, to the effect that, from a certain step on, further narrowing will no longer result in noticeable improvement. It is not necessary for the justification of this method that the limit of the probability, respectively, is $=1$ or is $=0$, as the hypothesis of causality [i.e., the hypothesis of determinism] assumes. Neither is this necessary *a priori*; modern quantum mechanics asserts the contrary. It is obvious that for the limit 1 or 0 the probability still refers to a class, not to an individual event, and that the probability 1 cannot exclude the possibility that in the particular case considered the prediction is false. Even in the limit the substitute for the probability of a single case will thus be a class probability, ...[18]

The evident reply on Reichenbach's analysis, therefore, is that the appropriate reference class K relative to a knowledge context $\mathscr{K}zt$ containing every sentence that is true of the world and no sentence that is false would be some reference class K^i where $T \in K^i$ and, for every class K^j such that $T \in K^j$ and $K^i \supset K^j$, it is not the case that $P(K^i, A) \neq P(K^j, A)$; i.e., with respect to attribute A for trial T, K should be an *ontically homogeneous reference class* in the sense that (i) $T \in K$; (ii) $P(K, A) = r$; and, (iii) for all subclasses K^j of K to which T belongs, $P(K^j, A) = r = P(K, A)$.

Notice that the class K itself is not necessarily unique with respect to the set of properties $\{F^i\}$ specified by the reference class description of K; for any class K^j such that $K \supseteq K^j$ and $K^j \supseteq K$ (where K is an ontically homogeneous reference class relative to attribute A for trial T) will likewise qualify as an ontically homogeneous reference class relative to attribute A for trial T even though the set of properties $\{F^j\}$ specified by the reference class description of K^j differs from that for K, i.e., $\{F^i\} \neq \{F^j\}$. Consequently, although any classes K and K^j which happen to be such that $K \supseteq K^j$ and $K^j \supseteq K$ will of course possess all and only the same trial members and will therefore yield the same frequencies for various specific attributes with respect to those same trial members, their reference class descriptions, nevertheless, will not invariably coincide. From this point of view, therefore, the resolution of the reference class problem by assigning a single case to an ontically homogeneous reference class does not provide a unique *description* solution but rather a unique *value* solution. Nevertheless, on Reichenbachian principles, it may be argued further that there *is* a unique description solution as well as a unique value solution, namely: that any single individual trial T should be assigned to the *narrowest* class, i.e., *the class whose description includes specification of the largest set of properties of that individual trial*, "for which reliable statistics can be compiled".

The largest set of properties of an individual trial T that might be useful for this purpose, of course, is not logically equivalent to the set of all of the properties of that individual trial; for any individual trial T may be described by means of predicates that violate the provision that reference classes may not be logically limited to a finite number of members on syntactical or semantical grounds alone. Let us therefore assume that *a predicate expression is logically impermissible for the specification of a reference class description* if (a) that predicate expression is necessarily satisfied by every object or (b) that predicate expression is necessarily satisfied by no more than a specific number N of objects during the course of the world's history.[19] Predicates that happen to be satisfied by only one individual object during the course of the world's history, therefore, are permissible predicates for reference class descriptions so long as their extensions are not finite on logical grounds alone, as, e.g., might be any predicate expression essentially requiring proper names for its definition or such that the satisfaction of that expression by some proper name would yield a logical truth.[20] Let us further assume that no predicate expression logically entailing the attribute predicate or its negation is permissible. Then the largest set of permissible predicates describing a single trial T is not logically equivalent to the set of all of the properties of that individual trial; but nevertheless it will remain the case that, in general, such *narrowest* reference class descriptions are satisfied by only one individual event during the course of the world's history *as a matter of logical necessity.*[21]

These considerations suggest an insuperable objection to the applicability, in principle, of the frequency interpretation of probability; for if it is true that each individual trial T is describable, in principle, by some set of predicates such that (i) every member of that set is a permissible predicate for the purpose of a reference class description and (ii) that reference class description itself is satisfied by that individual trial alone, then the indispensable criterion of statistical relevance is systematically inapplicable for the role it is intended to fulfill. For under these circumstances, every individual trial T is the solitary member of a reference class K^* described by a reference class description consisting of the conjunction of a set of permissible predicates $F^1, F^2, \ldots, F^n$, i.e., $\{F^n\}$, where, moreover, lacking any information concerning the statistical relevance or irrelevance of any property of any singular trial – other than that the attribute A did occur (or did not occur) on that particular trial – it is systematically impossible to specify which of the properties of the trial T are statistically relevant and which are not; for the occurrence of that outcome, whether A or not-A, in principle, *must be attributed to the totality of properties present at that individual trial.*[22] Reliable statistics, after all, are only as reliable as the individual statistics upon which they are based; so *if the only statistical data that may be ascertained, in principle, concern the occurrence of outcome attributes on singular trials where each singular trial is the solitary member of a reference class, there is no basis for accumulating the 'reliable statistics' necessary for the applicability of the frequency criterion.* Of each individual trial $T^1, T^2, \ldots,$ it is possible in principle to specify a homogeneous reference class description $K^{*1}, K^{*2}, \ldots$; but since each trial is the solitary member of its particular reference class, it is impossible to employ the frequency criterion to ascertain which properties, if any, among all of those present on each such singular trial, are statistically irrelevant to its outcome.

The theoretical problem of the single case, let us recall, requires, for any single individual trial T, (i) to determine the kind of trial K that T happens to be with respect to the occurrence of outcome attribute A; and, (ii) to ascertain the limiting frequency r with which attribute A

occurs in the infinite sequence of trials of kind K, when such a sequence and such a limit both happen to exist. With respect to any single individual trial T that actually occurs in the course of the world's history, however, these conditions are, in effect, automatically satisfied; for (i) any such trial T may be described by some set of permissible predicates $\{F^n\}$ specifying a kind of trial K^* of which T is the solitary member, where, nevertheless, (ii) the reference class K^* is not logically limited to any finite number of members and Reichenbach's limit concept for infinite sequences of this kind is trivially satisfied by 0 or 1. Any other such singular trial T^i, after all, may likewise be described by some set of permissible predicates, i.e., $\{F^i\}$, where $\{F^i\} \neq \{F^n\}$, a condition of individuation that distinct events surely have to satisfy. With respect to each such singular trial T, of course, outcome attribute A either occurs or does not occur. Assume that T belongs to reference class K for which the probability of A is $\neq 0$ and $\neq 1$; i.e., (a) $T \in K$; and, (b) $P(K, A) = r$, where $0 \neq r \neq 1$. Then there necessarily exists a subclass of that class K^* such that $P(K^*, A) = 0$ or $= 1$, namely: the class specified by any set of permissible predicates $\{F^n\}$ of which trial T is the solitary member. Hence, since $P(K^*, A) \neq P(K, A)$ and $K \supset K^*$, the properties differentiating between K^* and K are statistically relevant with respect to the occurrence of attribute A, necessarily, by the frequency criterion; moreover, for any such reference class K such that (a) $T \in K$ and (b) $P(K, A) = r$, where $0 \neq r \neq 1$, it is theoretically impossible that K is a homogeneous reference class for trial T with respect to attribute A in Reichenbach's sense. Indeed, on the basis of the frequency criterion of statistical relevance, such an assumption is always invariably false.

