Direct Inference and the Problem of Induction

It is truth very certain that, when it is not in

our power to determine what is true, we ought

to follow what is most probable.

-- Rene Descartes, *Discourse on the Method*, III, 4

It would be difficult to overestimate the influence Hume's problem of induction exercises on contemporary epistemology. At the same time, the problem of induction has not perceptibly slowed the progress of mathematics and science. This ironic state of affairs, immortalized by C. D. Broad's description of induction as "the glory of science" and "the scandal of philosophy,"[1] ought in all fairness to give both sides some pause. And on occasion, it does: the mathematicians stop to concede that Hume has not yet been answered,[2] the scientists worry about randomization of experiments,[3] and inductive skeptics stir uneasily in their chairs at the mention of certain mathematical theorems that seem palpably to have bearing on the problem.[4]

But even when there is some cross-polination between fields, there is depressingly little sign of consensus on the underlying problem. Part of the difficulty lies in the babble of conflicting interpretations of probability, which has grown markedly worse since Broad's time.[5] Part of it lies in the structure of Hume's original argument, which scarcely makes direct contact with the mathematically sophisticated approach of contemporary statisticians and probabilists.[6] And no small part of it lies in the conviction of a considerable number of philosophers that Hume's problem, or at any rate a refurbished modern version thereof, is quite simply and clearly insoluble.[7]

I aim to show that this pessimism is unfounded. To this end I will articulate and defend the epistemic legitimacy of a very simple form of direct inference; a version so minimal, indeed, that the celebrated questions of confirmational conditionalization do not arise.[8] This is tantamount to sidestepping the delicate issue of competing reference classes, surely one of the most difficult problems facing any comprehensive theory of inductive inference. This might at first blush seem to leave a project too modest to be of interest, but as we shall see even this minimal appeal to direct inference is enormously controversial. And small wonder. For I will argue that the form of direct inference I am defending provides the key to the refutation of Humean skepticism -- theoretical and practical, historical and modern -- regarding induction.

Inverse and Direct Inference

A long tradition, stretching from Bernoulli and Bayes to Howson and Urbach, identifies the inference from sample to population as an exercise in "inverse" reasoning. From this standpoint, the structure of our inference makes use of Bernoulli's theorem, also known as the "Law of Large Numbers," in reverse. A Bayesian reconstruction runs thus: we are

seeking the probability that the frequency with which feature X occurs in a population lies within a small interval e of the value p, given that an n-fold sample exhibits X with frequency p (where m is the number of members in the sample exhibiting X, and p=m/n).[9] More formally, we are seeking

$$P((Fx(Pop) = (p \pm e)) \,/\, (Fx(Smp) = p) \,\&\, (S(Smp) = n))$$

for pertinent values of p, e, and n. From a Bayesian standpoint, we find this by expanding in the usual fashion:

$$\frac{P((Fx(Pop) = (p \pm e)) \,/\, (S(Smp) = n)) \times P((Fx(Smp) = p) \,/\, (Fx(Pop) = (p \pm e)) \,\&\, (S(Smp)=n))}{P((Fx(Smp) = p) \,/\, (S(Smp)=n))}$$

Bernoulli's theorem allegedly supplies us with the right hand term in the numerator. Unfortunately, as early critics of inverse inference were quick to point out, the left term of the numerator and (tacitly) the denominator both invoke a prior probability that the proportion of X's in the population lies within e of p.[10] How such priors are to be acquired is the fundamental problem of Bayesian inference; its apparent intractability is doubtless the chief stone of stumbling for non-Bayesians.

Actually, matters are complicated even with regard to the least controversial aspect of the Bayesian expression. What Bernoulli's theorem actually supplies is not the right hand term of the numerator in the Bayesian expression,

$$P((Fx(Smp) = p) \,/\, (Fx(Pop) = (p \pm e)) \,\&\, (S(Smp)=n)),$$

but rather

$$P((Fx(Smp) = (p \pm e)) \,/\, (Fx(Pop) = p) \,\&\, (S(Smp)=n)).$$

The former expression, unlike the latter, requires us to specify a probability distribution over possible values of Fx(Pop) -- a function that might vary sharply within a small interval around p.

I do not propose here to survey Bayesian responses to these difficulties, much less to adjudicate disputes about their adequacy. What I want to investigate instead are the prospects for a very different approach to inductive extrapolation that does not invoke prior probabilities and inverse inference but rather utilizes *direct* inference and Bernoulli's theorem to calculate

$$P((Fx(Pop) = (p \pm e)) \,/\, (Fx(Smp) = p) \,\&\, (S(Smp) = n)).$$

The method, if defensible, should hold interest for all but the most committed subjectivists.

Direct inference is perhaps the simplest and most natural expression of a "degree of entailment" interpretation of probability. Given the that frequency of property X in a population G is p, and the knowledge that *a* is a random member of G with respect to possession of X, the probability that *a* is an X is p.[11] Put more exactly,

$$P(Xa \, / \, (Ga \, \& \, FX(G) = p)) = p.$$

Clearly, much depends on the clauses ensuring that *a* is a random member of G with respect to possession of X. We will return to this qualification in due course.

The intuitive appeal of direct inference comes out strongly in simple examples. Donald Williams, a passionate advocate of direct inference, describes it in terms of the "intermediate cogency of proportional syllogisms."[12] Just as the classical syllogism warrants our concluding, from

1. All G are X

2. *a* is a G

with full assurance, that

3. *a* is an X,

so the proportional syllogism, subject to the restrictions mentioned above, licences our inference from

1'. m/n G are X

2'. *a* is a G

with assurance m/n, that

3'. *a* is an X.

As Williams points out, we use the classical syllogism but rarely: our major premises are not of the form "All falling barometers portend storms" or "All red-meated watermelons are sweet" but rather the more modest form that falling barometers *generally* portend storms and *most* red-meated watermelons are sweet.

> In the cadres of the traditional deductive logic, these changes make a fatal difference: the propositions that falling barometers generally portend a storm and that the barometer is now falling entail, strangely enough, nothing whatever about an impending storm.... Impatient with this finicking of the logician, the native wit of mankind has jauntily transcended the textbook formulas, has found the principle self-evident that if *All M is P* makes it certain that any one M will be P,

> then *Nearly all M is P* makes it nearly certain, and has quite satisfactorily predicted its storms and purchased its melons accordingly.[(13)]

Indeed, the classical syllogisms Barbara and Celarent, with a singular minor premise, can readily be seen as limiting cases of the proportional syllogism when m=n and m=0, respectively.[(14)] From this point of view, statistical syllogisms constitute a spectrum of inferences, each moving from statistical information to singular statements about members of the relevant class. The conclusion, as in the traditional syllogism, is always categorical, but the strength of the argument varies with the proportion cited in the major premise.

