HUME, PROBABILITY AND INDUCTION.

MICHAEL ROWAN

Presented for the degree of Doctor of Philosophy,
Department of Philosophy,
UNIVERSITY OF ADELAIDE.

August 1985.

Awarded 28/1/86

CONTENTS.

ABSTRACT.

This thesis examines a number of responses to Hume´s sceptical critique of induction in order to provide a clear assessment of the current status of the attempt to provide a rational foundation for inductive inference.

Beginning with an examination of Hume´s original texts, intended to clarify his argument, the problem Hume posed is then defended as a genuine problem, and it is argued that the problem is not solved by Stove´s attempted disproof of Hume´s inductive scepticism. Several other possible lines of argument against Hume are mentioned, but on the basis of its dominance, and in the light of the strong criticisms of the alternatives given in the literature, the attempt to circumvent Hume´s inductive scepticism by basing inductive inference on probability is selected for extended analysis.

A brief analysis of the concept of probability identifies three paradigms of inductive (including statistical) inference for discussion: the testing paradigm (Neyman and Pearson); the confirmation paradigm (Fisher, Hacking, Carnap, and Kyburg); and the Bayesian paradigm (Jeffreys, Ramsey, de Finetti, Savage, Levi and Jaynes). The systems of inductive or statistical inference proposed by each author mentioned is subject to scrutiny, and it is concluded that none defeats Hume´s inductive scepticism.

The thesis concludes with a brief review of possible lines for future research on the problem.

I hereby certify that this thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and that, to the best of my knowledge, the thesis contains no material previously published, or written by another person, except where due reference is made in the text of the thesis.

If accepted for the degree of Doctor of Philosophy I consent to this thesis being made available for photocopying and for loan.

MICHAEL ROWAN.

FOR MY SON, KIM.

PREFACE

This thesis examines the dominant theories of inductive and statistical inference to see if any provide a solution to Hume's problem of induction. I conclude, regretfully, that none do.

While the thesis reaches the conclusion that Hume's problem remains unsolved, it is plain that the many authors studied herein have contributed much to our understanding of the problem. But, of course, we have no choice but to be thorough-going in our efforts to find the faults in even the most impressive philosophical work in this field, for it is important that we discover a satisfactory account of the rational foundations of scientific inference, so that we might be able to defend sound inductive practice against the attacks of the numerous irrationalities and dogmatisms ranged against it. I hope, therefore, that the reader will not mistake my attempt to be ceaselessly critical for lack of respect for the theories of induction criticised.

Apart from this acknowledged general debt to all whose work is considered in this thesis, there are three philosophers to whom I owe particular gratitude. Over the six years during which the thesis was written I found Salmon's carefully argued insistence that Hume's problem of induction is real and urgent a constant source of encouragement, which I valued no less than the particular arguments of his which, as duly noted below, I incorporated in my own view of the problem. During this long period of part-time study I also became aware of my enormous debt to Carnap and Hacking, both of whom I know to have influenced my perception of every facet of the attempt to find an answer to Hume, since no matter how frequently I cleared my desk Carnap's **Logical Foundations of Probability** and Hacking's **Logic of Statistical Inference** were always to be found upon it.

On a more personal level, I am pleased to record my gratitude to my supervisor, Michael Bradley, for his guidance and meticulous line by line criticism of my work. I must also thank the staff and students of the Department of Philosophy of the University of Adelaide for the helpful discussions of the various chapters when delivered to the Department's research seminar.

Despite being saved from very many mistakes by those whose help and inspiration has been acknowledged, the inference from errors having been found in the past in passages I originally thought to be unproblematic, to claim that further errors remain to be discovered in the thesis, is one induction which I am sure is correct. For all of the faults which remain I bear complete responsibility.

M R
Adelaide
July, 1985.

CHAPTER 1. AN EXAMINATION OF HUME´S PROBLEM OF INDUCTION

1. INTRODUCTION.

It is no unusual thing for philosophers to be deeply divided on some question. This characteristic of philosophy can be explained in a number of ways: philosophical problems are apt to go deep, dividing people according to their fundamental commitments beyond which there is no further appeal; according to one´s philosophical temperament the same thesis may appear in strongly contrasting lights, one philosopher´s scepticism being another´s critical rationalism; some problems in philosophy have such a venerable history and have become so familiar that they tend to become landmarks by which we map out intellectual territory, especially for the benefit of students new to the discipline, and thereby cease to be treated as live problems susceptible to solution.

Each of these factors no doubt contributes to an adequate explanation of the extraordinary diversity of learned opinion on Hume´s problem of induction, it being taken to be, in Kyburg´s felicitous phrase, everything from a solved problem to a pseudo-problem to a real but insoluble problem. But I think that the third of these factors is especially important in explaining the status of Hume's problem in contemporary philosophy, for the problem has become such a trusty piece of our intellectual history that it is only infrequently formulated in any detail. Typically it is tacitly assumed that everyone knows what the problem is and the debate then begins with the evaluation of rival solutions to the problem. Consider, for example, the collection of papers on the problem in Swinburne ed [1974]. Only once does any of these papers explicitly refer to Hume's writings, and the only detailed presentation of his problem is given in the Introduction to the volume by simply quoting Hume without comment. When work proceeds on the solution to a problem without close analysis of the problem itself it is not surprising that little agreement can be found on the success of the various purported solutions.

That the formulation of the problem has been paid, in comparision to the enormous effort that has been poured into finding a solution, relatively little attention requires some further explanation. At least three facts can plausibly be cited in this connection. First, philosophers have likely mistaken the force and simplicity of Hume's writing for precision and clarity in his argument. Second, once its main features are grasped, the problem Hume posed is so glaringly obvious that its general form is not readily forgotten, even if its fine details are not appreciated. Third, while Hume's problem has soaked up a great deal of effort many philosophers are of the view that no major victory has been scored against it, and this promotes the temptation to take as Hume's problem not the problem Hume actually posed but some related and refined problem on which present work proceeds.

We have then, three bases for being tempted to avoid detailed study of Hume´s actual texts. They ought to be resisted, for as Stove´s work has shown, while Hume´s style is straightforward his argument is complex, resting on significant and unstated assumptions. Moreover, in not tying Hume´s problem to the arguments he actually gave we run the risk of foisting upon him arguments which are in fact quite foreign to his analysis, thereby hiding from view the advances against his problem which have in fact been made. This is especially important in relation to the question of Hume´s supposed critique of inductive inferences founded upon the probability calculus.(1)

Thus it is desirable for a study of Hume´s problem of induction and the developments in inductive logic which have tried to deal with the problem, such as this thesis aims to provide, to begin with an examination of Hume´s texts to discover the actual argument he gave. That course will be followed here. We do not, however, need to begin our examination from scratch, since Stove´s study has provided us with a detailed and authoritative account. I accept the major features of Stove´s exegesis of Hume´s argument, and in my own discussion I will consider Hume´s own writings only in relation to settling points on which I disagree with Stove´s account.

2. HUME´S SCEPTICAL CRITIQUE OF INDUCTION.

a. The Main Features of Hume´s Argument.

I propose to accept without further discussion most of Stove´s claims about the main features of Hume´s argument on induction.(2) First, I agree with Stove that Hume´s argument, while it employs the term ´probability´, does not deal with probable inferences in any of our senses of that term. Thus Hume´s argument, as it is given in the passages for which he is famous as the scourge of inductive inference, does not provide any sceptical critique of inductive inferences based upon the probability calculus, contrary to the claims of Popper and others which Stove examines. Indeed, some while after this main discussion of induction Hume then turns to probability and what he says there, while it is rather muddled as I shall discuss below, does at least make plain that Hume did not consider the idea of probable inference familiar to recent philosophy. Second, I agree with Stove that Hume´s conclusion is intended to convey a sceptical appraisal of inductive inference (but I disagree with Stove´s analysis of that scepticism, as I shall explain below). Third, I accept that Hume´s argument proceeds in two sages, Hume first considering (what Stove calls) a **priori** inferences, inferences which proceed from the information about some object or phenomenon gained from the agent´s initial acquaintance with the object or phenomenon, and second turning to inferences based on a collection of experiences of the phenomenon or the behaviour of the object in question.(3) Fourth, I accept Stove´s claim that the second kind of inference which Hume launched a sceptical attack upon, that is the form of inference dealt with in the second stage of his argument, can be described as inferences from observed to unobserved instances of empirical predicates, that these inferences are what we would call ´predictive-inductive inferences´, and that Hume´s argument contains no element which restricts its scope to predictive-inductive as opposed to universal-inductive inferences. And finally I agree with Stove, and this is his most important point, that both of Hume´s arguments presuppose a suppressed premiss which gives Hume´s criterion for judging an inference to be justified - or as Stove says, wrongly in my view, for an inference to be reasonable - and thus constitutes the standard of reason upon which Hume´s sceptical critique of induction is based. Exactly what this standard of reason is, and consequently just what thesis represents Hume´s inductive scepticism, are my only significant disputes with Stove´s analysis of Hume, and these are the only points of Hume´s argument to which I will give any further attention.

b. **Hume´s Analysis of the a priori Inference.**

Let us consider first not inductive inferences proper, but the **a priori** inferences Hume discussed in the first stage of his argument. In the **Treatise** Hume dismisses these with the following argument:

> There is no object, which implies the existence of any other if we consider these objects in themselves, and never look beyond the ideas which we form of them. Such an inference wou´d amount to knowledge, and wou´d imply the absolute contradiction and impossibility of conceiving any thing different. But as all distinct ideas are separable, ´tis

evident that there can be no impossibility of that kind.(4)

The argument in the Abstract is very similar:

> The mind can always **conceive** any effect to follow from any cause, and indeed any event to follow upon another: whatever we **conceive** is possible, at least in a metaphysical sense: but wherever a demonstration takes place, the contrary is impossible, and implies a contradiction. There is no demonstration, therefore, for any conjunction of cause and effect.(5)

These passages require some examination. Note first that neither contains a sceptical appraisal of the **a priori** inference. This is given in the **Treatise** in these terms: ´´Tis therefore by EXPERIENCE only, that we can infer the existence of one object from that of another´; and the **Abstract** is similar. Hume´s conclusion concerning **a priori** inference is thus, as Stove points out, ostensibly a psychological thesis concerning the possibility of making a certain inference. This is understandable (rather than, as Stove claims, a confusion) given Hume´s conception of inference as the agent´s being propelled from belief in one proposition to belief in another under the pressure of reason or experience, but it is not acceptable to modern philosophy with its logical rather than psychological understanding of inference.(6) A plausible modern interpretation of Hume´s claim is easily found, however, namely that while it would be possible to make the inference it would be unreasonable to do so. This is Stove´s suggestion, and I accept it.

Employing this clarification it is then tempting to interpret Hume´s argument thus: Hume first shows that one can consistently accept the premises of an a **priori** inference and reject the conclusion, concludes from this, by a trivial step, that the inference is invalid, and then, presuming that all invalid inferences are unreasonable, concludes that **a priori** inferences are all unreasonable. This is Stove´s proposal.(7) It seems to me, however, that the introduction of the concept of invalidity is gratuitous, and likely to mislead. To explain this note that ´invalid´ has two aspects to its meaning, an informal aspect, according to which an argument is invalid if the truth of the premises leaves open the possibility that the conclusion is false, and a formal aspect, according to which an argument is invalid, strictly speaking we should say ´invalid in the logical system L´, if the logical form of the inference is not among the valid patterns of L. It is plain that Hume wanted only the first of these meanings. Indeed, he had little time for formal logic.(8) It seems wise, therefore, to avoid a possible confusion and take as Hume´s reason for thinking the **a priori** inference to be unreasonable to be that if an a **priori** inference is not provably truth-preserving then it is unreasonable, or, since Hume clearly did not consider the possibility that the **premises** of such an inference might be false and thus simply assumed their truth, that if an **a priori** inference does not prove its conclusion to be true then it is unreasonable. This is all Hume´s argument requires for it to be valid.

Note that I have suggested that we restrict the scope of Hume´s suppressed premiss to **a priori** inferences rather than adopt, as Stove suggests, an unrestricted claim, which on our account would be ´All inferences which do not prove their conclusions to be true are unreasonable´. I do not believe that Hume held this claim, and indeed we shall see that he did not think that inductive inferences are unreasonable even though he showed them to fail to prove the truth of their conclusions. Thus, I shall argue, there is a significant difference between the two stages of Hume's argument. Before I turn to this, however, let us examine the discussion of a priori inference in the **Enquiry**.

In the **Enquiry** Hume gives much greater attention to the **a priori** inference than he had in his earlier analyses, and he makes no mention of the inference being not provably truth-preserving. In the **Enquiry**, therefore, Hume clearly launches a fresh attack on the **a priori** inference, taking as the basis of his criticisms that the premises of the a **priori** inference offer no one of an arbitrary range of non-equivalent claims any more support than they offer any other. This is plain in a number of passages, of which his discussion of various possible conclusions for an a **priori** inference concerning collisions between billiard balls is the clearest:

> All these suppositions are consistent and conceivable. Why then should we give preference to one, which is not more consistent or conceivable than the rest? All our reasonings **a priori** will never be able to show us any foundation for this preference.(9)

As Hume might have put it, we are unable to make the a **priori** inference because, knowing nothing of the object under consideration other than what we can tell from its external appearance, we are not propelled to any particular conclusion concerning it. We shall interpret this to mean that to make an a **priori** inference is unreasonable, because the premises of the inference offer no stronger support to one claim than they do to any other non-equivalent possible conclusion. Thus in his discussion of the **a priori** inference in the **Enquiry** Hume comes close, even if only in strongly psychologistic terms, to entertaining the idea of support short of provable truth-preservation. But even here he does not get to the point of admitting the idea of degrees of support, for he assumes as his standard of reason that if one possible conclusion (of an **a priori** inference) is **no more** strongly supported than others then it **would be unreasonable** to make that inference, and this does not entail, of course, the thesis that if one possible conclusion **was** more strongly supported then it **would be reasonable** to make the inference.

This completes our analysis of Hume's critique of **a priori** inference. We turn now to his discussions of inductive inference, where we shall see, as already asserted, that Hume does not again come near to considering support in any other terms than provable truth-preservation, and that his conclusion concerning inductive inferences is not the same as his conclusion concerning **a priori** inferences.

c. Hume's Analysis of Inductive Inference.

While Hume's discussion of inductive inference is given in psychological terms it is possible to distinguish the various steps in his argument and recast them in philosophical terms. First he examines what in fact leads us to make inductive inferences, claiming that it is our past experience of a constant conjunction of a certain cause and effect which makes us, in a further instance of observing the cause, expect that the effect will also occur. Second, from this he concludes that if we form this expectation under the pressure of reason, ie. if our expectation is arrived at by inductive inference from our past experience, we must have assumed that what has happened in the past, viz the observed constant conjunction of cause and effect, will hold good in the future. Third, he then goes on to seek the rational foundation of this assumption, argues that it has no rational foundation, and thus, fourth and finally, concludes that the expectation is not formed under the pressure of reason, since reason does not support the inference, but rather is prompted by habit. The first two of these steps are clear in following passage:

> Since it appears, that the transition from an impression present to the memory or senses to the idea of an object which we call cause or effect, is founded on past **experience**, and on our remembrance of their **constant conjunction**, the next question is, Whether experience produces the idea by means of the understanding or the imagination; whether we are determined by reason to make the transition, or by a certain association and relation of perceptions. If reason determin´d us, it wou´d proceed upon that principle, **that instances, of which we have had no experience, must resemble those, of which we have had experience, and that the course of nature continues always uniformly the same.** (10)

What is of prime interest to us here is Hume´s reason for asserting that if the inductive inference is supported by reason then it must ´proceed upon the principle´ of the uniformity of nature. Here I agree with only one thing Stove claims, namely that when Hume says that the inference proceeds upon the principle of the uniformity of nature he means that the principle must be numbered among the premises of the inductive inference if the inference is to be supported by reason.(11) Before I can pursue this disagreement with Stove, however, it is necessary to gather together the evidence upon which the question must be settled. Why, then, did Hume think that if an inductive inference was to be supported by reason it must have the principle of the uniformity of nature among its premises?

We learn little from the remaining argument in the **Treatise**, for here Hume is content, after arguing that the uniformity thesis is not demonstrable and thus if it is to be supported by argument it must be supported by an inductive inference, to assert that inductive inference, or, to use Hume´s term, ´probability´, ´is founded on the presumption... [of the uniformity thesis - MR]... and therefore ´tis impossible this presumption can arise from probability´. He then concludes his discussion with the claim that in forming expectations of the unobserved, ie. in accepting the conclusions of inductive inferences, the mind ´is not determin´d by reason´. We are none the wiser, therefore, why inductive inference requires the uniformity thesis as a premiss if in making the inference the mind is to be ´determin´d by reason´, ie. if the inference is to be supported by reason.

The relevant passages in the **Abstract** and the **Enquiry** offer no further guidance on the point, merely asserting, for example, that ´all inferences from experience suppose, as their foundaton, that the future will resemble the past´, and that such inferences are ´founded upon´ this supposition. But we ought not conclude from this, however, that Hume has suppressed an important thesis. It might equally be that the thesis required as a premiss of his argument, viz his reason for holding that inductive inferences are supported by reason only if the uniformity thesis is among their premises, is trivial. Certainly there is an obvious candidate, for plainly Hume thought that in making an inductive inference we assert that the conclusion **will turn out to be true**, but it will only turn out to be true if the course of nature remains uniform, for it is on that expectation that we have come to our conclusion. Thus our inference, since it will will only be turn out to preserve the (assumed) truth of its premises if nature is uniform, can only be supported by reason if we are entitled to make that assumption, for only if the assumption is true will the conclusion turn out to be true, and in making the inference we have made that claim. The missing premiss is thus ´Only those inferences which prove the truth of their conclusions are supported by reason´.

We have good evidence for this account both in its making sense of Hume´s discussions of induction, and in its conformity with his earlier analysis of **a priori** inference. But the premiss we have attributed to Hume in his discussion of induction is not exactly the same as that which we claimed was employed in his critique of **a priori** inference, for we have not placed any restriction upon the kinds of inferences to which the principle of reason upon which the critique of induction is based applies. The reason for this will become apparent with further examination of Hume´s texts, for this will allow us to give a more definite meaning to the phrase we has so far adopted to express his sceptical judgement of inductive inference, namely that such inferences are ´not supported by reason´. Before we consider any further the question of the meaning of

Hume´s sceptical conclusion concerning induction which is embodied in the passages we have been considering, however, we should examine his discussion of probable inferences, to see if any light is thrown upon the question by Hume´s examination of these inferences.

d. **Hume on Probability and Induction.**

The argument considered above was directed at arguments from uniform experience, that is to say from the experience of a constant conjunction. Hume later considers inductive inference based upon non-uniform experience, such as a past history of only 19 of 20 ships sailing from a certain port returning safely. In his discussion of this second kind of experience he introduces a concept of probability in our sense of the term, whereby the probability of an event is its relative frequency in some actual reference class. This probability he further subdivides into the probability of chances, and the probability of causes, a distinction I shall not try to get clear. For all we need do is consider his argument on probabilities of each type, and note that he provides no thorough analysis of probable inference; indeed, his discussion is aimed mainly at reconciling his psychology of inference, which requires that inferences be forced upon us by the strength of association of ideas, with the fact of inferences proceeding from experiences which should give rise to conflicting associations, and he adds only the following new points to his earlier examination of the logical foundations induction from uniform experience.

Hume begins his discussion of probable inference with a critique of inference based on ´the probability of chances´. Here he dismisses the possibility that we can infer with certainty that the event backed by the ´superior number of chances´ will occur in some as yet unexamined case, arguing that if we were able to do so that would ´overthrow what we have establish´d concerning the opposition of chances´, ie. be inconsistent with our belief that we are dealing with a chance setup. He also rejects the claim that we can infer with certainty that the expected event is the more likely or probable, on the basis that this adds nothing to the premises since saying that an event is likely or probable is merely to affirm that it has on its side the greater number of equal chances, which is what the premises of the inference have already given us.(12) This later point is of considerable interest, for we can plausibly read it as a precursor of the modern criticism of the Bayesian account of inductive inference (viz, the view that induction merely establishes probabilities for hypotheses, not the hypotheses themselves). But the argument cannot be accepted as anything like the crushing critique of induction contained in the main part of his examination of inductive inference.

On the problems of inferences based on the ´probability of causes´ Hume is less clear, merely asserting that they are founded on the assumption that the course of nature will be uniform, but giving no explanation of why this must be the case. To be sure Hume´s own conception of induction based on probability requires, as he puts it, that we ´transfer the past to the future, the known to the unknown´, such that ´every past experiment has the same weight´, and that ´tis only a superior number of them, which can throw the balance on any side´.(13). Hume, that is, adopts as his main conception of probable inference, or inductive inference from non-uniform experience, that such inferences merely place on the unobserved the observed relative frequency of the various events considered, thus giving an unqualified prediction. He then harks back to his proof that we cannot prove that this uniformity will hold. But his critique is weak since he provides no analysis of probability or induction to back up his assumption that probable inductive inference must take this form.

Clearly Hume did not have any deep appreciation of the notion of probability. The concepts of probability as degree of belief, or as degree of inferential strength, were not considered by him in relation to his critique of inductive inference. Thus we must conclude that Hume´s analysis of probable inference provides no convincing basis for scepticism with respect to probable inductive inference.

e. **Hume´s Inductive Scepticism.**

Thus far I have left Hume´s sceptical conclusion concerning induction unexamined, using a phrase close to Hume´s actual words to express it by taking it to be that inductive inferences are not supported by reason. This vague phrase must now be clarified.

Stove takes Hume´s sceptical conclusion concerning induction to be that all inductive inferences are unreasonable. This is certainly one possible interpretation of the phrases Hume used, but it is not accurate, and it misses Hume´s point. Stove´s error is easy to make, however, because Hume does conclude that the a **priori** inference is unreasonable, and he believes that Hume´s judgement on the one inference is his same as that on the other. I shall defend these points in turn.

It is wrong to take Hume´s conclusion to be that inductive inferences are unreasonable, since this thesis entails that a reasonable man should refrain from such inferences, or that making an inductive inference constitutes some kind of error. But Hume made plain that this was not his view, commenting on his assessment of the relationship between premiss and conclusion of an inductive inference in these terms:

> I shall allow, if you please, that the one proposition may justly be inferred for the other: I know, in fact, that it is always inferred.(14)

This passage is found in the **Enquiry**, and it might therefore be thought that it is an attempt to make his youthful scepticism more palatable to an unreceptive audience. But while there is no passage known to me in the **Treatise**, nor in the **Abstract**, which is quite so clear, the point can be found in this earlier work. Consider, for example, the claim in the **Abstract** that

> ´Tis not, therefore, reason, which is the guide of life, but custom. That alone determines the mind, in all instances, to suppose the future conformable to the past.(15)

The point to note here is that Hume is not deriding inductive inference, as he would be by describing it as unreasonable, but making a claim about its foundation, a claim he described in the ensuing passage as a ´very curious discovery´. It is plain, therefore, that while this claim about the foundation, or mechanism - a term more suited to Hume´s conception of inference - of induction rests upon a prior appraisal of the standing of induction in reason, that prior appraisal cannot be such as to entail that inductive inference is faulty. Clearly we need to get that appraisal clear, for it is this that interests us, not the antiquated psychological conjecture which Hume took to be a discovery supported by his analysis of the foundation of induction in reason.

The appraisal we wish to get at is expressed in terms of which these are typical: ´´tis impossible for us to satisfy ourselves by our reason, why we shou´d extend that experience beyond those particular instances, which have fallen under our observation´; ´When the mind, therefore, passes from the idea... of one object to the idea... of another, it is not determin´d by reason´; ´there is no argument which determines me to suppose, that the effect will be conformable to past experience´; ´even after we have experience of the operations of cause and effect, our conclusions from that experience are **not** founded on reasoning´.(16) I think it is plain that by these phrases Hume meant that there is no rational justification for making inductive inferences, and not that such inferences are unreasonable, for the first interpretation leaves open the possibility of some other justification and does not entail that the inference is in any way improper, whereas the second entails that to make an inductive inference is to commit an error. The first is thus both a more accurate translation of Hume´s own words, and directs us to the point he was trying to make, which is best caught in this famous passage from the **Enquiry**:

> My practice, you say, refutes my doubts. But you mistake the purport of my question. As a agent, I am quite satisfied in the point; but as a philosopher, who has some share of curiosity, I will not say scepticism, I want to learn the foundation of this inference.(17)

Clearly we ought to take Hume´s sceptical conclusion concerning inductive inference to be that inductive inference has no foundation in reason, or more simply that no inductive inference is rationally justified. This contrasts with his judgement on the **a priori** inference, however, and it is instructive to clarify the difference.

Hume did conclude, as I have already accepted, that the **a priori** inference is unreasonable, for, unlike inductive inference, which he admits as a fact to be explained, Hume asserts that the **a priori** inference is not to be taken seriously, it is a vain pretension.(18) It is not hard to find an explanation of Hume´s differing assessments of **a priori** and inductive inference, for the **a priori** inference has no mental mechanism of association between ideas to act as a psychological foundation for the inference, such as Hume held inductive inference to have. That is why he completely dismisses **a priori** inference, and is also, no doubt, the basis of his fresh analysis of the a **priori** inference in the **Enquiry**, for this new examination aims to show the **impossibility**, and not merely the lack of foundation in reason, of the **a priori** inference. While Hume concluded, therefore, that the a priori inference is unreasonable, his conclusion concerning inductive inference was that such inferences are not justified by any rational foundation. This provides a clear account of the differences in Hume´s formulation of his assessment of the two kinds of inferences, and in particular allows for a satisfactory explanation of the originality of Hume´s analysis of the a **priori** inference in the **Enquiry**, which otherwise would require us to suppose to suppose that on this one occasion he tried, and only just failed, to admit the concept of degrees of support short of truth-preservation.

f. **Hume´s Conception of Reason.**

Before we conclude our discussion of Hume´s text we must be certain that we have got clear the conception of what it would take for an inductive inference to be justified which underlies Hume´s sceptical critique of induction.

I have suggested that a premiss of Hume´s argument on induction is that only inferences which prove the truth of their conclusions are justified by reason. This involves two claims, for it requires that the inference have a logical form which is provably truth-preserving, and that its premises all be proven truths. It is important to grasp the consequences of both of these aspects of Hume´s conception of reason.

The first aspect rules unjustified any inference form which is not provably truth-preserving. Thus the idea of justifiable probable inference is dismissed without consideration. The very idea of justifiable probable inference cannot arise on Hume's conception of reason.

The second aspect rules unjustifiable any inference which contains a premiss which itself requires support by an inductive inference, for a consequence of showing that inductive inferences are not provably truth-preserving it that any claim supported by such an inference is not a proven truth.

Both of these elements of Hume's conception of reason have been challenged by later authors, though the focus of attention has been the first. My own view is that the second ought to be abandoned, and indeed, in view of Popper's demonstration of the theory-ladenness of all observations I think we have no choice in this. As we shall see, however, jettisoning Hume's insistence that all of the premises of an inductive inference be independently justified, which conception I shall call (following Giere), 'foundationism', will not by itself solve the problem of induction. Such a change to Hume's conception of reason, moreover, raises the threat of solving the problem of induction at the price of sanctioning relativism. There would be little if any benefit to the defence of rationality in such an exchange. I shall not comment further in this thesis on the question of foundationism.

g. Hume´s Problem of Induction.

We have now set out Hume´s argument for concluding that induction is not justified by reason. Since on our account, if not on Hume´s, an inference which is not justified by reason cannot be defended as reasonable, this argument of Hume´s poses for us the problem of showing that there is some rationally compelling reason to assent to (correct) inductive inferences. The problem Hume left us, that is, is to defeat the sceptic´s assertion that there is no rational foundation for induction.

As we shall see, the approach most favoured by philosophers in this century is to argue that Hume´s conception of reason ought to be rejected in favour of a conception which admits rationally justifiable probable inferences. We shall be concerned to discover whether any of these attempts provides a rationally compelling form of inductive inference. Prior to this, however, we must consider the suggestion that the problem Hume posed is in fact a pseudo-problem, and also Stove´s claim to have proved Hume´s sceptical claim on induction to be a false proposition.

CHAPTER 2. HUME´S PROBLEM: SOLUBLE OR INSOLUBLE?; REAL OR PSEUDO?

1. INTRODUCTION.

Before we proceed to examine various attempts to find a solution to Hume´s problem of induction it is appropriate to pause to consider whether that effort is capable of bearing any fruit. We would be unwise to expend effort on trying to find a solution to a problem unless we were sure that the problem was a genuine problem, and that it was soluble, and a considerable body of philosophical opinion has it that Hume´s problem is either a pseudo-problem, or is insoluble, and even that it is a pseudo-problem because it is insoluble.

Such an analysis of Hume´s problem came to prominence in the middle part of this century, and has now receded a little under the impact of considerable criticism. Yet the matter has not been resolved; rather, from my experience of discussing the problem with a range of philosophers, the situation is that the disputants have settled into their various positions and the debate has gone cold rather than that the dispute has been ironed out to the satisfaction of the opposing camps. No doubt the reason for this truce is that both sides have in their own view proved their case to their satisfaction, but this has not satisfied the other side at least partly because the two opposing camps have different conceptions of the nature of philosophical problems and the basis on which they can be resolved. All those who attempt to get around Hume´s problem by declaring it a pseudo-problem I shall call ´dissolutionists´. Most of these philosophers subscribe to the ordinary-language view of philosophy, and these philosophers typically think that philosophical problems are properly **dissolved** by identifying the conceptual error in our talk that has lead us to think we have a problem; while the opponents of the attempt to dissolve Hume´s problem belong to other philosophical schools which take philosophical problems to be genuine problems requiring real solutions, such as are produced by abandoning some of our former ideas. Each has found the arguments which satisfy their opponents, therefore, to be unconvincing.

I propose to attack the issue afresh by considering six theses which are the basic materials from which various authors have, in differing ways, fashioned arguments purporting to show that Hume´s argument is deserving only of a dismissive solution.(1) I shall extract these theses from the various arguments as we briefly run through them. Once we have the six theses before us we can subject them to criticism and finally return to consider the impact on Hume´s problem of the theses which survive criticism.

2. THE SIX THESES OF THE DISSOLUTIONISTS.

In an early attack on the genuineness of Hume´s problem of induction Will pointed out that an inductive sceptic, in denying that induction gives us knowledge, must have in mind some conception of knowledge and some assumed standard which an inference must met if it is to yield knowledge (when so defined). He claimed that Hume´s assumed standard for knowledge yielding inference was validity, and that it was in consequence of his adopting this standard that he (Hume) was led to a sceptical appraisal of inductive inference. Will wrote:

> the only kind of knowledge [Hume - MR] recognised as genuine was either direct perceptual knowledge or demonstrative knowledge based on the necessity of avoiding contradiction. The laws of nature, and indeed all statements about the future, fall into neither of these classes. Therefore none of these things can be known.(2)

Let us slightly restate this point and list it as the first thesis for later discussion:(3)

1) Hume´s sceptical evaluation of induction rests upon applying to inductive inferences standards of appraisal appropriate to deductive inference.

Obviously this thesis forms the basis of an attack on the genuineness of Hume´s problem of induction only if it is further argued that the deductive standard is **inappropriate** to inductive inference. If both of these assertions could be proved then we would have shown that Hume´s sceptical appraisal of inductive inference rests upon a false assumption, namely that inductive inferences must be invalid to be justified, and would thus have dissolved, or given a negative solution to, Hume's problem of induction, by rebutting his argument. A positive solution to his problem would be provided if we could also show that on the standard which it is appropriate to apply to inductive inferences, at least some such inferences are rationally justified.

Will argues for both of these further claims, but in a rather muddled way. Among those who have done better is Edwards, who argues that we do not, in fact, employ deductive standards in our actual appraisals of inductive inferences. Rather, he claims, we accept inductive inference as a form of inference entirely separate from deduction and that this is shown by the language we ordinarily employ to evaluate inductive inferences. Indeed, he asserts that the question Hume asked about induction, as he (Edwards) understands Hume, namely 'Are there any inductive inductive inferences which are deductively justified, ie. valid?', is not a 'genuine' question, for:

> since part of the definition of 'inductive inference' is inference from the observed to the unobserved, it is a **contradiction** to say that an inference is both inductive and at the same time in the same respect deductively conclusive.(4)

This claim is, I think, false, but I do not which to enter into debate on that, nor on the virtues of Edwards' argument for his claims, at this point. That will come later when we turn our attention to the various theses we are now engaged in identifying as major claims against the genuineness of Hume´s problem of induction. Thus we rest content for the moment with extracting from Edwards´ discussion these two theses, which we add to the list begun above:

2) Since induction is not a species of deduction, but is rather a separate type of inference, it is incorrect to employ deductive standards to appraise inductive inferences.(5)

3) Since it is by definition impossible for an inductive inference to be valid, and thus impossible for inductive inference to meet the deductive standard of justification Hume employed to appraise inductive inferences, Hume´s problem is in principle insoluble and is thus a pseudo-problem.(6)

Edwards´ positive solution to Hume´s problem is less persuasive than his argument for its being considered a pseudo-problem. He seems to claim that simply because we speak of inductive inferences as though they (ie. those which are correct) are justified, it follows that they are justified - or perhaps only that we cannot coherently speak as though they were not, and thus cannot call them into question. The defects of this ´ordinary language´ solution of Hume´s problem - which aside from Edwards´ efforts, forms part of the arguments of Barker ([1965]:273), Black ([1949]:77), Pap ([1947]:193), Strawson ([1952]:256), and Will ([1952]:510f) - have been adequately dealt with, in my view, by Salmon ([1957] and [1965]) and Urmson ([1974]). I have nothing to add to Salmon´s and Urmson´s criticisms of the proposed ´ordinary language´ **solution** to Hume´s problem of induction, though I shall use their arguments in my analysis of the attempt by the dissolutionists to dismiss Hume´s problem as a pseudo-problem.

Part of Strawson's ordinary language dissolution of Hume's problem of induction was that while we can justify particular inductive inferences we cannot provide any wholesale justification of induction. For, he claimed, we justify particular inductive inferences by showing that they conform to our basic canons of correct inductive inference, and we cannot justify these canons themselves since there are no standards to which we might appeal for a justification.(7) This gives a further thesis.

4. We cannot give a justification of induction as a whole, for to do so we would need to justify the standards which we take to define correct inductive inference and this is not possible since there are no further standards to which we might appeal.

Other authors have been unhappy to leave the matter there and have sought to provide at least some rationalization for our having adopted the standard of inductive correctness we have adopted, turning for a justification to sources such as intuition (Carnap [1968]), critical reflection on what we can legitimately expect induction to achieve (Black [1949]:84 and Kyburg [1965]:275), or accepted inductive practice (Goodman [1955]:64).

These arguments rest upon the following thesis:

5. Particular inductive inferences are justified by reference to our basic canons of inductive inference which, being ultimate, cannot be justified in the same manner. Whatever justification these canons need can come only from intuition, or reflection on the purposes of induction and the outcomes of inductive inferences, or from the ultimate rules sanctioning only such inferences as we are prepared to accept.

Each of the authors who have taken the allied positions referred to in the preceding paragraph have defended their views by comparing the justification of ultimate rules of induction with the justification of ultimate rules of deduction, claiming that the two situations are analogous and thus, since we take deduction to be unproblematic (insofar as we do so take it) we ought to accord induction the same status. This has been argued most forcefully by Haack, and it forms the last of our six theses.

6. Since the fundamental rules of both deductive and inductive inference are logically ultimate (in their respective fields) the possibilities for justification are the same in the two cases; consequently, induction is no more (and no less) in need of special justification than is deduction; and thus, in the absence of such a justification, there is no better reason for inductive scepticism than there is for deductive scepticism.

I shall now discuss each thesis in turn, examining in detail the arguments which might be given in support of it. I shall take them in the order listed.

3. CRITICISM OF THE SIX THESES.

1) Hume´s sceptical appraisal of inductive inference rests upon applying to inductive inference standards of appraisal appropriate to deductive inference.

As it stands this thesis is rather vague. Let us clear it up by examining the arguments given in support of it. Unfortunately, they are generally rather brief and presumptive, often not quoting Hume as evidence for the claim or even referring to a specific part of his argument. Ambrose is an exception here, at least quoting Hume on the point, even if only two sentences - the first giving Hume´s reason for thinking that we cannot demonstrate that nature is uniform (viz that inductive inferences are such that the denial of the conclusion is consistent with the premises), the second recording Hume´s sceptical appraisal of inductive inference.(8) From this is appears that Hume took it, ie. assumed as a premiss of his argument, that an inference is unjustified if invalid. Thus it would appear that the following claim is true:

1´) A premiss of Hume´s argument for his inductive scepticism is that only valid inferences are justified.

But Ambrose´s quotation is extremely misleading. The two sentences come from a book of selections from Hume, and in that book are separated by fully eight pages. In the original the major part of Hume´s argument separates the two claims. Thus her quotation from Hume is poor evidence of (1´), and since that is her only evidence for (1), that too is not proved.

Black simply claims that 'Hume himself made it abundantly clear that his scepticism arose from the impossibility of a 'demonstration' or **deductive proof** of assertions concerning matters of fact'.(9) But this, as we saw in Chapter 1, is not the case. Indeed, if Black, like Ambrose, is asserting (1'), then his claim is false, as we shall shortly recall.

Edwards claims that Hume (in most places) meant by ''reason', a **logically conclusive** reason, and by 'evidence' **deductively conclusive** evidence, but he (Edwards) gives no source for this claim.(10) And, finally, while neither Pap nor Strawson refer to Hume by name they clearly hold that he assumed, ie. took as a premiss of his argument, that only valid inferences are justified.

None of our authors provide adequate grounds for accepting (1). Let us therefore recall our own discussion of Hume's argument given in the previous chapter. Does that analysis support (1)? Not if (1) was intended to be equivalent to (1'), for that is false, Hume assuming as a premiss of his argument not that only **valid** arguments are justified, but that only arguments which **show their conclusions to be true** are justified. The difference is important, for while it follows trivially from Hume's premiss, ie. from his standard of reason, that valid inferences are justified, it does not follow trivially from his standard that **only** valid inferences are justified; as it would, for example, if his standard of reason had been that only those inferences whose premises are jointly inconsistent with the denial of their conclusions are justified. Hume did not, therefore, adopt a standard of reason from which it follows trivially that inductive inferences are not justified, for even if it were obvious that inductive inferences are one and all invalid it would still not be obvious that inductive inferences cannot show their conclusions to be true. Rather this must be proved, and it was Hume who proved it. He did not assume it, nor did he adopt a standard of justified inference from which it follows trivially.

I do not, however, wish to dismiss (1) completely, for even though it is false there is something in it. For plainly Hume adopted a high standard of reason, and arguably this standard is higher than is needed to ensure that justified inferences yield only rational beliefs, and thus is unnecessarily high. But to argue for this case entails the acceptance of the claim that Hume posed us a genuine problem, namely to find an alternative to his standard of reason, and this is not the spirit in which (1) has been ventured. I therefore reject the dissolutionists' criticism of Hume based upon the defence of (1).

2) Since induction is not a form of deduction, but is rather a separate type of inference, it is incorrect to employ deductive standards to appraise inductive inferences.

This thesis cannot stand as valid criticism of Hume unless (1) is true. If we have been correct in claiming that (1) is false, then, the sting is taken out of (2). But let us put aside the question of the truth of (1) and examine (2) in its own right.

There are two points arising from (2) which we will need to clear up in the course of our discussion: what makes an inference of this or that logical type; and what makes it correct or incorrect to assess an inference by this or that standard? Bearing these questions in mind, let us turn to the arguments given in support of (2).