The point of the preceding criticism, therefore, may be stated as follows: The only data available for ascertaining the reliable statistics necessary for arriving at determinations of the statistical relevance or irrelevance of any property F^i with respect to any outcome attribute A is that the frequency with which outcome A accompanies trials of kind $K \cdot F^i$ differs from that with which A accompanies trials of the kind K. Every single individual trial T^i occurring during the course of the world's history may be described as a member of some K^* reference class specified by at least one conjunction of permissible predicates K^{*i} that is satisfied by that individual trial. As a necessary logical truth, every such trial happens to belong to one and only one *narrowest* class of this kind;

however, since the K^* reference classes are not likewise limited to any finite number of members, the occurrence of the attribute A may still be assigned a probability, which, in this case, will be $=1$ or $=0$, depending upon whether that outcome occurred or failed to occur on that individual trial. The frequency criterion of statistical relevance, presumably, should permit distinguishing between the statistically relevant and irrelevant properties of each such trial. However, *since each trial T^i happens to be different in kind*, the only conclusion supported by that criterion of relevance is that every single property distinguishing those individual trials with respect to the occurrence of a specific attribute A is statistically relevant to the occurrence of that outcome; for those properties certainly 'made a difference', insofar as in one such case the attribute A occurred, while in another such case A did not occur. Thus, under circumstances of this kind, it is systematically impossible to obtain the reliable statistics necessary to support determinations of statistical relevance; for in the absence of that statistical evidence, such 'conclusions' merely take for granted what that evidence alone is capable of demonstrating.

2. SALMON'S 'RELEVANCE' ACCOUNT OF STATISTICAL EXPLANATION

It is interesting to observe that Reichenbach himself tended to focus upon the problems involved in the *prediction* of singular events rather than those involved in their *explanation*. The differences that distinguish explanations from predictions, however, may figure in significant ways within the present context; for if it happens to be the case that the purpose of a prediction is to establish grounds for *believing that* a certain statement (describing an event) is true, and the purpose of an explanation is to establish grounds for *explaining why* an event (described by a certain statement) occurs, i.e., if predictions are appropriately interpreted as *reason-seeking* why-questions, while explanations are appropriately interpreted as *explanation-seeking* why-questions, it might well turn out that at least part of the problem with the frequency criterion of statistical relevance is that it is based upon an insufficient differentiation between explanation and prediction contexts.[23] Consider the following:

(a) Reason-seeking why-questions are relative to a particular epistemic context, i.e., a knowledge context $\mathcal{K}zt$ as previously specified, where, with respect to an individual hypothesis H whose truth is not known, *a requirement of total evidence* may appropriately be employed according to which, for any statement S belonging to $\mathcal{K}zt$, *S is inductively relevant to the truth of H* if and only if the inductive (or epistemic) probability EP of H relative to $\mathcal{K}zt \cdot S$ differs from the inductive probability of H relative to $\mathcal{K}zt \cdot -S$, i.e.,

$$\text{(III)} \qquad EP(\mathcal{K}zt \cdot S, H) \neq EP(\mathcal{K}zt \cdot -S, H);$$

where, in effect, any statement whose truth or falsity within $\mathcal{K}zt$ makes a difference to the inductive probability of an hypothesis H must be taken into consideration in determining the inductive probability of that hypothesis.[24]

(b) Explanation-seeking why-questions are relative to a specific ontic context, i.e., the nomological regularities and particular facts of the physical world W, where, with respect to the explanandum-event described by an explanandum statement E whose truth is presumably known, *a requirement of causal relevance* should appropriately be employed according to which, for any property F belonging to W, *F is explanatorily relevant to the occurrence of E* (relative to reference class K) if and only if the physical (or ontic) probability PP of E relative to $K \cdot F$ is not the same as the physical probability of E relative to $K \cdot -F$, i.e.,

$$\text{(IV)} \qquad PP(K \cdot F, E) \neq PP(K \cdot -F, E);$$

where, in effect, any property whose presence or absence relative to reference class K makes a difference to the physical probability of an explanandum-event E must be taken into consideration in constructing an adequate explanation for that event.[25]

If predictions belong to the epistemic reason-seeking context, then it is entirely plausible to take into account any property whose presence or absence changes the inductive probability of an hypothesis H with respect to a knowledge context $\mathcal{K}zt$; indeed, a requirement of total evidence would require that any such property G of any individual object x belonging to the knowledge context $\mathcal{K}zt$ has to be taken into consideration in calculating the inductive probability of any H where the statement G^* attributing G to x is such that

$$\text{(V)} \qquad EP(\mathcal{K}zt \cdot G^*, H) \neq EP(\mathcal{K}zt \cdot -G^*, H);$$

where, as it might be expressed, the statistical relationship between property G and the attribute property A described by the hypothesis H may be merely one of statistical correlation rather than one of causal connection, i.e., it is not necessary that G be a property whose presence or absence contributes to bringing about the occurrence of an outcome of kind A relative to some reference class K.

It is not at all obvious, however, that the frequency criterion provides any theoretical latitude for differentiating between statistical relations of these quite distinctive kinds; on the contrary, it appears as though any factor at all with respect to which frequencies differ is on that account alone 'statistically relevant' to the probability of an hypothesis H or an explanandum E, respectively, without reflecting any theoretical difference between the explanation-seeking and the reason-seeking situations themselves. If every member of the class of twenty-one year old men K were also a member of the class of tuberculous persons T, for example, then the property of being tuberculous would be statistically irrelevant to the outcome attribute of death D. But although this property might reasonably be ignored for the purpose of prediction (as a matter of statistical correlation), it would not be reasonable to ignore this property for the purpose of explanation (when it makes a causal contribution to such an individual's death). If there is a significant difference between these kinds of situations, therefore, then those principles appropriate for establishing relevance relations within one of these contexts may be theoretically inappropriate for establishing relevance relations within the other. Perhaps the crucial test of the utility of Reichenbach's criterion of statistical relevance is found within the context of explanation rather than the context of prediction; for an examination of the statistical relevance account of explanation advanced by Salmon may provide an opportunity to evaluate the criticisms of that principle thus far advanced.

Salmon departs from Reichenbach's formulations, not in deviating from his criterion of statistical relevance, but rather in assigning a single case not to the *narrowest* relevant class but to the *broadest* relevant class instead:

If every property that determines a place selection is statistically irrelevant to A in K, I shall say that K is a homogeneous reference class for A. A reference class is homogeneous if there is no way, even in principle, to effect a statistically relevant partition without already knowing which elements have the attribute in question and which do not.... The aim in selecting a reference class to which to assign a single case is not to select the narrowest, but the widest, available class.... I would reformulate Reichenbach's method of selection of a reference class as follows: choose the broadest homogeneous reference class to which the single event belongs.[26]

Precisely because Salmon preserves the frequency criterion of relevance, his own formulations encounter difficulties analogous to those previously specified; but the introduction of the concept of a partition and of a 'screening-off' rule may be viewed as significant contributions to the frequency theory of explanation as follows:

(i) On Salmon's analysis, *a partition of a reference class K* is established by a division of that class into a set of mutually exclusive and jointly exhaustive subsets by means of a set of properties $F^1, F^2, \ldots, F^m$ and their complements where each ultimate subset of that class $K \cdot C^1, K \cdot C^2, \ldots, K \cdot C^n$ is homogeneous with respect to the outcome attribute A.[27] This procedure may be regarded as effecting a refinement in the application of Reichenbach's criterion, since it thus assumes that a property C is *statistically relevant* to the occurrence of attribute A with respect to a reference class K if and only if:

$$\text{(II*)} \qquad P(K \cdot C, A) \neq P(K \cdot -C, A);$$

that is, the limiting frequency for the outcome attribute A within the infinite sequence of $K \cdot C$ trials differs from the limiting frequency for that same outcome within the infinite sequence of $K \cdot -C$ trials.