The notion of the statistical syllogism as a generalized form of the traditional one admitting intermediate grades of logical cogency is attractive, and a substantial number of philosophers have incorporated something like it in their treatment (though not always their justification) of inductive inference.[(15)] But granting for the moment the rationality of such direct inference, we have still to account for the truth of the major premise. How can we come by the knowledge that m/n ravens are black? In particular, how are we to come by it in a fashion that does not examine all ravens *seriatem*, including the one named *a*, so that in the last analysis direct inference falls prey to an analogue of Sextus Empiricus's complaint about the traditional syllogism -- that to complete the enumeration required to establish the major premise, we will have to make use of the conclusion, thus rendering the subsequent argument circular?[(16)]

Strictly speaking, we cannot guarantee the major premise without examining all of its instances. But as Williams points out, we can circumvent this problem by a clever combination of Bernoulli's "law of large numbers" and a second direct inference. Crudely but briefly put, Bernoulli's theorem says that most large samples differ but little from the population out of which they are drawn -- where "most" indicates a satisfyingly high percentage and "little" a gratifyingly small deviation from the true value, provided that "large" is sufficiently great. We may stipulate a small margin e such that a sample is said to "match" a population just in case

$$|p - m/n| \le e,$$

which is to say, the difference between the true proportion p and the observed sample proportion m/n is less than or equal to e. It is a simple matter then to choose a sample size n great enough that at least a proportion a of the possible n-fold samples will match the population to the stipulated degree of precision. The formula

$$n \ge .25 \, (e^2(1-a))^{-1}$$

gives the desired sample size.[(17)] For example, with e=.05 and a=.95, a sample of 2000 suffices. Given the sample size and the width of the interval, on the other hand, we can calculate the degree of confidence a simply by rearranging terms:

$$a = 1 - .25 \, (ne^2)^{-1}.$$

A lovely feature of this equation is that it does not mention p: we can calculate the confidence level without knowing the actual proportion in the population. It is easy to show that the likelihood of a "match" is worst when p = .5; so the constant .25, the maximum value of the product p(1 - p), represents the "worst case scenario" -- a will clearly be higher for any lower value of this term. By using this value, we can insure that our confidence levels are never overly optimistic.

Armed with an n-fold sample of balls (from the statistician's ubiquitous urn), 95% of which are red, we are in a position to reason as follows:

1^*. At least a of n-fold samples exhibit a proportion that matches the population

2^*. S is a random n-fold sample with respect to matching the population

===== [with probability a]

3^*. S matches the population

4^*. S has a proportion of .95 red balls

_____

5^*. p lies in the interval [.95 - e, .95 + e]

The move from 1^* and 2^* to 3^* is a direct inference, its major premise being underwritten by Bernoulli's theorem.[18] The move from 3^* and 4^* to 5^* incorporates the information regarding the sample proportion and the definition of matching. But 5^* is not quite the simple statistical statement we are accustomed to dealing with: rather, it states that the proportion of red balls in the urn lies in the interval [.95 - e, .95 + e]. Provided that e is small, however, the lower boundary of this interval is still a healthy majority. We can now extend the argument to predictive inference regarding an as yet unsampled ball from the urn:

6^*. *a* is a random ball from this urn with respect to redness

===== [with probability in the interval [.95 - e, .95 + e]]

7^*. *a* is red

Prima facie, this is a cogent response to Hume's challenge. There is no use caviling at 1^*, which is a mathematical truism. From 2^* and 4^*, which state the size and composition of our sample, and 6^* (which merely identifies *a*), we may draw a conclusion regarding an as-yet-unexamined member of the population with a reasonably high level of confidence. And by increasing the size of the sample, we can render the interval arbitrarily small without reducing the confidence level. Hence, an increase in sample size will allow us to take the sample proportion as an arbitrarily good estimate for p.

This solution is of more than academic interest. Hume himself grants that we have experience of bread nourishing us and the sun rising. If we may take our experience to be a sample, then it appears that we possess all the tools necessary to make a rational defense of everyday extrapolations against Humean skepticism. But philosophical battles are not so easily won. Virtually every aspect of the argument just presented has been called into question. To these objections we now turn.

Linear Attrition

A surprisingly common objection to direct inference, particularly in sampling examples, is that it reflects merely a linear elimination of alternatives and therefore offers no information regarding unexamined cases. A. J. Ayer suggests this in his description of a sampling experiment without replacement:

> If there are a hundred marbles in a bag and ninety-nine of them are drawn and found to be green, we have strong confirmation for the hypothesis that all of the marbles in the bag are green, but we have equally strong confirmation for the hypothesis that 99 per cent of the marbles are green and 1 per cent some other colour.[19]

In other words, drawing 99 balls from this bag gives us information precisely regarding the 99 balls in question, nothing more, nothing less. No matter how extensive our sample, the veil of ignorance always stands between us and the unsampled remainder.

John Foster, in his excellent book on Ayer, faithfully reproduces this criticism and explicates it with a clarity that leaves no interpretive questions. Mathematical arguments designed to show that favorable instances increase the probability of a generalization, says Foster, reflect

> merely the trivial fact that, with each new favorable instance, there are fewer cases left in which the generalization could fail. The point is seen most easily if we focus on the example of drawing balls from a bag. Let us assume, for simplicity, that we already know that there are ten balls in all and that each is either black or white. When we have drawn one ball and found it to be black, we have more reason to believe that all the balls are black, simply because there are now only nine possible counter-instances remaining.... This has nothing to do with induction, since it does not involve using the examined cases as a basis for predicting the properties of the unexamined cases. It tells us that the probability that *all* the balls are black increases, as the number of black balls drawn increases, but not that the probability that the *next* ball drawn will be black increases, as this number increases. Thus it does not tell us that, having drawn nine balls, we are entitled to be more confident about the colour of the tenth ball than when we first began the experiment.[20]

Fred Dretske raises a similar concern about inferring regularities from uniform instantial data:

> Confirmation is not simply raising the probability that a hypothesis is true: it is raising the probability that the unexamined cases resemble (in the relevant respects) the examined cases.[21]

His suggestion is that we ought to infer laws of nature directly, since they simultaneously explain the data we already have and imply something about, so to speak, the data we do not have.