Strawson claims:

> Of course, inductive arguments are not deductively valid; if they were they would be deductive arguments. Inductive arguments must be assessed, for soundness, by inductive standards.(11)

Clearly this an argument for (2), and equally clearly it is not a good one. Strawson seems to argue that because inductive arguments are **invalid** they are **non-deductive**, and since they are non-deductive it is incorrect to assess them by deductive standards. But it does not follow from an inference´s being invalid that it is non-deductive, for an invalid inference might either be non-deductive or a form of invalid deduction. Nor, of course, does it follow directly from an inference´s being invalid that it is incorrect to assess it by deductive standards; indeed, we can only determine that an inference is invalid by assessing it according to deductive standards, since ´invalid´ **means** incorrect according to deductive standards. Strawson´s argument, therefore, does not prove (2), and his attempt to dissolve Hume´s problem by proving (1) and (2) therefore fails.

Black´s argument for (2) suffers from the same defect. He asserts that

> it is a distinguishing sign of all the arguments properly described as ´inductive´ that their premises are compatible with the logical negation of their conclusion. It is true **by definition** that induction is not deduction.(12)

But what follows from the claim that the premises of any inductive inference are compatible with the negation of its conclusion is not that it is impossible for inductive inference to be **deductive**, but rather that it is impossible for inductive inference to be **valid**. Black thus commits the same error as Strawson.

Clearly the invalidity of an argument is not by itself a sufficient reason for declaring it to be incorrect to assess an inference by deductive standards. But what might be an adequate reason for this claim? A reason compatible with the general line of the dissolutionists´ argument can be found in a passage from Edwards:

> what we mean when we claim that we have a reason for a prediction is that the past observations of this phenomenon or of analogical phenomena are of a certain kind: they are exclusively or predominantly positive, the number of positive instances is at least fairly large, and they come from extensively varied sets of circumstances.(13)

Now Edwards uses this passage to support his claim that when we say that the premises of an inductive inference give us a reason for the conclusion we do not mean that they give us a deductively conclusive reason, and thus that it is inappropriate to assess inductive inference by deductive standards. He does not argue his case explicitly, however, and thus we are left to guess at his train of thought. It could be that he is arguing in a similar manner to Strawson and Black, taking it to be obvious from the character of inductive inferences they are not valid, and then arguing that in virtue of this that what we might call ´inductive reasons´ are not deductive reasons, and thus should not be assessed by deductive standards. Certainly this line of argument would lead one to conclude, as Edwards does, that it is a **contradiction** to say that an inference is both inductive and valid, and thus that to ask whether inductive inferences are justified, when one takes validity as one´s standard of justification, is not a sensible question. But his point is susceptible to another interpretation which is free of the error we have isolated in the first - though, as I shall argue below, not entirely free of defects - namely that since it is not our **intention** that inductive inferences be valid, and thus justified according to deductive standards, it is not appropriate to assess them by deductive standards. Rather, so the argument would go, our inductive inferences must be assessed against the standard they are in fact intended to meet, which standard, on this interpretation of Edwards´ point, he outlined in the passage quoted.

Now whether this was Edwards´ intended argument or not, I think that it is the right way to go about defending (2), a view I shall explain before going on to examine whether it provides us with an adequate argument for (2).

I have stated (2) in the form that I have since in that form it represents the way Strawson, Black and (if his point is the first of the two alternatives set out above) Edwards have sought to dissolve Hume's problem of induction. But this is not the right way to express the thought behind (2), since as it stands (2) commits a serious error in making it seem as though inferences are **in themselves** of a certain type and that it is being of this or that type that makes it appropriate to assess an inference by the corresponding standard and inappropriate to assess it by another. This view is incorrect. Rather, when considered in itself, ie. as a mere logical structure, an inference does not have a type, eg. is neither deductive or non-deductive. Rather we assign a type to an inference by applying to it some standard which we think appropriate for the inference under examination, and thus it is the **standard actually applied** that determines the type of the inference. Any inference can be assessed by any standard and thus assigned any type, though it will seem natural to assess an inference by a standard that it in fact meets, or by a standard it fails to meet if in so failing it commits a recognised fallacy. But while either of these circumstances make it natural to assign to the inference the type in fact assigned, they do not make it correct to assign the inference that type.

What does make it correct to assign an inference a particular type, ie. to assess an inference by a particular standard, is the intentions of the author in putting the inference forward (and thus, depending on the intentions of the author, a single logical form of inference could be assigned different types). For when we assess or evaluate an inference we are assessing whether the inference was successful in what it was trying to do, and what an inference was trying to do is determined by what its author was trying to do with it. If an author puts forward an inference intending that if the premises are true then there should be no logical possibility that the conclusion is false, then we rightly assess the inference by deductive standards no matter how unlike a valid deduction the inference in fact is.

Clearly, then, it is the application of certain standards that determines the type of an inference, not the type of the inference which determines the standard appropriately assigned to it. And which standard is appropriately applied to a given inference depends upon the role which its author assigned it in his argument.

When it is said, then, as (2) says, that a certain standard is inappropriately applied to an inference in virtue of its type, the relationship between an inference's being of a certain type, and its being assessed by a certain standard, is got round the wrong way. Once this confusion is brought to light, (2) must be rejected.

The thought behind (2) must be differently expressed if it is to give rise to a claim which may be able to be proved. As a start we might offer the following (which is what our second interpretation of Edwards, given above, had him say):

2′) Since it is not our intention that inductive inferences be valid (ie, correct on deductive standards) but rather that they be correct on inductive standards, it is inappropriate to assess inductive inference by deductive standards.

This will not quite do, however, since we do not choose the standard of inductive correctness arbitrarily. Rather our choice is motivated by the belief that the chosen standard is high enough to ensure that if an inductive inference meets the standard then its premises justify rational belief in the conclusion.(13) It is that the (ie. some specified) standard of inductive correctness meets this requirement which makes it the appropriate standard by which to distinguish between correct and incorrect inductive inferences.

The error in Edwards´ account (on the stronger interpretation of it) is that he neglected this point. For it is not our intention in putting forward an inductive inference to merely meet the standard he specified; rather we intend to justify belief in the conclusion of our inference, and are only concerned to meet Edwards´ standard (if that is our standard) because we believe that an inference meeting that standard does justify belief in the conclusion of the inference.

Assume that we have identified the standard at which we aim in putting forward an inductive inference, when we wish such an inference to justify belief in its conclusion. Call this standard ´the inductive standard´; my point is that (2´) should be amended to read:

2´´) Since it is our intention that inductive inferences should justify belief in their conclusions, and since they need not be valid but only correct on the inductive standard in order to do so, it is inappropriate to assess inductive inferences by deductive standards.

Now if (1) and (2´´) could be proved, then obviously we would have dissolved Hume's problem of induction. But quite apart from the problems we have identified with (1), this will be no easy matter if possible at all. For to prove (2'') we need to specify a standard for correct inference which can (at least in principle) be met by inferences from the observed to the unobserved and which is such that an inference meeting it would justify belief in its conclusion. But such a proof would **justify** inductive inference.

Of course the kind of justification of inductive inference which is required to prove (2'') refutes Hume's inductive scepticism not by accepting his conception of what would count as a justified inductive inference, and then finding a counter-example to his claim that there is no such inference, but by **proving** that he misconceived the nature of justified inductive inference. But the difference between the two approaches to Hume's problem is merely that they attack different premises of his argument. The former takes issue with Hume's claim that no inductive inference can be shown to be truth-preserving, the latter with his claim that an inference has to be shown to be truth-preserving to justify belief in its conclusion. If either of these attacks on Hume's argument, supposing they were to be successful, deserved to be called a 'solution' to Hume's problem of induction then they both would, and thus if one managed to prove either charge against Hume one would have solved his problem of induction.

Now since the dissolution of Hume's problem of induction requires the proof of (2''), and since the proof of (2'') would be a solution to Hume's problem of induction, it follows that we cannot prove (2´´) without solving Hume´s problem of induction, and thus that we cannot **dissolve** Hume´s problem without first **solving** it. Therefore those who have claimed, like Edwards, Black, and Strawson, have done, to have dissolved Hume´s problem of induction, must, if they have been successful, have either given or assumed a solution to his problem. If they have not solved it then they have not dissolved it either, while if they have solved it then their dissolution loses most of its interest.

3) Since it is by definition impossible for an inductive inference to be valid, and thus impossible for an inductive inference to meet Hume´s standard of justification, Hume´s problem is in principle insoluble and is thus a pseudo-problem.

Let us not object to (3) on the grounds that Hume did not, in fact, assume that an inference needs to be valid to be justified, for while we are entitled to reject (3) on this ground alone, it has other problems which are worthy of examination in their own right.

First, let me comment on the definition of ´inductive inference´ embodied in (3). While we are, of course, free to define our terms as we see fit, it is undesirable to include in the definition of ´inductive inference´ that inductive inferences are invalid, for it is not obvious that they are so. Rather it must be proved that inductive inferences, that is inferences from observed to unobserved instances of empirical predicates, are invalid, and it was Hume who proved it in consequence of proving that inductive inferences are not provably truth preserving. It would be unjust to Hume to incorporate what he proved of inductive inference into the definition thereof, and then chide him for proving what is obvious, indeed true by definition. It would be even worse to proceed, as (3) does, to denounce Hume for drawing a sceptical conclusion from his proof of the invalidity of inductive inference on the grounds that no scepticism is warranted by induction´s being what it is and not what it is not.(15)

Moreover, there would be little point in defining ´inductive inference´ in a way that made it true by definition that induction could not meet the standard for justified inference Hume laid down, for while on such an amended definition it would be trivially true that inductive inference cannot meet Hume´s standard and thus also trivial that if we hold to Hume´s standard then we must give a sceptical appraisal of inductive inference, this would not trivialize Hume´s inductive scepticism. Only Hume´s argument would be trivialized, not his conclusion, since it would remain a serious matter that inductive inference does not meet the standard he laid down for justified inference unless we can show, which we cannot meaningfully do by playing with definitions, that Hume´s standard for justified inference is inappropriate to inferences intended to legitimate belief in their conclusions.

If we want to defeat Hume´s inductive scepticism we can do so only by showing either that inductive inferences can meet his standard for justified inference, or that they need not do so to justify belief in their conclusions, or that even though they do not justify belief in their conclusions they can do something else which is sufficient to justify a non-sceptical appraisal of them. Adopting a definition of ´inductive inference´ which made it true by definition that inductive inferences could not meet Hume´s standard would not defeat his inductive scepticism, rather it would close off the first of these options by fiat, leaving open only the second and third possibilities.

Aside from the problem of the definition of ´inductive inference´ underlying (3), there are two further problems with it. First, it does not follow from the incapacity of inductive inference to meet Hume´s standard of justified inference (whether this is a matter or definition or proof) that his problem is insoluble, and second, it does not follow from his problem being taken to be insoluble (even insoluble in principle, or necessarily, or by definition impossible) that it can be dismissed as a pseudo-problem. I shall defend these claims in turn.

We have already said all that is necessary to prove the first of these claims. Let us briefly recapitulate. Hume´s problem is this: there is a certain standard which inferences must meet if they are to be able to be considered rational (this claim and his specified standard being given as his first premiss); no inductive argument can meet this standard (his second premiss); therefore no inductive inference is rationally justified. Now it would be logically possible to attack the sceptical claim which is Hume´s conclusion in one of four ways: one could prove it false directly; or one could undermine it by proving the falsity of Hume´s first, or his second premiss, or one could show that his argument is invalid. Hence a proof that just one, or indeed any three, of these possibilities for solution cannot be carried through would leave some possibility still open, and thus fail to show the insolubility of Hume´s problem. Certainly, for example, a proof that Hume´s second premiss is unassailable, or a definition which makes the premiss trivially true, is far from a proof that Hume´s conclusion cannot be refuted and thus far from a proof that his problem cannot be solved. Thus those who argue, like Edwards and Strawson, for the triviality of Hume´s problem as a consequence of his second premiss being true by definition, and those who offer, like Black, a proof of Hume´s second premiss, and then take this to show Hume´s problem to be insoluble, fail to appreciate that there are other possible avenues of solution which remain open despite their arguments. Such analyses fail to advance the discussion of Hume´s problem by perpetuating the misconception that the solution of Hume´s problem requires finding a counter-example to his second premiss and thus to his conclusion.(16)

It might be said in these authors´ defence that they **inherited** the idea that the problem Hume bequeathed to us is to find a counter-example to his second premiss, since this is found in the earlier tradition that tried to solve Hume´s problem in this manner (by defending, for example, the uniformity thesis). Certainly they did inherit this restricted conception of Hume´s problem, but they also perpetuated it by accepting the restricted view as part of their attempt to show that the problem is insoluble and thus a pseudo-problem.

Much the same can be said of Goodman´s attempt to ´dissolve the old problem´ of induction by showing that it is impossible to find a solution.

Goodman writes

> If the problem is to explain how we know that certain predictions will turn out to be correct, the sufficient answer is that we don´t know any such thing. If the problem is to **find** some way of distinguishing antecedently between true and false predictions, we are asking for prevision rather than philosophical explanation.... Now obviously the genuine problem cannot be one of attaining unobtainable knowledge or of accounting for knowledge which we do not in fact have.(17)

But unfortunately for us the genuine problem can be one of ´attaining unobtainable knowledge or of accounting for knowledge we do not in fact have´ if, as Hume believed, we need the kind of knowledge he took himself to have proved unobtainable (by rationally justified inference), to justify rational belief regarding the unobserved. His problem therefore remains, and it is to show either that we do not require rational belief in the unobserved in order to avoid scepticism, or that we do not require (what Goodman accepts Hume showed to be) unattainable knowledge to support rationally justified inferences from the observed to the unobserved. Nothing Goodman says shows that either line of attack on Hume´s problem, either way of attempting to refute his sceptical conclusion concerning inductive inference, will be unavailing. Thus Goodman has not shown that Hume´s problem is insoluble.

But even if we were inclined to accept that it was shown that there was no possible solution to Hume´s problem, that it is by definition, or in principle, or necessarily insoluble, it would not follow immediately that the problem could be dismissed as a pseudo-problem. The entire dissolutionist tradition has made this assumption, but Ayer puts the point flatly:

> [I]t appears that there is no possible way of solving the problem of induction, as it is ordinarily conceived. And this means that it is a fictitious problem, since all genuine problems are at least theoretically capable of being solved.(18)

While this claim is seemingly reasonable it is actually problematic. This is, I think, best brought out by noting that Ayer ought to have restricted himself to claiming that Hume´s problem **appears** to be a pseudo-problem, since he did not claim that there **is** no way of solving it, but only that this is the way it **appears.** How then could we **know** that a problem was a pseudo-problem? Only by knowing that it was incapable of solution. And how could we know this to be the case? ´When we have a proof that it is so´ is the obvious answer. But now we can ask ´How do we know when we have such a proof´?, and plainly the answer ´When we have a valid argument from true premises to this conclusion´ will not end the questioning, for the validity of the argument is always open to challenge even though we may not doubt it, and, likewise, even seemingly secure premises may be called into question. Clearly, if there was a reason to think that a problem was real, say, that it kept cropping up, it would be impossible to prove **conclusively** that it was a pseudo-problem by a proof which purports to show it to be insoluble, for every time we are led to form the problem again we will be lead to question the purported proof of its insolubility.

The way to show that a problem is a pseudo-problem is not to show that it is insoluble, but to show that it need not arise, to show how we can achieve our desired ends and yet avoid posing the problem. While we do not know how to do this we will rightly doubt any proposed proof of the claimed insolubility of the problem, for while we have a reason for posing the problem we have a reason for thinking it to be genuine. If the dissolutionists would save us from needless pains in wrestling with Hume´s problem of induction, therefore, they must show us how to stop having the problem, not try to get us to accept it as insoluble and thus as something which there is no point in worrying about. We don´t want to become well-adjusted to living with Hume´s problem, we want to solve it, and we need to solve it because we keep coming up against it when we try to provide a rational basis for the supposed epistemological primacy of science.

Hume´s problem will remain real, and urgent, until we can show either that inductions can meet his standard of reason or that they need not do so in order to support the kinds of claims concerning the unobserved which it is necessary for science to make. Thus (3) ought to be rejected, since it entirely misrepresents the situation. And with it ought to go the arguments of Black, Edwards, Goodman and Strawson which have sought to lay to rest Hume´s problem and get on with other things, since each of these rests upon (3).

4) We cannot justify induction as a whole, for to do so we would need to justify the standards which we take to define correct inductive inference and this is not possible since there are no further standards to which we might appeal.

The most famous argument for (4) is Strawson's. I shall briefly review it and offer criticisms based on points raised by Salmon and Urmson.

Strawson argued that our basic rules of inductive inference are basic in the sense that they give content to our notions of **reason** and **justified belief** (concerning matters of fact). Thus, he claimed, it makes no sense to seek a justification of the fundamental rules of inductive inference, asking, for example, whether it is reasonable to employ such rules, since employing such rules 'is what 'being reasonable' means in such a context'.(19)

He went on to make a further point, drawing an analogy between inductive inference and the law, arguing that while we can call into question and seek justification for various laws, 'it makes no sense to inquire in general whether the law of the land, the legal system as a whole, is or is not legal'. Strawson's reason for this claim was that while in seeking a justification for particular laws we can appeal to more basic legal principles, and thus make sense of the demand for justification, we cannot appeal beyond the last principle, and thus we cannot appeal beyond the law as a whole for a wholesale justification of it. We can therefore make no sense, Strawson concludes, of the demand for a wholesale justification for the law since there is no way of satisfying this demand.

In brief, then, Strawson considers the fundamental rule (or rules) of inductive inference to be ultimate in two senses: they define the meanings of the terms we use to assess inductive inferences ('correct', 'justified', reasonable'); and they are the foundation of a system of reasoning. He thinks that it follows from the first point that we cannot coherently question the correctness, justifiability, or reasonableness of these rules; and from the second that even if we could pose such a question we could not answer it, ie. that we could not provide a justification for the fundamental rules, since we would have no basis for justification.

Both of these arguments are fallacious. In the first case, and this is a point made by both Salmon and Urmson, 'correct', 'justified' and 'reasonable' are not merely **descriptive,** as they would be if their meanings were entirely given by fundamental rules of inductive inference, they are also **normative or evaluative.**(20) If we did define such terms by reference to fundamental rules of inductive inference - by saying, for example, that an inductive inference is to be called 'reasonable' if it is governed by rule **x** - then it would be no **justification** of that inference to show that it is reasonable in this **technical** sense of the term. For we could then ask, with clear sense, whether an inference reasonable in this technical sense was reasonable in the usual sense of the term - ie. the question of justification would be left open after a proof of the (technical or formal) reasonableness of the inference.

In the second case it does not follow that a rule which is ultimate in some system is incapable of justification. Let us pursue Strawson's legal analogy a little further to bring the point out.(21) Certainly we cannot hope to answer the question Strawson posed as analogous to the question of the wholesale justification of induction, viz, ´Is the law legal?´. But we can ask ´Is the law just?´, or ´Ought one to obey the laws of this land?´, provided, of course, one has some independent conception of justice and moral duty. And similarly, while we could not hope to answer the question ´Is our basic standard of inductive correctness correct?´ we might hope to find an answer to ´Is our basic standard of inductive correctness reasonable?´. Since it is this second question which Hume, in effect, posed, seeking what Strawson calls ´a wholesale justification of induction´, the legal analogy actually supports the legitimacy of Hume´s question rather than, as Strawson asserted, exposing it as non-sensical.

We cannot dismiss Strawson´s point entirely, however, for while he failed to show that we cannot sensibly seek a wholesale justification of induction in more basic principles of reasonable belief or rationality, his discussion does forcefully raise the question of just what these principles might be. That is a matter we must, however, leave till we examine the attempts to offer positive solutions to Hume´s problem.

5) Particular inductive inferences are justified by reference to our basic canons of inductive inference which, being ultimate, cannot be justified in the same manner. Whatever justification these canons need can come only from intuition, reflection on the purposes induction serves and the outcomes of inductive inferences, or from the ultimate rule of induction sanctioning only such inferences as we are prepared to accept.

By itself (5) need not constitute an attack on the genuineness of Hume´s problem of induction, but it does so when tied to (6) for we are then invited to conclude that induction and deduction are of a piece when it comes to justification, and, since there is no problem of deduction which is equivalent to Hume´s, that the latter problem is only a morbid fancy of some misguided philosophers of science. Since we shall reject (6), therefore, we need not consider (5) in any detail, except for the version of it defended by Goodman. This one of the listed variations on (5), that inferences are justified by the property of the system that the fundamental rules sanction only acceptable inferences, does constitute an attack on the genuineness of Hume´s problem by virtue of the thinness of the justification offered.

That Goodman´s proffered solution to Hume´s problem is dismissive of it is plain enough from the following passage, which deserves to quoted in full:

> The point is that rules and particular inferences alike are justified by being brought into agreement with each other. **A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.** The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either.
>
> All this applies equally well to induction. An inductive inference, too, is justified by conformity to general rules, and a general rule by conformity to accepted inductive inferences. Predictions are justified if they conform to valid canons of induction; and the canons are valid if they accurately codify accepted inductive practice.(22)

Now it must be admitted that we do reject a rule if it sanctions an inference we are unwilling to accept, and that we reject an inference if it violates a rule we are unwilling to amend, but this **justifies** neither rule nor inference. Rather, recognition of this fact is a step in explicating our system of (inductive) inference. What **justifies** the rule, and thus the sanctioned inferences, is our **reason** for being unwilling to accept an inference or amend a rule. If we offer as our reason for accepting a rule that it justifies acceptable inferences, and then seek to justify our accepting the inferences by citing the rule, then clearly our justification for the system of rule plus inferences is flagrantly circular. Such a justification would plainly be inadequate, and thus it is clear that in offering it Goodman is simply dismissing Hume´s problem without good reason.

Now, of course, dismissing Hume´s problem, or rather the need for our taking any special pains to deal with Hume´s problem, would be unobjectionable if it could be shown that exactly the same situation **qua** justification arises in other systems of inference, but in these other systems we (properly) we do not feel ourselves to be in any special difficulty, nor search for any justification of ultimate rules. Goodman argues this case, as have a number of other authors. We shall turn our attention to this.

6) Since the fundamental rules of both deductive and inductive inference are logically ultimate (in their respective fields), the possibilities for justification are the same in both cases; consequently, induction is no more (and no less) problematic and in need of special justification than is deduction, and thus, in the absence of justification, there is no better reason for inductive scepticism than there would be for deductive scepticism.

I have already mentioned Goodman as holding (6). He is not alone, the point being adopted by Carnap, Strawson, and Kyburg, among others.

Recently Haack has carefully defended the thesis. All of these arguments share the same fatal defect, namely that it does not follow from the possibilities for the justification of deduction and induction being identical, that induction is no more in need of justification than deduction - or so I shall argue.

Carnap developed his argument on the point as a step in his defence of inductive logic. In his [1968] he employed what he called ´inductive intuition´ in the defence of certain axioms for inductive logic, and aware that this reliance on intuition would be controversial, he sought to head off objections by arguing that we must place a similar reliance on intuition if we wish to justify the axioms of deductive logic. Essentially his point was that since any argument for the fundamental rule of deductive inference, or proof the the virtues of this rule, would be question-begging, the rule must be assumed, with intuition called upon to defend the assumption directly. Thus, after considering how one might convince another to adopt the fundamental rule of deductive inference, Carnap concluded:

> the essential point is this: you have to appeal to his deductive intuition, in order to teach him deduction. Thus we see that the epistemological situation in inductive logic ... is not worse than that in deductive logic, but is quite analogous to it.(23)

Haack comes to a similar position, after a more thorough analysis of the impossibility (accepting, for the sake of the argument, that it is such) of proving the reasonableness of deductive inference. She wrote:

> The moral of this paper might be put, pessimistically, as that deduction is no less in need of justification than induction: or optimistically, as that induction is no more in need of justification than deduction. But however we put it, the presumption, that induction is shaky but deduction is firm, is impugned.(24)

But Haack's argument falls far short of proving this - which is what needs to be proved if seeking a solution to Hume's problem of induction as a matter of some urgency is to be derided. For if all that Haack musters in defence of her conclusion is true, it shows no more than that it is impossible to prove that valid deductive inferences will not fail to be truth-preserving. But this, by itself, does not constitute a problem of deduction analogous to Hume's problem of induction. For, in addition to proving that we cannot show inductive inferences to be truth-preserving, Hume showed further that it is possible for an inductive inference, which we would judge to be correct, not to be truth-preserving. And we can cite examples of very well based inductions which in fact turn out not to be truth-preserving.

If we grant both Hume's and Haack's points, then, we have the following situation. Inductive inference is such that:

a) it cannot be shown to be invariably truth-preserving;

b) it is possible that an inference judged correct will fail;

c) actual inferences can be constructed which, while we would judge them to be correct, in fact fail.

Only (a), however, is shown by Haack to apply to deductive inference. The two situations are not, therefore, analogous. And the differences are crucial.

In virtue of (a), we have it that it is not possible to show that there can be no counter-example to the thesis that correct inferences (both deductive and inductive) will be truth-preserving. But while this might cause us some unease, it does not constitute a problem by itself. It merely shows that the ground upon which we stand when making (deductive or inductive) inferences cannot be **guaranteed** solid, but it gives us no reason to **doubt** that it is solid, and thus need initiate no investigation into the matter.

A proof of (b), however, which we grant that Hume gave for inductive inference, but which has not been given for deductive inference - discounting the claims of the so-called ´deviant logics´, which we will discuss below - does raise a problem. For it shows that not only is it not possible to prove that the ground on which we stand when making a deductive inference is solid, but, additionally, that it is possible that it **will** give way. Now we do have a problem, for we now have a **reason** to seek assurance that the ground will not give way beneath us, ie. that an inductive inference which we judge correct will not fail to be truth-preserving. But we must stress that in the case of deductive inference we have no reason to seek such assurance, and thus no need to be perturbed by a proof of (a) in respect of deductive inference. The absence of a guarantee is only a cause for concern if there is some reason to need one, and we have been shown no such need in the case of deductive inference.

Finally, a proof of (c) shows that for inductive inference the ground does in fact occasionally give way under the agent's feet. And now, of course, we are faced with a real and urgent problem, namely to find (to paraphrase Hume), some logic, some argument, to secure us against the supposition that it is we who will be led, by induction, into the adoption of a false, and commonly, for that reason, costly belief.

Now it might be objected that there is a problem of deduction simply in virtue of the proof of (a), since even though there is no suggestion that (b) may be proved for deductive inference, (a)'s being true implies that there can be no disproof of (b) either. But to think that the impossibility of disproving (b) raises a problem of deduction analogous to Hume's problem of induction is to incorrectly assess where the burden of proof lies. That (b) can be proved for inductive inference raises the question of induction's reliability, and the proof of (c) establishes a **prima facie** case for its unreliability, and thus for the lack of rational foundation for making inductive inferences. Inductive inference, once (b) and (c) are proved, has a case to answer. But deductive inference has no such case to answer; rather the situation is that, in virtue of the proof of (a), we cannot prove that it is impossible that a case against deduction could be constructed; but this does not constitute such a case.

It might be objected further that there is a problem of deduction inasmuch as there are rival systems of deductive logic. Unless it is possible, it might be argued, to eliminate all but one of these rivals then we cannot know whether it is safe to make an inference which is sanctioned by one system, but not by a rival. Now if this point were granted, it would make the status of deduction, both classical and rival, problematic (though not in the same way as induction). But that would not make it improper or unnecessary to seek a justification for induction. For the assumption that underlies (6) as an attack on the genuineness of Hume´s problem is that even though deduction cannot be justified it is nonetheless unproblematic. Finding that deduction is problematic, therefore, while it places induction and deduction on something nearer the same footing, does so by weakening deduction, and from this we can derive nothing to shore up induction. Indeed, to return to Carnap´s discussion, the situation we are here envisioning would show not that intuition is a reliable guide, since we employ it to justify our unproblematic deductive logic, but rather the opposite - that intuition can lead us astray even in such a seemingly hospitable environment as deduction; which undercuts Carnap´s argument for the reasonableness of relying on inductive intuition.

4. REVIEW OF HUME´S PROBLEM.

This completes our discussion of the six theses from which various authors have constructed arguments designed to show that there is no need to seek a solution to Hume´s problem of induction. We have found that none of these arguments establishes what was intended, since the various theses are either false, or true only when modified in such a way that they no longer lend themselves to impugning the genuineness of Hume´s problem. We are left then with the problem of considering possible solutions to Hume´s problem.

We have, however, learned something from our examination of the dissolutionists´ cases, namely that Hume´s problem arises because Hume adopted a high standard of justification for inductive inference. Attention could profitably be focused, as indeed it has focused, on whether this standard is avoidably high. The major part of the remainder of this thesis will consider just that question, dealing with one of the two main responses, viz, the attempt to found induction on probability. The other main response, the attempt to provide a non-inductive foundation for scientific inference, championed by Popper, will not be considered here, and various other lines of attack on Hume´s problem - in particular the attempt to find an inductive justification of induction, and a vindication of induction - will also be excluded from consideration. Even if these other strategies for dealing with Hume´s problem had not, in my view, been convincingly criticised by others, in particular by Salmon, limitations of space and time would prevent any justice being done to them in this thesis.

Before we turn to probability, however, we shall consider Stove's claim to have disproved Hume's inductive scepticism. This will throw further light on Hume's standard of reason.

CHAPTER 3. STOVE´S PURPORTED DISPROOF OF HUME´S INDUCTIVE SCEPTICISM.

1. INTRODUCTION.

In his [1973] Stove proposed a certain interpretation of Hume´s argument on induction, claimed that this analysis revealed that Hume´s conclusion, his inductive scepticism, was equivalent to a certain statement of logical probability, and then, employing an argument due to von Thun, claimed to have disproved Hume´s inductive scepticism.

In my view, however, Stove´s purported disproof of Hume´s inductive scepticism fails, for Hume´s inductive scepticism is not the thesis Stove takes it to be, and it could not be refuted by an argument of the kind given by von Thun. To support this conclusion I shall argue that even before translating any of Hume´s claims as statements of logical probability, Stove´s account of Hume´s argument is inaccurate; that it is not plausible to attribute any statement of logical probability to Hume; that, in particular, Hume´s inductive scepticism is not correctly represented as a statement of logical probability; and, finally, that if Hume´s inductive scepticism were accurately represented by a statement of logical probability it would not be the statement Stove adopts, and it would not allow the von Thun argument to proceed. Evidence for these claims will be adduced from my exegesis of Hume's argument, given in Chapter 1 above, and from problems with Stove's own account of Hume.

2. STOVE´S MISTAKEN ACCOUNT OF HUME´S ARGUMENT.

Before I turn to the evidence for my claims to be found in the problems which Stove´s analysis of Hume´s argument creates, I shall first set out the argument based on the account of Hume´s argument I have given. This will then allow a clear separation between those parts of my critique of Stove which rest upon accepting my account of Hume´s argument and those that do not.

a. Stove´s Inaccurate Reading of Hume.

There are two significant points of detail on which my presentation of Hume´s argument differs from Stove´s, and one more major issue where I agree with what Stove says in one place, but not with what he says in another.

The major issue is whether Hume´s inductive scepticism covers inductive arguments based upon some probability logic. I do not think that it does, since I hold that Hume excluded the very idea of reasonable probable inference in adopting, as his standard of reason, that inferences are justified only if they show their conclusions to be true. Thus Hume offered no argument against probable inductive inferences, or inductive probabilism, and therefore to solve his problem of induction one need not find fault with his argument directly, but merely show why his conception of reason ought to be abandoned, which one could do by exhibiting a reasonable probable inference. I shall not argue this case on the basis of my analysis, however, for Stove himself has given strong reasons why we should not consider Hume to have offered a sceptical thesis on - ie. putatively refuted, rather than rejected - inductive probabilism. I will delay further consideration of this issue, therefore, till the next section, in order to maintain the separation between those parts of my critique of Stove that depend upon the details of my exegesis of Hume's texts and those that do not.

Both significant points of detail concern Hume's conception of reason. As I understand Hume his conception of reason was that inferences which **do not show their conclusions to be true** are **not rationally justified**, whereas Stove has Hume take the view that all **invalid** inferences are **unreasonable**. In my discussion of Hume in Chapter 1 I noted these points of disagreement with Stove, and gave evidence in support of my account. I do not intend to rehearse that discussion here, but it will be helpful merely to recall that my suggestion allows us both to account for the differences between Hume's conclusions concerning a **priori** and inductive inference, and rationalizes his acceptance of the force of inductive inference as a fact to be explained rather than a judgement to be controverted. On this basis I believe that the reading I have given of Hume's conception of reason, and thus also his sceptical conclusion, ought to be preferred to Stove's.

Both of the differences between our accounts are significant, and in relation to the same point. Each of the alterations I have made to Stove's account presents Hume as giving an **external** critique of induction, whereas Stove´s account presents Hume as giving an **internal** critique of induction. Consider first the difference between rejecting a form of inference because it is not provably truth-preserving, as against rejecting it because it is invalid. The first focuses attention on what the inference can, or rather, cannot, do for us, the second on its logical structure. The first is immediately practical, and has an obvious basis for justification, whereas the second is a formal complaint, a complaint which is only rationalized by another, non-obvious, argument, connecting invalidity with undesirable traits, presumably with lack of provable truth-preservation. I admit that this is only suggestive, but it seems that to opt for Stove´s assessment, that Hume´s objection was to inferences which are invalid, rather than my suggestion, that he rejected inferences which do not show their conclusions to be true, changes Hume´s focus from a user-oriented, or external appraisal of the inference forms, to an interest in the internal logical status of the inference.

This suggestion is very much strengthened by consideration of the second difference between Stove´s and my account of Hume´s argument. For to say that Hume rejected the forms of inference he did because he took them to be not rationally justified is clearly to have him concerned not with the **strength** or **conclusiveness** of the inferences, which is suggested by Stove´s term ´unreasonable´, but with their **legitimacy**. Indeed, to say that an inference is ´unreasonable´ is complaint about an inference which is only given definite meaning **within** some system for measuring the reasonableness, or the degree of conclusiveness, of the inference - a fact Stove makes use of in his attempt to commit Hume to some principles of logical probability, once he (Stove) has identified degrees of reasonableness with logical probabilities. To say, however, that an inference is not justified, or that it is not legitimated, is plausibly not a reflection on the strength of the inference within some system for measuring inferential strength, but a comment upon the whole inference system.

This, I think, is the key, viz, that Hume's conception of inference excluded as unjustified, or as Stove would have it, unreasonable, not individual inferences, but **whole forms of inference.** Thus Hume's conception of reason, and consequently his sceptical appraisal of inductive inference, is, as a number of Stove's critics have insisted, external rather than internal.(1) And as such Hume's inductive scepticism is not translatable as a statement of logical probability, since all such statements are assessments of the strength of inferences **within** a system of assessment, rather than, as Hume's scepticism plainly is, an assessment, or denial of the possibility of, **whole systems** of inference.

This completes my argument that Stove has misread Hume, and that when corrected Hume's conclusion is not plausibly translated as a statement of logical probability, so far as that argument is based on my own exegesis of Hume. I shall now turn to arguing that Stove's analysis of Hume ought to be rejected because of its own internal problems.

b. Hume´s Argument and the Idea of Reasonable Probable Inference.

I have already claimed that Stove´s account of Hume´s argument is inaccurate in one major respect, namely that contrary to Stove´s reading of Hume´s sceptical conclusion concerning inductive inference, Hume did not give a sceptical argument against inductive inferences based on probability. Rather, in adopting as his standard of reason that only inferences which show their conclusions to be true are justified, Hume simply ruled out of court the idea of justified probable inference. Support for this view may be found in a surprising quarter, namely Stove´s own exegesis of Hume´s argument.

In his [1966] Stove shows convincingly that, contrary to the claims of Popper and others, Hume did not refute the thesis of inductive probabilism - that there are some inductive arguments which, while invalid, lend their conclusions sufficiently strong support to justify the acceptance of their conclusions. The basis of his argument here is that Hume did not even consider the possibility of there being reasonable probable inferences.(2) In his [1973], however, Stove claims to have shown that Hume´s inductive scepticism is the claim that the premises of an inductive inference do not confirm the conclusion, from which it follows (granted only the standard empiricist thesis that the conclusion of no inductive argument is acceptable **a priori**) that Hume did, if his conclusion stands, refute inductive probabilism. Thus my account of Hume as ignoring the whole idea of probable inference is supported by Stove´s [1966] exegesis of Hume´s argument - indeed, it was that paper that convinced me that Popper had misread Hume. But there is a clear conflict between my account and Stove's [1973] analysis of Hume's inductive scepticism, and there is also a conflict between Stove's [1966] exegesis of Hume's argument, and his [1973] analysis of the same.

We shall review the problems arising from Stove's [1973] position below. For the moment, however, let us try to throw some light on his internal disagreement. Obviously, we should consider first whether we can account for his divergent opinions as a simple change of mind in the intervening years. But this is not the case, though Stove does shift his ground on an important point, which changes the problem posed for him by the conflict between his exegesis of Hume and his desire to disprove Hume's inductive scepticism. For in his [1973] he explained Hume's failure to refute inductive probabilism, as follows:

> Hume could not have been intending to refute the thesis that some inductive inferences are of a high degree of conclusiveness, **though invalid**, [since] throughout his argument he simply **assumed** that invalidity suffices to make an inference ... unreasonable.(3)

Now, as we shall see, this avoids the outright inconsistency just considered, but the old problem is merely exchanged for a new one. For Stove in his [1973] gives as Hume's inductive scepticism the claim that the premises of an inductive inference do not **confirm** the conclusion, while in the passage quoted above it is plain that scepticism with respect to probable inferences is expressed not by this claim, ie. not by Hume's conclusion as Stove interprets it, but by the claim that no inductive inferences are of a high degree of conclusiveness.(4) The new problem, then, is that as Stove now (ie. in his 1973 discussion of Popper's account of Hume's argument) defines inductive scepticism with respect to probable inferences, what he has given as Hume's conclusion is, contrary to his claim in the main part of his [1973], not a sceptical appraisal of inductive inferences. Thus what is now given as the reason why Hume did not refute the thesis that there are reasonable, though merely probable, inductive inferences is that he did not come to a sceptical conclusion concerning inductive inferences.

But clearly Hume did defend a sceptical thesis concerning at least some inductive inferences; and if there is to be any point to Stove's purported refutation of Hume's inductive scepticism, Hume's argument has to apply to probable inductive inferences, as indeed Stove interprets it as applying in the main part of his discussion. For if Hume's argument for his inductive scepticism does not cover probable inferences then the task, so far as an examination of the rationality of probable inductive inference is concerned, is not to find fault with his argument, but to justify some form of probable inductive inference, doing the former only being a necessary condition of achieving the latter if Hume's inductive scepticism applies to inductive inference founded upon probability. Stove therefore has a problem in defending the relevance of his critique of Hume, since, as his exegesis of Hume's texts in the discussion of Popper's exaggerated claims on Hume's behalf shows, Hume simply did not consider probable inferences, and thus cannot plausibly be thought to have offered a reasoned sceptical thesis concerning them.

In his discussion of Popper´s unjustified assertion that Hume established the untenability of probable inductive inference Stove tries to avoid this problem, as we have seen, by representing Hume´s conclusion as though it was not a sceptical claim concerning inductive inferences, despite his earlier claim that it is so. In the main part of his discussion he employs another strategy, namely interpreting the claim of Hume´s whereby he excludes probable inferences from consideration, which Stove calls ´Hume´s deductivism´, not as **excluding the possibility** of reasonable probable inferences, but as **providing the basis for a sceptical appraisal** of them. Stove then goes on to interpret this claim as a statement of logical probability. But a claim rejecting the whole idea of reasonable probable inference cannot be expected to make much sense as a claim within the system of logical probability, and so it proves to be, as I shall now argue.

c. Hume´s Sceptical Thesis and the Theory of Logical Probability.