(ii) Furthermore, a property *F screens off a property G* with respect to an outcome attribute A within the reference class K if and only if:

$$\text{(VI)} \qquad P(K \cdot F \cdot G, A) = P(K \cdot F \cdot -G, A) \neq P(K \cdot -F \cdot G, A);$$

where the equality between the limiting frequency for A within the classes $K \cdot F \cdot G$ and $K \cdot F \cdot -G$, on the one hand, establishes that the property G is not statistically relevant with respect to attribute A within the reference class $K \cdot F$; and the inequality between the limiting frequency for A within the class $K \cdot -F \cdot G$ and classes $K \cdot F \cdot G$ and

$K \cdot F \cdot -G$, on the other hand, establishes that the property F is statistically relevant with respect to the attribute A within the reference class $K \cdot G$.[28]

Salmon consolidates these ingredients as the foundation for his explication of explanation on the basis of the principle that screening-off properties should take precedence over properties which they screen-off within an explanation situation.[29] According to Salmon, an explanation of the fact that x, a member of the reference class K, is a member of the attribute class A as well, may be provided by fulfilling the following set of conditions, i.e.,

(1) $K \cdot C^1, K \cdot C^2, \ldots, K \cdot C^n$ is a homogeneous partition of K relative to A;

(2) $P(K \cdot C^1, A) = r^1, P(K \cdot C^2, A) = r^2, \ldots, P(K \cdot C^n, A) = r^n$;

(3) $r^i = r^j$ only if $i = j$; and,

(4) $x \in K \cdot C^m$ (where $m \in \{1, 2, \ldots, n\}$).[30]

Consequently, the appropriate reference class to specify in order to explain an outcome A for the trial T is the ontically homogeneous reference class $K \cdot C^m$ such that (a) $T \in K \cdot C^m$; (b) $P(K \cdot C^m, A) = r^m$; and, (c) for all homogeneous reference classes $K \cdot C^1, K \cdot C^2, \ldots, K \cdot C^n$ relative to outcome A, $r^i = r^j$ only if $i = j$, which, presumably, is intended to insure that there is one and only one reference class to which T may appropriately be assigned, namely: the *broadest* one of that kind. From this point of view, therefore, Salmon provides a unique description solution as well as a unique value solution to the single case problem.

Salmon's conditions of adequacy, let us note, are sufficient for the purpose of assigning individual trials to broadest homogeneous reference classes *only if* broadest homogeneous reference classes may be described by means of what may be referred to as *disjunctive properties*, i.e., predicate expressions of the form, '$K \cdot F^1 \vee F^2 \vee \ldots \vee F^n$', where a statement of the form, '$P(K \cdot F^1 \vee F^2 \vee \ldots \vee F^n, A) = r$', is true if and only if

'$P(K \cdot F^1, A) = P(K \cdot F^2, A) = \ldots = P(K \cdot F^n, A) = r$' is true, as Salmon himself has explicitly pointed out.[31] Otherwise, Salmon's conditions would be theoretically objectionable, insofar as it might actually be the case that there is an infinite set of reference classes, $\{K \cdot C^i\}$, whose members satisfy conditions (1) and (2) only if they do not satisfy condition (3), and conversely. The difficulty with this maneuver as a method of preserving Salmon's conditions (1)–(4) as sufficient conditions of explanatory adequacy, however, is that it entails the adoption of *a degenerating explanation paradigm*; for the occurrence of attribute A on trial T may be explained by referring that trial to a successively more and more *causally heterogeneous* reference class under the guise of the principle of reference class homogeneity. For if condition (3) is retained, then if, for example, the division of the class of twenty-one year old men K on the basis of the properties of having tuberculosis F^1 or heart disease F^2 or a brain tumor F^3 establishes a subclass such that, with respect to the attribute of death D, $P(K \cdot F^1 \vee F^2 \vee F^3, D) = r$ is an ultimate subset of the homogeneous partition of that original class, i.e., $\{K \cdot F^1 \vee F^2 \vee F^3\} = \{K \cdot C^m\}$, then the explanation for the death of an individual member i of class K resulting from a brain tumor, perhaps, is only explainable by identifying i as a member of class $K \cdot C^m$, i.e., as a member of the class of twenty-one year old men who either have tuberculosis or heart disease or a brain tumor.

The significance of this criticism appears to depend upon how seriously one takes what may be referred to as *the naive concept of scientific explanation for singular events*, namely: that the occurrence of an outcome A on a single trial T is to be explained by citing *all and only* those properties of that specific trial which contributed to bringing about that specific outcome, i.e., a property F is *explanatorily relevant* to attribute A on trial T if and only if F is a *causally relevant* property of trial T relative to attribute A.[32] From this perspective, Salmon's explication of explanation is theoretically objectionable for at least two distinct reasons:

First, *statistically relevant* properties are not necessarily *causally relevant* properties, and conversely.[33] If it happens that the limiting frequency r for the attribute of death D among twenty-one year old men K differs among those whose initials are the same F when that class is subject to a homogeneous partition, then that property is explanatorily relevant, necessarily, on the basis of the frequency criterion of statistical relevance;

and it might even happen that such a property screens-off another property G, such as having tuberculosis, in spite of the fact that property G is causally relevant to outcome D and property F is not. Indeed, Salmon himself suggests that "relations of statistical relevance must be explained on the basis of relations of causal relevance",[34] where relations of statistical relevance appear to fulfill the role of evidential indicators of relations of causal relevance.[35]

Second, the admission of disjunctive properties for the specification of an ultimate subset of a homogeneous partition of a reference class does not satisfy the desideratum of explaining the occurrence of an outcome A on a single trial T by citing *all and only* properties of that specific trial, whether causally relevant, statistically relevant, or otherwise. This difficulty, however, appears to be less serious, in principle, since a modification of Salmon's conditions serves for its resolution, namely: Let us assume that a reference class description 'K' is *stronger than* another reference class description '$K+$' if and only if 'K' logically entails '$K+$', but not conversely.[36] Then let us further assume as new condition (3*) Salmon's old condition (4), with the addition of a new condition (4*) in lieu of Salmon's old condition (3) as follows:

(4*) $K \cdot C^m$ is a strongest homogeneous reference class;

that is, the reference class description '$K \cdot C^m$' is stronger than any other reference class description '$K \cdot C^j$' (where $j \in \{1, 2, \ldots, n\}$) such that $x \in K \cdot C^j$.[37] The explanation of the death of an individual member i of class K resulting from a brain tumor, therefore, is thus only explainable by identifying i as a member of a *strongest* class $K \cdot C^m$, i.e., as a member of the class of twenty-one year old men who have brain tumors, which is a broadest homogeneous reference class of the explanatorily relevant kind, under this modification of the frequency conception.

It is important to observe, however, that none of these considerations mitigates the force of the preceding criticism directed toward the frequency criterion of statistical relevance itself. For it remains the case that any single trial T which belongs to a reference class K for which the probability of the occurrence of attribute A is $\neq 0$ and $\neq 1$ will likewise belong to innumerable subclasses $K \cdot C^j$ of K and, indeed, T itself will necessarily belong to some subclass K^* of K such that $P(K^*, A) = 0$ or $= 1$, which, on Salmon's own criteria, requires that trial T be assigned to

K^*, so long as K^* is a broadest (or a strongest) ultimate subclass of a homogeneous partition of the original class K. In effect, therefore, those properties that differentiate trial T as a member of class K from all other trial members of that class establish the basis for effecting a homogeneous partition of that class, a partition for which the probability for attribute A within every one of its ultimate subclasses (whether strongest or not) is $=0$ or $=1$, necessarily. And this result itself may be viewed as reflecting a failure to distinguish those principles suitable for employment within the context of explanation from those principles suitable for employment within the context of prediction, promoted by (what appears to be) the mistaken identification of statistical relevance with explanatory relevance.