In each case, the implication is that a direct inference does not do the job required. Ayer and Foster are concerned that the data conveyed by a sample speak only for themselves and not for the unexamined cases; Dretske is concerned that a direct inference, since it does not conclude with something stronger than a statistical generalization, will not allow the data to speak as they ought. In either case, the promised probabilities are a will-o-the-wisp.

All of this is half right. Surely, an inductive argument is of no value unless it gives us, on the basis of examined cases, a justification for our beliefs regarding unexamined ones. But as an explication of the mathematical rationale for direct inference, the thesis of linear attrition is demonstrably wrong. To see this, we need only shift to sampling from Ayer's bag with replacement -- creating, in effect, an indefinitely large population with a fixed frequency. No finite sample with replacement, no matter how large, ever amounts to a measurable fraction of this population. Yet as we have seen, using direct inference and Bernoulli's theorem it is simple to specify a sample size large enough to yield as high a confidence as one likes that the true population value lies within an arbitrarily small (but nondegenerate) real interval around the sample proportion.

The thesis of linear attrition resembles an intuitively plausible error to which many beginning students of statistics are prone, namely, the mistake of thinking that the value of information in a sample is a function of the proportion of the population sampled.[22] In fact, the relative proportion of the population sampled is not a significant factor in these sorts of estimation problems. It is the sheer amount of data, not the percentage of possible data, that determines the level of confidence and margins of error. This consideration sheds some light on the worry raised by Peter Caws:

> Scientific observations have been made with some accuracy for perhaps 5,000 years; they have been made in quantity and variety only for about 500 years, if as long. An extrapolation (on inductive grounds) into the past suggests that these periods represent an almost infinitessimal fraction of the whole life of the universe.[23]

Caws is certainly right to doubt whether every present regularity may properly be extrapolated into the misty past. But the grounds of such doubt have to do with our concrete evidence for differing conditions in the past rather than with the small fraction of time in which we have sampled the aeons. When we have no reason to believe conditions were relevantly different -- as in the case, say, of certain geological processes -

- we may quite rightly extrapolate backwards across periods many orders of magnitude greater than those enclosing our observations.

Randomness, Fairness and Representative Samples

Or may we? There is a sharp division of opinion on the question of randomness, and the defense of direct inference sketched above takes its stand on what is, admittedly, the more thinly populated side of the line. For four decades Henry Kyburg has stood almost *solus contra mundum* in his insistence that randomness is epistemic, that it is a primitive notion rather than something to be defined in terms of probability, and that in conjunction with statistical data it yields probabilities without "fair sampling" constraints. I think he is right; and an examination of the problems generated by the standard definition of randomness indicates why Kyburg's approach is so important.

The standard statistical approach defines "randomness" in terms of equiprobability: a selection of an n-fold set from a population is random just in case every n-fold set is as likely to have been drawn from that population as any other.[24]

"But surely," runs the argument, "it is incumbent upon the defenders of direct inference to make some sort of defense of the claim that the sample selected was no more likely to be chosen than any other. The assumption is not generally true. Elementary textbooks are replete with examples of bias in sampling. To assume without argument that one's sample is unbiased is more than imprudent: in effect, it attempts to manufacture valuable knowledge out of sheer ignorance."

No other single criticism is more widely canvassed or more highly regarded in the literature. Appropos of an example involving a sample of marbles selected one each from 1000 bags, each of which contains 900 red and 100 white balls, Ernest Nagel urges that while Bernoulli's theorem

> does specify the probability with which a combination belonging to M [the set of all possible 1000-fold samples, one from each bag] contains approximately 900 red marbles, it yields no information whatever concerning the proportion of combinations satisfying this statistical condition that may be *actually selected* from the 1000 bags -- unless, once more, the independent factual assumption is introduced that the ratio in which such combinations are *actually selected* is the same as the ratio of such combinations in the set M of all *logically possible* combinations.[25]

Without a special assumption of "fair sampling," we are vulnerable to the possibility that some samples may be much more likely to be selected than others; and perhaps the ones most likely to be selected are highly unrepresentative. Isaac Levi explicitly urges the need for such restrictions on direct inference in his critique of Kyburg.

> Suppose X knows that 90% of the Swedes living in 1975 are Protestants and that Petersen is such a Swede. Imagine that X knows nothing else about Petersen. On

Kyburg's view, X should assign a degree of credence equal to .9 to the hypothesis that Petersen is a Protestant.

I see no compelling reason why rational X should be obliged to make a credence judgment of that sort on the basis of the knowledge given. X does not know whether the way in which Petersen came to be selected for presentation to him is or is not in some way biased in favor of selecting Swedish Catholics with a statistical probability, or chance, different from the frequency with which Catholics appear in the Swedish population as a whole....

For those who take chance seriously, in order for X to be justified in assigning a degree of credence equal to .9 to the hypothesis that Petersen is a Protestant on the basis of direct inference alone, X should know that Petersen has been selected from the Swedish population according to some procedure F and also know that the chance of obtaining a Protestant on selecting a Swede according to procedure F is equal to the percentage of Swedes who are Protestants.[26]

Here is a pretty puzzle. We set out initially in search of a form of inference that would supply something we lacked: a rationally defensible ascription of probabilities to contingent claims on the basis of information that did not entail those claims. If we are required for the completion of this task to have in hand already the probability that this particular sample would be drawn (and indeed an identical probability for the drawing of each other possible sample), or information on the "chance" of obtaining a given sort of individual from the population (above and beyond frequency information), then the way is blocked. Direct inference is impaled on the empirical horn of Hume's dilemma.

The first step toward answering this criticism is to distinguish a "fair" sample from a "representative" one.[27] Fair samples are drawn by a process that gives an equal probability to the selection of each possible sample of that size; a representative sample exhibits the property of interest in approximately the same proportion as the overall population from which the sample is drawn. To insist on a guarantee that the sample be *representative* in this sense is to demand something that turns induction back into deduction, for if we are certain that the sample is representative, we know *eo ipso* approximately what the population proportion is.

If representativeness is too much to demand, however, fairness seems at first blush to be a just requirement. We should like to avoid biased (i.e., unrepresentative) samples; and since most large samples are representative, a selection method that gives each such sample equal probability of being selected yields an agreeably high probability that a given sample is representative. Fair sampling will on occasion turn up samples that are wildly unrepresentative. But constraints of the sort outlined by Levi, if we could be sure they held good, would assure us that in the long run these biased samples will make up only a small proportion of the total set of samples.