I have already disputed the accuracy of Stove´s interpretation of Hume´s inductive scepticism. But let me lay this objection aside, and accept for the sake of the argument that Hume did come to some conclusion concerning inductive inference which is reasonably interpreted as ´All inductive inferences are unreasonable´. Can we plausibly interpret this as a statement of logical probability? I shall argue that the problems in so doing are such that the proposition that we can is just not plausible. I have five objections to Stove´s translation proposal.

The first objection arises from the fact that if Hume´s conclusion is to be a statement of logical probability, then the premises of his argument must include statements of logical probability, and in particular Hume´s conception of reason, which according to Stove is that invalid arguments are unreasonable, must be interpreted as a statement of logical probability. This now raises a **prima facie** difficulty, for as we have just noted, this very statement is supposed also to constitute Hume´s rejection of the whole idea of reasonable probable inference, which entails the rejection of the whole theory of logical probability, and as such it presumably cannot be interpreted as a statement of logical probability. My fourth objection will develop this point.

The second objection is that to attribute to Hume a statement of logical probability will be to foist upon him a completely foreign element unless we can find in his texts some indication that he employed some concept which the theory of logical probability is now taken to explicate. But we can find no such element in Hume, aside, of course, from his sceptical appraisal of induction, and we cannot take this as evidence for the point without begging the question. Hume does not, for example, describe one inference as more conclusive than another, even though neither is valid. Indeed, as Stove ably put it ´the distinction between validity and invalidity is the only distinction among ´degrees of evidence´ that Hume takes notice of´.(5)

Thirdly, one cannot employ statements of logical probability without adopting also a specific conception of degrees of belief, or degrees of confirmation, or whatever one takes logical probabilities to be, such that the principles of the theory of logical probability - eg. axioms of the probability calculus - hold for one's logical probabilities. But there is no evidence that Hume adopted any of the conceptions of confirmation or belief which others have employed to justify their accounts of logical probability. There is no evidence, therefore, that Hume is entitled to a theory of logical probability.

Fourthly, it is by no means clear that one can coherently adopt a theory of logical probability and yet remain an inductive sceptic, since one cannot do so if one's inductive scepticism is expressed by the statement of logical probability Stove offers. For that statement is shown by von Thun to be inconsistent with the premises of his argument, and these are just Stove's principles of logical probability, augmented by the principle that invalid inferences lend their conclusions less than the maximal probability (which I shall call the 'regularity principle') plus two judgements of invalidity.(6) If these premises are true, therefore, one of the premises of Hume's argument must be false, since his argument is plainly valid.

Now according to Stove, Hume's argument has just two premises: that inductive inferences are invalid; and that invalid inferences are unreasonable. It is plainly true that inductive inferences are invalid, and thus to resolve the conflict between the premises of the von Thun argument, and the conjunction of Hume's conception of reason and the thesis that inductive arguments are invalid, either some of the premises of the von Thun argument, or Hume's conception of reason, will have to be given up. Perhaps it is Hume's conception of reason that ought to go, or perhaps it is Stove's principles of logical probability. Clearly, however, we have to choose between the two, as Stove happily admits, but that surely means that it is not correct to translate any of Hume's claims as statements of logical probability, if that it to commit Hume to a set of principles which are incompatible with his conception of reason. Indeed, we can almost hear Hume complain that far from Stove's translations preserving the meanings of his claims, Stove has foisted upon him a conception of reason which is incompatible with his own. To suggest these translations thus charges Hume with incoherence. We shall want some very compelling evidence for one of Stove's translations to secure a conviction.

Finally, it is doubtful that there are any theories of logical probability sufficiently well developed to make it plausible that translating a thesis of Hume's as a statement of that theory of logical probability constitutes a substantial clarification of the original statement.

If we now consider what Stove has said which could be taken to be a response to these objections we find, as it turns out, very little. We shall deal with the objections in turn.

In response to the first objection, Stove says nothing at all.

Presumably, then, all we can do is weigh this objection, and the others to which he gives no response, against the reasons Stove gives for his interpretation of Hume's conclusion as a statement of logical probability. We shall come to this shortly.

Stove gives no response to the second objection. It is therefore left as a mystery why Hume should have accepted a theory of logical probability and yet not employed it, or given any sign that he had adopted it, except by drawing distinctions between inferences on the basis of their validity or invalidity - for which one does not require a theory of logical probability.

In response to the third, Stove goes on the attack, claiming, in effect, that it is Hume's problem if he did not have the right conception of belief, or confirmation, to justify some formal principles of logical probability, for having made statements of logical probability Hume requires some such principles otherwise his statements will be devoid of meaning.(7) But the attack does not solve Stove's problem, rather it shows that whatever evidence there is for interpreting Hume as making statements of logical probability must be strong enough for us to accept that Hume adopted an appropriate conception of belief, or confirmation, without giving any sign of having done so, or alternatively, for us to convict Hume of having adopted theses which are devoid of meaning in consequence of there being no basis in his work for a formal account such as is required to give his claims meaning. Clearly the evidence in favour of Hume's having made statements of logical probability will have to be very compelling indeed.

In relation to the fourth, Stove does show that Hume´s standard of reason, interpreted as the statement of logical probability that if an inference is invalid then its premises are irrelevant to its conclusion, is consistent with the fundamental thesis of logical probability, viz that invalid arguments are not all of the same degree of conclusiveness, (since their degrees of conclusiveness will depend upon the a **priori** probabilities of their conclusions).(8) But this does not show that Hume´s standard of reason is consistent with the principles of logical probability, and thus it leaves open the possibility, which we have shown to be realized, that translating Hume´s inductive scepticism as a statement of logical probability such as Stove gives reduces Hume´s argument to an incoherent attack on induction. And the same problem would apply to any other statement of logical probability which allowed Hume´s conclusion to be proved false given no important assumption other than the principles of the theory of logical probability. Since, however, Hume´s argument is on the face of it coherent, the translation must be rejected, unless independent analysis provides evidence of an inconsistency in the original argument, or, at the very least, very compelling evidence in favour of the translation can be produced.

As to our fifth and final objection to translating statements of Hume´s as statements of logical probability, Stove does not consider in any detail the problems with the theory of logical probability. From his discussion of the theory it appears initially that there is just one theory of logical probability, and then, after Bradley pointed out that this is false, Stove argued that if there are problems with the theory of logical probability then that is Hume´s problem, he having made statements of logical probability and thus needing a theory of logical probability to give these claims meaning.(9) But this is a muddle, for we can only interpret Hume´s claims as statements of logical probability if **we** have the theory. Problems with the theory are thus problems for Stove, not Hume. (Despite this objection to Stove´s analysis, however, I will for the sake of ease of expression continue to follow him in speaking of **the** theory of logical probability.)

Consideration of our five points, the fourth being particularly important, makes it obvious that we are taking a position which is open to a number of serious objections if we insist that Hume made statements of logical probability. But of course we ought to be prepared to adopt this thesis if there is good evidence for it. But there is no compelling evidence, as a review of Stove´s argument for his claim quickly reveals.

d. Stove´s Argument for Translating Hume´s Conclusion as a Statement of Logical Probability.

Stove first establishes that Hume´s conclusion is not a statement of fact. He then writes, and I quote the whole of his argument on the point:

> Hume is considering the predictive-inductive inference, not from the point of view of fact, but from the point of view of reason. He is asking here, not how men **do** infer from ´This is a flame and all of he many flames observed in

> the past have been hot´, but how **would** they, if reason were ´the guide of life´. In other words, Hume was asking: what degree of belief, if any, in the conclusion of a predictive-inductive inference, would accompany knowledge of its premises in a completely rational inferrer?(10)

But do the ´other words´ Stove supplies here mean the same as Hume´s original question? Not if Hume was asking, as he seems to be, not what **degree** of belief would the completely rational inferrer have in the conclusion, but how would he **justify** his degree of belief. Evidence for this alternative interpretation of Hume is of two kinds. First, it provides a more natural interpretation of the question ´how would he reason...?´, for this clearly asks for the logical basis, and the justification of the inference, to be spelt out. Second, from the fact that Hume admitted that agents, including himself, in fact adopt strong beliefs on the basis of inductive inferences, and when Hume admits this he does not seem to be convicting himself of irrationality, as he would have to be if the sceptical judgement of the rational inferrer is to be embodied in his success in having his strength of belief unaffected by the inference. Thus I reject Stove´s attempt to have Hume´s inductive scepticism interpreted as a thesis about the rational degree of belief, in the conclusion of an inductive inference, which ought to be adopted by an agent given the premises of the inductive inference.(11)

Having rejected this first step in Stove´s argument, the second is left without its foundation. But the second would be faulty, in my view, even if we had accepted the first. For Stove says of his question concerning rational degrees of belief, which he has just foisted on Hume:

> But any answer to that question would be an assessment of the degree of conclusiveness of the predictive-inductive inference; that is a certain statement of logical probability.(12)

But while it is true that all statements of logical probability are assessments of the strength of inferences, not all assessments of the strength of inferences are statements of logical probability. Rather, they are only such if we adopt certain conventions concerning inferential strength, or, degree of rational belief. And indeed, just what those conventions are is by no means clear, since we do not have any adequate theory of logical probability. To establish his point Stove therefore needs, as we noted above, to show that when **Hume** assessed the strength of an inference he was making a statement of logical probability, and he also needs to **supply** a theory of logical probability to which Hume can be committed.

Nor is this the end of Stove's difficulties, for even if we grant him the first two steps in his argument, his argument will still not go through. For it is plain, I think, that the statement of logical probability Stove adopts as representing the sceptical content of Hume's appraisal of inductive inference does no such thing. According to Stove, Hume's sceptical conclusion is equivalent to a judgement of the irrelevance of the premises of an inductive inference to the conclusion.(13) But this statement of logical probability is no less a definite statement than any other, and in a full theory of logical probability it will have certain consequences. For example, on Carnap's (first) theory, to adopt the position that the evidence of an inductive inference is irrelevant to the conclusion entails certain factual assumptions about the phenomenon under investigation (provided lambda is assigned a finite value), for example that the evidence is not derived from a random sample. But Hume would rightly insist that his inductive scepticism does not presuppose any such factual claim. Indeed, the point is obvious: any system of logical probability will have to include instructions for deciding when the premises of an inductive inference are or are not relevant to the conclusion, and these will involve factual assumptions, as the setting of the gamma and eta parameters in Carnap's second system of induction, discussed in Chapter 7 below, makes plain; opting for irrelevance will therefore always entail some assertion about a matter of fact. But the inductive sceptic does not adopt his view because he knows the premises of an inductive inference to be irrelevant to its conclusion, nor even because he believes this to be so, but rather because he does not accept that there is any rational basis for opting for either relevance or irrelevance. If Hume's inductive scepticism could be interpreted as a statement of logical probability, therefore, it would have to commit the agent to no empirical claim, and the only statement which will achieve this is the statement of the maximally indeterminate probability, ie. assignment of the interval [0,1], to every inductive inference. And then the von Thun argument, and any argument like it, would not go through.

3. AN ALTERNATIVE CONCEPTION OF PROBABILITY AND HUME´S INDUCTIVE SCEPTICISM.

I have argued that Stove´s analysis of Hume´s argument, and in particular, his attempted disproof of Hume´s inductive scepticism, ought to be rejected. If this is done, however, we give up the most careful study we yet have of the relationship between Hume´s argument and modern conceptions of probability. What are we to replace it with?

To my mind the answer is clear. Stove´s [1966] exegesis of Hume had the point correct when it asserted that Hume simply failed to consider the whole idea of reasonable probable inference. There is a possibility here, therefore, for providing a rational foundation for induction, for there is a form of inductive inference which is not refuted by Hume´s argument but simply denied out of hand. Plainly, therefore, one fruitful line of argument would be to see if we can find in one of the various notions of probability the clear and compelling logic of induction which Hume sought. That will be the task to which the remainder of this thesis will be devoted. We shall attempt to discover whether any of the main systems of probable inference thus far offered provide a clear justification for assenting to the inferences sanctioned by the system.

Of course this is only one way of attempting to refute Hume´s inductive scepticism. Alternatively, as I have already recognized, one could try to find a flaw in Hume´s rejection of inductive justifications of induction, or provide a pragmatic vindication of induction, or consider the presuppositions of induction, or follow Popper in trying to provide a non-inductive foundation for science, or take some further path. But even if one's heart was behind one of these alternatives, the project I have set out would still merit attention. And in the light of published criticisms of the other possibilities, particularly those given in Salmon's [1966], as well as the influence of the idea of reasonable probable inference, particularly in statistics, the analysis of probability must rank as the most interesting of the options.

CHAPTER 4. PROBABILITY, INDUCTIVE LOGIC, AND STATISTICAL INFERENCE.

1. INTRODUCTION

Among philosophers the most vigorously developed and perhaps the most common response to Hume´s problem of induction has been the development of models of inductive inference which do not seek to show that the conclusion of an inductive inference is shown to be true by the premises, but merely to show that the premises of the inference may make the conclusion probable (in one of the several senses given to that term).

Almost completely independently of this enormous philosophical effort, statisticians have sought to clarify the foundations of statistical inference, and they too have relied heavily on the (various) concept(s) of probability to provide a model for inference from sample to population.

In examining the impact we might make on Hume´s inductive scepticism by taking probability as the basis of induction, we thus have to take account of two, quite separate, intellectual traditions. Our first problem will be to find an effective way of ordering our discussion in order to avoid being overwhelmed by what is in total a vast, and very varied, literature.

To propose such a framework will be the main task of this chapter's brief overview of the field with which we have to deal.

2. THE AIM OF OUR ANALYSIS OF STATISTICAL AND INDUCTIVE INFERENCE.

Before turning to the main business, however, some restrictions must be placed on the scope of the analysis of probable inference which is to come. My aim in this thesis is to see if basing our inferences from the observed to the unobserved (which inferences include what statisticians call inverse inference, as well as the philosophers' various models of inductive inference) upon probability allows us to provide a form of inference for which we can give a rational justification of the kind Hume sought for the form of inductive inference he considered.

This limits our discussion in two ways. First, it is beyond the scope of our investigation to pursue modifications to induction other than those occasioned by the introduction of probability. Thus, for example, the suggestion of Giere's that we should admit that induction cannot be provided with a foundationist justification (ie. cannot be justified on the model of justification which requires that an inference have no unjustified premiss), since every induction presupposes some other in having among its premises a factual claim which cannot be proved directly from experience, will not be considered here, even though the criticisms we shall make of several of the systems considered show the importance of attempts to find rational, non-foundationist forms of justification.(1) Nor will we attempt to evaluate the systems of inference we shall consider from any point of view other than the adequacy of the justification which can be provided for the inferences they sanction; thus there is no attempt to provide an exhaustive analysis of any system, and especially no attempt to compare the various systems to see which, if any, provides intuitively acceptable, or scientifically respected, answers to the various inference problems which it covers. Our point of view is not that of the agent wishing to select a brand of statistical or inductive inference to enable some particular problem to be solved, but that of the philosopher who wishes to know, should the agent be satisfied that the inference employed justifies the conclusions to which it leads, the rational basis of this satisfaction.

3. A FRAMEWORK FOR THE ANALYSIS OF STATISTICAL AND INDUCTIVE INFERENCE.

As already mentioned we have, in inductive logic and mathematical statistics, two attempts to provide a rational foundation for scientific inference, and these have developed almost independently of one another. In recent years the situation has greatly improved, due in no small measure to Braithwaite's **Scientific Explanation** and, later, Hacking's influential **Logic of Statistical Inference**. Evidence of the increasing contact between the fields is given in Giere's [1979], which surveys work in inductive and statistical inference as though they represented a single field of of enquiry. Lest we forget how recent a development this represents, however, recall that Carnap's [1950] devotes but a few of its 600 pages to the whole statistical tradition, while Fisher's thoroughly philosophical [1956] contains no reference to any philosopher.

The legacy of this separate development is the completely different internal structures of the two disciplines. In the field of inductive logic the big division is that between competing accounts of probability. Thus we have separate systems of inductive inference based on the frequency, logical and subjective theories, with the frequency theory the least influential and the logical theory the most strongly developed.(2) Among statisticians, however, the theory of probability involved in the various accounts of inverse inference has been seen as less significant than other features. There has been no major statistical school based upon the logical concept of probability (despite the respect given to Jeffreys' work), and widespread tolerance of the subjective theory is a late development. In mathematical statistics, the big division was between the the Neyman-Pearson school, and the followers of Fisher, with the former group dominating, until more recent revivals of Fisher's approach and the resurgence of Bayesianism; and in these disputes the source of the excitement was usually some difference over methodology, rather than a philosophical difference over the concept of probability.(3)

Thus the field of inductive logic is structured by the divisions between the frequency, logical, and subjective theories of probability, while in the land of mathematical statistics the traveller keeps his bearings by plotting his position in relation to the Neyman-Pearson, Fisherian, and Bayesian accounts of statistical inference. It seems, therefore, that having admitted the mutual desire to know each other more closely, the two traditions will be forced to take to Procrustes' bed if they wish to get on intimate terms.

The situation is not too grim, however, for as it turns out the two enquiries can be fitted into a single framework in a way that allows us to preserve the important divisions in both of the original fields. The classification I propose to achieve this combines systems advanced by Giere and Seidenfeld.

Giere distinguishes between what he calls the information and testing paradigms in statistical inference, sketching the main ideas of each in these striking terms:

> The leading idea of the information paradigm is that there should be a direct measure of the bearing of evidence on hypotheses. As new information comes in, whether in the form of experimental data or not, it is to be processed more or less continuously. The output of the information processing is, of course, a new evaluation of all relevant hypotheses.
>
> In the testing paradigm there is no direct evaluation of hypotheses relative to data. Rather the data are the output of a setup designed to test a particular hypothesis. The result of the testing process is not a re-evaluation of some continuous measure, but a simple dichotomous acceptance or rejection of the hypothesis under investigation.(4)

While, as Giere admits, such paradigms are ´inherently vague, exceedingly general, and largely invisible´, their use in organizing a large body of material for analysis is undoubted, placement of the differing theories either facilitating or frustrating attempts to get at the logic of divisions and disputes between the various theorists. Thus while it is hard to argue convincingly that one version of some paradigm is more accurate than another, or that an original thinker should be attached to one paradigm rather than to another, it is important to try to get the best fit of theories to frameworks, hoping that the choice will be vindicated by the fruitfulness of the analysis which the framework adopted promotes. Thus I shall not adopt Giere´s suggested paradigms as they stand, for while in my view Giere has very clearly identified one major division in the field of induction and statistical inference, he does not draw the line in the place I think it is best drawn, and he blurs what seems to me to be an important distinction.

The key here is the placement of Fisher in the testing paradigm. There is certainly some evidence for so placing him since Fisher did at least buttress his logic of tests of significance by referring to their long run error probabilities, which, as Giere points out, is typical of the testing paradigm. But he also heaped scorn on the testing paradigm, and attempted with his notion of significance, his defence of likelihood as a measure of evidential support, and his development of the concept of fiducial probability (all of which will be examined in Chapter 6), to provide a logic of evidential support. Moreover, the bitter controversy between Fisher and Neyman is surely made quite mysterious if they share a fundamental outlook.

Thus, while the matter is not clear cut, (but that is what we should expect, since these paradigms are, after all, our invention, based on what is clear now, not on what was clear in 1930), I conclude that Fisher ought not to be placed with Neyman and Pearson in the testing paradigm.(5)

But neither would it be appropriate to put Fisher into Giere's information paradigm, since this best fits the Bayesian approach to statistics, and Fisher's great achievement, as I understand it, was to provide a basis for statistics independent of the earlier Bayesian tradition which came down from Laplace.(6)

A solution to the difficulty with placing Fisher is suggested by Seidenfeld's division of inductive inference into 'frequentist' and 'Bayesian' views. Seidenfeld explains his terms thus:

> I call an inductive logic **bayesian** when (i) an agent's beliefs, at a time, are represented by a single coherent probability

> function, (ii) an agent´s confirmational commitments, at a time, to changes of belief that follow a growth in knowledge are governed by Bayes´ theorem (this is called confirmational conditionalization), and (iii) total evidence is respected. I shall call an inductive logic **frequentist** when it seeks to bypass confirmational conditionalization in inverse inference and, instead, relies upon... a frequency principle. That is... frequentists solve inverse inference problems by reduction to direct inference and problems of direct inference (the ´single case´) are solved with a frequency principle.(7)

On this division the Bayesian programme is split off from the rest of Giere´s information paradigm, with all the rest, including the defenders of likelihood as basis for a logic of support, and fiducial probability, being classed, with Neyman and Pearson, as frequentists. The former division I accept, but the latter is undesirable; we do not wish the attempt to provide an independent logic for induction to be confused with conditionalization (ie. use of Bayes´ theorem as the rule of induction) but neither do we want to lose the distinctiveness of this approach by placing it with the testing paradigm.

My comments on Giere´s and Seidenfeld´s proposals suggest a **three-fold** classification, distinguishing between: the **reliability paradigm** of Neyman and Pearson, which led to Wald´s decision theory; the **Bayesian paradigm**, which takes Bayes´ rule as the fundamental rule of induction; and the **confirmation paradigm**, which is defined by the attempt to find a distinct logical basis for inductive or statistical inference, ie. a logic of evidential support or confirmation.

The main feature of the reliability paradigm is the rejection of any logic of support, including the use of inverse probability (except in special cases), finding the basis of support for the conclusion of a statistical inference in the long run frequency with which the rule employed leads us to accept mistaken inferences. In addition to the account of statistical procedures given by Neyman and Pearson, I include Wald's decision theory in this paradigm. As already discussed it is arguable that certain of Fisher's proposals belong here too, but I shall place them with the confirmation paradigm, since it was to that paradigm that he made his major contribution.

The confirmation paradigm is distinguished by the attempt to find a logic for inverse inference independent of Bayes' theorem, and thus free of the need to defend **a priori** probability distributions, while retaining the Bayesians' ability to speak of probabilities for hypotheses, which the reliability paradigm does not allow. The main contributions here are Fisher's theory of significance testing, maximum likelihood estimation, and fiducial probability, and, stimulated by these idea of Fisher's, Hacking's logic of support and Kyburg's epistemological probability. I shall also include Carnap's development of confirmation functions here, since although Carnap was a Bayesian, his major contribution was not to the development and defence of the Bayesian methodology, but to the construction of a concept of probability tailored to make sense of the notion that an hypothesis can be considered more or less probable, depending on the strength of support offered it by the evidence upon which its probability is determined.

Finally the Bayesian paradigm needs to be briefly characterized. That is not so easy, however, for this tradition is internally divided in significant ways. First, there are objective and subjective Bayesians, divided by the concept of probability they employ, which then leads to considerable differences in the nature of the theory of inference to inverse probability which they defend. Jeffreys and Jaynes are the principal objective Bayesians, while Ramsey, de Finetti, Savage, and Jeffrey are the main figures in the subjective Bayesian camp. Since these two positions have become clearly established, Levi has added a new strand to the Bayesian paradigm with his original account of Bayesian decision-making, governed by convex sets of probability functions and epistemic utilities; and, pressing Jaynes' method of determining prior probabilities into more general service, a school of neo-Bayesians has emerged which replaces Bayes' theorem as the basis of inverse inference with the rule of maximum entropy conservation.

Each of the three paradigms will be the subject of a later Chapter in this thesis, and all of the theorists mentioned will be discussed. While this, so far as I am aware, will cover all of the prominent theories of induction and statistical inference, not every theory is listed for discussion. It has been my intention, however, to cover all of the versions of the three paradigms that have thus far been developed, and thus if I have been able to make good my intention then there is no attempt to provide an answer to Hume's problem of induction, by basing inductive inference upon probability, which is not closely related to a theory discussed, if it is not itself given consideration. Moreover, the coverage of my three paradigms can be confidently thought to be sufficiently comprehensive, since they jointly cover the fields defined by the suggestions of Giere and Seidenfeld.

4. ON THE MODAL THEORY OF PROBABILITY.

Certainly, so far as the philosophical tradition is concerned, the three paradigms cover all of the possibilities, for, broadly speaking, the frequency, logical, and subjective theories of probability form the basis for systems of inference in the reliability, confirmational, and Bayesian paradigms, respectively. Aside from these, the only major concept of probability is the modal concept. We shall deal with that here.

Hacking revives the modal theory of probability in a recent defence of the Neyman-Pearson statistics, retracting his criticism of 1965 that this form of statistics does not provide us with what we most want, namely a measure of support for hypotheses. Hacking now thinks that Neyman-Pearson statistics do give us what we need, viz a basis on which to make the (modal) assertion that an hypothesis is probable.(8) Thus, according to Hacking, inferences on the Neyman-Pearson model support statements of modal probability. I want to emphasize here that the support runs from the theory of inference to modal probability, and not the other way around; the modal theory of probability is not a complete theory, being dependent upon the success of another theory of probability, or rather a theory of inductive inference. (And we shall see in our discussion of Neyman-Pearson statistics below, that while Hacking's analysis of modal probability and statistical inference in the Neyman-Pearson school does blunt the force of his earlier criticism of those inferences mentioned above, this dependence leaves the Neyman-Pearson justification of statistical inference within the reliability paradigm open to the most damaging of Hacking's original attacks.)

The distinction between non-modal and modal theories was defined by Toulmin as the distinction between taking 'probably **p**' to differ in **content** from '**p**', as against taking 'probably **p**' to state the same fact as the simple assertion '**p**', but to state the claim in a guarded way, indicating that the assertion is not to be relied upon. Hacking neatly catches the main idea of the modal concept when he asserts, in its defence, that ''probably **H**' is not a statement about **H**. It is a way of stating **H**.' The way of stating **H** being a way in which the reader or listener sees that the speaker, while willing to assert H, wishes to make clear that the possibility of **H**'s being false is not ruled out by the evidence he possesses.

The modal theory thus stands opposed to both the frequency and logical theories of probability, since on both these accounts 'probably H' is a statement about H: on the frequency account, that in some real or hypothetical reference class the proportion of occurrences of states in which H is true dominate those in which it is false; and in the logical theory that on some evidence, the inference to **H** is stronger than the inference to **-H**. Whether the modal theory is incompatible with the subjective theory of probability, however, is not so clear, since on the subjective theory to assert 'probably **H**' is to announce that your belief in **H** is stronger than your belief in **-H**, which is plausibly what one is doing when one guardedly asserts H, ie. when one asserts 'probably **H**' intending 'probably' to be interpreted in the modal sense.

Toulmin does not, however, adopt a degree-of-belief analysis of modal probability, for he wants, as does Hacking, to retain, when one asserts 'probably H', the propriety of the listener's inference that one has good evidence for 'H', whereas this is not part of the meaning of 'probably H' on the subjective theory. For when one repeatedly asserts 'probably H' in the subjective sense of 'probably', one is not at fault in any way if 'H' should prove false; rather the repeated non-occurrence of the event predicted by H should, via Bayes' theorem, lower the probability of H until one asserts 'probably ¬H'. However, in distinguishing the modals 'probably' and 'perhaps' Toulmin notes that one can be unlucky enough for H always to turn out to be false and yet suffer no embarrassment for having repeatedly asserted 'perhaps H', whereas in asserting 'probably H' one makes oneself 'answerable for fulfilment, if not on all, at least on a reasonable proportion of occasions'.(9)

From this it is clear that one only has a warrant for asserting 'probably H', ie. one may only assert 'H' under the protection from occasional failure afforded by the modal 'probably', when one has evidence to support the prediction that the events described by 'H' will occur. Consequently, if one accepts the modal theory of probability one must also accept that before one can legitimately employ 'probably' one must be in possession of a satisfactory inductive logic. In short, before one can derive any benefit from probability on the modal account, Hume's problem must be solved. There can be, therefore, no solution to Hume's problem in the modal theory itself, and thus we need not discuss it further.

We turn now to the main task of this thesis, the examination of the models of probable inference provided by the three paradigms indentified in this Chapter.

CHAPTER 5. THE RELIABILITY PARADIGM.

1. INTRODUCTION.

In this Chapter I shall describe and examine the forms of inductive inference defended by those who belong to the reliability paradigm. This will require us to consider the theory of hypothesis testing due to Neyman and Pearson, the theory of estimation by the calculation of confidence intervals which Neyman developed, and the work of Wald on providing a decision-theoretical basis for these techniques.

Our aim will be to clarify whether the inference forms recommended by these authors provide an answer to Hume's problem in the terms already defined, namely the construction of a form of inductive inference for which a clear and compelling justification can be given. More precisely, we seek to clarify whether the inference forms promulgated within the reliability paradigm are such that the critic can reject the conclusion of an inductive inference, judged correct by the paradigm, without transgressing against a principle of rationality to which the critic can be committed. We shall see that, on this specification of what is required to refute the inductive sceptic's claim that induction has no foundation in reason, the reliability paradigm contains no answer to Hume, since the inference forms which have thus far been developed suffer from a number of defects, which allows the critic to reject each of them.

One problem with Neyman-Pearson statistical methods, as I shall from now on refer to the forms of statistical inference defended by these theorists, is that they all presuppose that the agent is in possession of certain factual knowledge which cannot be inferred directly from observation statements. Noting this, Reichenbach declared these methods to be secondary rules of induction, and concluded that they could not stand alone as an answer to Hume.(1) Recently, Giere has challenged Reichenbach's view, seeking a form of justification which does not require that all premises of a justifiable inference be independently justified. Using an obvious terminology, Giere describes his project as the rejection of Hume's foundationism in favour of a non-foundationist justification.(2) As already indicated, this response to Hume does not rest essentially on the concept of probability, and thus falls outside the scope of this thesis. Our failure to consider Giere's suggestion ought not be taken, therefore, to imply lack of interest in it; indeed, I think Giere is on the right track, but this is not the place to follow him.

2. INDUCTIVE INFERENCE AND INDUCTIVE BEHAVIOUR.

a. Neyman´s Behaviourism.

The main figure of the reliability paradigm, especially as an interpreter of the philosophical basis of the programme, is Neyman. Now since Neyman´s most notable philosophical contribution has been his repudiation of the idea of inductive **inference** and his championing of the idea of inductive **behaviour**, we must begin our analysis of the statistical methods of the reliability programme as methods of inductive inference with an analysis of Neyman´s view and a defence of our assessment of his statistical methods as methods of inductive inference. In so doing we shall find that despite Neyman´s protestations to the contrary, and his espousal of scepticism with regard to induction, he does offer statistical procedures which are intended to be a rational basis for the scientific practice which is ordinarily and properly described as accepting a claim on the warrant of an inductive inference. But even if Neyman had no intention of offering statistical procedures to support scientific inductions, the question whether Neyman-Pearson methods can be used as rational forms of inductive inference would still be an interesting and urgent problem; for even if Neyman did not view these methods in this light, it is indubitable that practising scientists and statisticians do so view them, and leading philosophers have defended this interpretation. Hacking, for example, now defends the Neyman-Pearson theories of testing hypotheses and confidence interval estimation as ´sound theories of probable inference´, declaring this to be ´worth arguing´ precisely because Neyman rejected the concept of inductive inference.(3)

Turning now to Neyman's claims, we shall see that there are two distinct conceptions of statistics confused in his account of inductive behaviour. On the one hand he presents inductive behaviour as just a particular way of reaching a conclusion; while on the other he bases the choice of the appropriate statistical method to apply to some problem on consideration of the consequences of various outcomes, here preventing, I shall argue, Neyman-Pearson statistics being taken as an account of scientific inference as opposed to practical decision-making. Once the two conceptions are clearly separated, the interesting questions are whether either of them can provide an independent basis for Neyman-Pearson statistics, or whether the logic of the tests incorporates elements of both; and, whichever of these alternative proves to be the case, whether the clarified theory can serve as a rationally justifiable form of inductive inference.

Neyman adopts the first conception of inductive behaviour in a paper written jointly with Pearson:

> We are inclined to think that as far as a particular hypothesis is concerned, no test based upon the theory of probability can by itself provide any valuable evidence of the truth or falsehood of that hypothesis.
>
> But we may look at the purpose of tests from another view-point. Without hoping to know whether each separate hypothesis is true or false, we may search for rules to govern our behaviour with regard to them... Here, for example, would be such a 'rule of behaviour': to decide whether a hypothesis, H, of a given type be rejected or not, calculate a specified character, **x**, of the

> observed facts; if $x > x_o$ reject **H**, if $x <= x_o$ accept H.(4)

Now provided nothing extraneous to the assessment of evidential support is involved in the choice of **H**, **x**, x_o, or any other element of the test, there seems no good reason why we should not conisder this rule of behaviour as a rule of inference and thus take this behaviourism to be only a terminological idiosyncrasy. That Neyman and Pearson used the term ´behaviour´ rather than ´inference´ could then be put down to their wish to distance their proposed basis for the acceptance or rejection of tested hypotheses from the prevailing conception of statistical inference, according to which an inference must involve deductive proof of the conclusion or at least confirmation of it by the premises. This is Hacking´s suggestion: rejecting what went by the name of reasoning in the field of inductive inference, Neyman and Pearson rejected the name as well when they came to setting out the logical basis of their tests of hypotheses.(5) Neyman gives support to this line of analysis in a paper from 1941, in which he wrote:

> In the author´s opinion, the word ´conclude´ has been wrongly used in that part of the statistical literature dealing with what has been termed ´inductive reasoning´. Moreover, the expression ´inductive reasoning´ itself seems to involve a contradictory adjective. The word ´reasoning´ generally seems to denote the mental process leading to knowledge. As such, it can only be deductive. Therefore, the description ´inductive´ seems to exclude both the ´reasoning´ and also its final step, the ´conclusion´. If we wish to use the word ´inductive´ to describe the results of statistical enquiries, then we should apply it to ´behaviour´ and not to ´reasoning´.(6)

Here Neyman reveals himself to be (to borrow Stove´s term), a deductivist, his behaviourism resulting from a restricted and mistaken view of inference. Thus we get the impression, urged on us by Hacking, that we can reinterpret Neyman´s rules of inductive behaviour as rules of inductive inference, provided only that we refuse to accept both Neyman´s deductivism, and the conception of inductive inference (shared by both the confirmation and Bayesian programmes) that inductive inferences essentially involve the calculation of (non-modal) probabilities for the conclusions of the inferences. We would then have that Neyman-Pearson tests are inductive inferences leading to the acceptance or rejection of hypotheses, which judgements are fallible and thus at best merely (modally) probably correct.

Before we could accept this characterization of Neyman-Pearson statistical methods as inductive inferences we would have to be sure, however, that Neyman´s behaviourism is only terminological, and I have already indicated that this is not the case. For part of Neyman´s adoption of his behaviouristic conception of induction involved identifying statistical methods as procedures for making decisions rather than reaching conclusions, and features proper to decision-making, but foreign to making inferences, may thus have become included in the very logic of the Neyman-Pearson statistical methods. The question that needs to be answered, therefore, is whether Neyman´s **decisions** can be reinterpreted as **conclusions**.(7)

b. **Neyman´s Naive Empiricism.**

In our discussion of Neyman-Pearson statistics, in order to determine whether the decisions Neyman spoke of in elaborating his concept of inductive behaviour can be taken as conclusions in our sense, we shall be primarily concerned to investigate whether the theory of hypothesis testing and estimation by the determination of confidence intervals relies upon any subjective element. Since Neyman held a frequentist concept of probability, the likely place to look for an element of subjectivity in his statistics is in the use of utilities in chosing appropriate statistical procedures for given problems. Before we begin that analysis, however, it will be helpful to examine Neyman´s conception of science, for this will give us an idea of the kinds of circumstances in which he envisaged his statistical procedures being employed; and since Neyman will be found to hold a naive empiricist account of science, this will explain how he could have thought that science could get by merely with statistical tests governing inductive behaviour, rather than, as is ordinarily supposed, requiring inductive inferences to support scientific conclusions.

In the main paper in which he developed his concept of inductive behaviour Neyman wrote:

> The general development of science proceeds very much as the process of children´s learning to walk. With a tremendous expenditure of time, energy, and funds, a great number of organized sets of observations and of

> experiments are performed daily. The results of these observations and experiments are recorded, and some of them are systematized and collated with material accumulated earlier. Then theoreticians ponder on the records, and work out theories.(8)

We have here a naive empiricist account of science, which at the time Neyman wrote was not, of course, uncommon, for it was prior to the publication of Kuhn's influential [1962], and Popper's [1959]. Lack of awareness of one of the main points of the latter, the theory-ladenness of observations, is particularly significant in Neyman's concept of science, for contrary to Popper's demonstration, Neyman assumes that theories depend for their support on observations directly and thus not on other theories. Thus Neyman mistakenly believed, at least at this time, that statistical procedures are only called for at the end stage of an enquiry, as he made plain:

> In the present paper I intend to sketch an anatomization of a particular phase of scientific research, namely the **concluding** phase, frequently called **inductive reasoning**.(9)

Neyman's picture of science, then, imagines it to be a two step procedure: first the observations are made, and then the inductive judgement completes the enquiry. Another piece of science will have this same structure. What is missing from Neyman's account is the dependence of one enquiry on the results of some prior investigation, and thus the realization that success in some inductive judgement presupposes accuracy in some prior judgement. Neyman thus fails to see that a sound **decision** in one enquiry, presupposes a correct **inference** in a prior investigation whose results are employed in the latter enquiry. And since any test, any step in the growth of scientific knowledge might become, and is expected to be capable of becoming, a presupposition of some later step, the results of every scientific judgement must be capable of standing as a conclusion. The process whereby a judgement is justified, therefore, cannot properly be characterized as the adjustment of an agent's behaviour; it must be a rationally justified inference in order that the outcome be reliable for another agent in another enquiry, where the results of the first are presupposed. In short, science cannot avoid theory-ladenness, and thus it cannot make do with decisions. Science cannot be fabricated from decisions, but only from conclusions.

Clearly any science that allowed judgements to be regularly and properly made on the basis of assessments which include consideration of subjective factors, such as the consequences for an agent's reputation if a negative result is published, would soon break down. If there is to be any consideration of consequences in the process by which judgements are justified, the consequences considered must be purely epistemic; for then the judgements remain objective and are thus (according to the distinction between conclusions and decisions adopted here) to be considered the conclusions of inferences, not decisions derived from application of decision-rules.

Before turning to the consideration of the details of Neyman-Pearson statistical procedures it would only be fair to Neyman to remark that he did not adopt his conception of inductive behaviour for no good reason.

Rather, he provided a penetrating analysis of some statistical procedures Fisher defended as inductive inferences, and argued that in each there appears to be a step which, in Neyman's terms, is 'an act of will', this act having the consequence that the hypothesis under consideration is either accepted or rejected. Neyman here comes very close to what Hume argued in his analysis of induction; indeed, it would be felicitous to say that Neyman's reason for rejecting any claim of Fisher's that he (Fisher) has provided rational methods of induction is that in each of his statistical procedures there is a step which 'is not supported by reason'; and Neyman's rules of inductive behaviour can profitably be viewed as an attempt to educate Hume's habits. We may not like Neyman's solution to his problem of statistical inference any more than we do Hume's solution of his sceptical doubts concerning induction, but that gives us no basis for ignoring either critique.

3. A SKETCH OF NEYMAN-PEARSON TESTS OF HYPOTHESES.

In their early joint papers Neyman and Pearson put forward a new theory of tests of statistical hypotheses, incorporating both a new logical structure, and a new defence of the rational force, or cogency, of the tests. Some years later Neyman added a new theory of estimation, estimation by the determination of confidence intervals.(10) As Neyman's presentation of his theory makes clear, confidence interval estimation is closely related to tests of hypotheses, or hypothesis testing (as I shall commonly call it), and thus we shall only need to discuss in detail one of the theories. I shall choose the theory of hypothesis testing for close examination, since that is the simpler of the two.