It should not be overlooked that the properties *taken to be* statistically relevant to the occurrence of attribute A on trial T relative to the knowledge context $\mathcal{K}zt$ are not necessarily those that actually *are* statistically relevant to that attribute within the physical world. For this reason, Salmon's analysis emphasizes the significance of the concepts of *epistemic* and of *practical* homogeneity, which, however, on Salmon's explication, actually turn out to be two different kinds of *inhomogeneous* reference classes, where, for reasons of ignorance or of impracticality, respectively, it is not possible to establish ontically homogeneous partitions for appropriate attributes and trials.[38] Salmon's analysis thus does not provide an explicit characterization of the conception of reference classes that are believed to be homogeneous whether or not they actually are; nevertheless, there is no difficulty in supplementing his conceptions as follows: Let us assume that a reference class K is an *epistemically homogeneous reference class* with respect to attribute A for trial T within the knowledge context $\mathcal{K}zt$ if and only if the set of statements $\{\mathcal{K}\}$ accepted or believed by z at t, considerations of truth all aside, logically implies some set of sentences, $\{S\}$, asserting (a) that $T \in K$; (b) that $P(K, A) = r$; (c) that for all subclasses K^j of K to which T belongs, $P(K^j, A) = r = P(K, A)$; and (d) that it is not the case that there exists some class $K^i \supset K$ such that $P(K^i, A) = r = P(K, A)$. The satisfaction of conditions (a), (b), and (c), therefore, is sufficient to fulfill the epistemic version of the concept of an *ontically homogeneous* reference class K for trial T with respect to attribute A, while satisfaction of (d) as well is sufficient to fulfill the epistemic version of the *broadest homogeneous* reference class for T relative to A, within the knowledge context $\mathcal{K}zt$. In order to differentiate this concept from Salmon's original, however, let us refer to this definition as the *revised* conception of epistemic homogeneity, while acknowledging the theoretical utility of both.

3. HEMPEL'S REVISED REQUIREMENT OF MAXIMAL SPECIFICITY

The conclusions that emerge from the preceding investigation of Salmon's own analysis of statistical explanation support the contention that, on the frequency criterion of statistical relevance, statistical explanations are only *statistical* relative to a knowledge context $\mathcal{K}zt$, i.e., as the matter might be expressed, "God would be unable to construct a statistical relevance explanation of any physical event, not as a limitation of His power but as a reflection of His omniscience".[39] It is therefore rather intriguing that Salmon himself has strongly endorsed this conclusion as a criticism of *Hempel's account* of statistical explanation, for as the following considerations are intended to display, the fundamental difference between them is that Hempel's explicit relativization of statistical explanations to a knowledge context $\mathcal{K}zt$, as it were, affirms *a priori* what, on Salmon's view, is *a posteriori* true, namely: that for probabilities r such that $0 \neq r \neq 1$, the only homogeneous reference classes are epistemically homogeneous. What Salmon's criticism fails to make clear, however, is that this difficulty is a necessary consequence of adoption of the frequency criterion of statistical relevance for any non-epistemic explication of explanation, including, of course, Salmon's own ontic explication. Additionally, there are at least two further important issues with respect to which Hempel's analysis and Salmon's analysis are distinctive, in spite of their initial appearance of marked similarity.

In order to establish the soundness of these claims, therefore, let us consider the three principal ingredients of Hempel's epistemic explication. First, Hempel introduces the concept of an *i-predicate* in $\mathcal{K}zt$ which, in effect, is any predicate 'F^m' such that a sentence '$F^m i$' asserting the satisfaction of 'F^m' by the individual i belongs to $\mathcal{K}zt$, regardless of the kind of property that may be thereby designated.[40] He then proceeds to define *statistical relevance* as follows:

'F^m' will be said to be *statistically relevant* to 'Ai' in $\mathcal{K}zt$ if (1) 'F^m' is an i-predicate that entails neither 'A' nor '$-A$' and (2) $\mathcal{K}zt$ contains a lawlike sentence '$P(F^m, A) = r$' specifying the probability of 'A' in the reference class characterized by 'F^m'.[41]

Insofar as sentences of the form, '$P(F^m, A) = r$', are supposed to be *lawlike*, it is clear that, on Hempel's analysis, (a) their reference class descriptions must be specified by means of permissible predicates and (b) these sentences are to be understood as supporting subjunctive (and counterfactual) conditionals.[42] Salmon likewise assumes that the limiting frequency statements that may serve as a basis for statistical explanations are lawlike, although the theoretical justification for attributing counterfactual (and subjunctive) force to these sentences should not be taken for granted, since it may entail the loss of their extensionality.[43]

The condition that the knowledge context $\mathcal{K}zt$ contain a set of probability statements, of course, might be subject to criticism on the grounds that it requires $\mathcal{K}zt$ to contain an enormous number of lawlike sentences.[44] This objection lacks forcefulness, however, when the following are considered, namely:

(i) the sentences belonging to $\mathcal{K}zt$ are accepted or believed by z at t, i.e., they represent what z takes to be the case;

(ii) this set of sentences, therefore, may be believed to be exhaustive with respect to attribute A and trial T, whether that is actually true or not; and,

(iii) presumably, this explication is intended to specify the conditions that must be fulfilled in order to provide an adequate explanation relative to a specified knowledge context without presuming that these conditions are always (or even generally) satisfied.

Second, Hempel defines the concept of *a maximally specific predicate 'M' related to 'Ai' in* $\mathcal{K}zt$ where 'M' is such a predicate if and only if (a) 'M' is logically equivalent to a conjunction of predicates that are statistically relevant to 'Ai' in $\mathcal{K}zt$; (b) 'M' entails neither 'A' nor '$-A$'; and, (c) no predicate expression stronger than 'M' satisfies (a) and (b), i.e., if 'M' is conjoined with a predicate that is statistically relevant to 'Ai' in $\mathcal{K}zt$, then the resulting expression entails 'A' or '$-A$' or else it is logically equivalent to 'M'.[45] Thus, for outcome A on trial T, 'M' is intended to provide a description of that trial as a member of a reference class K determined by the conjunction of every statistically relevant predicate 'F^m' of trial T with respect to outcome A, i.e., 'M' is the conjunction of all and only the statistically relevant *i*-predicates (or the negations of such *i*-predicates) satisfied by trial T in $\mathscr{K}zt$.

Third, Hempel formulates the *revised requirement of maximal specificity* as follows: An argument

$$\text{(VII)} \qquad \begin{array}{l} P(F, A) = r \\ Fi \\ \hline\hline Ai \end{array} [r],$$

where r is close to 1 and all constituent statements are contained in $\mathscr{K}zt$, qualifies as an explanation of the fact that i, a member of the reference class F, is also a member of the attribute class A, within the knowledge context $\mathscr{K}zt$, only if:

(RMS^*) For any predicate, say 'M', which either (a) is a maximally specific predicate related to 'Ai' in $\mathscr{K}zt$ or (b) is stronger than 'F' and statistically relevant to 'Ai' in $\mathscr{K}zt$, the class $\mathscr{K}$ contains a corresponding probability statement, '$P(M, A) = r$', where, as in (VII), $r = P(F, A)$.[46]

Since a predicate expression 'F^j' is stronger than a predicate expression 'F^i' if and only if 'F^j' entails but is not entailed by 'F^i', moreover, a predicate which is logically equivalent to the conjunction of a maximally specific predicate 'M' and *any other i-predicate* 'F^k' such that $P(M \cdot F^k, A) = r$ will entail that maximally specific predicate 'M' itself and will presumably satisfy condition (b) of (RMS^*).