Here again, however, the road to an *a priori* justification of induction appears closed. For under the demands of fair sampling, we cannot rely on the direct inference unless we

know that each possible sample was equally likely to be chosen. And that is itself a contingent claim about matters that transcend our observational data and stands, therefore, in need of inductive justification. An infinite regress looms.

The point can be put in another way. Levi requires that X know Petersen has been selected by a method F that has a .9 chance of selecting Protestants from the population of Swedes. But what does "chance" mean here? Surely it does not mean that 90% of the *actual* applications of F result in the selection of Protestants from among Swedes. For this would reduce the problem to another direct inference, this one about instances of F rather than about Swedes, and if this sort of answer were satisfactory there would have been no need to appeal to F in the first place.[28]

Perhaps sensible of this difficulty, Michael Friedman has tried to rescue a notion of objective chance by appealing to the set of actual and physically possible applications of a method, arguing that if the ratio of successes in such a set is favorable then we may say that the objective chance of success is high -- and hence, in Friedman's terminology, that the method is "reliable" -- regardless of the ratio of actual successes.[29] But this approach is open to three serious objections. First, it is unclear that there is a definite ratio of successes to trials for the set of all actual and physically possible applications of an inferential method: for there may be infinitely many physically possible applications, and hence an infinite number both of successes and of failures. Second, waiving this difficulty, many of our applications of inductive methodology may yield theories and empirical claims which are accepted at present but not *independently* certifiable as true. Hence, the ratio of successes to trials even among our actual applications may prove impossible to estimate without begging the question. Third, even if this problem can be circumvented, we are left with the question of how to estimate the proportion of successes among actual *and physically possible* applications of the method. To do so by deriving the "reliability" of our inductive methods from extant scientific theories tangles us up in epistemic circularity, for those theories have nothing to commend them except that we have arrived at them by our inductive methods. Friedman is, I think, overly optimistic about the epistemic worth of such appeals.[30] On the other hand, if we are to estimate the frequency of inferential successes from our actual experience, we are back to direct inference once again. If simple direct inference is not epistemically acceptable, we are back to fair sampling constraints. And fair sampling constraints will not rescue induction.

Contrary to common wisdom, however, an assumption of fairness is *not* necessary for the epistemic legitimacy of the inference from sample to population. What is required instead is the condition that, relative to what we know, there be nothing about this particular sample that makes it less likely to be representative of the population than any other sample of the same size. And this is just a particular case of the general requirement for direct inference that the individual about which the minor premise speaks be a random member of the population with respect to the property in question.[31]

Even some critics of direct inference have recognized the justice of this point. Wisdom, for example, points out that it accords well with practical statistical work.

We know in practical affairs that we must take random samples. But this is because we utilise existing knowledge. If we know of some circumstance that would influence a sample, we must look for a sample that would be uninfluenced by it.... Now all this is only to say that *we avoid using a sample that is influenced in a known way*.... If we demand that they should be random in some further sense, it is either a demand for knowledge of 'matching' or for additional knowledge about the influences that might affect the sample -- the one would render statistical inference superfluous, the other is worthy in the interests of efficiency but does not come into conflict with Williams' argument. After all, *probability is used when all available knowledge has been taken account of and found insufficient.*[(32)]

An example makes this plain. Every Friday afternoon at 3:30 p.m. sharp, Professor Maxwell emerges from his office, strides down the hall to the freshly-stocked vending machine, inserts the appropriate amount of coin of the realm, and punches the button for a Coke. Because of the way the machine is designed, he will of course get the can resting at the bottom of the column: it is *that* can, no other, that will emerge. Yet given the information that one of the fifty cans in the vertical column is a Mello Yello and the other forty-nine are Coke, he is still justified in placing the probability that he will get a Coke at 98%. True, Maxwell is not equally likely to get any of the various cans stacked within the machine: his selection is not fair. But the Mello Yello is, on his information, a random member of the stack of cans with respect to position. Consequently, the can at the bottom is, on his information, a random element of the stack with respect to being a Coke.

The contrary intuition that demands fairness depends, I submit, on a Cartesian worry rather than a Humean one: it conflates the presence of possibilities with the absence of probabilities.[(33)] If Maxwell sees that the machine has just been stocked by Damon, a resentful former logic student, he may harbor reasonable doubts that the can at the bottom is a Coke; it may not be a random member of the stack with respect to that property. (It may be a random member of the set of objects deliberately placed in someone's path by a practical joker intent on upsetting his victim's expectations -- a set in which the frequency of anticipated outcomes is rather different!) But in the absence of some definite contrary evidence, the mere *possibility* that some can or other of the fifty might have been chosen deliberately to be placed at the bottom does not, in itself, provide information that changes the *probabilities* obtained by direct inference. And the fact that possibilities do not eliminate probabilities is a point that Descartes himself, for all his skeptical arguments, recognized very clearly.

The same considerations apply, *mutatis mutandis*, to sampling. The possibility that we might be sampling unfairly, like the logical possibility that Maxwell's nemesis Damon has maliciously stacked the machine to trick him, cannot be eliminated *a priori*. But in the absence of concrete evidence that, e.g., places the about-to-be-selected sample in a different and more appropriate reference class, mere possibilities should not affect our evaluation of epistemic probabilities.

Appearances notwithstanding, this is not a retreat to the old principle of indifference; nor is it vulnerable to the charge, to which some advocates of that principle have exposed themselves, that it manufactures knowledge out of ignorance. Indifference assigns equal probabilities to each element of a set on the basis of symmetry considerations, and a drawing method from that set is baptized "random" in terms of that assignment. On the account advocated here, by contrast, randomness is not parasitic on probability. To say that *a* is a random member of class F with respect to having property G, relative to my corpus of knowledge K, does invoke symmetry considerations. But when combined with knowledge of the frequency of G's among the F's, epistemic symmetry yields probabilities that reflect the relevant empirical information rather than reflecting hunches, linguistic symmetries or preconceived predicate widths. It is a consequence of this view that, in situations of complete ignorance regarding the proportion of F's that are G's, symmetry by itself yields no useful probability information. This is an intuitively gratifying result. Epistemic symmetry conjoined with ignorance yields ignorance; conjoined with knowledge, it yields symmetrical epistemic probabilities.

Success versus Rationality

The foregoing defense of randomness as a basis for assigning probabilities raises a fresh difficulty. The sort of "probability" that can be gotten from randomness and statistical information regarding a reference class is relativized, in the very definition of "randomness," to the state of our knowledge; and this strikes some critics as too much of a retreat from the goal of arriving at true beliefs. As a consequence, so runs the objection, any defense of induction predicated on epistemic probability fails to address the true problem -- the problem of future success.