In essence, a test of an hypothesis is a rule for determining whether the hypothesis under test, H_o, is to be accepted, or whether it is to be rejected in favour of some specified alternative, H_a, according to the observed outcome of some relevant experiment, where the outcome observed is the result of some stochastic process and is represented by the value of a specified random variable. Some hypotheses require more complex tests than others, but the problems I wish to consider are almost all present in the simplest of tests, so we shall, for the most part, restrict our attention to these.(11)

The initial step in constructing a test of some hypothesis is to adopt a particular kind of mathematical model for the possible outcomes of an experiment designed to discriminate between the hypotheses under consideration, namely a model in which the outcome of the experiment is supposed to be a quantity which, on repeated trials of the experiment would take each of its possible values with a long run frequency determined by the model and hypothesis under test, but whose value cannot be predicted with certainty on any given trial. The outcome of the experiment, that is, is assumed to be a random variable distributed according to some one of the family of functions specified by the model, the precise function being picked out by the hypothesis under test.

The simplest example is the experiment designed to test whether a coin is biased, the experiment consisting of a set of n tosses and the outcome being the number x of tosses resulting in heads. Now on the assumption that the tosses are independent we adopt as the mathematical model of the experiment, M, that x is randomly distributed according to a binomial function $f(x)$, with one free parameter needing to be fixed to determine the distribution uniquely, this free parameter being p, the probability that the coin will land heads on a single toss (which Neyman and Pearson identify with the long run relative frequency of heads among all tosses).

Now this model, like any other mathematical model which would have allowed us to apply the theory of hypothesis testing to the question of the bias of the coin, defines a **class of admissible hypotheses** concerning the coin's bias. Hypotheses such as that the coin is biased to produce a predominance of whatever face is uppermost on the first throw, are inadmissible, being inconsistent with the assumption of independence of tosses incorporated in the model. In fact the class of admissible hypotheses is limited to hypotheses assigning **p** point values in the interval [0,1].(12)

Once the model and set of admissible hypotheses is defined, the next step is to select the **test** and **alternate** hypotheses. We shall not examine here on what basis this selection may be made; Neyman typically refers to elements of decision theory at this point in constructing a test, and we shall want to enquire whether this is avoidable, and if not, whether it introduces an element of subjectivity into Neyman-Pearson tests. For now we note simply that two hypotheses, H_o and H_a, are chosen from the set of admissible hypotheses to serve as the hypothesis to be tested and the alternate, respectively, and leave till later both the analysis of why we need **both** of these hypotheses to be specified, and the basis of their selection.

The outcome of the experiment is the number of heads observed among the **n** tosses. It is this variable, **x**, which is the random variable in the experiment, and the set of possible values for the random variable, the **sample space**, is thus the set of integers from 0 to **n**, inclusive.

With these preliminary definitions completed we come to the core of the construction of a test, the selection of a portion of the sample space as the **critical region**, so called because if the random variable should take a value within the critical region then the test hypothesis is to be rejected and the alternate accepted. The criteria Neyman and Pearson give for selecting the critical region constitute their most distinctive contribution to the theory of statistical tests.

Neyman and Pearson's recommendation was that in order to select the critical region we must consider two kinds of error which we might make in performing a test of an hypothesis. First, we might reject H_o even though it is true; this is an **error of the first kind.** Second, we might accept H_o even though it is false; this is an **error of the second kind.** We will want to avoid both of these errors, and thus choose our critical region to give us the best chance of so doing. The matter is complicated, however, both in respect of the determination of the error of the second kind, and in choosing a critical region which gives the optimum balance between the two kinds of errors, and we shall postpone further discussion till our criticism of hypothesis testing, below.

Having chosen a test and alternative hypothesis and a critical region to use in a test between them, we conduct the experiment and observe whether the random variable falls in the critical region or not. If it does the test directs us to reject H_o in favour of H_a, if not we retain H_o. The justification for acting in this way, or, rather, making these inferences, is that because the probability of x taking a value in the critical region when H_o is true is low, we will not often err if we reject H_o when a such a result occurs, ie., when x takes a value in the critical region; while, since the probability that x takes a value not in the critical region when H_o is false is also kept low, we will again err only infrequently if we reject H_a in this circumstance.

We have, then, a long run justification for the tests, viz, set the error probabilities for your test at a level where the errors represent acceptable risks, and then abide by the test knowing that you will be wrong only so seldom that the risk of being wrong on any particular occasion is acceptable.

This justification is ingenious, and is part of the great appeal of Neyman-Pearson tests. However, it requires further scrutiny, and we shall consider it in the next section, along with the other major elements of the theory, as identified above.

4. ANALYSIS OF TESTS OF HYPOTHESES.

a. The Empirical Presuppositions of a Neyman-Pearson Test.

All statistical tests begin from the assumption that the result of the experiment which generates the data to be used in the test is subject to the influence of some chance factor, that is to say, that there is ineliminable though limited variability in the experimental outcome, so that while the outcome of no single trial is predictable with certainty, the relative frequencies of the various types of outcomes is stable in the long run of trials. This assumption is required to enable the random variable, x, to be defined, and to allow a distribution for it, $f(x)$, to be calculated. The essence of a test of an hypothesis, H_o, is to determine whether the assumption about the experimental outcome made by H_o in determining $f_o(x)$ as the distribution of values of x, can account for the observed experimental outcome.

What we are concerned to clarify here is whether the stochastic model incorporated in H_o, is put to the test in the test of H_o, or whether the test has an empirical presupposition - which would raise the problem of the justification of this presupposition. This question was raised early on by Fisher, with subsequent investigations by Rogers and Giere.

To begin our analysis note first that in my discussion of the construction of a Neyman-Pearson test I have followed Neyman's example in his discussion of Fisher's lady tasting tea test and considered the model to be **shared between** the test and alternate hypothesis, and thus the model is, though Neyman does not point this out, an **empirical presupposition of the test.**(13) It might be thought that the need for an empirical presupposition can be avoided by taking the model to be incorporated in H_o itself, and thus put to the test by the test of the hypothesis. This, however, in consequence of problems with defining the probability of an error of the second kind on this interpretation of the nature of the hypothesis tested, requires another empirical claim to be adopted as a presupposition for the test, and so the problem is not avoided by the manoeuver suggested. Let us get this clear.

If the statistical model **M** is shared between H_o and H_a then it is simple to define both kinds of errors, and to determine their probabilities. Given any critical region we can compute the probability of making an error of the first kind, for given H_o it is simply the probability that **x** will take a value in the critical region even though H_o is true; and this is determinable because H_o, in conjunction with **M**, the mathematical model for the experiment, completely specifies **f(x)**, the hypothesized distribution of the random variable. Computing the probability of an error of the second kind is a little more complex, but still possible, because while $-H_o$ is not any definite hypothesis, **M** & $-H_o$ is - it is a complex hypothesis specifying that the real distribution belongs to the family **M** but is not the particular member of the family identified by H_o. Thus because **M** is not put to the test, and thus is not rejected along with H_o should that go, the probability of an error of the second kind can be calculated.

Suppose now that M is incorporated in H_o and is not a presupposition of the test. This will raise two problems, the first being that we no longer have any basis for the determination of the class of admissible hypotheses. Now this is a serious problem, for if the admissible class is not logically exhaustive, to choose a class of admissible hypotheses is to assert that the disjunction of the hypotheses, which is not logically necessary, is true - and this is an empirical claim, and thus an empirical presupposition of the test. Perhaps, then, the admissible class can be logically exhaustive, merely defined as H_o and its negation.

This will not work, however, or rather it gives rise to a second problem since we cannot now calculate the probability of an error of the second kind, ie. the probability that **x** will not take a value in the critical region even though H_o is false, for $-H_o$ is not a definite hypothesis and thus it determines no specific distribution for **x**, as Fisher pointed out, deriding the whole idea of an error of the second kind.(14) Moreover, the vast class of distributions compatible with $-H_o$ cannot be structured in the way necessary to allow a best test to be chosen for the test of H_o against $-H_o$, as Giere, following Rogers, proves.(15) Commonly, this problem is avoided by, without discussion or defence, **redefining** 'error of the second kind' as the error of accepting H_o when H_a **is true.**(16) Thus it is assumed, as we have seen it must be if the choice of the test is to be justified on the criteria Neyman and Pearson give for choice of critical regions, that if H_o is false H_a is true, and thus $H_o \vee H_a$ becomes an empirical presupposition of the test.

Justification of Neyman-Pearson tests of hypotheses, and also confidence interval estimation, for the logical situation is no different in this later case, will require that the empirical presupposition we have shown to be required by a Neyman-Pearson test, namely the specification of a manageable class of admissible hypotheses, be justified. That will not be easy, since the specification will require, in effect, the assumption of a claim requiring inductive support in its turn, viz that one of a non-exhaustive disjunction of hypotheses concerning the distribution of outcomes of an experiment is true. Indeed, it is this problem with stimulated Giere to consider a non-foundationist conception of justification for induction.

b. **The Problem of Partitioning the Class of Admissible Hypotheses.**

Thus far I have assumed that there is no problem in partitioning the class of admissible hypotheses into a test and alternate hypothesis. However, this selection cannot be arbitrary, for the test is not symmetrical with respect to these two. How then do we determine where to make the partition?

In their early joint papers Neyman and Pearson were concerned principally with developing the mathematical foundations and machinery for hypothesis testing. The theory provided the formal criteria to be applied when making the partition, but no guidance on their application. It was said only that a test should be chosen to minimize the probability of a type one error, and then the probability of a type two error should be minimized subject to the error of first type not exceeding its specified minimum (the two errors being so related that decreasing the probability of one increases that of the other). Thus the statistician was told only to determine which hypothesis it was most important not to reject if true, and to adopt this as the test hypothesis, H_o, the remainder of the class of admissible hypotheses becoming the (typically complex) alternate, H_a.(17)

Now it is obvious that consideration of the consequences of decisions taken according to whether one accepts this or that hypothesis will make the selection of H_o trivial, at least in principle. This can be done by attaching loss functions to the actions of accepting the two hypotheses when false, and then choosing H_o so as to minimize the probability of the most serious loss. However, before we slide into decision theory here, as Neyman was prone to do, we ought to reflect upon whether we can make sense of the talk of loss functions in purely theoretical contexts, bearing in mind that science is based upon conclusions, not decisions.(18)

Neyman gives one analysis of the application of purely epistemological criteria in the choice of H_o.(19) He notes that we want to avoid making erroneous claims, that we will proclaim H_a if H_o is rejected, and thus that we should take as H_a the hypothesis we wish to find evidence for, and take as H_0 the remainder of the admissible class of hypotheses. But this distinction is not at all clear. What is it about an hypothesis that makes it the kind of hypothesis we wish to find evidence for? The key here, I think, is in the name often used for the test hypotheses, the 'null' hypothesis; what makes an hypothesis a null hypothesis is that it adds nothing to present knowledge, it merely restates the current situation. The basis for the partitioning, therefore, is not in the logic of tests of hypotheses but in the analysis of the relationship between the admissible hypotheses and the scientific context in which the research is being conducted. What is required, therefore, to formalize and rationalize the partitioning, is a theoretical account of the growth of knowledge which accepts the theory of hypothesis testing as a form of induction and provides an account of how this tool may be best used (problems with the rational justification of induction aside for the moment - here we are concerned with practice, not its rationalization). The lack of such an account is sorely felt in the social sciences particularly, where the banality of papers bristling with statistical research, which has no obvious point, is often remarked. But the absence of such an account is not, as it sometimes seems to be thought, a failure on Neyman's, or Pearson's, or their followers' parts, for the distinction they rely upon between null and research hypotheses is a **philosophical** distinction, and it is up to philosophy of science to give a clear account of it. An attempt was made by Churchman, but it has been largely, and in my view unjustly, ignored.(20)

c. **The Problem of Defining a Best Test.**

From the way we have discussed Neyman-Pearson statistics so far, in particular because we have continually referred to **the** rejection region, it might seem as though a test is defined by the hypotheses considered. This is not the case, however, for given any pair of test and alternate hypotheses it is possible to define many, perhaps an infinite number, of potential rejection regions, thus making it possible to draw conflicting inferences from the same data. If Neyman-Pearson tests are to be rational methods of induction, or even useful rules of behaviour, this liberality will have to be severely curtailed; in particular, if the tests are to be rational methods of induction, then there must be, for given evidence and hypotheses, just one correct inference, and this requires the identification of the uniquely best rejection region, or best test (a test being defined by the rejection region chosen). On what criteria, then, might this choice be based?

The main criterion has already been introduced. A test is defined by choosing a region of the set of possible values of the random variable and deciding that if the observed value of the random variable falls in this region then H_o is to be rejected for H_a, the choice of rejection region being governed by the desire to have a very low probability of getting a result in that region if H_o is in fact true, while also ensuring that the probability of getting a result in the rejection region is quite high if H_o is false. If we get a result in the rejection region even though H_o is true we will falsely reject that hypothesis and thus commit an error of type one, while if the result is not in the rejection region even though H_o is false, we will fail to reject it (and thus fail to accept true H_a, according to the empirical presupposition of the test), and thus will commit an error of type two. Let us call the probability of an error of type one the **size** of the test; and, recalling that since $H_o \vee H_a$ is assumed to be true the probability of the type two error can be identified with the probability of rejecting H_a when it is true, and supposing that **b** is the probability of a type two error, we define the **power** of a test against the alternative H_a to be equal to 1 - **b**. Then the main criterion for choice of a test is to **choose the rejection region so that for a test of acceptable size the power of the test is maximized.** By adopting such a test we get the probability of making the most important error, a type one error, acceptably low, and then minimize the probability of committing a type two error.

Several difficulties can be noted at once. First, size and power are directly related, for reducing the size of the test also reduces its power - we can reduce the size to zero by including no possible values of the random variable in the rejection region, but then the probability of committing an error of type two would be unity and the power would be zero; while we can get maximum power by taking all of the possible values of the random variable to be the rejection region, making the probability of an error of type 1 unity. Obviously a balance between size and power must be struck, but just where it ought to be struck is left by Neyman and

Pearson to the good sense of the scientist. Moreover, since we can improve both size and power, ie. reduce size and increase power, by increasing the sample size, we have to strike a three way balance between test size and power, and sample size. Thus an already complex problem is further complicated by the introduction of the possibility of varying the characteristics of the test by varying the sample size. Now there is no obvious purely epistemic consideration which could arbitrate between competing proposals on the optimum balance between size and power in a given case, and the introduction of sample size makes the prospects for such a rule even more remote; but leaving it to the good sense of the scientist merely hands him a problem the experts have been unable to solve - a fact Wald exploited in introducing a decision-theoretic basis for choice of the optimum test.

The other main problem with the criterion that the best test is the one that maximizes the power for the desired size is that it only applies to relatively few sets of admissible hypotheses, viz those for which both H_o and H_a are simple hypotheses. It is only possible to calculate a probability distribution for a completely specified hypothesis, thus we must consider complex hypotheses as sets of simple hypotheses and determine a function to give a distribution for each of the simple hypotheses. Thus if H_o is complex we will have to consider the size of the test in relation to each of the simple hypotheses in complex H_o and then choose a test whose maximum size is acceptable. If H_a is complex we will have to calculate the power of the test against every simple hypothesis in H_a, which gives the power function of the test.

These things done, we can then try to choose a best test, which we will be relatively straightforward only if among all tests of a given (acceptable) size there is a test which is most powerful against all of the simple hypotheses in the complex alternative, ie. is uniformly most powerful, the superiority of this test being an immediate consequence of the main criterion for choosing a best test, viz, maximizing power for acceptably low size.

It may be, however, that there is no uniformly most powerful test, in which case we have no straightforward extension to the basic criterion to fall back on. In this situation a series of further criteria are proposed, but I shall not go into them, for as it turns out there is no complete set, and there are not even clear rationales offered for some of them, it being a ´matter of taste´, in Lindgren´s words, whether they are desirable properties for a test to have.(21) This is plainly unsatisfactory. If tests of hypotheses aimed to be inductive inferences which refute Hume´s inductive scepticism, then their choice would have to be defensible in the face of probings from a sceptical critic, and this they would not be if the critic were only committed to the choice of a test by his particular tastes in the matter of testing. For without a basis for insisting that a certain test is the best test for the hypotheses at hand, conflicting inferences will be able to be drawn from the same evidence, and the critic will be able to choose his test to protect his cherished beliefs, and in so doing expose the lack of rational foundation of the testing procedure.

d. The Long Run Justification.

We have now the final element of the Neyman-Pearson system of statistics to discuss, namely the justification for using tests which meet their criteria for optimality. As we have already seen the justification offered depends upon the long run error frequencies characterizing the tests; for example, if we are using a test with size **a** and power **b**, the rationale for employing the test is that in the long run of applications (for which the empirical precondition of the test is met, ie. either H_o or H_a is true) we will not reject true H_o, and thus assert false H_a, more than 100.a% of the time, and we will reject false H_o, and thus assert true H_a, at least 100.b% of the time.(22) Neyman-Pearson statistical tests are thus provided with a simple and appealing objective basis (provided, of course, the empirical presupposition can be justified).

There is a problem, however, since the justification refers to the long run, and thus not to any particular application or finite set of applications. Any actual employment of the test, therefore, must be backed up by an inference from the long run to the short run, and rests upon Bernoulli´s theorem, that theorem allowing us to calculate the probability that a set of **n** tests will contain no more than the expected number of tests which commit either of the two errors. And we know from this theorem that the probability that the tests in a finite set will collectively be as reliable as their long run characteristics promise, comes close to unity even with relatively small sets of tests. There are three problems, however, with the employment of Bernoulli's theorem in defence of the tests.

First, in order to apply the long run error frequencies to some short run of tests we have to assume that the tests we conduct are a random sample from the infinite set of tests. This cannot be literally true, since the tests conducted are not any kind of sample taken from an infinite set, but there does not seem to be any reason to refuse the point which is really required, namely that the tests actually conducted are not atypical in any respect. While the critic could point to an unjustified assumption here, he could make little of it, since there is no obvious basis for casting doubt upon its presumed truth.

Second, and much more important, the use of Bernoulli's theorem sketched above presumes that the tests are applied singly. But they are not; rather, in one enquiry, we will get chains of tests as one set is employed to test the laboratory equipment for sterility, another set to test the purity of reagents, another to test the purity of the strain of biological material being tested, another to check the accuracy of meters and measures, and another to determine the outcome of the experiment, the last being accurate only if all or the others are error free. Once the complex structure of science is recognised, it is plain that for a final test to be error free we must have had a run of quite a few error free tests, for the conclusions of these earlier tests are presupposed in the final test. What we need to calculate, therefore, is not the probability that no more than the expected number of test will contain an error, but rather the probability that **none of n** tests will contain an error, and this probability falls alarmingly as **n** grows.

Third, there is a problem in making sense of the justification via Bernoulli´s theorem within the reliability programme. The problem is that we do not get from Bernoulli´s theorem the assurance that there will be, in a group of n tests of size (say) 0.01, no more than m tests containing an error, rather we only get a **probability** for this error frequency. As Hacking pointed out, the very high probabilities which we can get for a test´s proving reliable are often referred to as ´practical certainty´, or some such phrase, but he rightly declared that such phrases merely conceal the fact that in the justification of Neyman-Pearson tests, which are supposed to involve no notion of probability for hypotheses or a logic of support for hypotheses, just such a notion is employed.(23) Hacking thus poses a dilemma for the reliability paradigm. For its supporters must either adopt a logic of support or confirmation, to give some way of making sense and use of the probabilities assigned to hypotheses by Bernoulli´s theorem - and thus give up the distinctive feature of the reliability paradigm, the rejection of probability for hypotheses in favour of objective error frequencies for statistical tests, and thereby undermine Neyman-Pearson statistics; or admit that the tests are without rational foundation, the long run justification being unavailable because, apart from any other problem, it involves the notion of probability for hypotheses. And Hacking is surely right in his attack.(24)

This is a particularly important problem for the reliability paradigm, for without a notion of probability for hypotheses the long run justification of Neyman-Pearson tests involves just the kind of obscure and unanalysed jump, here from 'low error frequency in the long run' to 'success on this occasion', which Neyman found in Fisher's jump from high probability to acceptance, a step Neyman described as 'taking a calculated risk', or 'an act of will to behave... in a particular manner'. Neyman argued that calling this act of will an inference obfuscated the true nature of learning from experimental results. That charge can be directed also, therefore, at the attempt to find a long run **justification** for Neyman-Pearson tests, whether interpreted as rules of behaviour, or rules of inference.

This completes my discussion of Neyman and Pearson's contribution to the reliability paradigm. For the sake of balance in my critique, it is worth pointing out that I have not dealt with all of the criticisms which have been levelled at Neyman-Pearson tests. Criticisms which are prominent in the literature but not touched upon here include Hacking's claim that Neyman-Pearson tests are essentially suited for before-trial analyses, while we are often able to employ, and are advantaged by employing, post-trial analyses; that counter-intuitive tests can be constructed using the possibility for mixed tests that the reliability programme allows; and that for just significant outcomes, the criteria of optimality allow tests which behave in paradoxical ways.(25) The problems identified above, however: the non-uniqueness of optimal inferences on Neyman-Pearson criteria, due to no uniquely best combination of sample size, and test size and power; the need for empirical assumptions; and the failure of the long run reliability justification; these are the main problems for the those who profess to find a solution to Hume's problem of induction in the statistical methods recommended by the reliability paradigm.

5. THE DECISION-THEORETICAL CONTRIBUTION TO THE RELIABILITY PARADIGM.

Although I have already rejected any characterization of inference as decision-making, at least in the absence of a satisfactory theory of epistemic utility, it will be useful to see if the criticisms we have made of Neyman-Pearson statistics could be overcome if we adopted a decision-theoretic approach to statistical tests. Thus we shall finish our analysis of the reliability paradigm with a brief discussion of Wald's development of Neyman's conception of statistics.

As we have already seen the Neyman-Pearson analysis of hypothesis testing does not select any one test as the uniquely best for any given situation. Rather the statistician is required to first decide what risk of an error of the first kind he can afford, and then to adjust the choice of critical region to achieve the desired test size, with as high a power as can be got without requiring too large a sample. Given any situation, then, the decision to set the size of the test at a certain level and not to reduce it by increasing sample size (thus also increasing the cost of sampling), or allow it to increase by making do with a smaller sample, and the similar balancing of test size against power, are not determined by the theory. How to fix the independent variables is a problem, as we have already remarked, that Neyman and Pearson left to the good sense of the scientist.

Neyman did begin a line of argument which offers hope of bringing the initial set of decisions under the rule of law. As Giere nicely puts it, once we think of tests of hypotheses as the basis of actions in practical contexts, once, that is, that we decide to act in a certain way given the acceptance of a certain hypothesis, it becomes natural to attach weights, perhaps even monetary weights, to the consequences of the various acts open to us in each of the possible states of nature.(26) But although the idea of inductive behaviour leads naturally to the introduction of this idea, and provides a simple solution to the partition problem - choose as H_o the hypothesis which, should it be erroneously believed false, would lead to the worst loss - it provides no theoretical framework in which to consider the optimality of a given test in terms of the gains and losses attached to the various possible actions in the possible states of nature. Once we adopt a set of parameters permitting the determination of these gains and losses, we need a method for weighing them, or rather their expectations, against one another, and this was provided by Wald, bringing together the game theory of von Neumann and Morgenstern, and Neyman-Pearson statistics.

If we follow this line of development we end up with a set of interrelated costs and gains dependent upon the unknown state of nature. There is thus a problem of defining the best decision function, ie. the best decision-making strategy. For it will not generally be the case that any one decision function, ie. any one sample size plus critical region, will maximize our gain or minimize our loss in all possible states of nature. Rather, it is usually the case that one choice will do best if one state of nature obtains, another in another state, and thus the introduction of gains and losses does not lead immediately to a basis for chosing an optimum test. Rather it poses the general problem of chosing the optimum decision function. Now one decision rule, the minimax rule, which requires choice of the decision function which has the smallest maximum loss, has gained wide acceptance, but there is doubt over whether it can be meaningfully applied in the scientific context. For seeking to guard against the worst possible outcome seems unduly cautious unless Nature is out to get us by waiting until we adopt a belief which will lead us badly astray if the state of the world is not such as we believe it to be, and then organizing things to realize this state. Why not believe, on the contrary, that Nature wishes to be understood, and thus choose the decision function which maximizes the possible gain? Clearly either choice, or any other which attributed to nature the psychology of the player of a competitive, or co-operative, or any other kind of game, badly strains whatever analogy there might be between science and playing games.(27)

I do not intend that these brief comments should be taken as any kind of thorough critique of the decision-theoretic approach to statistics. Rather my intention has been to quickly sketch out where the introduction of the consequences of the acts associated with adopting various beliefs leads us, and to indicate that even if it solved the problem with which it does give some aid, the problem of finding criteria to select an optimum test, it would leave the other problems with Neyman-Pearson statistics untouched. And recalling, of course, that the introduction of consequences of acts threatens the objectivity of science, and thus may introduce more problems than it solves, it seems clear that there is no solution to Hume's problem to be found in the reliability paradigm, whether as a distinctive approach to the nature of inductive inference, or as the basis for a thorough-going decision-theoretical interpretation of science.

That judgement requires qualification only by pointing out that it rests heavily on maintaining Hume's foundationism, his insistence that the premises of a justifying inference be justified in their turn. Should we give that up, as Giere recommends, then the attraction of the statistical methods of the reliability paradigm, particularly in view of the problems of the other paradigms we have yet to investigate, would be greatly increased.

CHAPTER 6. THE CONFIRMATION PARADIGM.

1. INTRODUCTION.

The philosophers and statisticians who have sought to provide forms of inductive inference which measure the levels of support offered the conclusions of inductive inferences by their premises, leaving the Bayesian tradition until the next Chapter, may be divided into two groups. In the first group are Fisher and Hacking, in the second Keynes and Carnap; and of course there are others who have made greater and lesser contributions to the lines of thought initiated by Fisher and Keynes respectively. But my discussion, since it deals with the common basis of the two variants of the programme, is restricted to the main figures. Indeed, my discussion will focus on Fisher and Carnap, the most influential figures in the two traditions (Carnap's work superseding Keynes'), and also on Kyburg, whose work combines elements of both traditions to form a new and powerful conception of inductive inference.

What I aim to show in this chapter is that each of the inference schemes recommended by the authors to be discussed is not logically compelling, or that it is logically compelling only if we grant some assumption which is equivalent to an empirical claim itself requiring the support of an inductive inference. This discussion thus extends to further theories of induction, theories based on logical conceptions of probability, the kind of critique of statistical inference directed at the reliability paradigm in the previous Chapter, and leads to the conclusion that there is no solution to Hume's problem of induction to be found in the statistical systems put forward in the confirmation paradigm. For the critic is left free to reject the conclusion, or to refuse to attach a specified probability to the conclusion, of all inductive inferences judged correct by the standards laid down by the confirmation paradigm, since he can adopt this position without transgressing against a principle of reason to which he can be committed.

2. FISHER'S THEORY OF INDUCTIVE INFERENCE.

a. Fisher on Induction

Fisher makes it plain that his aim in proposing various statistical and experimental procedures is to justify inductive inference (or rather inductive inferences of certain types). Indeed, he takes inductive inference to be at the very centre of statistics, finding in the history of the subject,

> the steps, some hesitating, some even false, by which men have come gradually to understand how their reason may be applied to uncertainties, yet applied with logical rigour, and how, in particular, it may be applied to observational facts with all their paucity in number and their imperfect precision, and yet draw from them precisely those inferences which the observations warrant.(1)

High on the list of false steps Fisher placed the attempt to use Bayes' theorem as the standard rule of inductive inference; like other frequentists, Fisher thought that **in general** Bayes' rule could not be employed as a basis for inductive inference because **in general** we cannot find any rational basis for assigning probabilities a **priori**, on which application of the theorem depends.(2) The frequentist's position is well captured in Fisher's claim that in using the rule generally,

> advocates of inverse probability seem forced to regard mathematical probability, not as an objective quantity measured by observable frequencies, but as measuring psychological tendencies, theorems respecting which are useless for scientific purposes.(3)

Thus while on occasions we may be able to find acceptable prior probabilities and then use Bayes´ theorem to give us the probability of an hypothesis under test, or an inverse probability as it is sometimes known, in general some other method for assessing inverse probabilities, or some form of inference from sample to population which does not aim to assign a probability to the hypothesis under examination, must be found. Fisher saw clearly what was at stake in rejecting Bayes´ rule as a basis for inductive inference in general, and set himself the task of supplying a substitute form of inference to do the work of the rejected procedure, warning that

> the rejection of the theory of inverse probability was for a time wrongly taken to imply that we cannot draw, from knowledge of a sample, inferences respecting the corresponding population. Such a view would entirely deny validity to all experimental science.(4)

In order to secure the validity of experimental science, or more briefly, induction, which is clearly the problem under discussion by him, Fisher proposed three distinct forms of statistical inference (in addition to the use of Bayes´ theorem) each appropriate to distinct conditions of information and ignorance, and each allowing inverse, ie. inductive, inference, to proceed. The three forms of inference are significance testing, estimation (according to the procedures Fisher laid down), and fiducial inference. Of these inference forms Fisher wrote:

> In fact, in the course of this book, I propose to

> consider a number of different types of experimentation with especial reference to their logical structure, and to show that... entirely valid inferences may be drawn from them... If this can be done, we shall, in the course of studies having directly practical aims, have overcome the theoretical difficulty of inductive inferences. (5)

This claim we shall be concerned to evaluate in our discussion.

b. Tests of Significance.

In one of his clearest passages on the point and purpose of tests of significance Fisher wrote:

> In general, if, in connection with a given observational record, a hypothesis is considered which is well defined in the sense that from it can be derived definite expectations, we may use the observations to test whether these expectations have been realized, or whether, on the contrary, the observations depart so far from expectation in some relevant respect that the hypothesis under consideration must be deemed to be contradicted by the data, and must be abandoned. In the latter case the deviations from expectation are deemed to be significant, while in the contrary case, if the observations are such that with reasonable probability they might have arisen on the hypothesis under test, this hypothesis, though not proved, has at least so far been confirmed, and, pending further and more stringent observations, may be accepted.(6)

I want to draw out two features of significance tests which are introduced in this description. First note that tests of significance are clearly presented as an initial stage of a perhaps longer process; if the hypothesis passes the test it becomes the working hypothesis and goes on to more elaborate and more stringent trials, while if it fails then the significance test is the terminus of that enquiry. Thus, so far as any final verdict is concerned, a significance test can only have one outcome, namely test failure resulting in the rejection of the hypothesis. A significance test, if successful, does not lead to the final acceptance of the hypothesis under test, nor to the acceptance of any rival hypothesis, since (unlike Neyman-Pearson tests) no other hypothesis is involved in the test.(7) (According to Fisher the consideration of rival hypotheses in bidding for confirmation comes only when an hypothesis embodying a statistical model has passed a test of significance and has thus been adopted as a working hypothesis, at which time we consider which of the various hypotheses assigning rival values to parameters of the model is best confirmed by the evidence.(8))

So far as the logical cogency of tests of significance is concerned, then, there are two questions to be decided: does failing a test of significance constitute a reasonable basis for the rejection of the hypothesis tested?; does passing a test of significance constitute a reasonable basis for adopting the statistical model embodied in the hypothesis as the basis for further research and testing? Before we can answer these questions we shall have to get clear the concept of significance.

Considerable confusion surrounds Fisher's concept of significance. On the one hand he used the term, as in the passage quoted above, to refer to experimental results which differ so markedly from the predicted result, ie. lie so far out on the tail of the hypothesized distribution, that we cannot reasonably conclude that the hypothesized distribution is the true distribution. This is the sense in which the term has been used by other statisticians influenced by Fisher.(9) On the other hand, however, when Fisher explicitly defined the term and, moreover, in his most self-conscious elaboration of the logical cogency of tests of significance, he identified the significance of a result with its low probability.(10) Now, as many authors have pointed out, this second definition leads to ridiculous results, since it will often be the case that every possible outcome of an experiment will have a probability below that of the significance level.(11) But when this was pointed out to Fisher, however, as Savage tells the story, he responded evasively.(12) We need not be concerned with this, however, for we have Fisher's practice to go by, and this, for example in his famous analysis of the test of the lady's powers to discriminate between cups of tea made in different ways, guides us to take the significance level of a test of significance to refer to the probability that a result in a tail area of the distribution will occur. (13) More precisely, using a suggestion of Seidenfeld's, let us say that an experimental outcome is more or less discrepant (from the expected outcome) depending on whether the reciprocal of its probability is greater or lesser; outcomes with the lower probability are thus said to be more discrepant.(14) An experimental outcome is then defined to be significant if the sum of the probabilities of that outcome plus any outcome more discrepant than the actual outcome does not exceed the significance level of the test.

Fisher claimed that an hypothesis which is subject to a test in which a significant result has (repeatedly- I ignore this qualification here) occurred may be reasonably rejected, citing as the basis for the inference that

> the force with which such a conclusion is supported is logically that of the simple disjunction: **either** an exceptionally rare chance has occured, **or** the theory of random distribution [ie. the hypothesis under test- MR] is not true. (15)

Now clearly this disjunction has no force if we identify significance with the rarity of the experimental outcome, for as Hacking noted, there are some distributions for which any particular experimental outcome will be rare. But while the claim is no longer obvious nonsense if we define ´significance´ in the manner suggested, it is not yet clear that the disjunction has the force Fisher intended. For it will be the case, after all, that significant outcomes will occasionally occur in the course of testing a true hypothesis - indeed, they will presumably occur in a fairly long run of tests with a frequency close to their probability. When a significant result occurs, therefore, why cannot we put it down to chance and retain our hypothesis? The logical cogency of tests of significance clearly requires further discussion.

What could be the rational basis of rejecting an hypothesis for the reason that its truth would entail the occurrence of a significant outcome? Let us examine Fisher´s own explanations, noting that whatever is the answer to this question constitutes whatever rational justification there might be for the use of tests of significance as Fisher expounded them. We begin with his earliest papers and work our way through to his most detailed analysis, given in his [1956].

There is not much to work through: surprisingly, Fisher provided, so far as I have been able to discover, only three accounts of the logical force of his tests of significance. In his [1935a] he wrote:

> It is usual and convenient for experimenters to take 5% as a standard level of significance, in the sense that they are prepared to ignore all results which fail to reach this standard, and, by this means, to eliminate from further discussion the greater part of the fluctuations which chance causes have introduced into their experimental results. No such selection can eliminate the whole of the possible effects of chance coincidence..... for the ´one chance in a million´ will undoubtedly occur, with no less and no more than its appropriate frequency, however surprised we may be that it should occur to us. (16)

It would seem here that Fisher identifies the significance level of a test with the long run relative frequency with which the test would yield a significant result purely by chance (ie. would yield an outcome very different from that expected even though the tested hypothesis is true), and suggests that by keeping the significance level low we will ensure that only rarely will we be misled by a test of significance to reject a true hypothesis. If this is not yet a reliability paradigm interpretation of the logical force of a statistical test it is well on the way to such an analysis, as Giere argued.(17) We need not, therefore, comment further upon this suggestion of Fisher´s, since we have already dealt with this analysis of the logical force of statistical tests in the previous Chapter, and thus have covered this interpretation of the logical force of tests of significance.

Fisher did not return to the question of the logical basis of tests of significance until 1955 when he claimed, in a discussion critical of the reliability interpretation of statistical procedures, that

> from a test of significance... we learn more than that the body of data at our disposal would have passed an acceptance test at some particular level; we may learn, if we wish to, and it is to this that we usually pay attention, at what level it would have been doubtful; doing this we have a genuine measure of the confidence with which any particular opinion may be held, in view of our particular data.(18)

While this passage is clear in its rejection of the reliability interpretation of tests of significance, arguing that significance levels ought to be interpreted as supplying a measure of the confidence we can rationally place in the hypothesis tested and thus giving a confirmational interpretation of tests of significance, it sheds no light at all on the rational basis of the tests when so interpreted. We therefore turn to Fisher's [1956] where he pursued this interpretation, first conducting a sketchy psychological analysis of the mind of an experimenter who has just recorded a significant result, in the hope of showing that this mental state is not what the reliability or Bayesian analysis of statistical tests would lead us to expect, and then turning to the rational basis of this psychological state:

> Though recognizable as a psychological condition of reluctance, or resistance to the acceptance of a proposition, the feeling induced by a test of significance has an objective basis in that the probability statement on which it is based is a fact communicable to, and verifiable by, other rational minds. The level of significance in such cases fulfils the conditions of a measure of the rational grounds for the disbelief it engenders.(19)

Here Fisher seems to take the position that the explanation of why having a significant result recorded against it constitutes an adequate ground for rejecting an hypothesis is that it is definitively reasonable to do so, ie., the fact that it is reasonable to reject an hypothesis which fails a significance test is **definitive** of reasonableness (at a level of stringency determined by the significance level of the test).

I am suggesting, that is, that Fisher proposed **as a principle of reason** that it is reasonable to reject (at the level of stringency of the test) an hypothesis which has failed a test of significance.

Fisher did not offer anything in defence of this principle of reason, at least not by way of argument- presumably he trusted that his examples of tests based upon the principle would, by appealing to the scientific good sense of his readers, show the principle to be worthy of adoption. But an application of a test of significance may be judged reasonable for some other reason - for example because the test can be given backing by an interpretation within the reliability paradigm; the principle itself thus requires examination and defence. We shall scrutinize it in the light of Spielman's careful critique of tests of significance.

Spielman interprets Fisher's claim about the logical basis of tests of significance in much the same way as we have proposed here, reconstructing Fisher as asserting that the level of significance at which a test rejects an hypothesis is an index of the level of its objective incredibility.(20) He then argues that we may not have the low level of credibility in an hypothesis rejected by a test that Fisher's principle requires, nor is there any compelling reason for us to do so. For, he continues, the observed result may be even more incredible on the assumption that the hypothesis in question is false, for we may have in mind no other plausible hypothesis which could account for the test result, and in this case the test provides no objective basis for assigning the hypothesis a low level of credibility; while if our level of credibility in the hypothesis prior to testing is very high we may judge the test outcome to be just a rare chance outcome and continue to assign the hypothesis a high level of credibility. On the basis of these criticisms (which I think well directed) Spielman goes on to try to salvage tests of significance by incorporating a requirement about the subjective prior credibility of the hypothesis, and the rarity of a significant result on the assumption that the hypothesis is false; both of which are, as he says, easier to incorporate in a Bayesian model. (21)

I do not think that Spielman's salvage operation is successful, however, for it does not answer our question about the logical force of tests of significance. His requirement concerning the rarity of a significant result in the test of an hypothesis is merely that significant results be sufficiently frequent if the tested hypothesis is false, in comparison to their frequency if the hypothesis is true, to enable us to rationally reject the claim that a significant result in some test is just a chance rare outcome in favour of the claim that the tested hypothesis is false.(22) Now this is just the suggestion Hacking made for the repair of tests of significance after his critique of Fisher's attempt at showing the logical force of the tests: drawing on a point of Laplace´s Hacking asserted that it is not the improbability of some event on some hypothesis that makes us reject the hypothesis, but rather that there is some alternative hypothesis on which the event in question has a significantly higher probability.(23) Thus Hacking suggested that tests of significance cannot be conducted on single hypotheses but rather must be conducted as tests to choose between rival hypotheses, Spielman weakening this requirement to the condition that the frequency of significant results on the assumption that the tested hypothesis is false must be part of the information on which the test is based. But neither proposal clarifies the logical force of tests of significance. For say that we know that if H is false then H´ is true, that the probability of a significant outcome if H´ is true is much greater than if H is true, and that we observe a significant outcome in a test of H: what is to stop us opting for the conclusion that a rare event has occurred by chance (as is bound to occur now and again, at least in the long run)?; what is the logical force of the test which might make it rational to reject H for H´?