The motivation of Hempel's introduction of a requirement of this kind, let us recall, was the discovery that *statistical explanations* suffer from a species of explanatory ambiguity arising from the possibility that an individual trial T might belong to innumerable different reference classes $K^1, K^2, \ldots$ for which, with respect to a specific attribute A, the probabilities for the occurrence of that outcome may vary widely, i.e., the reference class problem for single case explanations. In particular, Hempel has displayed concern with the possibility of the existence of alternative explanations consisting of alternative explanans which confer *high probabilities* upon both the occurrence of an attribute A and its non-occurrence $-A$, relative to the physical world itself or a knowledge context $\mathscr{K}zt$, a phenomenon which cannot arise in the case of explanations involving universal rather than statistical lawlike statements.[47] However, it is important to note that there are *two* distinct varieties of explanatory ambiguity, at least one of which is not resolved by Hempel's maximal specificity requirement. For although Hempel provides a unique *value* solution to the single case problem (which entails a resolution to the difficulty of conflicting explanations which confer high probabilities upon their explanandum sentences), Hempel's approach does not provide a unique *description* solution to the reference class problem, a difficulty which continues to afflict his conditions of adequacy for explanations invoking universal *or* statistical lawlike statements within an epistemic *or* an ontic context. Hempel's explication, therefore, appears to afford only a restricted resolution of one species of statistical ambiguity, which, however, does not provide a solution to the general problem of explanatory ambiguity for explanations of either kind. For explanations involving universal laws as well as statistical laws continue to suffer from the difficulties that arise from a failure to contend with the problem of providing a unique *description* solution for single case explanations.

Hempel resolves the problem of statistical ambiguity, in effect, by requiring that, within a knowledge context $\mathcal{K}zt$ as previously specified, the occurrence of outcome attribute A on trial T is adequately explained by identifying a reference class K such that (a) $T \in K$; (b) $P(K, A) = r$; and, (c) for all subclasses K^j of K to which T belongs, $P(K^j, A) = r = P(K, A)$; but he does not require as well that (d) it is not the case that there exists some class $K^i \supset K$ such that $P(K^i, A) = r = P(K, A)$. For on the basis of condition (b) of (*RMS**), if, for example, the probability for the outcome death D within the reference class of twenty-one year old men who have tuberculosis K is equal to r and the probabilities for that same outcome within the reference classes $K \cdot F^1, K \cdot F^1 \cdot F^2, \ldots$ of twenty-one year old men who have tuberculosis K and who have high blood pressure F^1, or who have high blood pressure F^1 and have blue eyes $F^2, \ldots$ are likewise equal to r, then the explanation for the occurrence of death for an individual who belongs not only to class K but to class $K \cdot F^1$ and to class $K \cdot F^1 \cdot F^2$ and so on is adequate, *regardless of which of these reference classes is specified by the explanans in that single case.*[48] Consequently, on Hempel's explication, there is not only no unique explanation for the occurrence of such an explanandum outcome but, if the density principle for single trial descriptions is sound, *there may be an infinite* *number of adequate explanations for any one such explanandum,* on Hempel's explication. And this surprising result applies alike for explanations invoking laws of universal form, since if salt K dissolves in water A as a matter of physical law (within a knowledge context or without), then table salt $K \cdot F^1$ dissolves in water, hexed table salt $K \cdot F^1 \cdot F^2$ dissolves in water, . . . as a matter of physical law as well; so if a single trial involves a sample of hexed table salt an adequate explanation for its dissolution in water may refer to any one of these reference classes or to any others of which it may happen to belong, providing only that, for all such properties F^i, it remains the case that all members of $K \cdot F^i$ possesses the attribute A as a matter of physical law.

Comparison with Salmon's explication suggests at least two respects in which Hempel's explication provides theoretically objectionable conditions of adequacy. The first, let us note, is that Hempel's requirement of maximal specificity (*RMS**) does not incorporate any appropriate relevance criteria that would differentiate statistically relevant from statistically irrelevant properties in the sense of principle (II*). For Hempel has defined the concept of statistical relevance so generally that a property F^i such that there exists some probability r with respect to the attribute A where $P(F^i, A) = r$ necessarily qualifies as 'statistically relevant' independently of any consideration for whether or not there may exist some class K such that $K \supset F^i$ and $P(K, A) = r = P(F^i, A)$; in other words, *Hempel's concept of statistical relevance is not a relevance requirement of the appropriate kind.*[49] In order to contend with this difficulty, therefore, major revision of Hempel's definition is required along the following lines:

> 'F^m' will be said to be *statistically relevant* to 'Ai' relative to 'K' in $\mathcal{K}zt$ if and only if (1) 'F^m' and 'K' are *i*-predicates that entail neither 'A' nor '$-A$'; (2) $\mathcal{K}zt$ contains the lawlike sentences, '$P(K \cdot F^m, A) = r^i$' and '$P(K \cdot -F^m, A) = r^j$'; and, (3) the sentence, '$r^i \neq r^j$', also belongs to $\mathcal{K}zt$.

The second is that Hempel's explication of explanation is incapable of yielding a unique description solution to the single case problem because it is logically equivalent to the *revised concept of epistemic homogeneity*, i.e., the epistemic version of the concept of an *ontically homogeneous* reference class, rather than the epistemic version of the concept of a *broadest homogeneous* reference class. There appears to be no reason, in principle, that precludes the reformulation of Hempel's requirement so as to incorporate the conditions necessary for fulfilling the desideratum of providing a unique description solution (for arguments having the form (VII) previously specified) as follows:

(*RMS***) For any predicate, say '*M*', such that '*M*' is a maximally specific predicate related to '*Ai*' in $\mathcal{K}zt$, if '*M*' is stronger than '*F*', then for any predicate '*J*' implied by '*M*' that is not implied by '*F*', '*F*' implies a predicate '*H*' such that '*H*' screens-off '*J*' from '*Ai*' in $\mathcal{K}zt$, where $\mathcal{K}zt$ contains a corresponding probability statement, '$P(M, A) = r$', and, as in (VII), $r = P(F, A)$.

Condition (*RMS***), therefore, not only requires that any property of trial *T* that is statistically relevant to '*A*' but nevertheless not explanatorily relevant must be excluded from an adequate explanation of that outcome on that trial in $\mathcal{K}zt$, but also requires that trial *T* be assigned to the *broadest* homogeneous reference class of which it is a member, in the sense that '*F*' is the *weakest* maximally specific predicate related to '*Ai*' in $\mathcal{K}zt$. Moreover, on the reasonable presumption that the definition of a maximally specific predicate precludes the specification of reference classes by *non-trivial* disjunctive properties, i.e., disjunctive properties that are not logically equivalent to some non-disjunctive property, these conditions actually require that trial *T* be assigned to the *strongest* homogeneous reference class of which it is a member. From this point of view, therefore, the reformulation of Hempel's requirement appears to provide a (strongest) unique description solution as well as a unique value solution to the single case problem within the spirit of Hempel's explication.

The revised formulation of Hempel's requirements (incorporating appropriate relevance conditions and (*RMS***) as well) and the revised formulation of Salmon's requirements (incorporating his original conditions (1), (2), and (3*), together with new condition (4*) as well), of course, both accommodate the naive concept of scientific explanation to the extent to which they satisfy the desideratum of explaining the occurrence of an outcome *A* on a single trial *T* by citing *all and only relevant properties of that specific trial.* Neither explication, however, fulfills the expectation that a property F is *explanatorily* relevant to outcome A on trial T if and only if F is *causally* relevant to outcome A on trial T, so long as they remain wedded to the frequency criterion of statistical relevance. Nevertheless, precisely because Hempel's explication is *epistemic*, i.e., related to a knowledge context $\mathcal{K}zt$ that may contain sentences satisfying the conditions specified, it is not subject to the criticism that the only adequate explanations are *non*-statistical, i.e., those for which probability $r = 0$ or $r = 1$. On the other hand, it *is* subject to the criticism that there are no *non*-epistemic adequate explanations for which probability $r \neq 0$ and $r \neq 1$, i.e., there are no ontic (or true) *statistical* explanations. As it happens, this specific criticism has been advanced by Alberto Coffa, who, while arguing that Hempel's epistemic explication is necessitated by (1) implicit reliance upon the frequency interpretation of physical probability, in conjunction with (2) implied acceptance of a certain 'not unlikely' reference class density principle, unfortunately neglects to emphasize that Salmon's *ontic* explication is similarly afflicted, precisely because the fatal flaw is not to be found in the *epistemic*-ness of Hempel's explication but rather in *reliance on the frequency criterion of statistical relevance for any ontic explication.*[50]