This criticism recalls our reconstructed version of Hume's dilemma: "Granted that these premises are true and that the conclusion is linked to them by a direct inference; why should that fact make the conclusion probable for me, in a sense that commends it to me if I prefer truth to falsehood?" By analogy with the natural answer regarding deductive inference, it would be at least *prima facie* satisfying to answer that direct inference guarantees a high proportion of future successes. But without fair sampling constraints, which as we have seen would only engender a regress, direct inference offers no such guarantee. Hao Wang puts the challenge succinctly when he notes that on an epistemic interpretation of probability

> we shall at no stage be able to pass from a certain frequency being overwhelmingly probable to it being overwhelmingly frequent. That is to say, on any non-frequency interpretation we have no guarantee that on the whole and in the long run the more probable alternative is the one that is more often realized.[34]

And again, criticizing Williams's *a priori* interpretation of probability, Wang asks:

> [W]hat guarantees induction to lead us more often to success than to disappointment,--granted that we can justify inductive generalizations with high probability on some a priori ground? ... a principle of induction which might

always lead to disappointment does not seem to be what is wanted....the conclusions reached in such fashion need not guarantee success, on the whole and in the long run, of our actions guided by them as predictions. In granting that we know a priori that a large sample very probably has nearly the same composition as the whole population, we must not forget that here what are known to be more probable need not be those which are on the whole and in the long run more often realized.[35]

Predictably, this line of criticism is advanced most vigorously by those who insist that both the definition of probability and the legitimacy of induction are bound up inextricably with contingent claims about the nature of the physical world. Nagel makes it clear that what makes Williams's justification of induction unacceptable to him is precisely this failure to guarantee success.

> For without the assumption, sometimes warranted by the facts and sometimes not, that a given method of sampling a population would actually select all samples of a specified size with roughly the same relative frequency, arithmetic can not assure us that we are bound to uncover more samples approximately matching the population than samples that do not.[36]

Why should such a "guarantee" or an "assurance" seem a compelling requirement for the justification of induction? Russell, in his defense of a finite frequency interpretation of probability, offers a clue. If we are obliged to admit that the improbable may happen, then

> a probability proposition tells us nothing about the course of nature. If this view is adopted, the inductive principle may be valid, and yet every inference made in accordance with it *may* turn out to be false; this is improbable, but not impossible. Consequently, a world in which induction is true is empirically indistinguishable from one in which it is false. It follows that there can never be any evidence for or against the principle, and that it cannot help us to infer what will happen. If the principle is to serve its purpose, we must interpret "probable" as meaning "what in fact usually happens"; that is to say, we must interpret a probability as a frequency.[37]

But the moral drawn here confuses success with rationality. What Russell means by a world in which induction is "true" is, apparently, one in which inductive reasoning works well. Since it might turn out that all of our samples are unrepresentative, our extrapolations from them might all be hopelessly wide of the mark. This is, however, a reversion to the Cartesian worry. It is possible to get a large but unrepresentative sample, just as it is possible to draw the one black ball from an urn of a million, 999,999 of which are white. But it would be irrational to expect this, given no further relevant information; and it is equally irrational to expect our samples to be unrepresentative and our inductions, in consequence, unsuccessful.

This conflation of Humean and Cartesian worries underlies Russell's complaint that such a principle "cannot help us to infer what will happen." If we demand a guarantee of success, or at any rate a guarantee of a high frequency of future successes, then we are indeed out of luck: that sort of "help" is not forthcoming. No amount of reasoning will turn contingent propositions into necessary ones. But rationality requires both less and more than this: less, because it is logically possible that a rational policy of nondemonstrative inference may always lead us astray; and more, because no accidental string of successes can in and of itself establish a policy of inference as rational. Perhaps the real value of surveying our success frequencies is that it gives us a rough gauge of the "uniformity of nature" in a sense that, while *post hoc* and therefore not useful for justifying induction, is at least tolerably clear.

Ironically, a guarantee of a high proportion of successes is not only unavailable but would be useless to the apostles of success without a subsequent appeal to unvarnished direct inference.[38] This is not merely because in the long run we are all dead: it applies even to an ironclad guarantee that 99% of all of the inductions we make in the next year will be true. For in applications, it is always *this* induction, this particular instance, that is of importance. Even if it were granted that the proportion of successes among our inductions in the next year is .99, and that this application of inductive methodology is, given our present evidence, a random member of the class of those inductions with respect to its success, why should these facts confer any particular epistemic credibility upon the notion that this induction will be successful? The rationality of direct inference is so fundamental that it cannot even be criticized in this fashion without a covert admission that it is rational.

Once we have seen this, we are freed from the trap of thinking that a proper justification of induction must necessitate future success. The correct response to the modern Humean challenge regarding probabilities is to distinguish it from Cartesian anxiety over possibilities and, having done so, to point out the way in which direct inference is underwritten by the symmetry of epistemically equivalent alternatives with respect to concrete frequency data. That symmetry offers no binding promises with respect to the future, no elimination of residual possibilities of failure. But our probabilistic extrapolations are apt to fail *only* if our samples have been unrepresentative; and despair over this bare possibility is, at bottom, an instance of the same fallacy that drives the credulous to purchase lottery tickets because of the *possibility* of winning. To see this fixation on possibilities aright is to understand the legitimacy of direct inference and to recognize that the probabilities it affords us are, in every sense of the term, rational.

Sampling the Future: the Modal Barrier

Granting that the rationality of direct inference is logically independent of its record of successes, it is subject to what appears at first sight to be a severe limitation: it applies only to the population *from which we are sampling*, and that population often seems much more restricted than the scope of our conclusions. C. D. Broad raises this consideration to cast doubt on any approach to the problem of induction that takes its cue from observed samples, both because of our "restricted area of observation in space" and

because of the "distinction of past and future cases" -- by which he means quite simply that the probability of our having met any future crow is zero.[39] It is impossible to sample the future. Wisdom picks up on Broad's criticism to supply a vivid image of the modal barrier that apparently blocks the use of direct inference from the past and present to the future:

> [I]f some balls in an urn were sewn into a pocket, we could not get a fair sample -- or rather we could not get a sample at all. Likewise the 'iron curtain' between the present and the future invalidates inductive extrapolation about the composition of things behind the curtain -- we cannot sample them from this side.[40]

This objection has a plausible ring, but it proves extraordinarily difficult to give a detailed explanation of just why the modal barrier should block direct inferences. There is a metaphysical thesis, going back to Aristotle, that future tense contingent statements have no present truth value.[41] This seems strong enough to scotch direct inference regarding the future, but it goes well beyond the modal barrier raised by Broad and Wisdom; indeed, it is difficult to see how the problem of induction could even arise with respect to the future if we run no risk of speaking falsely when we make contingent claims in the future tense. Traditionally, the chief motivation for this approach has been the fear that allowing contingent claims about the future to be true in the present would commit us to fatalism.[42] It is by now widely acknowledged that there are serious problems with the reasoning behind this charge.[43]

But even if fatalism did follow from the unrestricted law of excluded middle, the attempt to salvage human freedom by denying the truth of future contingents seems to be a cure nigh as evil as the disease. For in a great many contexts where freedom matters to us it is bound up with *deliberation*, and deliberation involves, ineliminably, the consideration of possible but avoidable future courses of action and their possible but avoidable future consequences. If future contingents have no truth values, then deliberation is a sham. This is not a plausible way to rescue human freedom.