As far as I can see knowing that the observed outcome is more probable on H´, ie. if H is false, than it is if H is true, does not advance us at all in our search for a rational basis for tests of significance; since it might, after all, be the case that H is true and the result is just a chance outcome. Which is not to say that Hacking´s and Spielman´s critiques are without point: either allows a reliability paradigm justification to be given for tests of significance, which cannot be given unless the tests are amended in one of the ways suggested; for, to use

Spielman's version, unless we know that the observed outcome would be a less rare event if H were false, we certainly cannot assert that in rejecting hypotheses which fail significance tests we minimize the long term frequency with which we adopt false hypotheses. But we have seen in the previous Chapter that such long run justifications are problematic; and besides, that is not the justification Fisher wanted to give for tests of significance.

So far, then, we have been unable to find any rational basis for rejecting an hypothesis following its failure to pass a test of significance. I expect that their widespread use could be traced to statisticians' willingness to argue that they have worked in the past, ie. that hypotheses rejected after failing significance tests have regularly been shown to be false by the later acceptance of a rival of the rejected hypothesis, and so the tests can be relied upon in the future; that is to say I expect that an inductive justification would commonly be offered for their use. But such a justification cannot be accepted here unless Hume's critique of such arguments is first answered.

How then do tests of significance fare as inferences supporting hypotheses which pass the test? First it must be said that they are ill-suited to the role of confirmation. Clearly it would not be reasonable to take the non-occurrence of a significant result to offer the tested hypothesis any worthwhile support, for the result may be quite discrepant even though non-significant. Nor would it be reasonable to raise the significance level to some high figure, say 95%, for we would then be forced to reject hypotheses which are quite likely true. One plausible interpretation of tests of significance as confirming inferences is to consider the test to be a procedure for choosing between a pair of hypotheses, one being confirmed by the rejection of the other at a suitably low level of significance - but of course Fisher forthrightly rejected such an analysis, and we have already dealt with it, since it leads to the Neyman-Pearson testing model.

The suggestion Fisher made, in the passage first quoted in this discussion of tests of significance, is that an hypothesis is confirmed by a test of significance ´if the observations are such that with reasonable probability they might have arisen on the hypothesis under test´, which makes the probability of the result given the hypothesis, or the likelihood of the hypothesis on the evidence, the basis for tentative acceptance, or rejection. Perhaps we could follow this hint and base a theory of confirmation upon likelihood. We shall not make such an attempt, however, for it can be shown that no such test provides a way of escaping Hume´s inductive scepticism.

To see this, consider a test which yielded a result giving maximum likelihood to the hypothesis tested. This would occur when the probability of the result given the hypothesis was unity, ie. when the hypothesis entailed the evidence. Such a test is just an ordinary hypothetico-deductive test of an hypothesis, and such tests do not have to be subject to Hume´s criticism to be found wanting; indeed, hypothetico-deductive inference is merely the fallacy of affirming the consequent ´more politely described´.(24)

Fisher´s first form of inference offered as a solution to the ´theoretical difficulty of inductive inference´ thus proves to offer no help at all to the the philosopher in search of relief from inductive scepticism, because the critic is under no compulsion of logic or clear reason to accept the verdict of such a test. For if an hypothesis passes a test of significance it is supported by nothing stronger than a hypothetico-deductive inference; while if an hypothesis fails to pass a test of significance, that does not constitute a reasonable basis for requiring abandonment of that hypothesis, since, after all, the hypothesis predicts, albeit with low probability, the result which is now offered as evidence against it. We now proceed to consider the second inductive method proposed by Fisher.

c. **Estimation.**

As noted in the discussion of tests of significance, Fisher intended estimation to be the second stage of the scientific examination of some phenomenon. This second stage presupposes that the result of the first stage test of significance has determined to which family the distribution of the phenomenon under study belongs, and the task now is to decide which member of this family represents the distribution in question, requiring that a parameter or parameters characterizing the family of distributions be specified. This specification is the task of estimation. Fisher made this plain:

> Problems of estimation arise when we know, or are willing to assume, the form of the frequency distribution of the population, as a mathematical function involving one or more parameters, and wish to estimate the values of these parameters by means of the observational record available.(25)

Estimation is thus a secondary method, in Reichenbach´s classification; and one which, as we have just seen, is in difficulty with its empirical presupposition (the determination of the family of distributions involved in the problem of estimation) which must, given the failure of tests of significance to rationally support tested hypotheses, be taken as an assumption.(26) As a complete method of induction, therefore, an initial problem with Fisher´s theory of estimation is that its rational justification will require a non-foundationist conception of justification to be shown to be adequate.

Placing the problem of the justification of presuppositions to one side, the interest in Fisher´s theory of estimation centres on the possibility that it incorporates a further challenge to Hume´s conception of reason, viz, an alternative to the idea that it is reasonable to make only those inferences known to be truth-preserving. We shall see that for a certain class of inferences Fisher does aim to provide just such an alternative standard of reason; or rather, bearing in mind the opacity of his writings, at least to the penetrative ability of this reader, we should say that such an alternative can plausibly be extracted from Fisher´s discussions of estimation.

As Fisher explicitly defines his notion of estimation it is hard to see what estimation has to do with induction. Fisher´s definitions vary from the utterly inscrutable to the circular version quoted above. But most of Fisher´s comments on the evaluation of estimates, especially in relation to his main innovation in criteria for good estimates, the criterion of efficiency, present estimates as summaries of the data. This reading of Fisher has been adopted by both Hacking and Seidenfeld, Hacking claiming that

> in Fisher´s opinion, an estimate aims at being an accurate and extremely brief summary of the data bearing on the true value of some magnitude. Closeness to the true value seems to be conceived of as a kind of incidental feature of estimates.(27)

The problem here is that unless we show the superiority of some other conception of estimation, the idea that an estimate is good or bad depending on how accurate it is defines the notion of estimation itself. Therefore what is required to show the logical structure and rational basis of estimation on Fisher´s model is either a connection between an estimate´s meeting Fisher´s criteria for good estimates and the accuracy of the estimate, or some analysis of inductive inference qua estimation which shows why it is reasonable to accept an estimate meeting Fisher´s criteria, rather than to insist that accuracy is an necessary criterion for estimation.

Hacking, rejecting Fisher´s conception of estimation as a misguided point of view, developes the first option. We shall review his efforts below. For the moment, however, let us try to extract from Fisher some alternative conception of the rational foundation of inductive inference in problems of statistical estimation.

Fisher does, I think, propose a novel conception of reason in relation to statistical estimation, namely, that those inferences are reasonable which maximize the relevant information available from the data. Though he never puts the matter quite so baldly as this, my claim is well supported by his usual discussion of estimation. The point is most clear in a comment in his main paper on estimation, where he summarized his aim in the following manner:

> Finally it may be possible to prove ... that a particular statistic summarizes the whole of the information relevant to the corresponding parameter, which the sample contains. In such a case the problem of estimation is completely solved.(28)

I do not think that my claim will be controversial, but in any case it is of some interest, quite apart from what Fisher's theory of estimation actually was, to see if we can extract from his work a novel principle of estimation which might not be susceptible to Hume's inductive scepticism.

What I take to be Fisher's principle looks attractive. He proposes, in effect, that we stop trying to find forms of inverse inference - and a rule of estimation constitutes such a form of inference - which we can guarantee will be truth preserving, and look for rules of inference which make the best use of the data available. That seems to supply a justification for the inference, for while we do not know that it will not lead us astray, we do know that there is no further guidance to be got from the data we have, or that the inference is the most soundly based we could make. But there is one proviso that must be noted here, and it turns out to vitiate Fisher's proposal. For that it is a virtue of an estimate to make the fullest use of the data available presumes that that data is a sound basis for the estimate. Therefore, for Fisher's criterion to be a reasonable basis for selecting an estimate we must be in a position to know that the evidence we have is not misleading; to know, that is, that we would not do better by using just a part of the data we have available, or by ignoring it altogether. Now Fisher did provide a guide here since one of his criteria for estimation is that the data upon which the estimate is based must be relevant. It is to this criterion that we first turn our attention.

Fisher's principle of relevance was consistency. He wrote:

> The fundamental criterion of estimation is known as the Criterion of Consistency, and is essentially a means of stipulating that the process of estimation is directed to the particular parameter under discussion, and not to some other function of the adjustable parameter or parameters.(29)

Consistency, and this is the most important point about it, is a property of the process of estimation, ie. of the function or statistic whose values we take as estimates; consistency is not a property of the individual estimates. This is made clear in the passage above, as it is in most places in Fisher's writings, but he slipped at least once into asserting that

> an inconsistent estimate is an estimate of something other than that which we want an estimate of.(30)

This claim may well leave the unwary with the impression that it makes sense to speak of a consistent estimate (rather than a consistent estimating function, or estimator) and that a consistent estimate **is** an estimate of the parameter we are trying to estimate. If we fall into that confusion, we are then led on to ask how good an estimate is some particular value we have mistakenly called ´consistent´, and when we discover that the estimate is produced by an estimator which has minimum possible variance for that problem of estimation (another of Fisher´s criteria for a good estimate), and especially if we understand this to mean that the estimate itself has the minimum possible error, then at the end of that line of thought it seems that we have the problem of induction solved for any problem to which the theory of estimation yields an answer. But an error was made in the first step, in applying the term ´consistent´ to the estimate rather than to the estimator.

I am not suggesting that Fisher himself ever offered that confused line of argument; rather my point is that it is by such an argument that one might mistakenly think that one has found in Fisher´s work on estimation a solution to Hume´s problem of induction for inferences to which Fisher´s theory of estimation applies. However, quite apart from any logical flaw in the argument cited, it is not supported by Fisher´s theory of estimation, as investigation of that theory will show.

Fisher was keenly aware that not all seemingly desirable properties of esimators are of any use as a guide to the reasonableness of the associated estimates. Thus he rejected a common formal definition of ´consistency´ - that the estimator should take the actual value of the parameter on a sample consisting of the entire population, or that as the size of the sample grows without limit the estimate should tend to the actual value of the parameter. As he pointed out, on this definition any estimator at all can be made consistent by constructing a function of both the estimator of our choice and some consistent estimator such that on small samples the chosen estimator dominates while on very large samples (or on the whole of a finite population) the position is reversed.(31) Fisher avoided this problem with the received definition by offering a definition which made consistency a property of an estimator on samples of all sizes, adopting as a definition that to be consistent a statistic must take the actual value of the parameter on every sample in which the observed frequencies on which the estimate is based take their expected values.(32)

Fisher´s innovation solves the problem of the relationship between consistency in the long run and consistency as a property of an estimator employed on small samples, making consistency a more useful characteristic of estimators. But it does not solve the problem of securing the relevance of the estimate, for what is required to ensure the relevance of a particular estimate is that consistency of the estimator be related to the estimate itself; in short it does not solve the problem of the relationship between the property of the estimator and the property of the estimate.

Why should we take an estimate provided by a consistent estimator to be relevant to determining the value of some parameter? There seems to be no reason, since we do not know how far our particular sample diverges from that expected and thus we do not know whether a consistent estimator ´directs us to the particular parameter under discussion´ or directs us away from it. Of course, if we adopt a conception of estimation according to which the best estimate is the one closest to the true value, and accept that while we cannot pick this estimate in any particular case we can employ an estimator which in the long run will frequently give us values near the true value, ie. accept a reliability paradigm conception of the rational basis for using a particular estimator, then we have some reason for accepting the estimate of a consistent estimator. But this approach is far from problem free, as we have seen in the preceding chapter; in particular, the need to **assume** the functional form of the distribution (given the failure of tests of significance to provide adequate support for successfully tested hypotheses) becomes crucial here, for a mistaken assumption will generally give us a false value for the long run error frequency associated with a particular test. We conclude, therefore, that Fisher failed to show that consistency, his principle of relevance, provides the justification needed to make the choice of an estimate based upon rational grounds.

As to Fisher´s second criterion, efficiency, this provides no basis for estimation by itself. Being defined in terms of minimum variance, a statistic which always took the same value regardless of the sample would be maximally efficient - though, of course, inconsistent, (but we have found no basis for insisting on consistency). Fisher's theory of estimation must therefore be rejected as an answer to Hume's inductive scepticism, regardless of its relative virtues as a statistical procedure when doubts about the rational foundations of induction in general are not on the agenda.

d. **Fiducial Inference.**

Among the statements chosen by Edwards to set the scene for his [1972] was a remark by Devenant (1698) in which what we would today describe as logico-mathematical methods in the sciences are referred to as 'The Art of reasoning upon Things by Figures'. The mystery, if not magic, conjured up by this wonderful phrase nowhere makes itself more keenly felt than in the statistical inference form Fisher called fiducial inference.

I think that the reason for the air of other-worldness that clings to fiducial inference is that there has thus far been, so far as I am aware, little attempt to reconstruct the intuition which led Fisher to put the inference scheme forward as a general method of inverse inference. Fisher's own discussions consist not of analyses but of examples; indeed, Seidenfeld declares that it is a misnomer to refer to the fiducial 'argument' on the grounds that all Fisher left us was a 'poorly sketched technique'.(33) This tradition extends to the critical literature which, judging from Seidenfeld's review, has proceeded to examine fiducial inference by proposing counter-examples to what were apparently Fisher's rules for sound fiducial inference. Notable exceptions to this trend are Hacking and Kyburg, both of whom have been at pains to set out the logic of the inference in clear detail, but the inference scheme which each lays out is his own product, not Fisher's fiducial inference, even if that be the inspiration for these later suggestions.

We can find the intuitive basis of fiducial inference by examining the logical structure of an an easily accessible inference Fisher himself put forward as a basis for 'examining the logical cogency' of his fiducial argument.(34) He clearly took this inference to reveal the intuitive basis of fiducial inference, declaring that his new inference form could only be rejected 'by those willing to reject this simple argument'. We shall find, however, that Fisher's 'simple argument' is by no means straightforward, and that, to judge from the line of argument Neyman pursued in relation to fiducial inference proper, he at least might have rejected the 'simple argument' for the kind of reason we shall give. Accepting, then, Fisher's claim that the 'simple argument' does reveal the logical basis of fiducial inference we shall not push on to examine that inference form proper, since the genuine inference form would share the problems of the 'simple inference', problems to which Fisher's theory of statistical inference provides no answers. On this basis we shall conclude that fiducial inference offers no solution to Hume's problem of induction.

The argument of Fisher's referred to in the previous paragraph may be put thus. Let there be a population of trials on a random variable x of which nothing is known save that its distribution function is continuous. Let m be the median of the distribution, and a be the number of outcomes of trials on x yielding values less than the median in a set of trials randomly chosen from the population. Consider a case in which this set of trials on x has only one member, x = 0. Fisher suggests that we may argue as follows:

	1) $P(a = 1 \mid x) = 1/2$
	2) $x = 0$
therefore	3) $P(a = 1 \mid 0) = 1/2$
therefore	4) $P(m > 0) = 1/2$

In order to clarify the importance of examining this inference note that from two premises, the first being an immediate consequence of the definitions of m and a, and the other a single observation, we are apparently entitled to conclude that the conclusion of an inductive inference (for the inference set out is inductive) has a certain definite probability. If this inference stands scrutiny, therefore, for some concept of probability which provides an adequate account of the link between probability and rational degree of belief, it would seem that Hume's problem is well on the way to solution for the class of cases to which such inferences would apply. But does the inference go through? This cannot be answered until the inference is substantially clarified.

What mainly needs clarification here is the meaning of the term 'probability' as it is used in the various lines of the inference. It seems clear that at least in the first premiss stochastic or factual probability is being referred to, for no justification is given for (1), and if it is a statement of factual probability it requires none, being an immediate consequence of the definition of 'median'. But any other concept of probability would make the assertion of (1) a less trivial matter. To make its meaning plain (1) should be rewritten as

$$(1')\ P_2[M_s](a = 1 \mid x) = 1/2$$

indicating that the probability is a probability$_2$, in Carnap's classification, and is based upon the stochastic model $\mathbf{M_s}$ representing the set of trials on the random variable $\mathbf{x}$.

Is (3) also to be interpreted as a probability$_2$ claim? It is clear that Neyman, a thoroughgoing frequentist who did not use the concept of probability$_1$, thought so, for the relevant section of his critique of fiducial inference proceeds much as follows.(35)

We cannot make sense of (3) as a probability$_2$ claim as easily as we were able to deal with (1), for in M_s $\mathbf{m}$ is a not a random variable but an unknown constant. The expression '$\mathbf{P_2[M_s](a = 1 \mid 0) = 1/2}$' therefore makes no sense, for once $\mathbf{x}$ is known to be 0 there is no random variable in the expression and thus no basis for a probability$_2$ claim (other than trivial claims like '$P_2[\mathbf{M_s}](\mathbf{x = 0 \mid x} = 0) = 1$'). If we are to make sense of (3) as a probability$_2$ claim we must construe $\mathbf{m}$, the median of $\mathbf{M_s}$, as a random variable, or rather invent a new random variable $\mathbf{m_x}$, the median of the distribution $M_{\mathbf{x}}$, where $M_{\mathbf{x}}$ is the distribution model for a population of trials on **x** randomly selected from a population of such populations (or super-population, to use Fisher's term) ~~M~~.(36) An example will facilitate further discussion of this idea. (Note that the example employs a discontinuous distribution, against Fisher's explicit instruction for performing fiducial inference; however, no criticism to be made here hangs upon the fact that the distribution is not continuous).

Let a large container be filled with dice of unknown shape and markings and let the dice be tossed from the container so that they fall randomly with respect to the face showing on any given die. Let **x** be a random variable whose ith value is determined by the number of spots on the uppermost face of the ith die to be examined. Consider M_s, the set of values of **x** generated by tossing the die once. Let the median of M_s be m_s and the number of values of **x** less than m_s in an n-membered subset of M_s be a_s. Clearly we may now assert, for $n = 1$ and a container filled with an odd or infinite number of dice,

$$(1)\ P_2[M_s](a_s = 1 |\ x) = 1/2$$

and of course there is no problem with

$$(2)\ x = 0$$

But what shall we write for (3), and can this claim be validly inferred from (1) & (2)?

To find a well formulated probability$_2$ claim to put for (3) we must first identify the random variable which is to be the subject of the claim. To this end consider **M+**, the set of sets of values of **x** arising from a series of tosses of the container full of dice, **m+**, a random variable whose ith value is given by the median of the ith set of values of **x**, and a_i, the number of values of **x** less than **m+** in an n-membered subset of the ith set of tosses of **x**. It now makes sense to claim, for **n** = 1,

$$(3') \; P_2[M+](a_i = 1 \mid x_i = 0) = 1/2$$

and from here the remainder of the inference goes through without trouble. But the inference from (1) & (2) to (3´) is invalid. Indeed, (3´) is a claim about a new random variable defined on a different population. There is no plausible route from (1) & (2) to (3´) at all; rather we have made further assumptions from which (3´) follows just as (1) follows from the initial model defining **x** and $\mathbf{M_s}$. (1) and (2) are completely independent of (3´).

Clearly for Fisher´s ´simple inference´ to go through we must interpret (3) as a probability$_1$ claim, so that we can maintain it as a probability claim about a and $\mathbf{M_s}$ when **x** is known. But we will then require a theory of probability$_1$ in at least enough detail to tell us whether the probability$_1$ claim (3) can be validly inferred from the probability$_2$ claim (1) and the observation(2). We could begin to attribute such a theory to Fisher by providing a rule of inference to get us from (1) & (2) to (3). The rule which is adequate to this task while attributing to Fisher as minimal a theory of probability$_1$ as is possible is the following:

> (2a) The probability$_2$ of events of kind **e** in an infinite set of trials on a random variable for which events of kind **e** are one possible outcome is to be used as the probability$_1$ that an event of kind **e** will occur on a single trial on the random variable.

Now Fisher did employ such a principle, though he gave it as a rule for determining probability$_2$ values for subsets of the set of events for which the probability$_2$ is defined, rather than as a rule connecting probability$_1$ and probability$_2$.(37) That is not surprising, since the rule can be used to perform both roles, and the distinction between single case probability$_2$ values and probability$_1$ values is required here only to make sense of after trial probabilities, a problem which Fisher does not seem to have addressed.

Our analysis of Fisher's inference concerning the probability of the median of a distribution exceeding the value of the random variable on a single trial has removed one of the barriers, albeit a minor one, to a clear understanding of fiducial inference. It shows, however, the inadequacy of the theory of fiducial inference as Fisher left it, for we have concluded that the inference requires the support of a theory of probability$_1$, and while it is possible to attribute the basis of such a theory to Fisher there is nothing in his writings to guide us in dealing with the many questions which immediately arise concerning probability$_1$. We want to know, for example, just what is a probability$_1$; whether probability$_1$ values are always equal to corresponding probability$_2$ values; the rules that determine which probability$_2$ corresponds to which probability$_1$; whether probability$_1$ values are unique; whether all probability$_1$ values are equally soundly based, and if not in what the difference in their bases consists; and so on. Lack of clear answers to such questions, particularly concerning the uniqueness of probability$_1$ values got by fiducial inference, accounts, I think, for some of the muddiness surrounding the critique of fiducial inference in the statistical literature, as Seidenfeld surveys it. For without a theory of probability$_1$ underpinning fiducial inference, debates over whether a certain inference constitutes an error in applying the rules of the inference or a genuine counter-example, or over the propriety of avoiding a counter-example by an addition to the rules of the inference, are bound to lack any clear basis for settlement. We shall, therefore, not proceed from this discussion to Fisher's fiducial inference proper, but turn to other authors who, being aware of the importance of distinguishing between probability$_1$ and probability$_2$, have given much clearer accounts of the form of inference which Fisher was aiming to elaborate.

Fiducial inference being the last of the three forms of inference Fisher offered as a solution to the 'theoretical difficulty of inductive inference', our discussion of his statistics can now be concluded with the claim that despite the originality and power of Fisher's thought it does not live up to his claims on its behalf in the matter of finding a rational basis for inductive inference.

3. HACKING´S ACCOUNT OF STATISTICAL INFERENCE.

In his [1965] Hacking aimed to set out a group of ideas which he thought could provide a coherent foundation for statistical inference. While he drew on Fisher´s work, Hacking contributed many new ideas to the conception of statistics Fisher defended, and contributed much even to the lines of thought taken fairly drectly from Fisher (particularly the idea of fiducial inference), raising Fisher´s conceptions to much higher levels of clarity. Consequently, the influence which Hacking´s work continues to exercise on those working in the field of foundations of statistics is considerable and independent of Fisher´s continuing influence.(38)

Very briefly, Hacking urged that relative frequency in the long run, or chance, is the fundamental concept in statistics; that when we have to come to some conclusion concerning a single case we ought to adopt, though we are not logically compelled to do so, the conclusion which would apply to the long run, ie. to expect that event whose chance is greatest; and that this leads us, when augmented by Koopman´s logic of support and a principle which says that one hypothesis is supported by evidence only by the failure of another hypothesis on that evidence (and thus that support is always choice between hypotheses), to what Hacking calls ´likelihood tests´, which are offered as the basic statistical procedure. He then illuminates fiducial inference with his revamped conception of likelihood, and proposes a way of assessing the goodness of estimators. Finishing with a discussion of Bayesian inference, Hacking covers the major topics in the foundations of statistics and leaves the reader with the impression that a coherent and solid foundation has been proposed for the main elements of statistical analysis. But Hacking restricts his discussion in one very important way, which I shall discuss first. We shall then review his suggestions for likelihood tests, estimation and fiducial inference, arguing that none provides any escape from Hume's inductive scepticism.

a. Statistical **Inference and Induction.**

Our fist task must be to investigate Hacking's **claim** that statistical inference is not inductive, which he puts forward in discussing the inference from a random sample to the population. The basis for his claim is his distinction between 'closed' and 'open' populations, which he draws thus:

> Some populations may have a definite and more or less practically ascertainable number of members; they will be called **closed.** In contrast are **open** populations, whose members have in some sense not been determined, or, more correctly, whose members could in no way be determined at present nor have been determined in the past.(39)

The point of introducing this distinction, Hacking argues, is that statistical inference concerns only closed populations, or at least an inference from a sample to a population is only **statistical** if the population sampled is closed, for 'there is no such thing as a random sample from the open population´.(40) For the same reason Hacking also denies that an inference from a sample to a population is statistical if some members of the population could not have been included in the sample.

Now clearly Hacking thinks that by restricting statistical inferences to cases in which the entire population was available to be included in the random sample he will be able to free the problem of the rational justification of statistical inference from the problem of induction.(41) But that turns out not to be the case. Before arguing this point, however, let us reflect on the severity of the restriction placed upon statistical inference by Hacking´s proposal.

The consquences of separating statistical and inductive inferences are very serious indeed, since it would reduce statistical inferences to inferences whose only utility would be as premises in subsequent inductions. Take for example the case of a voter survey. For an inference from a sample of the population to the whole population to be statistical in Hacking´s sense it must be the case that all of the population was available for selection and such that each had an equal chance of selection. Now suppose that some population could meet this requirement at time **t**. For the conclusion of the statistical inference to be of any use to us other than as a purely historical fact (from which is drawn no inference of the kind which gives history a point over and above satisfying our simple inquisitiveness about the past - or rather past instants) we require a further inference from our statistical conclusion to an inductive conclusion concerning voter intentions at time t+, say polling day.

It might be argued, however, that the severity of the restriction placed upon statistical inference by the distinction Hacking draws is salutary as it forces us to distinguish between what are really separate problems: the problem of ensuring a sound statistical inference, which is a problem of the relationship between the sample and the population sampled; and the problem of induction, which is a problem of the relationship between one population and another, or between a population at one time and itself at another time. But this argument requires a new definition of ´induction´ as well as Hacking´s new definition of ´statistical inference´, for as ordinarily defined induction includes statistical inference as one of its kinds; Hacking´s statistical inferences are, after all, inferences from observed to unobserved instances of empirical predicates. Hacking´s proposal ignores, that is, that even an ideal statistical inference from a random sample to a closed population is inductive, and is therefore subject to Hume´s critique of induction - though the problem might well be soluble for this special case even if for no other.

Finally, even if we could define ´statistical inference´ along Hacking´s lines and have it follow from this that statistical inferences are not inductive inferences, still it does not follow that we can keep statistical inference free of the problem of induction as Hacking seems to have thought. For even if we could force a neat separation between the **concepts** of statistical and inductive inference we could not achieve this same separation between the **inferences themselves**, since every statistical inference, on Hacking's definition, can only be known to be such if we presuppose the soundness of a prior inductive inference; for we can only arrive at the conclusion that the inference is based upon a random sample from a closed population on the basis of an inductive inference, randomness in selection being equal frequency of selection in the long (and unobserved) run.(41)

For these reasons we shall not abide by Hacking's proposal to distinguish statistical from inductive inference, and thus we shall treat his suggestions for statistical inference as though they were intended to provide a rational foundation for (at least some kinds of) inductive inference. Accordingly, we shall examine Hacking's statistical inference schemes to see if they offer any aid in our analysis of possible solutions to Hume's problem of induction, whilst noting that should the answer be negative then that is just what Hacking led us to expect, since he took the view that his proposals for statistical inference did not bear on the problem of induction at all.

b. Likelihood Tests.

As an alternative to Neyman-Pearson hypothesis tests, and having rejected Fisher's tests of significance as without logical force, Hacking proposes likelihood tests as a general method of assessing the acceptability of statistical hypotheses. A likelihood test is defined by a rule such as this:

(*) Reject hypothesis **h** in favour of hypothesis **i**, in the light of experimental result **e**, if

$$P(e|h) \,/\, P(e|i) < a$$

where a is the level of stringency of the test.(43)

I want to examine likelihood tests with two questions in mind. Do likelihood tests constitute an adequate basis for choosing between hypotheses in general? If not, under what circumstances do they constitute an adequate procedure for rational selection? I shall first discuss a number of restrictions on their general applicability, then turn to the question of their rational force once all problems of application are solved or put to one side for the sake of the investigation. Thus we begin by accepting the use of likelihood tests in general, consider some special cases which cause problems, and then return to the question of their suitability in general.

Consider first **h** and **i** such that on the evidence we have prior to getting **e**, **h** is very much better supported than **i**, the support for **h** against **i** has developed as a steady and unbroken trend, and yet **e** is such that alone, or in combination with the previous evidence, it tips the scales of likelihood against **h**, as measured by Hacking's rule, (*). Here it would seem reasonable that we should be loath to reject **h** for **i**, and would prefer to suspend judgement while further evidence is gathered, old data checked and so on. More generally, it seems unreasonable that the choice between hypotheses has to be based on a comparison of their likelihoods at a particular instant rather than also taking into account the history of the support for the two hypotheses.

There ought to be, I am asserting, some distinction drawn between the support offered an hypothesis by evidence which confirms an already established trend and evidence which does not fit a well established pattern of support. Likelihood tests, like any other theory of instant rationality, do not take account of historical factors; and in my view that is at least a minor failing - but since it is not a point which emerges from Hume's critique of induction I shall not follow it further. I mention in passing, however, that my point may be taken account of by distinguishing, as Barnard urges, between support and rational acceptance, Barnard giving as his reason for the suggestion that a simple **h** ought not always be rejected for a better supported but very much more complex **i**. Barnard's suggestion thus takes care of two special case objections to Hacking's theory, namely the problem of evidence which goes against the trend, and the problem of increased complexity outweighing increased support.

A third special case in which the use of a likelihood test would yield what is intuitively the wrong decision between two hypotheses arises when one of **h** or **i** is cooked up specially to be much more probable or much less probable on **e** than the other, thus ensuring the retention of a favoured hypothesis or the rejection of an unpopular one. Barnard makes this point as well. To prevent such an abuse of likelihood tests it would be necessary to insist that both **h** and i be genuine hypotheses rather than **ad hoc** hypotheses designed specifically to win or loose the contest. Unfortunately the logic of statistical inference then becomes entangled with the thorny problem of distinguishing genuine from **ad hoc** hypotheses, which may not be at all easy to solve, as Hacking himself remarks on another occasion.(43) But again, since this problem will affect any other proposal for statistical inference which, like Hacking's, relies only upon the probability of the hypotheses involved in the test, and since the problem has no special connection with the problem of induction, we will not dwell upon it here. Indeed, the problem is one which will arise in any methodology recommending comparisons between rival hypotheses, arising, therefore, in differing ways, in diverse accounts of empirical support - for example, the problem of determining what is to count as a severe test in Popper's methodology.

The final special case problem I want to raise will lead us directly to more general questions concerning the reasonableness of likelihood tests as inductive inferences. It is the question of the level of stringency required if a test is to give us a rational basis for rejecting particular **h** for particular **i**. Now provided we have some answer for why likelihood tests are reasonable inferences in standard cases, Hacking's answer to this question strikes me as satisfactory: he says we set the level of stringency in the light of the cost of further testing, the seriousness of a mistaken decision, and so on. But what is the rationale for employing a likelihood test in a standard case, ie. when, to mention only the special problems raised here, **h** and **i** are both genuine, equally simple, have similar histories of support and would give rise to similar costs were they mistakenly accepted or rejected? The answer to this question is to be found in Hacking's account of the logic of statistical inference. Let us trace it out.

c. Hacking's Theory of Chance and Support.

Hacking begins his attempt at building a coherent foundation for statistics with an analysis of the concept most agree is the basis of statistical inference, the idea of long run relative frequency, or chance. After contributing much to the clear expression of this idea Hacking turns to the question of support of hypotheses by data. Adopting a logic of support from Koopman, he notes that the logic requires supplementing by a support measure. This is to be provided by connecting support and chance, which Hacking begins to do by providing two rules for guessing, one for the long run and one for the unique case. The long run rule says: if there is a set of things which are S; and we know that each **S** is one of **A** or B but not both, and more **S**'s are **A** than **B**; and we want to guess whether each of a long sequence of **S**'s is **A** or B, such that we will be right as often as possible; then guess that each **S** is **A**. The unique case rule says: if we have to guess whether just one **S** is **A** or **B**; and the **S** in question is chosen in a way which is not influenced by its being **A** or **B**; then the best guess is A.(44)

Now if we could justify the long run rule and derive the unique case rule from it we would be well on our way to providing the coherent foundation sought for statistical inference. But, Hacking argues, there is no good justification for the long run rule, and in any case it does not entail the unique case rule.(45) How then does Hacking's development of rational foundations for statistical inference proceed? The answer is revealing.

To this point in Hacking´s book the reader has been led from first principles via careful argument towards suggested rules for sound statistical inference. Hacking´s argument thus holds the promise of a clear and convincing rationale for the abandonment of any statistical practice which conflicts with the methods recommended by his approach. But before we arrive at rules for statistical inference this line of development is abruptly halted; the unique case rule is not derived from the long run rule but is held to be ´validated´ by what Hacking calls ´the law of likelihood´. That law is as follows:

> If one hypothesis **h**, together with data **d**, implies that something which is an instance of what happens rarely happens on a particular occasion, while another hypothesis, **i**, consistent with **d**, implies, when taken together with **d**, that something which is an instance of what happens less rarely happens on the same occasion, then, lacking other information, **d** supports **i** better than **h**. (46)

Hacking claims that this law ´expresses a fundamental fact about frequency and support´, which is to be defended not by ´deducing it from other facts´ (as Fisher defended his principle of estimation by maximum likelihood) but by showing ´how it leads to many conclusions which are probably true, and which apparently could not be true if the law were false´.(47) So at this point the line of justification is reversed; the law of likelihood is justified by its consequences, and the unique case rule is justified by the law of likelihood. Thus Hacking´s later claim, in his discussion of ´Support and Likelihood Tests´, proves rather unhappy, for here Hacking writes:

> The theory of likelihood tests stands entirely on the logic of support. One might venture an abstract principle about testing in general: a good test for any hypothesis, statistical or otherwise, is one which rejects it only if another hypothesis is much better supported. This principle, plus the logic of support and the law of likelihood entails the theory of likelihood tests.(48)

While Hacking makes no claim here which is not true, the passage is nonetheless mystifying since it gives the **impression** that the theory of likelihood tests is justified by some previously given **theory** of support strong enough to be the theory of likelihood tests´ ´entire´ foundation – and here the reader cannot help but think of the theory of chance and the long run rule. However, the actual situation is the very opposite of the impression given by the first sentence: the justification for the principles which entail that likelihood tests are good tests is derived from the ability of the tests themselves to sanction only good statistical practices. In fact, that this was the line of justification to be given for Hacking´s proposed statistical tests was made clear in the first sentence of his book, but at that point the reader was unlikely to realize that the claim was to be taken literally. Hacking wrote

> The problem of the foundations of statistics is to state a set of principles which entail the validity of all correct statistical inference and which do not imply that any fallacious inference is valid.(49)

Of course Hacking is putting his point too strongly here; if we had principles which allowed us to distinguish correct statistical inference from statistical fallacy we should not be so sorely in need of a coherent account of the foundations of statistics. But all we have got is agreement here and there and disagreement elsewhere - a situation wittily presented by Kyburg in his major work on the foundations of statistics.(50) This does not mean, of course, that there is no point to Hacking's demonstration that his likelihood tests have many features thought desirable by many statisticians, while not sharing other features widely held to be undesirable. All I am wishing to get clear is that Hacking provides no foundation for his likelihood tests other than that they accord with what he sees, and others to varying degrees also accept as, sound statistical practice.

It is not our task to enter into any discussion of which features of various statistical inference schemes ought to be prized and which avoided. We are interested in only one feature: does the proposed form of statistical inference in question, here likelihood tests, provide any basis for avoiding Hume's inductive scepticism? The point of the above discussion was to show that the answer to this question in the present case cannot be found in Hacking's discussion of foundations of statistics, his discussion of chance and support, for nothing in this discussion which might justify likelihood tests is without need of justification in its own right. Indeed, Hacking saw the independent buoyancy of the edifice built on his foundations as the best way of keeping the whole structure afloat.

What criticism might the inductive sceptic level at likelihood tests, then, aside from those problems with special cases canvassed above? The most important criticism would surely be that there is no assurance that we will choose the true hypothesis over a false alternative on any particular occasion when we select an hypothesis on the basis of a likelihood test. This follows immediately from the failure of the attempt to deduce the unique case rule from the long run rule, and was what led Hacking to seek a justification for his likelihood tests in the respectability of the inferences they sanction.

What then of the possibility that likelihood ought to be taken as a primitive concept of empirical support, for example by accepting as part of what it means for an hypothesis to be supported by the data that the hypothesis passes likelihood tests against all genuine rivals? Ironically, the problems with this move are collected together by Hacking in a review of a book by Edwards which presented this proposal.(51) Hacking pointed out in his review that there are counter-examples to the rule that the best supported hypothesis is the one with the highest likelihood - for example a single observation from a normal distribution of unknown mean and variance yields the observed value as the best supported hypothesis concerning the value of the mean, and, utterly implausibly, 0 as the best supported hypothesis concerning the variance. Hacking continues his review with a further similar example, then turns to the problem that a single likelihood ratio intuitively corresponds to differing degrees of relative support for competing hypotheses depending on details about the hypotheses and data which are irrelevant to the likelihood ratio. Apparently it was difficulties of this kind which led Hacking to rethink his defence of likelihood tests.

But the greatest problem for likelihood as a primitive concept, in my view, is not any of the difficulties Hacking raises, though these certainly detract from its intuitive appeal, but that it is always open to us to declare that it was purely by chance that the evidence upon which the test has been conducted was unfavourable to our pet hypothesis. This is the same problem as that which led us to deny that tests of significance have any rational force. I do not see how it can be overcome unless we employ the argument that while it may happen every now and then that the data acting through a likelihood test have the result of requiring that we reject an hypothesis when a mere fluke result has occurred, this will be in the long run a rare event; and this defence, of course, gives up likelihood as a primitive concept and bases likelihood tests on a rationale derived from the reliability paradigm. Any answer to Hume would then come not from the logic of the tests themselves, ie. not from the defence of likelihood as a primitive concept of empirical support, but from the adequacy of the reliability justification; and such justifications are, as we have already seen, in any case problematic.

Without disparaging the relative merits of likelihood tests as a method of statistical inference, we conclude therefore that with respect to the problem of induction they offer nothing more towards a solution than that provided by the reliability paradigm and tests of significance.

d. **Hacking's Theory of Estimation.**

After a subtle and interesting discussion of the nature of estimation Hacking proposes a criterion for distinguishing a good estimate from a bad one, based upon an estimate's closeness to the true value of the parameter estimated. Briefly, he calls estimate **A** 'uniformly better' than **B** if, for every possible size of the error **e**, the data supports the hypothesis that A is within e of the true value at least as much as it supports the hypothesis that B is within **e** of the true value, and for some value of **e** it gives better support to the hypothesis that A is within that region around the true value than it does to the corresponding hypothesis about **B**. Hacking then terms an **estimate** 'admissible' only if no other estimate is uniformly better than it, and calls an **estimator** 'admissible' if and only if, for every body of data on which it is defined, the estimate given by the estimator is admissible.(52) (It does not follow from these definitions that only one estimate, or estimator, is admissible in any given case.)

Now with this analysis Hacking replaces the problem of finding the best estimate with the problem of determining which of several hypotheses is the best supported. Thus Hacking's general solution to the problem of estimation stands or falls with his theory of support, which we have already examined and found wanting.