If Coffa's argument happens to be sound with respect to Hempel's rationale, then it is important to observe that, provided (1) is satisfied and (2) is true, there are, *in principle*, no non-epistemic adequate explanations for which probability $r \neq 0$ and $r \neq 1$; that is, an epistemic explication is not only *the only theoretically adequate kind of an explication*, but an explication remarkably similar to Hempel's specific explication appears to be *the only theoretically adequate explication.* So if it is not the case that an epistemic explication of the Hempel kind is the only theoretically adequate construction, then either (1) is avoidable or (2) is not true. Intriguingly, the density principle Coffa endorses, i.e., the assumption that, for any specific reference class K and outcome A such that trial $T \in K$, there exists a subclass K^i of K such that (i) $T \in K^i$, and (ii) $P(K^i, A) \neq P(K, A)$, is demonstrably false, since for all subclasses K^i such that $K \supseteq K^i$ and $K^i \supseteq K$, this principle does not hold, i.e., it does not apply to *any* homogeneous reference class K, whether or not T is the only member of K.[51] Insofar as every distinct trial is describable, in principle, by a set of permissible predicates, $\{F^n\}$, such that that trial is the solitary member of the kind K^* thereby defined, however, evidently Coffa's density principle is not satisfied by even *one single trial* during the course of the world's history. Coffa's principle is plausible, therefore, only so long as there appear to be *no* homogeneous reference classes; once the existence of reference classes of kind K^* is theoretically identified, it is obvious that this density principle is false. Nevertheless, another density principle in lieu of Coffa's principle does generate the same conclusion, namely: the density principle for single trial descriptions previously introduced. Coffa has therefore advanced an unsound argument for a true conditional conclusion, where, as it happens, by retaining one of his premisses and replacing the other, that conclusion does indeed follow, albeit on different grounds.

In his endorsement of Coffa's contentions, Salmon himself explicitly agrees that, on Hempel's explication, there are no *non-epistemic* adequate explanations for which probability $r \neq 0$ and $\neq 1$; thus he observes, "There are no homogeneous reference classes except in those cases in which *either* every member of the reference class has the attribute in question *or else* no member of the reference class has the attribute in question".[52] With respect to his own explication, by contrast, Salmon remarks, "The interesting question, however, is whether under any other circumstances K can be homogeneous with respect of A – e.g., if one-half of all K are A".[53] It is Salmon's view, in other words, that Hempel's position entails an *a priori* commitment to determinism, while his does not. But the considerations adduced above demonstrate that determinism is a consequence attending the adoption of the frequency criterion of statistical relevance alone, i.e., determinism is as much a logical implication of Salmon's own position as it is of Hempel's. This result, moreover, underlines the necessity to draw a clear distinction between *statistical* relevance and *causal* relevance; for, although it is surely true that two distinct events are describable, in principle, by different sets of permissible predicates, it does not follow that they are necessarily not both members of a *causally homogeneous* reference class K for which, relative to attribute A, $P(K, A) = r$ where $0 \neq r \neq 1$, even though, as we have ascertained, they may *not* both be members of some *statistically homogeneous* reference class K^* for which, with respect to attribute A, $P(K^*, A) = r$, where $0 \neq r \neq 1$. Although it is not logically necessary *a priori* that the world is deterministic, therefore, adoption of the frequency criterion of statistical relevance is nevertheless sufficient to demonstrate determinism's truth.

When the reformulated versions of Salmon's and Hempel's explications which entail assigning each singular trial T to the 'strongest' homogeneous reference class K of which it is a member (with respect to attribute A) are compared, the revised Hempel explication provides, in effect, a meta-language formulation of the revised Salmon object-language explication, with the notable exception that the Hempel-style explication envisions explanations as arguments for which high probability requirements are appropriate, and the Salmon-style explication does not. The issue of whether or not explanations should be construed as arguments is somewhat elusive insofar as there is no problem, in principle, in separating *explanation-seeking* and *reason-seeking varieties of inductive arguments*, i.e., as sets of statements divided into premises and conclusions, where the premises provide the appropriate kind of grounds or evidence for their conclusions.[54] But however suitable a high probability requirement may be relative to the reason-seeking variety of inductive argument, there appear to be no suitable grounds for preserving such a requirement relative to the explanation-seeking variety of inductive argument in the face of the following consideration, namely: that *the imposition of a high probability requirement between the explanans and the explanandum of an adequate explanation renders the adequate explanation of attributes that occur only with low probability logically impossible, in principle.*[55] Indeed, it appears altogether reasonable to contend that no single consideration militates more strongly on behalf of conclusive differentiation between 'inductive arguments' of these two distinct varieties than this specific consideration.

The theoretical resolution of these significant problems, therefore, appears to lie in the direction of a more careful differentiation between principles suitable for employment within the explanation context specifically and those suitable for employment within the induction context generally. The problem of single case explanation, for example, receives an elegant resolution through the adoption of the *propensity* interpretation of physical probability; for on that explication,

(VIII) $P^*(E, A) = r =_{df}$ the strength of the dispositional tendency for any experimental arrangement of kind E to bring about an outcome of kind A on any single trial equals r;

where, on this statistical disposition construction, a clear distinction should be drawn between *probabilities* and *frequencies*, insofar as frequencies display but do not define propensities.[56] The *propensity criterion of causal relevance*, moreover, provides a basis for differentiating between causal and inductive relevance relations; for, on the propensity analysis, a property F is *causally relevant* to the occurrence of outcome A with respect to arrangements of kind E if and only if:

$$\text{(IX)} \qquad P^*(E \cdot F, A) \neq P^*(E \cdot -F, A);$$

that is, the strength of the dispositional tendency for an arrangement of kind $E \cdot F$ to bring about an outcome of kind A differs from the strength of the dispositional tendency for that same outcome with an arrangement of kind $E \cdot -F$ on any single trial. By virtue of a probabilistic, rather than deductive, connection between probabilities and frequencies on this interpretation, it is not logically necessary that probabilities vary if and only if the corresponding frequencies vary; but that

$$\text{(X)} \qquad P(E \cdot F, A) \neq P(E \cdot -F, A);$$

i.e., that long (and short) run frequencies for the attribute A vary over sets of trials with experimental arrangements of kind $E \cdot F$ and $E \cdot -F$, nevertheless, characteristically will provide information which, although neither necessary nor sufficient for the truth of the corresponding probability statement, may certainly qualify as inductively relevant to the truth of these propensity hypotheses.[57]

On the propensity conception, it is not the case that every distinct trial T must be classified as a member of a *causally homogeneous* reference class $\{K^*\}$ of which it happens to be the solitary member merely because T happens to be describable by a set of permissible predicates $\{F^n\}$ such that T is the only member of the corresponding *statistically homogeneous* reference class. Consequently, a single individual trial T may possess a statistical disposition of strength r to bring about the outcome A (a) whether that trial is the only one of its kind and (b) whether that outcome actually occurs on that trial or not. On this analysis, the question of determinism requires an *a posteriori* resolution, since it is not the case that, on the propensity criterion of causal relevance, any two distinct trials are therefore necessarily trials of two different causally relevant kinds. And, on the propensity criterion of causal relevance, it is not the case that the only adequate explanations for which the probability $r \neq 0$ and $\neq 1$ are invariably epistemic; for a set of statements satisfying the revised Salmon conditions (1)–(4*) or the corresponding revised Hempel conditions (in their ontic formulation) will explain the fact that x, a member of K, is also a member of A, provided, of course, those sentences are true.[58] For both explications, thus understood, may be envisioned as fulfilling the theoretical expectations of the naive concept of scientific explanation, where the only issue that remains is whether or not, and, if so, in what sense, statistical explanations should be supposed to be inductive arguments of a certain special kind.