The real attractions of the modal barrier lie elsewhere. Ayer, for example, grants as an arithmetical truism that an omniscient being who made every possible selection precisely once would necessarily find that most of his samples were typical.

> It hardly needs saying, however, that we are not in this position.... So far from its being the case that we are as likely to make any one selection as any other, there is a vast number of selections, indeed in most instances the large majority, that it is impossible for us to make. Our samples are drawn from a tiny section of the universe during a very short period of time. And even this minute portion of the whole four-dimensional continuum is not one that we can examine very thoroughly.[44]

To extricate ourselves from this predicament, says Ayer, we require

two quite strong empirical assumptions. They are first that the composition of our selections, the state of affairs which we observe and record, reflects the composition of all of the selections which are available to us, that is to say, all the states of affairs which we could observe if we took enough trouble; and secondly that the distribution of properties in the spatio-temporal region which is accessible to us reflects their distribution in the continuum as a whole.[(45)]

He is prepared to grant the first assumption, provided that we have taken some precautions to vary our samples and test our hypotheses under different conditions to safeguard against bias. But the second one he finds deeply problematic. The problem is not just that we are intuitively disinclined to extrapolate our local sample billions of years into the future or billions of light-years across the visible universe. That problem can be resolved by restricting the field of our conjectures to our local cosmic neighborhood and the relatively near future, and such a restriction may guarantee that our sample is typical of the local region of spacetime. If we approach the matter in this fashion, then

> we can be certain, and that without making any further assumptions, that in many cases the percentages with which the characters for which we are sampling will be distributed among [the populations in which we are interested] will not be very different at the end of the future period from what they are now. This will be true in all those casees in which we have built up such a backlog of instances that they are bound to swamp the new instances, however deviant these may be. But this conclusion is of no value to us. For we are interested in the maintenance of a percentage only in so far as it affects the new instances. We do not want to be assured that even if these instances are deviant the final result will be much the same. If we make the time short enough, we know this anyway. We want to be assured that the new instances will not be deviant. But for this we do require a non-trivial assumption of uniformity.[(46)]

Ayer's adroit exposition almost succeeds in concealing the fact that he has smuggled in the thesis of linear attrition once again. The problem arises not because the unsampled instances are future, but rather because they are unsampled, and we want to be assured that the unsampled instances are not deviant. "New instances" are the ones about which we have no information. If this objection works at all, it will work regardless of their temporal position. The modal barrier is simply the veil of ignorance seen from a particular point of view.

This analysis of the objection casts doubt on Ayer's distinction between the two assumptions he thinks we need. If we are going to be worried about the unrepresentativeness of our sample regarding the far reaches of spacetime on the grounds that those far reaches may be deviant, then why not also be worried about unexamined ravens in the local wood at the dawn of the twenty-first century, since they may be deviant as well? That we have varied the conditions of our observations is no defense against this possibility, for we wish (following Ayer's example) to know not merely that our sample is representative of the whole spatiotemporally local population but that it is

representative of the unexamined instances within that population. And however uniform our sample heretofore, we cannot eliminate what Wisdom calls

> the theoretical [problem] of making an inference about unexamined things in view of the possibility that the universe might play some trick that would wreck our best calculated expectations.[(47)]

Thus the thesis of linear attrition, and with it the modal barrier, are grounded in the Cartesian worry about possibilities that we have already met; for the fear that the universe might "trick" us is plainly a reversion to Maxwell's apprehensions regarding Damon. Why, to use Wisdom's own analogy, should we believe that the balls sewn into a pocket in the bag are specially unrepresentative of the whole? To be sure if we had some information to that effect then epistemic randomness would be violated and we would not use direct inference. But Wisdom leaves no doubt that fear of the bare *possibility* that our samples might be unrepresentative lies at the root of his inductive skepticism, for in his critique of Williams he explicitly repeats the objection:

> It is true that in the absence of knowledge of factors influencing a sample we rightly use that sample as a guide and that with such knowledge we rightly reject a sample. But here the position is that we do not know whether or not there is an influence at work and we think it possible there may be. In view of this doubt we cannot regard the sample as a guide that has the required statistical reliability.[(48)]

Such is the moral of our extended examination of the problem of induction. In case after case, the challenges to direct inference reduce to the fundamental objection that the possibility of error has not been eliminated. The thesis of linear attrition, the demand for fairness constraints, the insistence on a guarantee of success and despair of breaching the modal barrier are all variants on the same underlying theme: the fear "that the universe might play some trick" on us. To such an objection there is in the final analysis only one answer, as old as Herodotus:

> There is nothing more profitable for a man than to take counsel with himself; for even if the event turns out contrary to one's hope, still one's decision was right, even though fortune has made it of no effect: whereas if a man acts contrary to good counsel, although by luck he gets what he had no right to expect, his decision was not any the less foolish.[(49)]

ENDNOTES

1. In a 1926 lecture on "The Philosophy of Francis Bacon," reprinted in Broad, C. D., *Ethics and the History of Philosophy* (New York: Humanities Press, 1952). The comment appears on p. 143.

2. For example, Harrold Jeffreys, *Theory of Probability*, 2nd ed. (Oxford: Oxford University Press, 1948), p. 395; I. J. Good, *The Estimation of Probabilities* (Cambridge, MA: MIT Press, 1965), p. 16.

3. For a classic reference, see R. A. Fisher, *The Design of Experiments* (New York: Hafner, 1971 (originally published in 1935)).

4. See, for instance, Edwin Hung's discussion of sampling and the Law of Large Numbers in his undergraduate textbook *The Nature of Science: Problems and Perspectives* (New York: Wadsworth, 1997), pp. 276-7, 292-4.