Before we leave his theory of estimation, however, we should note that Hacking makes one suggestion for a special case which merits our attention. According to Hacking, if there is among unbiased estimators with bell-shaped distributions (ie. among estimators which are such that the distribution of the estimates given by the estimator is normal or near normal, with the mean equal to the true value of the parameter being estimated) an estimator of minimum variance, then only this estimator is admissible. Now since knowing that an estimate has been produced by the sole admissible estimator in a certain class of estimators, here the class of unbiased estimators, informs us that no estimate from another estimator in the class is better than the estimate produced by the admissible estimator, we have a sound basis for accepting the estimate of the admissible estimator. Provided, that is, that we have a sound basis for insisting that any acceptable estimate must be produced by an estimator belonging to the class with respect to which our estimator is the sole admissible estimator, here the class of unbiased near normally distributed estimators. This requirement proves the downfall of the proposal, however, for, as Hacking points out, no good reason has ever been given for insisting upon unbiasedness, and Barnard goes further by presenting a case in which the use of an unbiased estimator is clearly inappropriate.(53) Thus it seems that there is no solution to the problem of induction in Hacking's theory of estimation.

e. **Fiducial Inference.**

Hacking, Seidenfeld claims, is to be credited with being the first to make the logical structure of fiducial inference plain. Certainly his discussion allows us to answer the two questions which any would-be fiducialist would want answered before he would commit himself to a fiducial inference, namely, what assumptions must be allowed if the inference is to go through?; and what, if the inference is cogent, does a fiducial inference give us? We shall deal with the second question first.

What we get from a fiducial inference, all questions concerning its cogency postponed for the moment, is a probability for an hypothesis concerning the population, given evidence from a random sample from the population. Since inductive inferences standardly take this form, or are assumed to be equivalent to such inferences (granted some liberty in the definition of ´random´), what we get from fiducial inference is a widely applicable inductive method. But the cogency of the inference requires to be demonstrated. We shall concentrate on the assumptions which must be granted for the inference to go through.

Working from the structure of the argument as Hacking gives it, we see that the very first step in a fiducial inference is the specification of some feature of trials on a chance setup.(54) Precisely what assumption is required varies from case to case. In Hacking´s simplified binomial example we are required to assume that the chance of heads is either 0.4 or 0.6; in another more complex example we are required to assume that the distribution of chances for the various outcomes of an experiment is normal, with unknown mean and variance.

How weak can this initial empirical assumption be made? Since in the second step of the argument we are to define a second kind of trial in terms of the first kind, such that the chance of an outcome of the second kind is known prior to the experiment, our initial assumption must be strong enough to allow this second step to be made. Except for very simple inferences concerning the median, or some other order statistic, this will require the functional form of the distribution to be assumed. Thus in general fiducial inference is not a primary inductive method, and therefore it relies upon the success of some prior induction to secure one of its premises, if the fiducial inference is to meet our standard conception of proof.(55)

Once this first required assumption is granted there is a mathematical problem to be solved in the construction of a pivotal variable. Though this can cause trouble, I shall assume that a pivotal can always be found. This allows us to concentrate on the next step of the inference, the transition from the probability$_2$ distribution to a statement of support or probability$_1$, the aspect of Fisher´s explanation of fiducial inference we found unsatisfactory.(56) To effect this transition Hacking provides no elaborate theory of support, but simply asserts :

> One principle about support and chance seems so universally to to be accepted that it is hardly ever stated. It is not so much a principle as a convention to which everyone is firmly wedded. Roughly, if all we know are chances, then the known chances measure degrees of support. That is, if all we know is that the chance of **E** on trials of kind **K** is **p**, then our knowledge supports to degree **p** the proposition that **E** will occur on some designated trial of kind K.(57)

Such an untheoretical proposal for getting a statement of support from the probability$_2$ claim generated by the fiducial inference will not do, however, for the rule given is controversial and thus requires some argument in its defence. Indeed, the major philosopher to have worked in the confirmation paradigm, Carnap, who devoted much attention to the relationship between knowledge of chances (or probability$_2$) and support (or probability$_1$), rejected the rule Hacking puts forward as an agreed common position.(58)

Hacking's attempt to get support from chance thus leads us to Carnap's more detailed investigation of this relationship. In the meantime, however, we conclude that Hacking, in being content to leave as a convention the fundamental principle required for fiducial inference to provide a reasonable inductive method, does not advance us towards a solution to the problem of induction - though no reader of his book can fail to learn much about the other subtle and interesting problems of statistical inference which Hacking made the focus of his analysis.

4. CARNAP'S INDUCTIVE LOGIC.

Carnap's system of inductive logic, following Keynes' and Jeffreys', is based upon a logical conception of probability. This conception has been developed by others since Carnap, notably Hintikka, and this later work has solved some of the problems in Carnap's system, as Carnap in turn had resolved many of the difficulties which beset Keynes' work in particular. It is fair to say, however, that neither his predecessors nor those who have joined his camp and carried on his work present the fundamental elements of the logical system of probability in a way which differs from Carnap with respect to Hume's problem of induction. In any case this is the assumption on which I shall work, restricting my discussion of the logical conception of probability to Carnap's system.

This initial restriction reduces somewhat the size of the literature relevant to our discussion, but it remains sizable. I propose two further restrictions, as follows. First, I shall ignore all critiques which proceed from some alleged incompatibility between Carnap's formal system and the presystematic notions they were supposed to reconstruct. It seems to me that Carnap's system has been developed in sufficient detail to live independently of its original host, and should any conflict occur between the presystematic ideas and the formal system, it is an open question whether the formal system or our presystematic ideas ought to give ground. Second, I shall ignore all criticisms which are based on the unfinished state of the system, that is on its inability to deal with this or that aspect of actual scientific practice. Time and again Carnap confounded his critics by extending his system to permit the development whose absence had been the basis of criticism, and it seems reasonable to suppose that those who have continued his line of thinking in probability and induction will be able to add further sophistication to the analysis. Thus, while such criticisms play an important role in suggesting problems for further work, I shall not deal with any here, restricting attention to the criticisms which, as I understand it, go to the heart of Carnap's system of probability and induction.

Having restricted our critique first to Carnap and then in these two further ways, we are left with what I view as the deepest of the critical discussions of Carnap's work, so far as the problem of induction is concerned. My aim here is to bring these discussions to bear on two questions which concern the fundamental conception Carnap employed to construct his theory of inductive logic:

> 1) What assumptions are required, and what justifications can be given for them, to allow the confirmation function **c(h|e)** to be defined for some h and **e** to which the system applies and such that the inference from **e** to **h** is inductive?
>
> 2) How can we employ values of the confirmation function **c(h|e)** to justify inductive inference?

These two questions raise the kinds of considerations which have been given our attention in discussions of other systems of induction, for (1) probes the matter of the status of confirmation functions (or c-functions, as I shall term them) as primary or secondary inductive methods, while (2) raises the question of the cogency of inductive inferences on Carnap's model. We shall discover that our answers to (1) and (2) lead to the conclusion that Carnap´s inductive logic is a secondary method of induction (though that was not Carnap´s intention). Carnap´s theory of c-functions can be counted as an advance towards a solution to Hume´s problem, therefore, only if accompanied by a reasoned rejection of foundationism. Furthermore, there is no simple amendment to Carnap´s system which could get around this difficulty, for as we shall see the cogency of inductive logic as Carnap reconstructs it depends upon its being a secondary method, since no value can be given for **c(h|e)**, where the inference from e to **h** is inductive, unless certain factual assumptions are made, which entails that the inference is secondary.

Before taking up the questions listed above one point about the development of Carnap´s system ought to be considered. So far I have spoken of ´Carnap´s system´ as though he proposed just one account of induction, perhaps modified in various ways over time. It is generally accepted, however, that Carnap produced two distinct systems, one given fullest expression in his **Logical Foundations of Probability** and amended in **The Continuum of Inductive Methods**, the other begun in his **Basic System of Inductive Logic** (hereafter **LFP, Cont** and **Basic,** respectively).(59) The most obvious difference between the two systems is the change in the mathematical formalism Carnap employed, but there are also substantial differences in the axioms included in the two systems, the justifications offered for the various axioms, and a major change in the factors influencing inductive inference of which the systems can take account.

There is also a development in Carnap's conception of the form of inductive inference, or methodology of induction, which by the time of the paper 'The Aim of Inductive Logic' (1962) brought inductive inference on Carnap's analysis closer to that proposed by the Bayesian tradition than to deductive inference, which had been Carnap's original model, at least for the determination of values of c-functions. It is my view, however, that this did not involve any break with his earlier work, but rather an elaboration of Carnap's conception of what constitutes a complete inductive inference, an aspect of his theory which is only sketched in the **LFP**, which concentrated almost exclusively on the explication of logical probability.

Therefore, while we have two theories of pure inductive logic, or two versions of the theory of c-functions, to examine, we have only one theory of the methodology of induction. We shall begin with Carnap's theories of pure inductive logic, and then pass on to his analysis of the proper use of degrees of confirmation in the making of inductive inferences.

a. The Theory of Inductive Logic in the LFP.

The **LFP** contains two studies on induction, one an impressive philosophical analysis of various problems in the theory of induction and the development of an account of inductive inference which systematically provides answers to the questions about induction raised in the preceding discussion, the other the formal construction of a confirmation function intended to provide an exact method for conducting inductive inferences. In order to get clear whether the **LFP** makes any contribution to the solution of Hume´s problem of induction it is necessary to get clear the relations between these two investigations.

The heart of the philosophical analysis of induction in the **LFP** is Chapter IV, ´The Problem of Induction´, in which rigorous discussion of many difficulties faced by induction leads to progressive refinement of the idea of logical probability, until, in the final section of the chapter, Carnap lays down conventions defining the concept ´degree of confirmation´. While these conventions are not strictly derived from what has gone before, the earlier analysis is sufficient to make the conventions offered seem natural to the reader. Then, on the basis of stipulative definitions in the succeeding chapter, Carnap constructs his c-function (or rather, his set of c-functions), and shows that it meets certain of the conventions offered earlier, including axioms for the probability calculus. Completion of the formal system then requires only that a free parameter individuating the members of the set of c-functions be given a particular value, or, what amounts to the same thing, that a rule be given for assigning a precise value to **c(h|e)**.

This analysis makes clear that if Carnap has presented any argument which might refute Hume´s inductive scepticism, it must be located in the philosophical discussion which leads to the conventions on c-functions, which are in turn captured in the stipulative definitions which are the basis of the formal system. And we do find in the informal analysis an argument which certainly reads as if it were intended to refute Hume's inductive scepticism - by showing that inductions based upon probability_1 can, if the inference is properly conducted, provide a rational basis for decision-making; in any case, that is how Carnap's section on the presuppositions of induction has been read by other authors, and I intend to adopt this view also.(60)

We shall subject this argument of Carnap's to extended examination, for, as we shall see, on a close analysis it becomes plain that the argument completely fails to provide any sound case for the rationality of basing inductive inferences on the probability_1 lent the conclusion of an inductive inference by its premises.(61)

Carnap begins his discussion of the presuppositions of induction by inviting us to consider the case of X, a person who would like to know if it will rain tomorrow. Carnap remarks that 'some reflection' will show X that 'for questions of this kind certainty is not attainable but only probability', and that he will thus be prepared to take as the basis of his decision:(62)

2. With respect to the available evidence, the probability_1 that it will rain tomorrow is high.(63)

With the problem of induction on our minds, we will want to know why this is a reasonable basis for a decision of X's which will turn out to be good or bad depending on tomorrow's weather. Carnap suggests that it is a reasonable basis, since although 'it may be that the event predicted with high probability will not occur´, there are other possible justifications. For example, if X knew (3) below ´then he would clearly be justified in following the inductive method´:

3. If X continues to make decisions with the help of the inductive method, that is to say, taking account of the values of probability$_1$ or estimation with respect to the available evidence, then he will be successful in the long run. More specifically, if X makes a sufficiently long series of bets, where the betting quotient is never higher than the probability$_1$ for the prediction in question, then the total balance for X will not be a loss.(64)

Now Carnap notes that (3) would be true if the world had a high degree of uniformity, but that any attempt to prove that it had this character would be, according to many philosophers, itself inductive and thus involve vicious circularity. Thus Carnap quickly sketches Hume´s argument, then goes on to pose the question whether the scepticism to which many philosophers think Hume´s argument leads can be avoided, suggesting that a solution is to be found in the theory of probability$_1$. This section of his argument is crucial, and accordingly I quote at some length:

> Let us examine what kind of assurance would justify X´s implicit habit or explicit general decision to determine all of his specific decisions with the help of the inductive method. We can easily see that he need not know with certainty that this procedure will be successful in the long run; it would be sufficient for him to have the assurance that success in the long run is **probable**. Just as in the case of the prediction of a single event it was clear that only probability but not certainty can be obtained and that probability gives a sufficient basis for the specific decision, thus analogously for the question of success in the long run it would suffice for X to obtain, instead of the earlier statement (3), an inductive statement in terms of probability$_1$.

I think that Carnap here makes a fundamental error. It was the role of (3) to justify the sufficiency of (2) as a basis for X's decision to be rational, ie. to show that a statement of probability_1 is an adequate basis for decision making, which (3) did by tying probability_1 to the long run frequency, or probability_2, of success. Once (3) is abandoned (2) is without justification unless some other connection is established between probability_1 and probability_2, or some new principle is offered to replace long run success rates as a rational basis for decision making according to the inductive method.

It is certainly not the case that we can substitute a claim in terms of probability_1 for (3) **in the way we substituted (2) for (1)**, for that substitution was **underwritten** by (3). The further substitution of some probability_1 claim for (3) would thus require some new guarantee to play the role which (3) played in the exchange of (1) for (2). What does Carnap offer in place of (3)? Just this:

6a. If X makes a long series of bets such that the betting quotient is never higher than the probability_1 for the prediction in question, then it is highly probable that the total balance for X will not be a loss.

Before we can gauge the strength of this support for basing one's beliefs upon one's assessments of probability_1 we must remove the ambiguity associated with the term 'probable': is it probability_1 or probability_2 that is referred to here? If it is probability_2 that is intended, and (6a) in this form can be proved, then (6a) does indeed support (2) as (3) supported (2) (subject to the difficulties mentioned in connection with that earlier inference). But if it is probability$_1$ that is intended then (6a) does not support (2), for they are notes written on the same security, namely the presumed reasonableness of making decisions on the basis of probability$_1$ assessments. If it was probability$_1$ that was intended and we have reason to doubt the soundness of (2) as a basis for decision making, and thus reason to seek some assurance of its adequacy in that regard, we will not be at all assured by (6a) since we shall have the same doubts about (6a) as we had about (2).

Now from the passage quoted prior to (6a) above it is plain that Carnap did intend ´probable´ in (6a) to refer to probability$_1$ - indeed, his text continues after the passage quoted to say this explicitly. The justification of induction given in Carnap´s analysis of the presuppositions of induction is thus shown to be inadequate.

The problem with the justification Carnap offers, continuing our assumption that the discussion was intended to justify his inductive method, can be brought out more simply and clearly by considering the final step in Carnap´s argument. He asks for the truth condition of (6a) and offers this:

8a. On the basis of the evidence that the relative frequency of a property in a long initial segment of a series is high (say, r), it is very **probable** that it will likewise be high (approximately equal to r) in a long continuation of the series.

Here 'probable' is definitely meant to be interpreted in the sense of probability$_1$, for Carnap remarks that (8a) is analytic and thus, unlike (3), can be shown to be true without need of an inductive inference, thus avoiding the circularity in the inductive justification of induction Hume exposed. But no less devastating a problem emerges for Carnap's justification, for it now clearly consists of just this: if you **define** the term 'probable' such that it is very probable that any uniformity will be continued, and then bet on the continuation of uniformities, then it is very probable (in the sense of that term just given) that you will win most bets - **though you might in fact lose all of them, and losing most of them is not (in the frequency sense of the term) known to be improbable.** Clearly the definitional truisms Carnap offers provide no rational foundation for induction.

Of course this result does not damage Carnap's inductive logic, since, as I have already remarked, he did not stand by the argument as a justification of induction. But it is worth noting that there is no other defence of the rationality of induction (on Carnap's model) given in the informal discussions in the **LFP**. The other section which might be expected to offer relevant arguments, Carnap's discussion of probability as 'a guide in life' is devoted to establishing the negative result that probability$_2$ cannot function as a guide in life because, like the true value of some quantity as against an estimate of the true value, probability$_2$ values are typically impossible to determine and thus we must employ estimates of their values if we are to have any guide at all.

Then Carnap reminds us that a value of probability$_1$ can be interpreted as an estimate of a corresponding probability$_2$, and takes this as an argument to advance his claim that probability$_1$ can serve as a guide in life. But it shows only that probability$_2$ is problematic if taken as a guide in life, while probability$_1$ does not share this particular problem; and this does not solve any other problem which probability$_1$ might have in relation to being employed as a guide in life.

As to the formal development of Carnap's theory, he does show in the **LFP** that his c-functions possess some intuitively appealing properties, but he does not develop the formal theory to the point of choosing a unique c-function, and thus the formal theory does not provide a complete basis for justifying an inductive inference as a rational procedure. Moreover, the formal development is set out much more clearly in **Cont**, so we shall take that as the text for our analysis of Carnap's formal account of induction, referring back to the **LFP**, for justifications of stipulations on c-functions, as required. Indeed, we shall find this often necessary, for **Cont** is a transitional work; while the style of its presentation and the axiomatic development prefigures the study of induction in **Basic**, the formalism is still that of the **LFP**, and justifications for the many stipulations defining c-functions are typically only sketches of the arguments peviously given in the **LFP**. The main critical analysis of induction in **Cont** concerns the justification which might be given for the choice of a single inductive method from the continuum of c-functions there defined, and it is to this that we shall give most attention.

b. The Theory of Inductive **Logic in The Continuum** of **Inductive Methods.**

Cont generalizes the formal system developed in the **LFP** by presenting a continuum of c-functions in which those discussed in the **LFP** are associated with special values of a parameter Carnap calls ´lambda´, which I shall denote ´L´. The continuum is constructed by building up a set of adequacy requirements for confirmation functions, and thus the assumptions underlying the continuum can be easily examined since they are required for the justifications of the various adequacy requirements. We shall examine these in turn.

Requirements C1 to C5 constrain **c** (we shall adopt ´**c**´ to denote an unspecified c-function) to obey a set of axioms for the probability calculus, ie. to be a probability function, and to take values which are not changed if one or the other or both of the arguments of **c** are replaced by a logically equivalent sentence.(65) Of these five requirements only the assumption that **c** be a probability function is open to dispute, though Carnap does not think even this a serious possibility if the weakness of his justification is any guide, for he remarks in defence of this requirement only that it is ´generally accepted´ and ´in accordance with what reasonable people think in terms of probability$_1$ and, in particular, what they are willing to bet on certain assumptions´.(66) Obviously the real justification here is the assertion about betting behaviour, but Carnap had established the connection between probability$_1$ and betting quotients by fiat in the construction of his explicandum for probability$_1$, so that does not constitute an adequate basis for the requirements defining c as a probability function.(67) We shall not, however, pursue this problem further since Carnap offered a stronger justification for that requirement in **Basic** by employing the Dutch book argument, which we shall discuss in the following Chapter.

C6 is uncontroversial, but this is not the case for the next two requirements which express what Carnap called ´the valid part of the principle of indifference´.(68) C7 requires that if **h** and **e** are exactly like h´and e´ except that they refer to different individuals, then $c(h|e) = c(h'|e')$; in short, C7 requires that c be symmetrical with respect to individual constants.(69) C8 is a similar principle, asserting that c must be symmetrical with respect to the Q-predicates.(70) These requirements have been criticised by Nagel on the grounds that they represent disguised factual claims, while Salmon remarks on the weakness of the justification offered them, unfavourably contrasting it with the strength of the pragmatic justification given by the Dutch book argument for the requirement that **c** be a probability function.(71) I shall not add anything to Salmon´s point, but Nagel´s requires further discussion. This will not be entered into here, however, since the whole question of symmetry, which C7 and C8 address, as well as the question of what evidence is relevant to the determination of a degree of confirmation, which is the concern of C9, is subject to a completely fresh treatment in **Basic**. There is little point in rehearsing a critical attack on these axioms given that Carnap rethought his position. We shall take up Nagel´s point, therefore, in the discussion of **Basic**, below.

Following the introduction of C9 Carnap shows that adequacy requirements 1 to 9 entail that $c(h_i|e_i)$ - where h_i asserts that some individual or group of individuals have the Q-property Q_i, and e_i asserts that of s observed individuals s_i had Q_i - is a function of s, s_i and k, where k is the number of Q-properties expressible in the language in which h_i and e_i are expressed. The problem now is to find further constraints on the set of possible functions of s, s_i and k (which Carnap calls the G-functions) such that each G-function determines a unique c-function meeting C1-9 and representing a reasonable assessment of $c(h_i|e_i)$. Once this has been done we can tackle the problem of finding an optimum c-function, perhaps making optimality depend on a small number of factors which might vary for different applications of a confirmation function.

C10 narrows the range of functions to be considered reasonable by asserting that $c(h_i|e_i)$ must be in the closed interval $[s_i/s, 1/k]$, where 1/k is the measure of the logical width of the property with which h_i is concerned, namely Q_i, which is one among k Q-properties.(72) Carnap´s justification for this requirement is accepted practice, and his own judgement that if one assesses $c(h_i|e_i)$ as anything other than s_i/s it will be by making the value depend in part on the logical width of the predicate Q_i, and thus the value of $c(h_i|e_i)$ will be shifted from s_i/s towards 1/k. Values outside this interval Carnap declares to be ´entirely unacceptable´. However, since one significant proposal does give values outside this interval, namely the minimax method of estimation proposed by Wald, Carnap's justification leaves much to be desired.(73)

Leaving to one side the weakness of Carnap's justification for restricting the values of c-functions in the way just described, we now arrive at the final steps in constructing the continuum of inductive methods. For a confirmation function is defined by how it weights the two factors s_i/s and w/k to produce a value for $c(h|e)$ (w being the number of Q-predicates in h; in the special case of h_i considered in the previous paragraph, $w = 1$). If we standardize the weight given to one factor, and Carnap opts to standardize the weight of s_i/s giving it the weight s, the sample size, then the weight given to w/k, the logical factor, selects a unique c-function from the continuum; thus the continuum is indexed by a single parameter, L, representing the weight given to w/k. For a sample of s observations, $L = s$ gives the empirical and logical factors equal weight, $L < s$ gives the empirical factor dominance, and $L > s$ gives the logical factor dominance, while setting L at a certain value gives the logical factor decreasing relative weight as the sample size grows.

Carnap opts to make L a fixed value, rather than a function of s or of s and s_i, and also to make it independent of h and e. Thus L is presented as an enduring disposition on the part of the inductive logician to give the logical factor the same weight as a fixed number of observations, regardless of what has so far been observed and regardless of which hypothesis is seeking support. The question thus posed is which such disposition, or class of dispositions, is optimal; and, of course, what definition of optimality applies here. The major critical attention given to **Cont** has focused on Carnap's answers to these questions.

Carnap first directs his attention to whether **L** ought to be made a function of **k**, the number of Q-predicates in the language employed. This would be disadvantageous, since values of **c(h|e)** would then be relative to a given language, but would have the advantage of mathematical simplicity if **L** were an uncomplicated function of **k**, say **L = k**, which yields **c***, the c-function favoured in **LFP**.(74) Of course we could choose **L** in very many ways - eg. we could make it a function of the age of the scientist who is being 'fitted' with an inductive method, on the grounds that young scientists tend to be rash and need slowing down while older ones get staid and need gingering up - but whatever way we determine L we will still be choosing one of the methods in the continuum. Once the constraints defining the continuum are accepted, then, the fundamental question becomes a decision about the region of the continuum, and if possible the single valued interval, in which **L** should fall.

Carnap first eliminates the extreme methods, corresponding to **L** = 0 and L = infinity. The first of these values gives the so called straight rule, and it is rejected because it gives too strong a confirmation to an hypothesis tested on a small sample to be acceptable to Carnap.(75) The second, which corresponds to the Wittgensteinian c-function **c+**, is rejected for being in gross contrast to what is generally accepted in inductive reasoning.(76)

The extreme methods are dealt with rather summarily, but that is characteristic of the approach to the choice of L that Carnap takes in **Cont**, for after dispensing with these first two suggestions, and reviewing the question whether L ought to be a function of **k**, he adopts a thoroughly pragmatic approach to the problem, on the basis that

> the adoption of an inductive method is neither an expression of belief nor an act of faith, though both may come in as motivating factors. An inductive method is rather an instrument for the task of constructing a picture of the world on the basis of observational data and especially of forming expectations of future events as a guidance for practical conduct. X may change his instrument just as he changes a saw or an automobile, and for similar reasons... after working with a particular method for a time he may not be quite satisfied and therefore look around for an another method. He will take into consideration the performance of a method, that is the values it supplies and their relations to later empirical results, eg. the truth-frequency of predictions and the error of estimates; further, the economy in use, measured by the simplicity of the calculations required; maybe also aesthetic features, like the logical elegance of the definitions and rules involved.(77)

Now it would doubtless be accepted by all that performance, that is, truth frequency of predictions, should over-ride any other criterion for choice. Indeed, should an inductive method be chosen on the basis of elegance of formalism and simplicity of use it would be unlikely that any elaborate method like those of Carnap's continuum would be chosen. Let us therefore examine the criterion of performance to see what support it might give an inductive method, ie. a choice of one particular value of **L**.

There is an obvious difficulty in using performance as a basis for choice of an inductive method, namely that we must find some justification for using past performance as an indicator of future performance. Clearly the rational basis for using performance as a basis for choice presumes that this induction is a reasonable inference; but that we can reasonably make inductive inferences was just what the theory of c-functions was originally intended to show. Now that is required as a presupposition, if choice of L is to be based on past performance. The problem is admirably captured by Carnap himself in his discussion of how we might compare the success of inductive methods in the actual world:

> Questions concerning the success of a given inductive method in the actual world would be of a factual, non-logical nature. And if they concerned not merely that part of the world which is known to us by past observations but also a part or whole of the future, then the answer could be given with certainty only after all observation reports were in, if that were ever possible. And if our question concerned not the actual success but the probability or an estimate of success, then it would make sense only on the basis of a chosen inductive method. The purpose of the intended study, however, is to examine the various inductive methods on a neutral basis without presupposing the acceptance of any one of them... Our investigation must necessarily abstain from making any judgement concerning the success of an inductive method in the total actual world. A judgement of the later kind is obviously impossible from an inductively neutral standpoint.(78)

Evidently Carnap thought that this difficulty could be surmounted, however, for he set as his aim for Part II of **Cont** that it should ´supply to anyone who wishes to choose an inductive method a rational foundation for his choice´.(79)

In fact Part II of **Cont** offers the agent little help in choosing a confirmation function, for at most it offers a lower bound for values of L such that we know that if we choose a value of **L** below this bound then our inferences will belong to a less successful set of inferences than if we had chosen a confirmation function defined by a higher value of L. There remains, however, an infinity of functions from which to choose. Furthermore, we do not know that as individuals we should not be more successful if we chose a lower value for **L**, for Carnap's result concerns the long run of inferences sanctioned by the various values of L, and thus, as Carnap puts it

> X cannot infer from this result that he will be more successful if he uses e´ than if he uses $\mathbf{e}_o$. Whether or not this is the case depends upon which particular sequence of individuals will happen to come his way. What X learns from the result is something which concerns, not of course his own life in particular, but rather the universe as a whole and hence, so to speak, the average observer.(80)

Carnap thus has the problem with which Hacking wrestled, namely to find some connection between the best rule for the long run and the best rule for some actual human time-frame, a problem he does not address.

We conclude, then, that while Carnap's contributions to our understanding of both philosophical and technical questions surrounding induction are undisputed and indeed indisputable, he made no headway on Hume's problem of induction in his first system of pure inductive logic. (Recall that his methodology of induction is yet to be examined.) For he leaves open the choice of **L** and thus fails to reasonably restrict the conclusion which may be drawn via an inductive inference from given premises. We turn now to the second system, wherein Carnap made a fresh start on the question of justification.

c. Carnap´s Basic System of Inductive Logic.

Carnap´s second period of work on inductive logic produced three texts relevant to our discussion: ´Inductive Logic and Inductive Intuition´, which discussed the basis for justification of various axioms of inductive logic; ´Replies and Systematic Expositions´; and his **Basic System of Inductive Logic**, both of which gave axiom systems for inductive logic.(81) The axiom system in **Replies** is superseded by that in **Basic**, and for that reason I shall only discuss the latter system, though **Replies** will be referred to for the justification of one axiom.

Basic presents just eight axioms, plus a set of principles concerning the determination of L. The first five axioms ensure that the C-function is a normalized probability measure, which requirement is held to be justified by considerations to do with rational betting behaviour.(82) As already indicated I intend to postpone all discussion of this line of justification until the following Chapter, and thus no comment is made here on this first group of axioms.

The sixth axiom concerns the symmetry of C-functions with respect to individuals, which I shall take up shortly along with the discussion of symmetry with respect to attributes, which is the focus of the philosophical development given in **Basic**. Axiom seven, which we shall also discuss (mistakenly called ´axiom 9´ in much of **Basic**) asserts that C-functions must show the positive relevance of confirming instances; and finally axiom eight lays down requirements on the extension of measure functions (which define the C-functions) from restricted to universal generalizations, which we shall not discuss since no criticism of Carnap´s system will be made here on the basis of its difficulties with universal generalizations. Thus we have only to discuss symmetry, the principles governing the choice of **L**, and instantial relevance.

I shall consider axiom seven, or A7, first, since I have only a brief comment to make concerning its justification. The axiom is a development of A13, the axiom of convergence in **Replies**, and states, in informal terms, that the limit of the degree of confirmation of the hypothesis that the next individual will have property **j** minus the frequency of individuals with **j**, relative to a sample of size **s**, tends to zero as **s** grows without limit.(83) A7 is used to knock **C+** out of consideration as a reasonable inductive method and to justify the principle of instantial relevance. This principle asserts, informally, that the degree of confirmation of the hypothesis that individual **s+2** is **j**, given that **s+1** was **j**, and a sample of size **s** of which **r** are **j**, is greater than the degree of confirmation that **s+1** is **j** given only the sample **s**.(84)

In **Basic** Carnap does not give anything in support of either his axiom or the principle it supports, noting only its plausibility. In **Replies**, however, he referred to Kemeny´s adequacy requirement for confirmation functions CA4 for support, but that does not seem to me to offer the axiom any worthwhile support since in this connection it flatly begs the question. For Kemeny's CA4 states that 'The definition of c must enable us to learn from experience'.(85) But the whole problem is whether there is any rational foundation for our inductions, and thus whether, for example, past instances of successful prediction constitute a rational basis for our making the same prediction again - whether taking a history of successful prediction to give grounds for future prediction constitutes learning or unwarranted assumption.

Now there would be nothing objectionable in laying down Kemeny's requirement if after the confirmation function was constructed it was to be subject to investigation concerning its reasonableness as a mode of reasoning, but this is not how Carnap and Kemeny proceed. Rather they were aiming to build a C-function from elements each approved before they were added to the construction, with the final product requiring no justification other than that it has been entirely manufactured from guaranteed materials by known trouble-free methods of combination. Thus in adopting CA4 as a justification for one of his axioms Carnap assumes that induction is a reasonable procedure and thus begs the question which the theory of confirmation functions is here being construed as an attempt to answer, viz 'Does the theory of logical probability provide a theory of inductive support which is not subject to Hume's sceptical critique of induction?'. Clearly if the rationality of induction is required to be granted (when CA4 is adopted) so that the construction of a C-function can be completed (with the adoption of A7) then any positive answer to the rationality of induction got from this theory of confirmation will be no useful justification of induction.(86)

This argument causes a great deal of difficulty for Carnap´s theory of C-functions, for if the theory is deprived of A7 then **C+** remains a choice of C-function allowable within the system, but chosing **C+** is tantamount to adopting a sceptical attitude to induction, since it constitutes a refusal to take past observations as inductively relevant to future predictions. If **C+** is adopted as a reasonable C-function then not only is inductive scepticism not countered, it is endorsed. If, however, it is argued that the construction of the C-functions was not intended to provide an inductive method immune to Hume´s sceptical critique, thus legitimating the adoption of Kemeny´s CA4, we shall find ourselves in the strange situation of (all other problems with C-functions aside) being able to determine that certain evidence strongly confirms a certain conclusion and **yet unable to assert what seems to be entailed by this, namely that the evidence constitutes grounds for reasoned confidence in the hypothesis in question.** It is surely unacceptable that our theory of confirmation should put us in a situation where our reasoned expectations concerning the future are held to be independent of the levels of confirmation of our hypotheses concerning the future, but that would be the consequence of adopting the view that the theory of C-functions is not intended to answer Hume´s inductive scepticism. Therefore either the problem is avoided, or the question is begged. Neither alternative is satisfactory.

Though the first of the three issues we have listed for discussion in relation to the theory of C-functions raises a serious problem, it is not the only problem, and we shall now place it to one side in order to proceed to the remaining axioms identified as topics for discussion, first taking up the question of symmetry, and concluding with the principles governing the choice of L.

In **Cont** there were three axioms in which symmetry considerations were set out: that **c** should be symmetrical with respect to the individual constants; and with respect to the Q-predicates; and that if $\mathbf{h_m}$ asserts that the next individual will be **m** and $\mathbf{e_m}$ records how many individuals thus far sampled were **m**, while $\mathbf{e_q}$ records the Q-predicates of each individual in the sample (thus giving an exhaustive description rather than merely recording how many were **m**), then **c** should be such that $\mathbf{c(h_m|e_m) = c(h_m|e_q)}$.(87) In **Basic** only the first of these survives, as it is not generally assumed that the **basic attributes**, corresponding to the primitive predicates from which the Q-predicates were constructed in **Cont**, are symmetric. Of Nagel's criticisms of the symmetry axioms, therefore, only his attack on the symmetry of individuals remains relevant. Let us examine this, Carnap's reply, and the development in **Basic**.

Nagel objected to the assumption of the symmetry of individuals in the LFP and **Basic** on the ground that it settled **a priori** a matter which in actual scientific practice requires experimental evidence, viz that several individuals, for example, samples of water, are similar in all relevant respects - in the sense that an experiment using any one of them is the evidential equal of an experiment using any of the others.(88) Symmetry of individuals is thus seen to lay down **a priori** that in the absence of information to the contrary, perfect experimental control is to be assumed, whereas in actual scientific practice, Nagel observed, the assumption that control has been achieved (in the standardization of reagents, for example) is a matter for careful test, and our inductive logic ought to reflect this rather than make an unwarranted assumption **a priori**.(89)

In his **Replies** Carnap rejected Nagel's criticisms, asserting that Nagel had committed the 'fallacy of incomplete evidence', which involves describing a situation with one body of information, raising the possibility that a relevant distinction is to be drawn between individuals (or predicates - though I ignore this part of their dispute since Carnap gave up his earlier position in **Basic**), and then asserting that such distinctions ought not to be ruled out a **priori** even if the information we actually possess gives us no basis for discrimination.(90)

I think Carnap failed to meet Nagel's criticism. The assumption that inductive logic ought to be symmetrical with respect to individuals is objectionable in **Cont**, for in the languages considered in that system it is possible for individuals to differ in some way relevant to an hypothesis under examination, and yet, because we have no information on whether the individuals are relevantly similar or different, the logic requires us to treat them as identicals. This is an arbitrary and unjustified assumption, as Nagel argued: why should it require evidence to establish a difference while identity - and thus uniformity among individuals, for it is this that the assumption of symmetry amounts to - is assumed unless there is evidence against it? Why place the burden of proof on establishing difference rather than uniformity? Clearly this distribution of the burden of proof is favourable to induction, and it requires some justification.

The justification Carnap offered, however, is completely inadequate, for he argued that

> to put it in very general terms, we require that logic should not discriminate between the individuals but treat all of them on a par; although we know that individuals are not alike, they ought to be given equal rights before the tribunal of logic. This is never questioned in deductive logic, although it is seldom stated explicitly.(91)

Surely the analogy does not hold, however; for if a deductive inference (with true premises) is valid then it is that fact alone which establishes the irrelevance of all properties of the individuals concerned other than those on which the validity of the inference depends. But the same is not the case for a correct inductive inference (to accept that there are such inferences for the sake of this discussion). If it is the case that all men are mortal and that Aristotle is a man we may be sure that he is a mortal regardless of whatever else he might be - say short or tall; but if it is the case that some men are immortal and that Aristotle is a tall man then it is by no means clear, in the absence of any statistical information on the proportion of tall men that are mortal, that we are nonetheless entitled to infer that Aristotle is a mortal (even if we would be allowed the inference if we knew that height was not a relevant variable in relation to mortality). Moreover, even if we do not know whether Aristotle is short or tall, and also do not know whether this would be relevant, still it is not clear that we would be entitled to the inference. I am not saying that we would not be, but only that there are problems to be resolved here which are peculiar to induction. Carnap's analogy between deduction and induction in relation to the symmetry of individuals does not hold; in deduction symmetry is entailed by the validity of the inference; in induction is it problematic.(92)

Our discussion of this problem with the axiom of the symmetry of individuals has thus far been based upon the system of inductive logic in **Cont**. In **Basic** the situation changes completely, for here symmetry with respect to individuals is only taken to be axiomatic for languages which do not have the resources to draw any distinction between individuals, languages which have only a single family of predicates only one of which attaches to any individual. Here symmetry is unavoidable; apart from a difference between individuals arising from their having different predicates from the family attached to them, they are taken to be the same in all other respects, there being, as it were, no other respects of which one can speak.(93) Carnap makes it plain, moreover, that if another family of predicates were added to the language, even a family representing spatial or temporal location, symmetry of individuals could no longer be assumed but would have to be experimentally justified.(94) So symmetry with respect to individuals in any significant sense is given up in **Basic**.

A similar development occurs between **Cont** and **Basic** in relation to the assumption of symmetry with respect to Q-predicates, or attributes, to adopt Basic's conception. I shall briefly review the treatment of attribute symmetry, and asymmetry, in **Basic**, and this will lead us into a discussion which bears on a problem in **Basic** concerning the choice of L, thus completing our discussion of Carnap's inductive logic.

In **Basic** attribute symmetry obtains only as a special case. Particular attributes are thought of as belonging to families of related attributes and each attribute is identified with a region of the space occupied by the whole family. Carnap conjectured that only the metrical and topological properties of these attribute spaces are relevant to inductive logic.(95) Carnap's conjecture amounts to identifying an attribute, for the purposes of inductive logic, with a certain region of the attribute space and considering the relations of attribute size and similarity or influence between attributes, insofar as they are relevant to induction, to be represented by the size of the region associated with an attribute and the distance between the centres of this region and the regions associated with other attributes, respectively.(96) Thus each attribute has associated with it a parameter, gamma (denoted here by '**g**'), representing the size or width of the attribute (which corresponds to **w/k** in the **LFP**'s formalism) and another parameter, eta (denoted here by '**n**'), representing the distance or dissimilarity between attributes (which was not considered in the **LFP** at all).(97)

The parameters g and **n** become the fundamental parameters of the continuum of inductive methods in **Basic**, for once they are given for a family of attributes then **C(h|e)** is determined for **h** and **e** involving only that family. The parameter **n** is associated with the major philosophical development in the system of inductive logic in **Basic**, for it introduces the influences of analogy into Carnap's inductive logic, and it turns out that the parameter **L** of **Cont** can be defined in terms of **n**.(98) Let us, however, first get **g** out of the way before we turn to these exciting developments.

The only detailed discussion of inductive logic in **Basic** concerns a language with just one family of attributes. In such a language the determination of g for each of the attributes in the family corresponds to the determination of the **a priori** probability that an individual will have one of the attributes rather than any of the others, and thus, though the basis of such a judgement is unclear to me, the meaning of the choice is plain.(99) In a language with more than one family presumably the relative widths of the whole families would be determined in the same way, that is by considering the a **priori** probability that an individual would have an attribute of one family, say a colour, rather than an attribute of another, say a scent. But there is mystery enough in trying to sort out the basis of the choice of widths for attributes belonging to the one family, without speculating on how the problem created by multiple families might be resolved. All that Carnap says in relation to setting **g** values is that the matter can be dealt with subjectively - which would turn logical probabilities into subjective probabilities - or objectively, but the suggestion he makes here is susceptible to the criticism to be made below in relation to the values of **n** within a family, namely that such choices are theoretical and thus presuppose a developed inductive logic. The same problem would arise for the most obvious basis for determination of the values of g for the various attributes in a family, namely to base our determination on the relative frequency with which each attribute has been met in the past; such a procedure is obviously inductive, and thus requires a rule of inference to license our projection from past encounters to the actual frequency with which the attribute is instantiated in the region of space and time concerning which we wish to be able to make inductive inferences. We shall draw out the consequences of the problem of the obscurity of the basis for the choice of **g** in the discussion of the determination of **n**.