Perhaps most important of all, therefore, is that Reichenbach's frequency interpretation of physical probability, which was intended to resolve the problem of providing an objective conception of physical probability, has indeed contributed toward that philosophical desideratum, not through any demonstration of its own adequacy for that role, but rather through a clarification of the conditions that an adequate explication must satisfy. For the arguments presented above suggest:

(a) that *the concept of explanatory relevance* should be explicated relative to a requirement of causal relevance;

(b) that *the requirement of causal relevance* should be explicated relative to a concept of physical probability; and,

(c) that *the concept of physical probability* should be explicated by means of the single case propensity construction;[59] and, moreover,

(d) that *the concept of inductive relevance* should be explicated relative to a requirement of total evidence;

(e) that *the requirement of total evidence* should be explicated relative to a concept of epistemic probability; and,

(f) that *the concept of epistemic probability* should be explicated, at least in part, by means of the long run frequency construction.[60]

If these considerations are sound, therefore, then it is altogether reasonable to suppose that the recognition of the inadequacy of the frequency conception as an explication of *physical* probability may ultimately contribute toward the development of an adequate explication of *epistemic* probability, where the frequency criterion of statistical relevance is likely to fulfill its most important theoretical role.

University of Kentucky

NOTES

* The author is grateful to Wesley C. Salmon for his valuable criticism of an earlier version of this paper.

[1] Hans Reichenbach, *The Theory of Probability*, Berkeley, University of California Press, 1949, esp. pp. 366–378.

[2] For the same attribute A on the same trial T in the same world W or knowledge context $\mathcal{K}zt$. Recent discussion of the symmetry thesis is provided by Adolf Grünbaum, *Philosophical Problems of Space and Time*, New York, Alfred A, Knopf, 1963, 2nd ed. 1973, Ch. 9; and in James H. Fetzer, 'Grünbaum's "Defense" of the Symmetry Thesis', *Philosophical Studies* (April 1974), 173–187.

[3] Reichenbach apparently never explicitly considered the question of the logical structure of statistical explanations; cf. Carl G. Hempel, 'Lawlikeness and Maximal Specificity in Probabilistic Explanation', *Philosophy of Science* (June 1968), 122.

[4] Especially as set forth in Hempel, 'Maximal Specificity', pp. 116–133; and in Wesley C. Salmon, *Statistical Explanation and Statistical Relevance*, Pittsburgh, University of Pittsburgh Press, 1971.

[5] Reichenbach, *Probability*, p. 68.

[6] Reichenbach, *Probability*, p. 69.

[7] Reichenbach, *Probability*, p. 72.

[8] Cf. Hilary Putnam, *The Meaning of the Concept of Probability in Application to Finite Sequences*, University of California at Los Angeles, unpublished dissertation, 1951.

[9] Reichenbach, *Probability*, pp. 376–77.

[10] Reichenbach, *Probability*, p. 377.

[11] This condition may be viewed as circumventing the problems posed by the requirement of randomness that might otherwise be encountered; but issues of randomness (or normality) will not figure significantly in the following discussion.

[12] As will be explained subsequently, this condition imposes a unique *value* solution but does not enforce a unique *description* solution to the problem of the single case.

[13] Assuming, of course, an unlimited supply of predicates in the object language. A different *density principle for reference classes* is discussed in Section 3.

[14] Reichenbach, *Probability*, p. 374. Variables are exchanged for convenience.

[15] It is not assumed here that '$P(K, A)$' stands for '$P(K \cdot -C, A)$'; see Section 2.

[16] Reichenbach, *Probability*, p. 375.

[17] Reichenbach, *Probability*, p. 376.

[18] Reichenbach, *Probability*, pp. 375–376.

[19] More precisely, perhaps, the term 'object' may be replaced by the term 'event' (or 'thing') to preserve generality; cf. Hempel, 'Maximal Specificity', p. 124.

[20] An extended discussion of this issue is provided by Hempel, 'Maximal Specificity', pp. 123–129.

[21] If a 'narrowest' reference class description is satisfied by more than one distinct event, then it is not a *narrowest* class description, since there is some predicate 'F^i' such that one such event satisfies 'F^i' and any other does not; otherwise, they would not be distinct events. This conclusion follows from the principle of identity for events and may therefore be characterized as a matter of logical necessity; cf. James H. Fetzer, 'A World of Dispositions', *Synthese* (forthcoming). A reference class description may be satisfied by no more than one distinct event, however, without being a narrowest reference class description. Thus, an

ontically homogeneous reference class may have only one member but is not therefore logically limited to a finite number of such members.

[22] There are therefore two theoretical alternatives in specifying a homogeneous reference class for the single individual trial T with respect to outcome A, namely: (a) a reference class description 'K^*' incorporating only *one* predicate 'F^i' (or a finite set of predicates {'F^n'}) satisfied by that individual trial alone; or, (b) a reference class description 'K^*' incorporating *every* predicate 'F^i', 'F^j',... satisfied by that individual trial alone. Since the set of predicates {'F^i', 'F^j', ...} satisfied by that individual trial alone may be an infinite set (and in any case will be a narrowest reference class description), alternative (a) shall be assumed unless otherwise stated. Note that $P(K^*, A) = P(F^i, A) = P(F^j, A) = \ldots = P(F^i \cdot F^j, A) = \ldots$, but nevertheless each of these predicates turns out to be statistically relevant.

[23] Cf. Carl G. Hempel, in *Aspects of Scientific Explanation*, New York, The Free Press, 1965, pp. 334–335; and Fetzer, 'Grünbaum's "Defense" ', pp. 184–186.

[24] This requirement is discussed, for example, in Hempel, *Aspects*, pp. 63–67.

[25] The relationship between the requirement of *total evidence* and a requirement of *explanatory relevance* (whether causal or not), moreover, is a fundamental issue. See, in particular, Hempel, *Aspects*, pp. 394–403; Salmon, *Statistical Explanation*, pp. 47–51 and pp. 77–78; and esp. Hempel, 'Maximal Specificity', pp. 120–123.

[26] Salmon, *Statistical Explanation*, p. 43. Variables are exchanged once again.

[27] Salmon, *Statistical Explanation*, pp. 42–45 and pp. 58–62.

[28] Salmon, *Statistical Explanation*, p. 55.

[29] Salmon, *Statistical Explanation*, p. 55.

[30] Salmon, *Statistical Explanation*, pp. 76–77.

[31] Wesley C. Salmon, 'Discussion: Reply to Lehman', *Philosophy of Science* (September 1975), 398.

[32] Where the role of laws is to certify the relevance of *causes* with respect to their *effects*; cf. James H. Fetzer, 'On the Historical Explanation of Unique Events', *Theory and Decision* (February 1975), esp. pp. 89–91.

[33] Cf. J. Alberto Coffa, 'Hempel's Ambiguity', *Synthese* (October 1974), 161–162.

[34] Salmon, 'Reply to Lehman', p. 400.

[35] Wesley C. Salmon, 'Theoretical Explanation', in *Explanation*, ed. by Stephan Körner, London, Basil Blackwell, 1975, esp. pp. 121–129 and pp. 141–145.

[36] Cf. Hempel, 'Maximal Specificity'; p. 130.

[37] This result may likewise be obtained by prohibiting the specification of any homogeneous reference class by *non-trivial* disjunctive properties, which is apparently implied by Hempel's definition of a maximally specific predicate, as discussed in Section 3.

[38] Salmon, *Statistical Explanation*, p. 44.

[39] Cf. Wesley C. Salmon, 'Comments on "Hempel's Ambiguity" by J. Alberto Coffa', *Synthese* (October 1974), 165.