5. Useful surveys of the conflicting schools regarding the interpretation of probability and its relation to statistics, coming from theorists of various persuasions, may be found in Howard Raiffa, *Decision Analysis* (Reading, MA: Addison-Wesley, 1968), Alex Michalos, *Principles of Logic* (Englewood Cliffs, NJ: Prentice Hall, 1969), J. R. Lucas, *The Concept of Probability* (Oxford: Oxford University Press, 1970), J. L. Mackie, *Truth Probability and Paradox* (Oxford: Oxford University Press, 1973), Henry Kyburg, *Logical Foundations of Statistical Inference* (Boston: D. Reidel, 1974), and Roy Weatherford, *Philosophical Foundations of Probability Theory* (Boston: Routledge & Kegan Paul, 1982).

6. See, for example, D. C. Stove, "Hume, Probability, and Induction," *Philosophical Review* 74 (1965): 160-77.

7. Notable representatives of this viewpoint are Karl Popper, *Conjectures and Refutations* (New York: Harper and Row, 1963) and three philosophers of science heavily influenced by Popper: J. O. Wisdom, *Foundations of Inference in Natural Science* (London: Meuthen & co., 1952), John Watkins, *Science and Skepticism* (Princeton: Princeton University Press, 1984), and David Miller, *Critical Rationalism* (Chicago: Open Court, 1994). But skeptical worries about induction also drive non-Popperians to pessimistic epistemological conclusions, as in A. J. Ayer's *Probability and Evidence* (London: Macmillan, 1973) and *The Central Questions of Philosophy* (New York: William Morrow and Co., 1973) and Richard Fumerton's recent book *Metaepistemology and Skepticism* (Littlefield Adams, 1995).

8. See Isaac Levi, "Direct Inference," *Journal of Philosophy* 74 (1977): 5-29, Kyburg's reply "Randomness and the Right Reference Class," *Journal of Philosohy* 74 (1977): 501-21, Levi's response "Confirmational Conditionalization," *Journal of Philosophy* 75 (1978): 730-37, and Kyburg's rebuttal "Conditionalization," *Journal of Philosophy* 77 (1980): 98-114. There are further details in Radu Bogdan, ed., *Profile of Kyburg and Levi* (Dordrecht: D. Reidel, 1981) and "Epistemology and Induction," in Kyburg's collection *Epistemology and Inference* (Minneapolis: Minnesota University Press, 1983). Not all version of direct inference violate confirmational conditionalization: see, e.g., John Pollock, *Nomic Probability and the Foundations of Induction* (New York: Oxford University Press, 1990), p. 137 n.16. The issue is important but does not affect the discussion here.

9. This statement simplifies slightly: the upper and lower boundaries need not be identical, as Keynes points out in his *Treatise on Probability* (London: Macmillan, 1963), pp. 338-9. Provided that $p(1-p)n$ is large enough, the asymmetry is negligible, but it will

not affect the discussion here if we select e so as to yield a conservative estimation of the probability that the true frequency of X in the population lies within that interval around p. In the finite case a unique shortest interval is always computable.

10. This prior may be assumed to be independent of the mere size of our sample, though some critics of direct inference, construing it as an inverse inference, have maintained that it may not be independent of the sample frequency of m/n. See Patrick Maher, "The Hole in the Ground of Induction," *Australasian Journal of Philosophy* 74 (1996): 423-32. But this criticism is blocked by the requirement of epistemic randomness discussed below.

11. There is some terminological variability in the use of the phrase "direct inference." Carnap, in *Logical Foundations of Probability* (Chicago: University of Chicago Press, 1950), sec. 94 uses it to denote the inference from the known constitution of a population to the most probable constitution of a *sample* drawn from that population. My usage in this paper resembles that in more recent discussions, e.g., Kyburg, "Epistemology and Induction," in *Epistemology and Inference* (Minneapolis: University of Minnesota Press, 1983), pp. 221-31. Unless otherwise noted I take direct inference to be *simple*, i.e., made without dependence on "fairness" constraints. This issue is taken up in detail below. Note that 'p' may take interval values of the form [a, b], where 0 a b 1. Point values for p may be construed as degenerate intervals where a=b.

12. Donald Williams, *The Ground of Induction* (New York: Russell & Russell, 1963), p. 39.

13. Williams, p. 8.

14. That general statements may be construed as limiting cases of probability statements is also stressed by R. B. Braithwaite, *Scientific Explanation* (Cambridge: Cambridge University Press, 1968), p. 152.

15. Roy Harrod indicates that he is in "substantial agreement" with Williams's position in *Foundations of Inductive Logic* (New York: Harcourt, Brace & Co., 1956), pp. xv, 103 ff, etc., though his own system is in some respects idiosyncratic. Max Black adopts a version of the statistical syllogism in "Self-Supporting Inductive Arguments," *The Journal of Philosophy* 55 (1958): 718-25. Stephen Toulmin advocates his own informal version of the statistical syllogism in *The Uses of Argument* (Cambridge: Cambridge University Press, 1958), pp. 109 ff, though he tends to confuse the strength of the inference with the strength of the conclusion (see p. 139). Simon Blackburn endorses a pair of weaker, qualitative claims analogous to results derivable from the statistical syllogism in *Reason and Prediction* (Cambridge: Cambridge University Press, 1973), pp. 126 ff. Paul Horwich, though a self-professed "therapeutic Bayesian," suggests supplementing coherence with a form of direct inference (couched, of course, in terms of "degree of belief") in *Probability and Evidence* (Cambridge: Cambridge University Press, 1982), pp. 33-4. J. L. Mackie exploits direct inference in his contribution to a 1979 Festschrift for A. J. Ayer, "A Defence of Induction," reprinted in Mackie, *Logic and Knowledge* (Oxford:

Oxford University Press, 1985), pp. 159-77. D. C. Stove endorses and elaborates upon Williams's position in the first half of *The Rationality of Induction* (Oxford: Oxford University Press, 1986). John Pollock develops a detailed theory of direct inference, incorporating sevaral variations on the statistical syllogism, in *Nomic Probability and the Foundations of Induction* (Oxford: Oxford University Press, 1990). The most extensive, probing, and systematic exploitation of direct inference is found in Henry Kyburg's work, spanning more than four decades from "The Justification of Induction," *Journal of Philosophy* 53 (1956): 394-400 and *Probability and the Logic of Rational Belief* (Middletown, CT: Wesleyan University Press, 1961) to the present.