Two steps are involved in determining **n**. The first is empirical, the second Carnap has variously considered to be empirical, logical and pragmatic, for it corresponds to the choice of **L** in Cont. I shall examine these steps in turn.

Carnap first supposes that we might be able to directly apprehend degrees of difference and similarity between the attributes of a family, which would determine relative values of **n** for the pairs of attributes of the family - and thus determine whether the family has equal values of **n** for all attributes, or n-equality. Then, if this is not possible, he considers as a second possibility that psychologists might develop a metric for similarity based upon whatever fragments of a system they can elicit from our actual classifications.(100) But both of these proposals are completely undercut by an earlier point which Carnap made very forcefully: the degrees of confirmation based upon one theoretical system or one language may be completely reordered by a change to a new theoretical system because the new account of the nature of the phenomena under study can easily recommend a different similarity ordering leading to new values for n (and new conceptions of predicate width, leading to new values for g). The example Carnap gave was that of the family of colours; perceptually, purple is closer to red than blue and between the two; but from the theoretical point of view which identifies a colour with its wavelength, purple is closer to blue than red, and blue comes between red and purple.(101) Clearly judgements about similarity within the family of colours is theory dependent (as is our judgement of the width of an attribute: consider how we might change our evaluation of the width of white, perceptually identified as just one of the hues, on adopting the theory of colour based on wavelength). Thus there is no possibility that a value for **n** (or **g**) could in any case be irrevocably determined by direct perception, for we surely cannot maintain an account of the direct perceptual determination of the relative values for **n** (or **g**) for the attributes of any family in view of the failure of this proposal for the family of colours, which might have been put forward as a paradigm case of such a procedure.

The importance of this point is that it shows that the value of $\mathbf{C}(\mathbf{h}|\mathbf{e})$, for given **h** and **e**, is determined in part by the theoretical framework adopted for the expression of our hypothesis and the experiences constituting our evidence; more precisely, the value of **C(h|e)** is relative to the language chosen and the values of **g** and **n** for the attributes involved in **h** and **e**, which attributes are determined by the language employed. Thus the determination of the support offered **h** by **e** is relative to a choice of theoretical framework, which must be made prior to induction, and thus must be made on grounds other than inductive support for the high level theory or theories generating the framework, since to invoke inductive support for the choice of framework would require a higher-level framework to support the determination of second order C-functions, and so on **ad infinitum**. Recalling Reichenbach´s distinction once more, Carnap´s theory of C-functions is thus shown to be a secondary method of induction.

The significance of this feature of C-functions is well brought out by Hilpinen, who claims, correctly in my view, that it shows that ´direct probabilistic comparisons between theories are in some cases in principle impossible´, and he goes on to claim that this ´vindicates the scepticism expressed by some philosophers about the possibility of making such comparisons´.(102) Now while I do not disagree with this assessment I think that it is important to keep clear just what the problem is. We ought not to conclude from this problem with Carnap´s inductive logic that there is no useful function performed by C-functions just because their values are relative the theoretical frameworks, for there may be compelling reasons, though not to do with inductive support, for the adoption of one framework for a particular field at a particular stage of its development. That is to say, all other problems aside, Carnap's C-functions, while they cannot be taken as absolute measures of confirmation, may still be useful as measures of confirmation within a framework. This assessment, however, reduces the C-functions to one half of a theory of inductive support adequate as a complete logic of scientific inference, the other half required to be a theory of rational criteria for framework choice.(103)

It must be stressed, however, that this rehabilitation of C-functions is contingent upon solving the other problems they presently face. Of those discussed here the most important is that on the axiom system defining $C(h|e)$ the question of the reasonableness of induction is begged by the assumption of Kemeny's adequacy requirement CA4, while without this requirement the function C+, which prevents learning from experience, is admitted as a reasonable C-function.

We turn now to the final element of the theory of C-functions we have identified for discussion, the second step in setting n values, namely the adoption of an absolute value for n for one pair of attributes in a family, their relative values being determined earlier on.

Setting the absolute value of n for one pair determines all n-values for a family, and for families with n-equality this is equivalent to setting a value for L. To complete our discussion of Carnap's second system of inductive logic, therefore, we need to examine the factors governing the selection of an absolute value for **n** for some **pair of attributes** of a family. We shall concentrate on a family for which n-equality holds since we may then take advantage of Carnap's analysis of this case, which he discusses in terms of selecting a value for **L**.

In **Basic** Carnap moves away from the empirical and pragmatic approaches to the selection of a value for **L** that he had adopted in **Cont**, and discusses various adequacy requirements for C-functions which he hopes will lead to a reduced selection from which choice is to be made, and perhaps leave only one admissible value. He argues first that L ought to be less than k, the number of attributes in the family with which **e** and **h** are concerned; next that L ought to be greater than 1/2, on the basis of the unreasonableness of values for **C(h|e)** given by smaller values for **L**, here extending the earlier arguments given for rejecting the straight rule (**L** = **0**); and finally, on the basis that **L** ought to be chosen for general use, not for use with a particular family, Carnap restricts **L** to values between 1/2 and 2 (since if **L** were greater than 2 it would be greater than **k** for some possible family). I shall first discuss Carnap's reason for making **L** independent of **k**, and then turn to his major new argument in **Basic** on the choice of L, namely that we should have **L** < **k**.

As we have already noted, **L** is defined in terms of **n**. Thus for families of different sizes (differing **k** values) but the same distance between attributes, **L** should be the same and therefore independent of **k**. Carnap proves another result to the same end, namely that if any family with n-equality and k value $k > 2$ is given, then we can construct a new family based on a coarser partition of the attribute space, and thus with **k** value k´ between k and 2, but such that the effective distance between attributes, and thus the value of **n**, does not change. L, then, cannot depend upon **k**.

Now if we can take a further step and restrict the range of possible values of L to values less than **k**, we will have achieved a remarkable reduction in the size of the continuum; and since this result alone would rule out out **C+**, it is even more welcome, for it was to rule out **C+** that Kemeny´s CA4 was adopted, and that caused the major problem we have identified with Carnap´s theory of C-functions. However, things do not turn as might be hoped: something very like CA4 is needed to yield the result that L should be less than **k**.

Carnap´s proof of this result relies on the the assumption that if E_1 and E_2 are two samples of s individuals, and in E_1 all s individuals in the sample have a single attribute from a given family whereas in E_2 the numbers of individuals with each of the **k** attributes in the family are nearly equal, then the measure of E_1 ought to be greater than the measure of E_2 - roughly, the **a priori** probability of E_1 ought to be greater than that of E_2 - because E_1 represents a higher degree of order than E_2. But why should a **priori** probability increase with the degree of order or uniformity of the sample considered? Why, that is, is it reasonable to expect an event (sample occurrence) more strongly the higher the degree of order exhibited by the event predicted? Carnap's answer to this was to argue that the proposal merely assigns equal probabilities **a priori** to the various possible structures, structures with increasingly lower degrees of uniformity being instantiated in larger numbers of distinct samples, thus making the **a priori** probabilities of relatively non-uniform samples lower. But why assign equal **a priori** probabilities to the structure descriptions rather than to the state descriptions, to go back to the terminology of LFP wherein this proposal was first made?

I do not think that Carnap supplied any reasonable answer to this question, and thus accept the view put by Ayer, that assigning equal probabilities **a priori** to structure descriptions constitutes an unjustified assumption of uniformity which is factual in nature and thus requires an empirical justification.(104) Moreover, since assigning structure descriptions equal **a priori** probabilities is an assumption favourable to the success of induction, as Carnap made plain, in the context of discussing justifications for induction it must be rejected as question-begging.

This completes our critique of Carnap's pure inductive logic. Laying aside the criticisms we have made, and also ignoring for the sake of further discussion the other criticisms made in the literature which we did not take up, we now turn to the question whether it might be possible to use C-functions to sanction inductive inferences according to rules of inference which would not fall foul of Hume's inductive scepticism. For thus far we have only considered the problem of getting an objective degree of confirmation. We now pose the question whether, if we had such a measure of empirical support, could it be put to use in justifying inductive inferences without running into problems justifying the inference from the degree of confirmation to the adoption of the conclusion of the inference?

d. Carnap's Methodology of Induction.

Carnap's account of the structure of inductive inference is set out most fully in his 'Inductive Logic and Rational Decisions', a revised and expanded version of his 'Aim of Inductive Logic'.(105) In these papers he adopts the view that the aim of inductive logic is to help us make rational decisions, and defines a decision to be rational if it has the maximum value for the actor, where the value of an act **m** for X at t is given by its expected utility. The expected utility of act A_m for X at t, $V_{x,t}(A_m)$ is defined thus:

$$V_{x,t}(A_m) = \sum_n [U_x(O_{m,n}) \cdot P(W_n)]$$

where $U_x(O_{m,n})$ is the utility for X of outcome **m** in the state of nature W_n, $P(W_n)$ is the probability of W_n, and n is the number of possible states.(106) The task of inductive logic, ie. the theory of C-functions, is to provide the values of $P(W_n)$ for the various possible states of nature.

Now in the traditional analysis of induction, inductive inference, like deductive inference, leads from certain premises to the acceptance of the conclusion according to some rule of inference. According to Carnap's conception of inductive inference, however, inductive inference does not lead to the acceptance of the conclusion of the inference, but only to the assignment of a degree of confirmation to the conclusion given the premises. Carnap's motivation for adopting this conception of inductive inference is to avoid Hume's problem of induction, as he makes plain in discussing an inference from today's weather to the possibility of rain tomorrow:

> According to the customary view, on the basis of this evidence the 'inductive method' entitles us to accept the prediction that it will rain tomorrow morning. But then Hume is certainly right in protesting that we have no rational reason for the acceptance of this prediction, since, as everyone will agree, it is still possible that it will not rain tomorrow.(107)

Now there is a considerable body of literature on the adequacy of this analysis of inductive inference, focusing on the question whether we can make sense of the cognitive claims of science, both pure and applied, without having an inductive acceptance rule. I do not intend to review that literature here.(108) Rather I wish to note just one difficulty with Carnap's conception of induction, and then proceed to Kyburg's theory.

The problem I wish to emphasize does not concern the adequacy of the relation Carnap suggests holds between inductive logic and rational decision-making, nor the idea that the grand theories of science are never proved or accepted, even tentatively, but only made more or less probable as the evidence favours or undermines them - both of which are open to challenge. For it seems to me that the greatest problem for a theory of induction which does not include the possibility of acceptance - and such theories are not peculiar to Carnap, being a view common to all Bayesians - is not the problem of making sense of the application of the conclusions of inductive inferences in applied science, but in making sense of their use as premises in further inductions. Now since the role of premiss in a further enquiry is potentially played by all factual claims of any generality at all, the problem of the **acceptance of evidence** is basic to any theory of induction which does not provide for acceptance, as Carnap recognised. In IL&RD he noted the problem, and offered a brief comment:

> for the purpose of a special investigation, an investigator will usually accept some general assumptions... or specific assumptions, for example about the reliability of his measuring instruments. But I would not say that such assumptions are conclusions obtained by 'inductive inference'.(109)

But this is no solution to the problem. Rather, Carnap must either show how the factual claims which need to be assumed for some experiment to proceed come to be available, or provide an account of experimentation which shows experiments to be **in general** free of empirical presuppositions. The second option would surely be impossible, given the convincing arguments by Popper and others against the idea that some factual assertions are theory free, and Carnap clearly opts for the first. But he leaves us completely in the dark concerning how we come by factual assertions which are not justified by immediate experience nor justifiably accepted on the basis of some induction. He does, however, refer the reader to an earlier discussion of the point, and on this previous occasion he had suggested that we may be able to construe induction as an inference which does not require that its premises be accepted, but only that their probabilities be known. Here Carnap refers to Jeffrey's generalization of conditionalization, which was put forward to deal with the problem of giving a Bayesian account of learning from uncertain experiences - which is just the problem that Carnap's theory of induction faces, for he requires an account of learning from representations of experience where the empirical claims involved are not accepted, but merely confirmed.(110)

I think it can be shown that Jeffrey's generalization of conditionalization cannot perform the role that Carnap assigns it, and thus that Carnap's methodology of induction fails for want of explaining how it is that we can learn from experimentation. I shall clarify the problem below in discussion of Kyburg's theory of acceptance, but I shall postpone further discussion of Carnap's preferred solution until the following chapter where Jeffrey conditionalization (as it has come to be called) and related ideas will be examined.

Since I take the failure of Carnap's methodology, at least as he left it, to be an immediate consequence of his incorrect assessment of the suitability of Jeffrey conditionalization as a Bayesian representation of experimentation, I shall not further discuss Carnap's methodology, asserting here that it fails as a reasonable formulation of induction for lack of accounting for experimentation, a lacuna which it is later to be shown cannot be filled by Jeffrey's generalization of conditionalization.

We turn now to Kyburg, who takes from Carnap the idea that probability is a logical relationship between a body of knowledge and some hypothesis, but who attempts to provide an account of induction closer to the traditional view, by incorporating acceptance into his inductive logic.

5. KYBURG´S INDUCTIVE LOGIC.

Though he was not a co-worker of Carnap´s, Kyburg has a good claim to be counted as the intellectual heir to the leadership of the confirmation paradigm. For although he does not work with Carnap´s logical concept of probability, Kyburg has realized more nearly than any other author Carnap´s original aim of developing an analytic confirmation measure adequate to the practice of science.(111) While his programme has many features of interest our task here is to restrict our comments to the major features of the system, those which determine whether it is possible for the programme to meet the aim Kyburg laid down, viz to develop a system of probability

> in which the notion of an empirical body of knowledge, based on a body of experimental evidence, makes sense; and [is such that] up to an arbitrary choice of theories of equal empirical content, that body of knowledge is uniquely rationally determined by that body of experimental evidence.(112)

In his attempt to provide a theory which can meet this aim Kyburg´s system has generated a number of interesting puzzles, and while his aim cannot be met while these remain unsolved, we shall not deal with these but concentrate on the major elements of his theory. I believe that it can be shown that serious philosophical difficulties are inherent in Kyburg´s approach, and that major changes to his system would be required if the faults were to be overcome, changes for which there is no obvious means of implementation.

a. Kyburg's Theory of Acceptance.

The first problem I wish to raise concerns Kyburg's theory of acceptance in relation to the last problem we identified in Carnap's system, namely the problem of securing the premises of inductive inferences when the premises consist of data gathered from scientific experiments. It seems to me that there are problems with Kyburg's intended solution to this problem, inasmuch as it is not clear that it can account for the ability of standard types of experiments to yield data suitable to play the role of premises in intuitively cogent inductive inferences. All I shall attempt to do here, however, is raise the problem, for the solution requires historical as well as philosophical investigation.

The relevant feature of Kyburg's theory of acceptance is that he allows statements which are probable rather than certain to be accepted by proposing a hierarchy of levels of rational corpora, a statement being accepted into a corpus which has a level not higher than the lower bound of the probability interval associated with the statement. These corpora are not deductively closed, thus if p and q are members of a corpus it does not follow that their conjunction is a member of that corpus. Kyburg defends this feature by noting that the force of a deductive argument ought to depend not upon the acceptability of its individual premises but on the acceptability of their conjunction, and that the conjunction of a set of probable statements has a lower probability than any one of the individual statements.(113)

Now consider a complex experiment, which has many elements designed to ensure that the experiment is not disturbed by any extraneous factor - background fields, physical shocks, changes of temperature, impurities in reagents, fluctuations of power supplies and the like are all common disturbances against which we require controls. Typically these controls are set up prior to the experiment, and then inferred to be properly working while the experiment is in progress unless an unexpected result occurs, whereupon we try to check our controls, and if one fails at the time of this test then we may infer that the experiment was not properly controlled. Typically, then, our evidence for the claim that a control functioned properly is inductive, and on Kyburg's account of induction it follows that it will be at best probable that the control has functioned correctly. An immediate consequence of this is that there is some number of controls for which the probability that all worked correctly, ie. the probability of the conjunction of the claims that the separate controls worked, will be less than any figure we care to name no matter how near the maximum is the probability that the least reliable control functioned correctly.

Now of course it is intuitively reasonable that the degree of confidence we should have in the reliability of the outcome of an experiment should fall off as the complexity of the experiment increases (other things being equal), just as our confidence in the conclusion of an inference falls off as the number of premises increases (other things being equal). The question I wish to pose is whether the law of multiplication of probabilities gets the rate of fall off right or whether, as I suspect, the rate of falling confidence forced on us by the probability calculus is unreasonably rapid.

One way to approach this question is historically. We could examine a number of interlocking experimental setups to see if the result of the experiment which is the last in the chain - or perhaps at the apex of the pyramid would be a better metaphor - can be held to be reliable with a reasonable probability that this claim is correct, for realistic assessments of the probabilities of conclusions based on experiments earlier in the chain. For example, if 0.99 was judged to be a reasonable probability for all of the 70 controls contributing to some final experiment, these 70 being arranged in the structure of 10 controls in the final experiment with each of these controls tested in an experiment which itself has 7 controls, the first level experiments requiring no controls that cannot be observed to be operative at the time of the experiment (an unusual case, I would think), then the probability that the outcome of the final experiment is a reliable result is just less than 0.5.

Is this reasonable? Consider the inductive inference to, for example, the success of a control designed to ensure the purity of a reagent. This inference itself would be based upon data got from a controlled experiment - we may, for example, have conducted some other experiment which it is known fails to give a certain result unless the reagent is pure, a correct inference from this experiment requiring that a number of laboratory apparatuses be functioning correctly and that certain equipment be free of any contamination at the time of the trial, each of which might themselves need to be established by prior tests. Once we start to trace out the hierarchy of tests involved it does begin to look as though the number of controls which are involved in a chain of experiments will creep sufficiently high to reduce the probability that the final experiment was successful to an unacceptably low level. But for a convincing case either for or against Kyburg's theory of acceptance to be made out on such grounds would require detailed historical work, taking, for example, a reputable series of experiments to see if, when they are examined in this light they yield a final result which must be judged unreliable. If this is so then that counts against Kyburg's theory of acceptance and presumably suggests that acceptance must be acceptance-as-true, not acceptance-to-the-extent-of-the-probability of the claim accepted; if not, that corpora ought to be deductively closed (which is a stronger claim, since it entails the first if the conception of probability involved is not to violate the conjunction axiom).

There are, therefore, problems with the idea that a claim can be accepted and yet remain only probable. I do not claim here to have shown that the idea is unworkable, however, but merely that it requires some further discussion of the problem we have identified with Kyburg's account of acceptance.

b. Kyburg's Theory of Probability.

According to Kyburg's theory of epistemological probability, all probabilities are interval valued; probabilities are relative to a language L and corpus K; a statement **S** in **L** has a probability only if relative to K it is equivalent to a statement of the form $\mathbf{z(x) \in w}$, where z is a random variable defined on a set **y**, **w** is a subset of the set of possible values of **z** on **y**, and **x** is a random member of **y** with respect to the value of **z(x)** being a member of **w**; and the probability of **S** relative to **L** and K is **(p,q)** if **K** contains the statistical statement that the measure of the subset of **y** on which **z** takes values in **w** is in the interval **(p,q)**.(114)

For our purposes this brief definition will suffice to set out Kyburg's theory of probability, for it is plain even from this which features of the theory require our scrutiny, namely: what analysis does the theory provide of the assignment of probabilities to generalizations?; how do we come by the knowledge required for the the statistical statements on which probabilities are based?; and what is the basis for the randomness claim which licenses using this statistical statement as the basis of the probability assessment? To answer these questions, and eventually to come to some judgement on the cogency of inductive inference on Kyburg's model, we shall have to discuss a number of questions, beginning with Kyburg's account of support for generalizations, first for those which are universal, then for statistical generalizations, for Kyburg distinguishes sharply between the two cases.

c. Kyburg's Theory of Support for Universal Generalizations.

Kyburg originally put forward an account of the acceptance of universal generalizations which allows a universal generalization to be accepted into a corpus of level p if it is known that the generalization holds for at least 100.p% of members of the set covered by the generalization. Such acceptance of a universal generalization, say that all A are B, at the level p allows us to assert with probability (p,1) that a randomly chosen A is B. In his [1977b], however, Kyburg informs us that he rejects this account of how we come by universal generalizations because it 'never happens' this way, and he offers a new and novel theory to account for the matter.(115) The old and new theories are described and contrasted thus:

> The two cases [theories] may be described as follows: We discover by enumerative induction that practically all F's are G's. Other things being equal, this entitles us to believe that a given F is a G; but counter-examples do exist. Alternatively, we introduce a new term F* of which it is incorrigibly true that all F*'s are G*'s, and which is such that we can be practically certain that what we think to be an F* is in fact an F*. Of course if we find a lot of F's that aren't G's, then we will have to revise our statistical generalization in the first case. **In the second case we have to revise our views about the probability of error in making F* attributions.** And of course if we find that we are wrong in making such attributions a good deal of the time, it may well be very tempting to suppose that our language reform was misguided to start with.(116)

Now as I understand the story, Kyburg's motivation for his change of heart towards his intended philosophical spoof was that the received theory, which he employed in what I have called his first theory of the acceptance of universal generalizations, fails to take account of the fundamental role played by error in measurement and genuine scientific observation generally, though not in the armchair ornithology of philosophical tracts.

For the old theory takes observations to be unproblematic, at least in the paradigm case, and incorrigible, while generalizations are contingent and thus liable to be proved false by observations. In real science, however, the paradigm case of accurate observation requires not incorrigibility but a well understood distribution of errors. The now seriously developed new theory reflects this more realistic account of observation and takes generalizations to be analytic with their former empirical content being shifted into the theory of the distribution of errors of observation.(117) (The reader might think that this is an avoidably strong reaction to the former idealization of measurement; but we take Kyburg´s theory at its face value, nonethelesss.)

While much of the interest in this new theory centres on the new account of observation introduced by the analogy with measurement, I want to focus on another aspect, namely the logic of inductive support, not now for the universal generalization, since this is analytic, but for the linguistic reforms which create analytic generalizations. This is of considerable interest since, as we can see in the passage above, while the contingent generalizations of the old theory required support from enumerative inductions (or so it was held), no mention of inductive inference occurs in the description of the new theory. We shall see what has happened to inductive inference by paying attention to the reasons for adopting a suggested linguistic reform which makes one generalization, rather than some rival, analytic.

Kyburg´s account of these reasons is given in the analysis of the following example.(118) Let there be theories T_1 and T_2, in languages L_1 and L_2, respectively. Let S_1 be the set of statistical statements associated with T_1, such that each member of S_1 asserts of some observation statement entailed by T_1 that its frequency of erroneous application is in an interval near 1. Let S_2 be similarly defined. Now let the interval associated with the ith member of S_1 be (p_i, q_i), and the interval associated with the ith member of S_2 be (t_i, s_i). Compare

$$d_1^2 = \sum (q_i - p_i)^2 \quad \text{and} \quad d_2^2 = \sum (t_i - s_i)^2$$

In the old theory, d_1 would have been interpreted as a (natural - I leave out Kyburg´s argument for this) measure of how well T_1 fits the data, and thus have been taken to measure the inductive support for T_1. In the new theory, however, we would not say that one theory fits the data better than another, but rather that it makes better sense of it, that it involves a lower frequency of observation errors, or that it has ´more empirical content´ than a rival; and it is this, according to Kyburg, that in his new theory ´gives us inductive grounds for preferring one convention to another´.(119)

We seem, therefore, in Kyburg´s new theory, to have inductive support without inductive inference, but of course that is not the case. The inductive inference occurs in one of two places: either we take the members of S_1 and S_2 to record the observed error frequencies of the observation statements entailed by T_1 and T_2 respectively, and thus require an induction to get from ´$d_1 < d_2$´ to ´T_1 is superior to T_2´, ie. superior in general, not merely in the field of our present experience; or each of the members of S_1 and S_2 asserts not the observed error frequency of a particular observation statement, but its theoretical error frequency, which is obtained, of course, by an induction from the observed error frequency.

If we take the second option then the inductive support for a language shift, which has the effect of making some hypothesis analytic, derives from a series of inverse statistical inferences from the error frequency in a sample of applications of an observation statement to the error frequency in the population of applications of that kind. The logic of that kind of inference will be examined shortly, for it is an inductive inference to a statistical generalization.

If we take the first option we must provide a basis for inference from ´On the set of observations made thus far, T_1 entails a lower frequency of observational error than T_2´ to ´In general T_1 entails a lower frequency of error than T_2´, for it is the latter claim which can be plausibly interpreted as entailing that T_1 is the superior theory, not the former. Now if Kyburg were to take this option for explaining the inductive support for language shifts, he might deal with the inference we have shown to be required in any number of ways, for it is an ordinary induction by enumeration that is wanting justification. Thus he might, for example, seek an inductive justification for this induction, viz that since it has thus far been the case that most theories which are related as are T_1 and T_2 have continued to be so related as more observations accumulated, this pattern will be continued in the future. My point is that Kyburg's new theory for the acceptance of universal generalizations does not avoid the problem of induction.

But might we not be able to apply the new theory to the problem at hand, ie. take it to be an analytic second-order generalization that if the sums of the squares of the diagonals associated with a first-order generalization is less than the sums of the squares of the diagonals associated with some other first-order generalization (to give the relationship between T_1 and T_2 defined above a convenient description (120)) over a wide field of observations, then the relationship will hold good generally, and the superior theory is then to be described as a 'law' rather than 'a well confirmed hypothesis'. We could do this, but there would be little benefit. First, it is doubtful that we have a characterization of the kinds of theories to which we would want to apply the generalization, and also that we could formalize the reference to a 'wide field of observations' in a manner which kept the frequency of mis-identifications within reasonable bounds. But the main problem is that we are back where we began: suppose we employ the second-order generalization for a time and find an acceptably low frequency of misidentification of instances of the second-order generalization, and therefore conclude that our linguistic reform, namely the acceptance of certain hypotheses as laws, has been desirable. We must then ask ourselves what reason we have for expecting this low error frequency to be continued in the future, and it is plain that any reasoned confidence we might have in this continuation must be based upon an inductive inference, and thus the proposed solution to the difficulty presupposes a solution to the very problem to which a solution was sought.

Clearly, Kyburg´s new theory of the relationship between observations and generalizations, whatever its other merits, does not advance us in the attempt to place induction upon a secure foundation in reason.

d. Support for Statistical Generalizations: Kyburg´s Theory of Direct Inference.

Now let us turn to Kyburg´s account of how the evidence from a sample can support a statistical generalization about the population (or a claim about some other sample not identical with the observed sample). Kyburg´s theory here is a development of Fisher´s theory of fiducial inference, with the predicate ´rationally representative´ being employed to turn a direct inference into an inverse inference.(121) Now just how this transformation is brought about, and under what conditions, if any, it is a reasonable procedure for making inverse inferences, are questions of considerable interest. However, I shall not deal with them here, for Kyburg´s analysis of the direct inference itself is controversial, and we shall concentrate on that. For it can be shown, I shall argue, that Kyburg´s direct inferences are not rationally compelling. From this it follows immediately that his inverse inferences are not rationally compelling quite apart from any difficulties associated with the fiducial step which converts a direct inference into an inverse inference. With our brief being to examine inference schemes to see how they stand with respect to Hume's problem of induction, we thus have no need, if my assessment of Kyburg's account direct inference is correct, to consider the merits of his inverse inference itself.

Kyburg's system of direct inference has been the subject of an extended debate between Kyburg and Levi, a discussion which focused on the failure of Kyburg's direct inference to obey, at least in some cases, a principle Levi calls 'confirmational conditionalization' (hereafter denoted 'CC').(122) This debate touched upon features of Kyburg's system of direct inference which I shall argue prevent direct inferences from being rationally compelling, but the controversy dealt mainly with other, more technical, questions, and ended, I think, without the significance of the issues debated being got clear. Thus I shall only briefly review the debate, and then go on to draw out those points which it raised which I think to be of most interest in the context of our discussion of the cogency of direct inference.

In his [1974] Kyburg noted that we can think about conditional probability in two distinct ways. In the first case, we can identify the probability that an A is a C, given that it is a B, with the measure of the C's among the things which belong to the intersection of A and B.(123) In the second case we can think of conditional probability as an epistemic concept, as the probability that an A is a C relative to a rational corpus that embodies the information that it is also a B.(124) Kyburg calls the first of these concepts 'conditional measure', while the second he does not explicitly name but it is obviously intended to be known as ´conditional probability´. The point of introducing the distinction is that there are occasions on which the two functions take differing values, as Kyburg goes on to show. For conditional measure is given by Bayes´ theorem, but conditional epistemological probability can only be got in this way when

> relative to the body of knowledge constructed by including as part of that knowledge the conditioning statement, our original item **x** is a random member of the appropriate part of the original reference class.(125)

In his [1977] Levi challenged Kyburg´s analysis of conditional probability, proposing the principle that, for a fixed rule for determining probabilities (eg. a given c-function), the probability of an hypothesis **h** conditional on evidence **e**, relative to a corpus **K**, which is given by Bayes´ theorem, ought to be the same as the probability of **h** relative to a corpus K´, where **K´** is just **K** augmented by **e**. This is his principle of ´confirmational conditionalization´, or, as we have decided to abbreviate it, ´CC´. Thus Levi rejected Kyburg´s concept of conditional epistemological probability, taking Kyburg´s conditional measure to be the only legitimate concept of conditional probability. He went on to argue that Kyburg´s rule for conducting direct inference, which is also his rule for calculating probability statements, violates CC under conditions which, Levi claimed, were more natural than the conditions under which Kyburg had admitted that conditional epistemological probability would not be given by conditional measure.(126) Kyburg, however, was unmoved by this since he found Levi´s example of a case for which CC fails to be more ´bizarre´ than the cases for which he had already admitted that CC fails; and he set out more clearly just when CC does fail in his system of direct inference, and argued that such failures were well motivated.(127)

Thus, so far as the fact of violation of CC was concerned, the only point in dispute was the question whose example of a violation was the more bizarre. As to whether violation of CC is anything to cause alarm, Levi initially gave no argument for this view, and when challenged to give reasons for it provided a proof that because it violates CC Kyburg´s system allows an agent to reverse his preference ranking for a pair of bets on reclassifying an irrelevant but serious possibility as a non-serious possibility. He did not, however, explain why this is a problem for Kyburg´s system of probability.(128) Kyburg took the view that the real attraction of CC is that it is necessary to hang on to it if you wish to adopt a universal betting interpretation for probability; if CC is given up then the fair betting quotient for h conditional on e, relative to K, may differ from the fair betting quotient on h relative to K augmented by e.(129) If you are willing to give up a universal betting interpretation for probabilities, Kyburg implies, you need have no qualms about dispensing with CC in favour of a rule for direct inference such as Kyburg offers. Levi´s latest paper on the subject does not offer any reply to this, though he reiterates his view that the main problem with Kyburg´s system of probability and inductive inference is its violation of CC, rather than the problems exposed by Seidenfeld and others.(130)

e. Randomness and the Cogency of **Direct Inference.**

Much of the controversy between Kyburg and Levi centred on the conflicting accounts they gave of an inference from the information that a man, Peterson, is a Swedish resident of Malmo, that a certain proportion of Swedes are Protestant, and that a less precisely specified proportion of Swedish residents of Malmo are Protestant, to the probability that Peterson is a Protestant. Levi suggests, if I have understood his claim, that for the direct inference (from the religious affiliations of the population we chose as a reference set for the inference, to the probability that Peterson shares the dominant affiliation), to be compelling for X, X must know that Peterson has been **randomly chosen** from the population adopted as a reference set. This is, in any case, the claim I wish to make out against Kyburg's view that for the inference to be compelling for X it suffices that he **does not know** that Peterson belongs to any class which is, according to Kyburg's rule for choosing the reference class for a direct inference, a more appropriate reference class. Now the rule Kyburg gives for choosing the reference class ties the choice of a reference class to the question whether the individual in question (here Peterson) is a **random member** of the class considered, the appropriate reference class being the class of which Peterson is a random member, according to Kyburg's definition of randomness. Kyburg's conception of randomness must therefore be the focus of our discussion.

Kyburg's account of randomness has undergone considerable change of detail since he first put it forward, and recently he has accepted a change in the status of his analysis. Initially he intended to provide a generally applicable account of when an individual should be considered to be a random member of some potential reference class for a direct inference. Lately, however, he has retreated from this plan in the hope of offering clear rules for certain common cases, and for the problematic cases thus far thrown up by his analysis. The first point to be noted, therefore, is that at present Kyburg does not offer an algorithm for determining randomness claims. Now since a potential reference set stands unless an alternative is judged superior by Kyburg's account of randomness, the outcome of the inference, ie. the probability assigned to the statement in question, depends upon the alternative potential reference sets we offer for consideration. Without an algorithm for determining randomness claims we have no way of knowing that we have considered all the possible reference sets and thus no way of knowing that we have picked the appropriate set.

A problem of this kind would remain, however, even if an algorithm for determining randomness claims could be developed, for even if we had a procedure for checking all possible reference sets, which one the algorithm directed us to select would depend upon what was known about each set. The fundamental problem is that on Kyburg's theory of randomness a potential reference set stands unless displaced by some other set, ie. a set stands as an appropriate reference set provided only that it is a potential reference set and no other potential reference set is judged superior to it. But this means that some inferences will go through simply because we are not in possession of information which is potentially available to us and which would upset the inference by forcing the choice of another reference set, while other inferences will go through because they are based on the most appropriate reference set. If direct inferences on Kyburg´s account are to be rational we must surely be able to justify the claim that the success of any particular example is a consequence of the appropriateness of the reference set on which the inference is based, not the paucity of our information about other potential reference sets.

In effect Kyburg proposes a new and unattractive principle of rational inference: **proceed to assign probabilities according to what you know regardless of how little you know.** However, it will surely be agreed that we have no basis for assigning any particular probability to a proposition unless we have reason to be sure that the assessment is well founded; assessments of probability are not reasonably made in ignorance if they are intended to be legislative for rational belief.(134)

In order to deal more precisely with the question of the information base necessary for the reasonable determination of a probability, consider a class **A** and a class **B** such that it is known that the relative frequency of **B** in **A** is p, and let **x** be known to be an **A**. We want to determine the probability that **x** is a **B**.

According to Kyburg, we may infer that $p(\mathbf{x} \in \mathbf{B}) = \mathbf{p}$ provided that we do not know that **x** belongs to any class other than **A**; or if we do know that **x** belongs to some other class **D** then the measure of **B** in $A\cap D$ is known to be an interval that includes **p**.(135) But this is plainly too slim a basis of information for the probability assigned to $x \in B$, by direct inference from the measure of **B** in **A**, to be legislative for the degree of rational belief in $x \in B$. I doubt whether many would be prepared to proceed with the inference on the basis of not knowing that there is any class **D** such that $A \cap D$ might be the more appropriate reference class **before checking**, say, that we have not had past experience with a class like A, say **A´**, which was found to share members with a class **D´** such that $A' \cap D'$ was the appropriate reference class - that is to say without first considering the possibility that $p(x \in B)$ ought to be influenced by an argument from an analogous case.

Of course a sophistication of Kyburg´s system might be developed to deal with this or any other similar possibility. But no such development could deal with the problem which is at the root of the point we are discussing. For it does not matter how exhaustive is the list of possible reference classes we must check and discard before accepting some candidate, we shall not be prepared to accept the candidate class as the appropriate reference class, and thus not be prepared to accept the inference as legislative for rational degree of belief, unless we can show that the candidate is the appropriate class, and not merely the best of those considered, or at the very least that the tests conducted to determine the appropriate reference class are competently conducted and could reasonably be thought to have yielded the appropriate reference class. In short, our willingness to proceed with a direct inference according to the model

Kyburg gives depends, I suggest, on our either knowing or accepting as a reasonable judgement that if the reference class employed was inferior to some other possible reference class then we would have discovered this and employed the superior class. The problem now is to find some formal criterion to take the place of these vague and intuitive demands.

It is not hard to be fairly precise about the stronger requirement, that we should know that the reference class employed was the appropriate class. We would be prepared to accept the class A as the appropriate reference class for a direct inference if we knew that none of the predicates which occur in an exact and exhaustive description of x define a subset of A in which the measure of B (the property with which we are concerned) differs from the measure of B in A, and that x's selection from A was random with respect to x's being a B. These two conditions ensure that the long run relative frequency with which the inference turns out to have a true conclusion equals the probability the inference gives the conclusion.

I take these to be the conditions Levi lays down for a direct inference to meet CC and also, which I have taken to be more important, for the inference to be rationally compelling.(136) But it is very doubtful that these conditions could ever be met, and it is not even clear what the second of them means. Take the problem of meaning first. Levi would say that each member of A should have an equal chance of selection at each trial, but that throws little light on it. Hacking would resolve equal chances into equal frequency of selection in the long run, but we would then have the problem of applying what is true in the long run to the particular selection we are interested in, and would also have to admit that few if any interesting inferences, and none where the entire membership of the reference class is not available for selection, will meet the criteria. Moreover, supposing that we got the idea of random selection clear in a way that allowed it to play the role intuitively intended for it, there would still be the problem of ensuring that the conditions were met for some inference under scrutiny. This alone is a sufficiently weighty problem to sink direct inference, as both Kyburg and Levi point out, for how are we to ensure that either condition is met? Levi writes:

> one cannot hope to be in a position to justify the admission of chance statements into evidence on the basis of testimony of the senses, the records of memory, an acceptable conceptual framework, and my account of direct inference, without further substantive assumptions.(137)

Levi is untroubled by this, however, claiming that it is a difficulty only for an empiricist such as Kyburg. Kyburg responds by declaring that Levi´s pragmatism substitutes ´postulating shared agreement about matters of empirical fact´ for ´constructing shared agreement on the basis of empirical data´, and rejects this approach as having the ´same advantages as theft over honest toil´, borrowing Russell´s famous phrase.(138) Now we shall examine Levi´s theory in the next chapter; what is to be concluded here, however, is that either we have, following Levi, set unnecessarily strict conditions for direct inference on Kyburg´s model to be legislative for rational belief, or that, since these conditions cannot be met, Kyburg's programme for providing a rational basis for statistical, and more particularly inductive inference, is unsuccessful.