[40] Hempel, 'Maximal Spceificity', p. 131.

[41] Hempel, 'Maximal Spceificity', p. 131.

[42] Cf. Hempel, *Aspects*, pp. 338–343; and Hempel, 'Maximal Specificity', pp. 123–129.

[43] Salmon, *Statistical Explanation*, p. 81. Reichenbach's epistemological program, including (a) the verifiability criterion of meaning, (b) the rule of induction by enumeration, and, (c) the pragmatic justification of induction, suggest a profound commitment to establishing wholly extensional truth conditions for probability statements, an effort that may be regarded as culminating in his introduction of the 'practical limit' construct. For discussion of certain difficulties attending the reconciliation of these desiderata, see James H. Fetzer,

'Statistical Probabilities: Single Case Propensities vs Long Run Frequencies', in *Developments in the Methodology of Social Science*, ed. by W. Leinfellner and E. Köhler, Dordrecht, Holland, D. Reidel Publishing Co., 1974, esp. pp. 394–396.
[44] Cf. Henry E. Kyburg, Jr., 'Discussion: More on Maximal Specificity', *Philosophy of Science* (June 1970), 295–300.
[45] Hempel, 'Maximal Specificity', p. 131.
[46] Hempel, 'Maximal Specificity', p. 131.
[47] Hempel, *Aspects*, pp. 394–396; and Hempel, 'Maximal Specificity', p. 118.
[48] As his own examples illustrate; cf. Hempel, 'Maximal Specificity', pp. 131–132.
[49] This contention may be taken to be Salmon's basic criticism of Hempel's view; cf. Salmon, *Statistical Explanation*, pp. 7–12 and esp. p. 35.
[50] Coffa, 'Hempel's Ambiguity', esp. pp. 147–148 and p. 154.
[51] Coffa, 'Hempel's Ambiguity', p. 154. Niiniluoto interprets Coffa's density principle analogously, concluding that it is *true* provided all the subclasses K^i of K required to generate these reference classes happen to exist. Ilkka Niiniluoto, 'Inductive Explanation, Propensity, and Action', in *Essays on Explanation and Understanding*, ed. J. Manninen and R. Tuomela, Dordrecht, Holland, D. Reidel Publishing Co., 1976, p. 346 and pp. 348–349.
[52] Salmon, 'Comments on Coffa', p. 167.
[53] Salmon, 'Comments on Coffa', p. 167.
[54] Cf. James H. Fetzer, 'Statistical Explanations', in *Boston Studies in the Philosophy of Science*, Vol. XX, ed. by K. Schaffner and R. Cohen, Dordrecht, Holland, D. Reidel Publishing Co., 1974, esp. pp. 343–344.
[55] Cf. Salmon, *Statistical Explanation*, pp. 9–10; Fetzer, 'Statistical Explanations', pp. 342–343; and esp. Richard C. Jeffrey, 'Statistical Explanation vs Statistical Inference', in Salmon, *Statistical Explanation*, pp. 19–28.
[56] Among those adhering to one version or another are Peirce, Popper, Hacking, Mellor, Giere, and Gillies. See, for example, Karl R. Popper, 'The Propensity Interpretation of Probability', *British Journal for the Philosophy of Science* **10** (1959), 25–42; James H. Fetzer, 'Dispositional Probabilities', in *Boston Studies in the Philosophy of Science*, Vol. VIII, ed. by R. Buck and R. Cohen, Dordrecht, Holland, D. Reidel Publishing Co., 1971, pp. 473–482; R. N. Giere, 'Objective Single-Case Probabilities and the Foundations of Statistics', in *Logic, Methodology, and Philosophy of Science*, ed. by P. Suppes *et al.*, Amsterdam, North Holland Publishing Co., 1973, pp. 467–483; and the review article by Henry E. Kyburg, 'Propensities and Probabilities', *British Journal for the Philosophy of Science* (December 1974), 358–375.
[57] See, for example, Ian Hacking, *Logic of Statistical Inference*, Cambridge, Cambridge University Press, 1965; Ronald N. Giere, 'The Epistemological Roots of Scientific Knowledge', in *Minnesota Studies in the Philosophy of Science*, Vol. VI, ed. by G. Maxwell and R. Anderson, Jr., Minneapolis, University of Minnesota Press, 1975, pp. 212–261; and James H. Fetzer, 'Elements of Induction', in *Local Induction*, ed. by R. Bogdan, Dordrecht, Holland, D. Reidel Publishing Co., 1976, pp. 145–170.
[58] Hempel's specific explication is not the only adequate explication (or the only adequate kind of an explication), therefore, precisely because condition (1) – reliance upon the frequency interpretation of physical probability – is not the only theoretical option. While Hempel's and Salmon's explications are both logically compatible with either the frequency or the propensity concepts, the choice between them is not a matter of philosophical preference but rather one of theoretical necessity (although Salmon suggests otherwise; cf. Salmon, *Statistical Explanation*, p. 82). For related efforts in this direction, see James H.

Fetzer, 'A Single Case Propensity Theory of Explanation', *Synthese* (October 1974), 171–198; and James H. Fetzer, 'The Likeness of Lawlikeness', *Boston Studies in the Philosophy of Science*, Vol. XXXII, ed. by A. Michalos and R. Cohen, Dordrecht, Holland, D. Reidel Publishing Co., 1976. Also, note that even when Hempel's conception is provided an *ontic* formulation (by deleting the conditions that render statistical explanations inductive arguments for which a high probability requirement within a knowledge context $\mathscr{K}zt$ is appropriate), it is still not the case that Hempel's explication and Salmon's explication are then logically equivalent. For Salmon's (original) conditions may be criticized as requiring the explanans of (at least some) explanations to be 'too broad', while Hempel's conditions even then may be criticized as permitting the explanans of (at least some) explanations to be 'too narrow'. Consequently, the relationship between Hempel's and Salmon's requirements is more complex than I previously supposed; cf. Fetzer, 'A Propensity Theory of Explanation', pp. 197–198, note 25, and Fetzer, 'Statistical Explanations', p. 342. For convenience of reference, finally, we may refer to the model of explanation elaborated here as the *causal-relevance* (or C–R) explication, by contrast with Hempel's (original) *inductive-statistical* (or I–S) model and with Salmon's (original) *statistical-relevance* (or S–R) model examined above.

[59] It is significant to note, therefore, that Hempel himself has endorsed the propensity interpretation of physical probability for statistical 'lawlike' sentences; Hempel, *Aspects*, pp. 376–380, esp. p. 378, Note 1. Hempel's view on this issue is critically examined by Isaac Levi, 'Are Statistical Hypotheses Covering Laws?', *Synthese* **20** (1969), 297–307; and reviewed further in Fetzer, 'A Propensity Theory of Explanation', esp. pp. 171–179. Salmon, in particular, does not believe that the choice between the propensity and the frequency conceptions of physical probability is crucial to the adequacy of his explication of explanation, contrary to the conclusions drawn above; Salmon, *Statistical Explanation*, p. 82.

[60] It is also significant to note that Hempel has recently expressed the view that Reichenbach's policy of assigning singular occurrences to 'the narrowest reference class' should be understood as Reichenbach's version of the requirement of total evidence *and that the requirement of total evidence is not an explanatory relevance condition*; 'Maximal Specificity', pp. 121–122. From the present perspective, moreover, Hempel's own revised requirement of maximal specificity (*RMS**), when employed in conjunction with the frequency interpretation of statistical probability, may itself be envisioned as applying within the *prediction context* generally rather than the *explanation context* specifically; indeed, for this purpose, Hempel's original definition of statistical relevance would appear to be theoretically unobjectionable, provided, of course, that these probability statements are no longer required to be *lawlike*. See also Fetzer, 'Elements of Induction', esp. pp. 150–160.