16. Arthur Prior gives a useful sketch of this controversy in his article "Logic, Traditional," in P. Edwards, ed., *The Encyclopedia of Philosophy* (New York: Macmillan and Free Press, 1968), vol. 5, pp. 41-2.

17. Note that there is a minor slip in the otherwise excellent discussion of this formula in Debora Mayo, *Error and the Growth of Knowledge* (Chicago: University of Chicago Press, 1996), p. 170: it is a, not (1-a), that represents the desired confidence level.

18. If the need arises, we can get an even more generally applicable result by replacing Bernoulli's theorem with Tchebyshev's inequality: given *any* distribution of data, not less than $1-(1/n^2)$ of the distribution lies with n standard deviations of the mean. The estimates yielded by Tchebyshev's inequality are generally more cautious than those derived using Bernoulli's theorem, and in a wide range of cases needlessly so. But they have the advantage that they are essentially independent of constraints on the distribution. See William Feller, *An Introduction to Probability Theory and its Applications*, Vol 1, 2nd ed. (New York: John Wiley & Sons, 1957), pp. 219-21.

19. Ayer, *The Central Questions of Philosophy* (New York: Wm. Marrow and Co., 1973), p. 178.

20. Foster, *A. J. Ayer*, p. 211. Foster brings this example up to counter a version of Bayes's Theorem, but it has more direct bearing on direct inference.

21. Dretske, "Laws of Nature," *Philosophy of Science* 44 (1977): 258.

22. Those interested in speculations on the history of probability might want to investigate the possibility that Keynes, by introducing his Principle of Limited Variety and thereby attempting to ground enumerative induction in eliminative inference, fostered the confusion visible in the thesis of linear attrition.

23. Caws, *The Philosophy of Science* (New York: D. Van Nostrand & Co., 1965), p. 265.

24. See, e.g., Bhattacharyya and Johnson, *Statistical Concepts and Methods* (New York: Wiley, 1977), pp. 86-7. Similar definitions can be found in almost any statistics text.

25. Nagel's review appears in *Journal of Philosophy* 44 (1947): 685-93. The quoted remark appears on p. 691.

26. Levi, "Direct Inference," pp. 9-10.

27. There is an unfortunate tendency in some introductory textbooks to use these terms interchangeably. See, e.g., Hung, *The Nature of Science*, p. 277, and Robert M. Martin, *Scientific Thinking* (Orchard Park, NY: Broadview Press, 1997), where the following definition appears on p. 55: "A **REPRESENTATIVE SAMPLE** is a sample that is likely to have close to the same proportion of the property as the population." The introduction of "is likely" here blurs the distinction between fairness and representativeness.

28. This point is raised in a slightly different form by Kyburg in "Randomness and the Right Reference Class," p. 515.

29. Michael Friedman, "Truth and Confirmation," *Journal of Philosophy* 76 (1979): 361-82. Reprinted in Hillary Kornblith, ed., *Naturalizing Epistemology* (Cambridge, MA: MIT Press, 1985), pp. 147-167. See pp. 153-4.

30. Friedman, "Truth and Confirmation," pp. 154-7. The appeal to epistemically circular arguments is characteristic both of inductive defenses of induction and of externalist epistemologies: the arguments, both pro and con, show remarkable similarities. For attempted inductive justifications of induction, see Braithwaite, *Scientific Explanation*, and Max Black, *Problems of Analysis* (Ithaca: Cornell University Press, 1954), ch. 11. Braithwaite's approach is criticized in Kyburg, "R. B. Braithwaite on Probability and Induction," *British Journal for the Philosophy of Science* 35 (1958-9): 203-20, particularly pp. 207-8, and Wesley Salmon critiques Black's use of epistemic circularity, which he terms "rule circularity," in *Foundations of Scientific Inference* (Pittsburgh:University of Pittsburgh Press, 1969), pp. 12-17. For appeals to epistemic circularity on behalf of epistemic externalism, in addition to Friedman see William Alston, "Epistemic Circularity," *Philosophy and Phenomenological Research* 47 (1986), reprinted in Alston's collection *Epistemic Justification* (Ithaca: Cornell University Press, 1989), pp. 319-349, and his recent book *The Reliability of Sense Perception* (Ithaca: Cornell University Press, 1993). For criticism, see Timothy and Lydia McGrew, "Level Connections in Epistemology," *American Philosophical Quarterly* 34 (1997): 85-94, and "What's Wrong with Epistemic Circularity," *Dialogue* (forthcoming), and Fumerton, *Metaepistemology and Skepticism*.

31. Strictly speaking, we should say "or better than random." The inference demands simply that the individual not be *less* likely to be representative than any other individual. But typically our evidence that a given individual is no less likely to be representative than any other is simply that it is a random member of the set and hence no more likely to be representative either.

32. Wisdom, *Foundations of Inference in Natural Science*, p. 216.

33. This point was noted by Williams, pp. 69, 149, though he unfortunately expounded it in a manner that did not sharply distinguish direct from inverse inference (see especially p. 149).

34. Wang, "Notes on the Justification of Induction," *Journal of Philosophy* 44 (1947): 701-10. The quotation appears on p. 703.

35. Ibid., pp. 705-6.

36. Nagel, p. 693.

37. Bertrand Russell, *Human Knowledge: Its Scope and Limits* (New York: Simon and Schuster, 1948), p. 402.

38. This point is made forcefully in Kyburg, "The Justification of Induction," *Journal of Philosophy* 53 (1956): 394-400.

39. Broad, *Induction, Probability, and Causation: Selected Papers* (Dordrecht: D. Reidel, 1968), pp. 7-8.

40. Wisdom, *Foundations of Inference in Natural Science*, pp. 218.

41. *De Interpretatione*, ch. 9.

42. See Steven Cahn, *Fate, Logic, and Time* (New Haven: Yale University Press, 1967) for a useful historical discussion and a defense of the view that fatalism is an inevitable consequence of the ordinary (temporally indifferent) formulation of the law of excluded middle.

43. L. Nathan Oaklander provides a careful and persuasive evaluation of the fatalism problem in *Temporal Relations and Temporal Becoming: A Defense of a Russellian Theory of Time* (Lanham, MD: University Press of America, 1984), pp. 195-220.

44. Ayer, *Probability and Evidence*, pp. 41-2. The reference here to the four-dimensional continuum is, of course, incompatible with the denial of future contingents.

45. Ibid., p. 42.

46. Ibid., p. 43.

47. Wisdom, *Foundations of Inference in Natural Science*, p. 217.

48. Ibid., p. 218.

49. Herodotus vii, 10. Quoted in Keynes, p. 307.