No doubt the requirements Levi lays down for direct inference to be rationally compelling are stringent, but what weaker set of requirements would do the job? Above we suggested that the minimum standard would be that it be a reasonable judgement that if there was a superior reference class to that used in a direct inference then we would have discovered this class. But we do not even have the formal theory of randomness that would allow us to precisely characterize what is meant by 'superiority' in this standard, and additionally there is the problem of finding some standard to assess the quality of the search for alternate reference classes so that we might reasonably declare a fruitless search to be fruitless because there is no better reference class than the one presently considered the best available. In short this suggested weaker standard is not an informal standard which now merely requires formal expression; rather it is the expression of the hope that we can find some standard for rational belief which is weaker than the standard Levi proposes. Which is what we should expect, for Levi's standard for justified inference is in effect Hume's standard of reason, viz that it is rational to accept only those inferences known to be truth preserving, Levi adapting this to probable inferences by insisting upon a standard which entails that the long run relative frequency with which an inference proves to be truth preserving should equal the probability lent the conclusion by the inference.

f. Kyburg on the Justification of **Induction.**

In a paper on the justification of inductive inference Kyburg set out his view of what it takes for an inductive inference to be rationally compelling, claiming that

> The appropriate thing to ask about an inductive rule (or a definition of probability) is... whether we can conceive of a universe in which (for example) (1) all of the A's that an individual has seen have been **B**'s, (2) there is absolutely nothing else that the individual in that universe knows, and yet (3) it would be **irrational** for him to expect the next A to be **B**.(139)

Kyburg obviously takes the answer to his implied question to be 'No!', but I, and I imagine many others, do not share his intuition. We can conduct a small test to see who is right, moreover, and I think that it counts decisively against Kyburg's view of the matter. For we can construct universes for ourselves at the drop of a hat: for example, imagine that you have observed one sayjum and it was bojan; are you then rationally justified in expecting the next sayjum (if another exists) to be bojan (or, at least, rationally justified in assigning this prediction a determinate probability)?. Or alternatively, should you reason that since a limited amount of bojan might have to be spread among the sayjums then it is reasonable to expect the next encountered sayjum not to be bojan? Or, finally, should you refuse to opt for either alternative and withhold judgement until more information is available? (The kind of information on which we would likely wait would be observations of sayjams at different times and at different places and when in the vicinity of the different things and beings we might find in this universe, and not merely the encountering of more sayjums.)

My point is that the line of argument Kyburg used to defend induction takes us back to the major difficulty we held to exist in his theory of direct inference; for an inductive inference to be rationally compelling we need to know that its premises state not merely all the relevant information we in fact have, but that the premises were arrived at by a process of observation and experiment which it is reasonable to assume would have yielded any further relevant evidence which was potentially available to us.

If we are to meet this standard, it will need to be given a much more definite formulation. We could do this either by specifying the kinds of processes of observation and experiment which must be exhausted before we hold an induction to be soundly based, or by specifying in broad terms what the outcome of our investigations must be. The problems of the first approach are familiar from Popper´s difficulties with the notion of a sincere test of an hypothesis. The second approach is easier to follow, but we have seen above that it plausibly leads to Levi´s conditions for inductive inference to be rationally compelling, which Levi admitted could not be met without abandoning an empiricist epistemology.

We conclude at this point, then, that Kyburg´s account of direct inference fails to provide a system of inductive inference which would compel the rational agent to accept that, via direct inference, our ´body of empirical knowledge´ is ´uniquely rationally determined´ by our ´body of experimental evidence´, and thus that Kyburg´s system does not achieve the aim he set for it.

CHAPTER 7 THE BAYESIAN PARADIGM

1. INTRODUCTION.

In this chapter we shall discuss the last of the three paradigms in inductive and statistical inference, the Bayesian paradigm. Our intention in this discussion is to determine whether the Bayesian paradigm has spawned a system of induction which is not susceptible to arguments for inductive scepticism akin to that given by Hume. In line with our policy hitherto we shall accept that Hume's inductive scepticism has been defeated only if a form of inductive inference is presented which does not allow the critic, who accepts the premises of an inductive inference, to reject the conclusion of the inference or to refuse to attach a specified probability to the conclusion of the inference, without transgressing against a principle or principles of rationality to which the critic can be committed. We shall be led to the conclusion that the arguments of the Bayesians do not defeat inductive scepticism.

As indicated in our initial analysis of statistical inference, there are a number of divergent models of inference within the Bayesian camp. The first task of this Chapter is to give a brief account of the divisions between these variants of the Bayesian paradigm. Four versions of Bayesianism will be distinguished: the objective theory (Jeffreys, with a distinctive contribution by Jaynes); the subjective theory (Ramsey, de Finetti, Savage and Jeffrey); Levi´s theory of inductive inference; and finally the recently developed neo-Bayesian theories proliferated by a number of authors who have been influenced by Jaynes´ work on the objective Bayesian theory and his rule of maximum entropy, which is the basis for the new account of induction. Having separated these versions of the Bayesian paradigm we shall then consider the analysis of induction given by each group and consider its adequacy as an attempt to defeat inductive scepticism, dealing first with the objective theory, then with the subjective theory, next Levi, and finally with the neo-Bayesians.

Placing the neo-Bayesians to one side for the moment, we may say that while the remainder of the group is divided in many respects, they do form a single if internally fissured school, united by the central place given to Bayes´ theorem in inductive inference. For all but Levi, inductive inference **is** the calculation of posterior probabilities from prior probabilities and likelihoods via Bayes´ theorem or some generalization of it. (Levi allows that an agent might arrive at a posterior distribution not sanctioned by Bayes´ theorem, given the agent´s prior, if the agent changes his ´confirmational commitment´. He insists, however, that unless such a change takes place Bayes´ theorem provides the only allowable rule of inductive inference - indeed, this is just his principle of Confirmational Conditionalization, which we discussed in the previous Chapter.) The neo-Bayesians take a further step away from the traditional

Bayesian account, for they adopt a new basis for inductive inference, a principle of maximum entropy or minimum information; but nonetheless their work is still in the spirit of the Bayesian paradigm, identifying induction with the determination of posterior probability distributions on the basis of prior distributions. Moreover the new inference principles offered can be construed as generalizations of Bayes´ theorem, for under certain conditions Bayes´ theorem, and Jeffreys´ rule of conditionalization, can be given as special cases of the neo-Bayesian rule of inductive inference - though there are problems with this construal, as we shall see.

But leaving discussion of Levi and the neo-Bayesians till later in the Chapter, we are left with the traditional Bayesian theories, which divide neatly into two groups. This division is effected by the divergence of opinion among Bayesians concerning the nature of probability. According to the objectivists, who place Jeffreys at the head of their school, probabilities are objective, and thus the inductive inferences conducted by application of Bayes´ theorem are, when correctly carried through, rationally compelling for all. The subjectivists, on the other hand, followers of Ramsey and de Finetti, opt for a subjective analysis of probability, aptly titled ´personal probability´ by Savage. They concede, therefore, that an inductive inference, that is, the calculation of a posterior distribution from some prior distribution, is only compelling for those who share the prior distribution; for there are no objectively correct probabilities (except in extreme cases where values are derivable from the axioms alone), and thus no inductive inferences to which all are logically required to assent. Let us consider these two Bayesian models in turn, returning to Levi and the neo-Bayesians after spelling out the more traditional Bayesian theories and considering their many problems.

2. INDUCTION ON THE OBJECTIVE BAYESIAN MODEL.

While it can reasonably lay claim to older roots I shall take the origin of the objective Bayesian theory to be Jeffreys´ work, for in his **Therory of Probability** the theory was for the first time presented at the level of rigour and clarity introduced into philosophy by the successes of formal logic in the early part of this century. That is particularly significant for the objective Bayesian theory, for the major sticking point for all Bayesian accounts of induction has been the question of just what factual assumptions must be accepted if the theory is to be applied, and the axiomatic method Jeffreys adopted under the inspiration of the **Principia** had the potential for making that question as easy to answer as it could possibly be. I shall briefly outline Jeffreys´ theory, and then discuss those aspects of the theory which seem to me to be of most interest in connection with the attempt to find an answer to Hume´s problem of induction.

a. An Outline of Jeffreys´ Theory of Induction.

First let us note that Jeffreys certainly took his theory of scientific inference to supply an answer to Hume´s problem, not by providing any proof of the reasonableness of induction **a priori**, but by stating a set of rules for conducting inductions which were intended to be justified by their fruitfulness. Thus he **begins** with the presumption that

> there is a valid primitive idea expressing the degree of confidence that we may reasonably have in a proposition, even though we may not be able to give either a deductive proof or disproof of it

and takes his task to be to ´express its [the primitive idea´s-MR] rules´.(1) We shall examine this notion of justification later, in connection with the justification Jeffreys gave his rules and postulates. For the moment, however, it is important to set out the main features of Jeffreys´ rules.

Jeffreys aims to provide a formal theory of probability, where probabilities are interpreted as rational and objective degrees of belief, or ´reasonable degree[s] of confidence´.(2) However, he does not offer any formal definition of his ´reasonable degrees of confidence´, remarking that

> it is intended to express a kind of relation between data and consequence that habitually arises in science and in everyday life, and the reader should be able to recognize the relation from examples of the circumstances when it arises.(3)

Jeffreys thus throws down the gauntlet to those who, like Ramsey in his review of Keynes´ theory of degrees of confirmation, profess themselves unable to discern the primitive relation the theory is intended to formalize. In so doing he weakens his system, however, for there is a real problem here which cannot be dismissed so lightly; the presumption that there is **one** logic underlying the weighing of evidence in the sciences requires some defence against the view that there is no unitary method responsible for the success of scientific practice. Borrowing Carnap's terminology, we note, therefore, that Jeffreys does not provide an analysis of his explicandum sufficient to convince the reader that there is any genuine concept of 'reasonable degree of confidence' suitable for explication. But since it is not our intention here to question the presumption that there is a coherent inductive logic implicit in scientific practice, we shall not pursue this line of criticism.

Presuming that there is a coherent inductive logic underlying the evaluation of theories in science, and proposing to construe these evaluations as assessments of probabilities (defined as rational degrees of belief), Jeffreys' task is to provide rules for these evaluations such that the evaluations will turn out to obey the probability calculus as well as having other features he desires rational degrees of confidence to possess. To this end he proposes a number of axioms and conventions which are provided with varying depths of justification - the thinness of the justifications being a consequence of the lack of initial clarification of the explicandum as well as his idea that fundamental axioms, be they for induction or deduction, cannot themselves be proved; rather their suitability is measured by the usefulness of the system which they found. Indeed, the seven axioms which constrain reasonable degrees of confidence to be probabilities are not defended at all. Clearly Jeffreys took it to be uncontroversial that degrees of confidence should be probabilities, citing with approval Maxwell's statement that

> the actual science of logic is conversant at present only with things certain, impossible, or entirely doubtful,

> none of which (fortunately) we have to reason on. Therefore the true logic for this world is the calculus of Probabilities..(4)

That this statement contains such an obvious non-sequitur shows how far it was from Jeffreys´ mind that a justification needs to be provided for the assumption that reasonable degrees of confidence are probabilities. We, however, having seen the rejection of that assumption by Cohen; being mindful of the problems with acceptance arising from degrees of belief being probabilities which are exposed by Kyburg´s lottery paradox; and noting the arguments against having point-valued degrees of belief conforming to the the calculus such as that given by Kyburg; can now see that these axioms do not lead to wholly uncontroversial results.(5) Any justification based on the fruitfulness of the axioms is therefore problematic, since there is reasoned disagreement concerning, as it were, the quality of the fruit. Some direct justification is thus required for these axioms, but none is provided.

If we waive this objection and pursue the development of the theory, its main rule is not far to seek: given that our assessment of support for a theory consists in our evaluating the degree of confidence rationally warranted by the evidence at hand, and that these degrees of confidence are probabilities, it follows immediately that theories are assessed by calculating their probabilities given the evidence, where such calculations are possible. And since probabilities given evidence are determined by Bayes´ theorem, the ´principle of inverse probability´, Jeffreys asserts this to be the ´chief rule involved in the process of learning by experience´, without which ´a general theory of induction is impossible´.(6)

Note that in accepting Bayes´ theorem as a **rule of inductive inference** Jeffreys identifies the probability of an hypothesis on certain evidence and conditional upon some further proposition with the probability of the hypothesis on the original body of evidence augmented by the further proposition, ie. he assumes $\mathbf{P_m(h|e) = P_{m+e}(h)}$ (call this (*)). This equation, however, may be open to challenge - as indeed we have already seen in the debate between Kyburg and Levi over the principle of confirmational conditionalization. That debate, however, concerned the status of the principle in a theory of probability (Kyburg´s) for which there is not a unique probability associated with every pair of propositions, there being the possibility that adding **e** to our body of knowledge **m** might lead to our adopting a reference class for the calculation of $\mathbf{P_{m+e}(h)}$ different from that which served for the calculation of $\mathbf{P_m(h|e)}$. Jeffreys´ theory, however, as Hacking points out in his discussion of the distinction between the two probability functions in question, does associate a unique probability with each pair of propositions **(h,e)**, and thus (*) is a tautology in Jeffreys´ theory of probability - as it is for Carnap, in whose theory the necessity of (*) is apparent, since in his theory, and also in Jeffreys´ though there not so clearly spelt out, the probability of a pair of propositions is given immediately by their content.(7)

For (*) to be a tautology, however, places heavy demands on Jeffreys´

theory, for it requires that the probability of all pairs of propositions be given analytically, and thus that prior probabilities, legislative for all, be given analytically. Just how heavy a demand that is will be considered in the next section. What is relevant here is that it avoids the cost of (*)´s not being a tautology, namely the requirement of finding a proof for it. For as Hacking points out, (*) is a dynamic assumption not entailed by the static assumption that rationality requires that degrees of belief held at any one time conform to the probability calculus. This leads to considerable dificulties for personalist Bayesianism, for which (*) is not a tautology, as we shall see.

b. The Determination of Prior Probabilities.

i. Jeffreys´ theory.

Finding a rationally compelling rule for the determination of prior probabilities has been the major problem of the objective Bayesian theory. Jeffreys, like most modern authors, rejects the solution put forward by Bayes, which was employed extensively by Laplace, but considers the problem soluble. He offers his reasons for rejecting the Bayes/Laplace solution, and provides his own, after a brief discussion of the meaning of a prior distribution.

Noting that the problem is to find a rule for assigning prior probabilities when we know nothing about the true value of the parameter for which a prior distribution is sought, Jeffreys claims that the answer is clear if we recall that the probability is ´merely a number associated with a degree of confidence´, whose purpose is to give one´s degrees of confidence ´formal expression´. From this he takes it to follow that

> If we have no information relevant to the actual value of a parameter, the probability must be chosen so as to express the fact that we have none. It must say nothing about the value of the parameter, except the bare fact that it may possibly, by its very nature, be restricted to lie within certain definite limits.(8)

Now this line of thought leads naturally, via the notion that if we have no good reason to draw a distinction between any two values of the parameter then we ought to assign the same prior probability to these two values, to the Bayes/Laplace idea that ignorance concerning the true value of a parameter is properly represented by attaching the same prior probability to every possible value of the parameter. We thus arrive at the view that unless we have a reason to prefer some possible values of a parameter to others, a uniform prior density ought be adopted by setting $P(dx) = c.dx$, where **dx** represents the size of the interval containing the value of the parameter and c is a normalizing constant. But the uniform prior is beset with difficulties, two in particular. First, for parameters which can take an infinite set of values, and thus for all which are continuously valued, normalization cannot be achieved and thus the distribution is improper, ie. the integral of the density function is unbounded, for otherwise the probability of the parameter taking any given possible value would be zero. Second, the uniform prior leads to well known paradox, for a prior which is uniform over values of the parameter P will not in general be uniform over values of functions of **P**.

Rejecting the Laplacian uniform prior as a **general** solution to the problem of determining prior distributions Jeffreys thus faces two problems; first, to supply alternative priors for problems where the rejected distribution had been adopted; and second, to provide a principle to defend his choice of priors as the principle of insufficient reason defended the Laplacian uniform prior.

While Jeffreys makes considerable headway in relation to the **technical problem** of finding mathematically well behaved priors, so far as a non-mathematician is able to judge, he is less successful in finding a plausible philosophical rationale to defend the priors he recommends in the various cases for which he offers solutions. Jeffreys' innovation consists in finding a general rule for determining priors which will be invariant under various transformations of the parameter which we are trying to estimate, or which is the subject of an hypothesis we wish to test. The rule is not entirely successful, however, yielding 'inappropriate' priors for some cases, and thus requiring **ad hoc** amendments. Further work by Huzurbazar provided more comprehensively adequate rules, but to a philosopher fundamental problems remain, particularly the fact that while a prior on Jeffreys' theory of probability constitutes a supposedly uniquely rational distribution of belief, the basis for choice of a prior is mathematical convenience (invariance) plus agreement with common sense values.(9)

Defending his rules Huzurbazar argues that 'it is not logically necessary to produce a single invariance rule which will be satisfactorily applicable to all distributions', and that it is no more likely that we should find such a single rule than that we should find a single scientific law to explain satisfactorily all physical phenomena.(10) But this argument ignores what is the vital point about prior probabilities, namely that they must be legislative for rational belief, and that failure to find a single rule to generate such priors may justly lead one to conclude that the single phenomenon the rule would elucidate, the rational basis for determining prior probabilities, does not exist. At the very least we require that a single theory be given to explain why different priors are appropriate, and not merely convenient, in different situations, otherwise we may be elucidating not one phenomenon but many, contrary to Jeffreys´ supposition that there is one logic underlying the weighing of evidence in the sciences.

On the basis of a similar argument, and also because the recommended priors are typically improper and thus lead to values for some probabilities that are infinitely large, values which must somehow be meaningfully combined with probabilities on the normal scale, Hacking rejects Jeffreys´ rule for assigning prior probabilities.(11) I agree with Hacking; in Jeffreys´ system a probability function represents the rational distribution of belief, but I do not see that the critic is rationally compelled to adopt Jeffreys´ prior in any of the cases he examines, particularly in the absence of a proof that some such prior must be selected, for if not forced to make such a choice the critic can convincingly argue that there is no determinate or unique representation of his ignorance concerning the value of some parameter. The critic could offer the maximally indeterminate interval valued probability [0,1] as his prior probability for every possible value of an unknown parameter. If I have read him correctly, Jeffreys´ reply to this could only be that such an imprecisely specified prior prevents employing Bayes´ theorem to give a posterior distribution, but to take that as a reason for adopting a more determinate prior would obviously be to beg the question at issue here, viz the reasonableness of induction on the Bayesian account.

Finally, there is a powerful criticism of Jeffreys´ theory given by

Seidenfeld, who argues that when our posterior probability depends upon information from more than one experiment, the posterior probability depends upon which experiment is considered first, since on Jeffreys´ account the one experiment may not lead to the same prior as the other. The point is put most forcefully by an example he gives. Suppose we wish to determine the unknown volume of a hollow cube, and have at our disposal two means of gathering relevant data (which we intend to feed into Bayes´ theorem to yield a distribution over the various possible values ot the volume). First, we may fill the cube with liquid of known density, say 1 unit weight/unit volume, and then weigh this quantity of liquid, repeating the experiment a number of times, and then using the information gathered to estimate the true value of the parameter **v**, the volume of the cube. This estimate ought, according to Jeffreys´ theory, be based on a prior which is uniform over all possible values of v. Second, we may lay a rod of known density, say 1 unit weight/unit length, against the cube and cut off a section the length of a side of the cube, weighing this section to determine its length, giving us a set of data which we can use to estimate the volume of the cube, the estimate being based, if Jeffreys´ rule is followed, on a prior which is uniform over the various possible values for the length of the segment, ie over $(\mathbf{v})^{1/3}$. But this prior is not the same as that which we would adopt for the first experiment, so which experiment we begin with will determine which of the two possible, and **inconsistent** posteriors we will arrive at. Jeffreys´ theory, therefore, is bedevilled with the possibility of inconsistency arising from the order in which information is assimilated, which is no less obnoxious than the possibility of inconsistencies in the Bayes/Laplace theory arising from adopting different priors for the same experiment. Seidenfeld urges that the root of the trouble is the notion, central to Jeffreys´ theory, that there is some precise, unique prior which best represents our ignorance in a given situation.(12)

In the face of such criticisms, we have to conclude, I think, that Jeffreys did not solve the sharply focused problem of finding objective prior probabilities set for him by his conception of probability as a logically determined and thus objective constraint on rational belief. But others have taken up the task, most notably Jaynes.

ii. Jaynes´ theory.

Seidenfeld also discusses Jaynes´ account of the determination of objective prior distributions, which is an extension of Jeffreys´ theory of invariants and is based on the proposition that a prior need not be invariant under just any transformation of the parameters involved, but must be invariant under any transformation which yields an equivalent problem. The theory is that if we can identify a set of transformations of a problem which do not change the problem in any significant way, then we must adopt a prior which remains invariant if the parameters for which the prior is defined are transformed in any or all of the ways yielding equivalent problems. Jaynes gives a clear account of his theory in a discussion of the Bertrand problem, which, as Jaynes gives it, is the problem of determining an objective prior for the distribution of sticks randomly dropped into a circle such that the sticks form chords of the circle. Jaynes argues that the problem is unchanged if we do any or all of rotating the circle, changing its size, or slightly shifting its centre, and thus that the prior ought to be invariant under transformations representing any or all of these changes, but not for other transformations.(13)

Now it may be that this extension of Jeffreys´ theory will yield more satisfactory priors than did Jeffreys´ own theory, but it is plain that it will be subject to the difficulty Seidenfeld raised against Jeffreys´ theory, for the posterior distribution we arrive at after a series of experiments will generally depend upon the choice of a first experiment. Thus Seidenfeld argues that Jaynes provides no adequate solution to the problem of determining objective prior probabilities.

Seidenfeld has shown, I think, that the whole theory of objective prior probabilities is caught in a dilemma. If the prior is chosen on the basis of general considerations which are arbitrary with respect to any particular problem, there is the possibility that the general rule will be applicable in more than one way allowing incompatible priors to be generated, while if rules are given such that each problem defines a unique prior then the posterior distribution supported by a set of experiments will in general depend upon which of the set was the first considered, since this will set the prior or ´informationless distribution´. Neither horn of this dilemma is a comfortable seat for a theory which takes objective probabilities to be rationally enforceable as degrees of belief, for it is deeply held both that posterior degrees of belief ought not be affected by the way the relevant problem is described, nor by the order in which information is processed.

3. THE SUBJECTIVE BAYESIAN THEORY.

In this section we shall discuss the subjectivists´ characterization of their relationship to Hume´s problem of induction and the justifications given for the axioms of subjective probability. The other main topic which needs to be considered in relation to the subjective Bayesian theory of induction is the adequacy of their methodology for science, ie. the adequacy of their analysis of inductive inference, which will be left until (4) below where the Bayesian theory can be dealt with as a whole, there being many issues common to both the subjective and objective accounts of the methodology of induction.

a. Subjectivism and Scepticism.

One feature of the subjective Bayesian tradition which requires immediate clarification is the aim of the theory. Jeffreys, as we have seen, sets out to refute inductive scepticism by providing rules for the rational evaluation of our degrees of confidence in hypotheses when our evidence for them falls short of deductive proof or disproof. The leading members of the subjectivist school, de Finetti and Savage, however, profess agreement with Hume. Savage notes that ´The riddle of induction can be put thus; What rational basis is there for any of our beliefs about the unobserved?´, and comments that as far as his theory of personal probability is concerned

> The theory as such is silent, but I am lead by study of it to doubt that there is a rational basis for what we believe about the unobserved. In fact, Hume´s arguments... appeal to me as correct and realistic.(14)

But despite this passage, and others like it, Savage is not an inductive sceptic in the sense philosophers have generally used that term. There are two important non-sceptical parts to Savage´s view of induction. First, while induction cannot provide a rational foundation for any **particular** belief or strength of belief, there are rational and irrational **changes** of belief; for Savage claims that while one may have no rational foundation for one´s belief, it can be rationally got from one´s previous beliefs, and would be rationally got from one´s previous beliefs by, and only by, use of Bayes´ theorem in connection with some new evidence. (Just prior to the passage quoted above Savage describes this as a ´partial answer´ to Hume´s problem; we shall evaluate that claim below.) Second, Savage reads Hume as a naturalist rather than a sceptic, as asserting that it is not reason but psychology which is the foundation of induction. This claim is made out most clearly by de Finetti, who takes ´Hume´s dictum´ as ´forerunning the present subjectivistic views about probability and induction´, rather than finding in Hume´s argument ´only its negative aspects´.(15) He claims that

> In connexion with induction, the tendency to overestimate reason- often in an exclusive spirit- is particularly harmful. Reason, to my mind, is invaluable as a supplement to the other psycho-intuitive faculties, but is never a substitute for them.(16)

Neither of these claims concerning the foundations of induction clear the matter up, however, for as is plain from the subjectivists writings they are not proposing a psychology of induction, but a philosophy of induction; the theory is normative, not descriptive. In fact their inductive scepticism amounts to the twin claims that there is **less to induction than is ordinarily supposed** - induction is the influencing of opinion by evidence via Bayes´ theorem, not the establishment of truths concerning the unobserved; and that there is correspondingly **less to inductive logic than is ordinarily supposed** - inductive logic consisting only of the weak constraints on rational changes of belief needed to ensure consistency of changes of belief with the probability calculus and with Bayes´ theorem. Jeffrey makes the matter plain:

> The radical claim that de Finetti makes, and characterizes as a translation of Hume´s ideas into logico-mathematical terms, is that this subjectivistic concept of probability is all we need for science, for statistics, and for decision making under uncertainty.(17)

It is this claim which we shall attribute to Savage and de Finetti, and also to subjectivists such as Ramsey and Jeffrey, and it is this claim we shall subject to scrutiny in this section. That scrutiny must obviously be directed to two questions: whether the subjectivists provide adequate arguments for the consistency constraints they seek to place on all changes of belief to be admitted as rational; and whether these constraints, defining induction on the subjectivist model, in fact suffice for the practice of science. We shall take up the first of these questions immediately, the other in (4) below.

b. Justification of the Axioms for Subjective Probability.

There are two approaches to the justification of the axioms of the probability calculus when probabilities are interpreted subjectively, as degrees of belief or personal probabilities: the axioms are derived from axioms governing preference between options of some kind (prizes, acts, or even propositions - the details need not concern us); or the axioms are derived from considerations about reasonable betting. The first of these needs to be supplemented with a rationale connecting belief (or whatever is taken as the basis of the interpretation of probability) and preference, the second with a connection between belief and betting. Both of these approaches were identified by the two fathers of the subjective school, Ramsey and de Finetti, and both have been developed by others. We shall review the various proposals.

i. Ramsey´s theory

Ramsey noted the ´old-established´ connection between measuring a person´s beliefs and observing their willingness to bet, and observed that if a person´s degrees of belief violated the probability calculus then (given a definition of degree of belief in terms of willingness to bet at certain odds) ´He could have a book made against him by a cunning bettor and would then stand to lose in any event´.(18) But he did not make this connection with betting the foundation of the probability calculus as others have done, preferring to ´work out a system with as few assumptions as possible´. His new system was based upon a set of axioms for coherent preferences between various options, but the connection between the axioms for preference and his definition of degree of belief is far from clear - at least to me. For after listing his axioms for preference, Ramsey merely informs the reader that

> If the option of **a** for certain is indifferent with that of **b** if **p** is true and **g** if **p** is false, we can define the subject's degree of belief in p as the ratio of the difference between **a** and **g** to that between **b** and **g**... This amounts roughly to defining the degree of belief in **p** by the odds at which the subject would bet on **p**, the bet being conducted in terms of differences of value as defined.(19)

Before we consider any further the grounding Ramsey gives this definition let us clear up just what it means, for it does not seem to make any sense as Ramsey gives it. Though Ramsey does not give a mathematical expression of his definition its literal interpretation is given by

$$(a - g):(b - g) = (a - g)/(a + b - 2g)$$

which I can make no sense of as betting odds. Perhaps Ramsey meant to write 'quotient' instead of 'ratio', for

$$(a - g)/(b - g)$$

is equivalent to the ratio

$$(a - g):(b - a)$$

which does make sense, the agent being willing to pay a to get **b** if p and g if not, thus risking the price of the bet **(a)** minus the payoff if notp (g) for the possible gain of the payoff if **p (b)** minus the price of the bet (a) - which is a bet at the odds suggested.(20) But while this amendment to Ramsey's definition gives a rationale for it in betting terms, it is of no help in establishing a connection between his logic of preference and the definition of degree of belief. For I think Ramsey intended the connection to be as follows.

He takes the axioms for preference as basic, defining what it means to prefer one option to another. Then he assumes that one's degree of belief in **p** is given by the **proportion** of g, where **g** is one's gain if **p** is true, which one takes as one's **expectation** of the yield of following a course of action whose success is dependent upon the truth of **p**. It then follows from Ramsey's axioms for choice, or valuation, that one's degrees of belief abide by the probability calculus if one's choices abide by the axioms for valuation.(21)

This does not, however, provide an adequate justification for taking the axioms of the probability calculus to govern reasonable degrees of belief, for Ramsey's definition of 'degree of belief' is itself in need of backing. For it is not obvious that degrees of belief ought to be defined in the way Ramsey proposes, since, to take just one problem, an agent might pay a price for a lottery ticket which is above the expected value he assigns to it, in order to have a chance of becoming very rich. In order to eliminate this and similar deviations from the agent's behaviour, we need the constraints on valuation which come in train of being able to force the agent to make a prescribed set of choices, such as is given by adopting the notion of choice in a forced betting system, employed in the Dutch book argument.

So far as I am able to penetrate it, therefore, Ramsey's theory provides no basis for the axioms of the probability calculus, for while those axioms follow from his definition of degree of belief, that definition itself is wanting support. And thus, while Ramsey proves that if we accept his definition of degree of belief then degrees of belief which violated the probability calculus would also violate the 'laws of preference between options', he gives no argument which would prevent a critic from rejecting the axioms of the probability calculus as a constraint upon degrees of belief, without thereby violating Ramsey's axioms for coherent choice between options, for the critic can refuse to accept Ramsey's definition of 'degree of belief'.

Of course it might be that we have missed some part of Ramsey's argument, and to check this we ought obviously to seek help in understanding Ramsey's analysis from those who have developed his approach to probability and preference, among whom Jeffrey's work is acknowledged as both clear and authoritative. But little light is thrown on Ramsey's argument by Jeffrey's discussion, for he generates Ramsey's definition of degree of belief by considering how we might measure probabilities in connection with the agent's calculations of the expected desirabilities of certain gambles. For example, if we offer the agent a gamble with payoffs **e** if **M** and **f** if not, which the agent assesses as having desirability d, then (writing **p** for the probability of **M**)

$$d = pe + (1 - p)f \quad (*)$$

from which it follows that

$$p = (d - f)/(e - f)$$

which is the definition of degree of belief we attributed to Ramsey.

But this is certainly no good as a basis for **proving** the axioms of the probability calculus, for the axiom of total probability is assumed in the derivation of (*). In fact Jeffrey later provides a proof of this axiom, but it is defended by a variant of the Dutch book argument, and the other axioms are simply assumed. Despite its other virtues, therefore, there is no proof of the axioms of the probability calculus in Jeffrey's interpretation and extension of Ramsey's theory of preference.(22) Nor have I been able to find any better exposition in the literature, and thus I turn to other arguments for the axioms of the probability calculus as constraints upon reasonable belief.

ii. De Finetti's theory.

Now let us consider the foundation for the axioms of the probability calculus to be found in the work of the other (and independent) founder of the subjective Bayesian theory, de Finetti. In his early essay 'Foresight: Its Logical Laws, its Subjective Sources', de Finetti, like Ramsey, notes that the axioms might be got from considerations on rational betting or more directly. His more direct way is to state a set of informal rules for probability, such as that one event can only appear as more, less, or equally probable than another, and then show that the usual axioms can be derived from these informal rules. However, his informal rules are stated in terms of **probability**, not degree of belief.(23) This is significant, for while his informal rules have strong intuitive backing as principles of probability, they lose much of this support when stated as rules governing degrees of belief. Certainly it cannot simply be assumed that one belief is as weak, as strong, or stronger than another, as some beliefs seem to be unsuited to such comparisons - indeed Keynes among others has argued that non-comparability is a pervasive feature of belief. Therefore while this first route to the axioms is certainly direct, it works only for the axioms of the calculus when 'probability' is taken as a primitive term, rather than when probabilities are taken to be the measures of an agent's degrees of belief, and thus the argument provides no basis to assert that a subject's degrees of belief ought to obey the axioms of the calculus.

De Finetti's second proof of the axioms suffers from something of the same problem. In this case he proves that a subject's betting quotients ought to obey the probability calculus, but does not provide any discussion of the legitimacy of taking betting quotients as **normatively** equivalent to degrees of belief, ie. for the assumption that one's degrees of belief ought to be the same as one's betting quotients. Nor does he justify taking betting quotients rather than degrees of belief as the foundation of a theory of inductive support for hypotheses, which would be his other alternative. Thus he fails to secure the axioms of the probability calculus as a constraint on rational sets of degrees of belief.(24)

Those who followed Ramsey and de Finetti sought to provide clearer arguments for taking degrees of belief to be probabilities. In particular, the logicians who developed the Dutch book theorem, as the defence of the axioms of the probability calculus in terms of rational betting behaviour has come to be called, argued that betting behaviour is an appropriate basis for assigning degrees of belief to rational agents; while Savage sought to replace degrees of belief with preferences between acts as the basis of the theory of subjective probability. We shall examine these remaining defences of the axioms in turn.

iii. The Dutch book theorem.

Although suggested by Ramsey, and sketched by de Finetti, the reliance on an analysis of rational betting strategies as a basis for the axioms of the probability calculus, under the interpretation that probabilities are degrees of belief, is relatively recent, and has been the concern of philosophers rather than statisticians. The great merit of the approach is its brevity and apparent simplicity, seemingly requiring only that it be accepted from the outset that it is irrational to follow a course of action that will lose one money (or goods of some kind), come what may, for it to follow that one must base one´s actions on degrees of belief conforming to the probability calculus. As has become apparent from the philosophical scrutiny the argument has received, however, it is not convincing - at least that is my assessment of the matter.

That assessment is based principally on the paper by Kennedy and Chihara, replying to Jackson and Pargetter, which I shall briefly summarize. In their presentation of the Dutch book argument, intended to avoid the criticisms of Baillie, Jackson and Pargetter identified what they called the ´universalizability principle´ as a major presumption of the Dutch book argument. Kennedy and Chihara accept that the universalizability principle is indeed a presupposition of the Dutch book argument, as one must, the need for this principle being made plain by Jackson and Pargetter´s careful analysis. But Kennedy and Chihara argue convincingly that that principle cannot be satisfied jointly with another presupposition of the argument, which they call principle ´R´, thus showing that the argument requires premises which cannot all be granted, and thus the argument does not go through.(25) Even if that were the only argument against the Dutch book argument, and Kennedy and Chihara present quite a few more, it should, I think, be given up. But it is of some interest, in any case, to consider further argument against it, for there are two further important problems with the argument which have so far, to my knowledge, gone unnoticed.

Consider first the knowledge which an agent must possess if he is to be able to ensure the coherence of his beliefs. We can illuminate this as

follows. Let X be considering laying a series of bets on the propositions p, q and r, each of which he considers to be just as probable as the other two. Suppose also that he considers that only one of the three can be true, and that at least one must be true. According to the Dutch book theorem X should therefore be willing to bet against each of the three at odds of 2 to 1, ie. his degree of belief in each should be 1/3. But before laying his bets X recalls the importance of avoiding laying himself open to having a book made against him, and also notes that he will have laid himself open to such a book if he has a non-zero degree of belief in any **s** which is such that exactly one of **p**, **q**, **r** and **s** is true.
Now X has a non-zero degree of belief in many propositions, and thus is in the position, should he go ahead with his initial three bets as planned, of laying himself open to a Dutch book if any of these other propositions are, **unknown** to X, bound to be true if all of p, **q** and r are false. For example, if X bets on all of **p**, **q**, r, and **s**, and one of p, q, and **r** is true, X will break even on the bets on p, **q** and **r**, since these are all at 2:1, but will lose his bet on **s**; while if **s** is true, X will lose all of his bets on p, **q**, and **r**, and the bookie can set the stakes on **s** so as to pay out less than he has won on p, **q** and **r**. Thus X loses, come what may.

To avoid such a book, X must be in possession of certain knowledge, namely, **knowledge of what is possible**. It is important to note that this knowledge is **factual**, not logical, for what determines, for example, whether it is true that the coin must land heads or tails is whether it can land on its edge, or disintegrate on impact, or do any other such thing; and if the coin cannot do anything but land heads or tails, that is plainly a matter of fact. Thus we cannot avoid the problem by investing the agent with infallible logical judgement, for an unerring sense of logical possibility will not suffice to delimit what is possible in the relevant sense. Nor will it avoid the problem for X´s misconceptions concerning what is possible to be shared by his bookie, for while the bookie would then fail to take advantage of X´s mistaken belief that his betting quotients conform to the probability calculus, it would then not be true that X can only avoid a Dutch book by ensuring that his beliefs conform to the probability calculus; it will suffice that he find a bookie that fails to take advantage of his lapses from coherence. And once we admit this possibility, why not assume that Nature is such an incompetent, or benevolent, bookie? Clearly, for the Dutch book argument to have force we must assume that we will always have a book made against us if we lay ourselves open to that possibility, and we make that possible whenever, in consequence of a mistaken belief concerning what is seriously possible (to borrow Levi´s apt term), we would be willing to accept bets which violate the axiom of total probability.

The problem now, however, is to account for X´s having the knowledge which we have shown to be a precondition of his avoiding a Dutch book. The problem is significant, for the Dutch book argument is of use primarily to Bayesian theories of induction, for they require that an agent´s degrees of belief be constrained by the probability calculus. But Bayesian theories do not provide for the acceptance of hypotheses, even though the Dutch book argument presumes that the agent has accepted statements concerning what is possible.

More particularly, it is a problem for **subjective** Bayesianism, since frequency and logical theories of probability do not rely on the Dutch book theorem for securing the axioms of the probability calculus as constraints on admissible degrees of belief. Moreover, the problem is particularly acute for a subjective theory of probability, for the statement which the agent accepts as defining what is possible must be **true**, if his beliefs are to be coherent, and the subjective theory of probability provides the agent with no way of testing whether his beliefs are true. Thus setting coherence as the standard of rationality, and holding to the subjective theory of probability, entail that the agent cannot know whether his degrees of belief are rational. The requirement of coherence, so far as the subjective theory of probability is concerned, is therefore a norm which cannot be applied, and as such provides no rational foundation for induction, not even on the subjectivists' minimal conception of rationality.

This problem takes away some of the attraction of subjective probability. But, to digress from our discussion of the Dutch book theorem for a moment, that does not recommend either the logical or frequency theories of probability, for an analogous problem arises for each. According to the logical theory of probability, probabilities are analytic, and thus it is not required of the agent that he determine whether the probability of heads plus the probability of tails should equal unity. But, plainly, the problem of the agent's determining what is possible is avoided by the logical theory of probability at the cost of incorporating factual assumptions into the language and thus posing a real problem of language choice. Finally, for objective, but non-logical, theories of probability, that is to say for frequency theories of probability, the problem we are discussing is familiar, for it is equivalent to the problem of determining the set of possible outcomes of the stochastic system in relation to which the probabilities in question are defined. In our discussion of the reliability paradigm we have already seen that the need for specifying these various possible outcomes causes trouble, requiring Neyman-Pearson hypotheses tests to admit a factual assumption. A similar difficulty arises for objectivist Bayesian theories, for the determination of the objective probabilities characterizing some set of possible outcomes of an experiment would require the assumption of a statement of factual possibility which could not, without launching an infinite regress, be justified by a prior induction. Clearly, therefore, objective Bayesianism will have to be supplemented by an attack on foundationism (which we discussed in connection with Giere's work in the reliability paradigm).

Returning now to the Dutch book theorem as a basis for requiring that degrees of belief be probabilities, consider the implausibly individualistic model of science on which the Dutch book argument is based. If there is good reason for the individual to avoid a Dutch book, so too is there good reason for a collective of individuals to ensure that they do not collectively have a book made against them, for that would entail certain diminution of the collective's fortune. But science is very clearly a collective enterprise; what one researcher can achieve depends very directly and obviously on the work of another. Suppose for example that you are trying to culture some newly discovered micro-organism, and I am responsible for ensuring the sterility of the laboratory equipment; that your degree of belief in the proposition that the experiment will fail if the equipment is not sterile is high, while mine is low; that I do my job badly, while you, knowing me to be committed to the team, assume I do it well; and that I assume that you share my low opinion of the importance of sterility in the laboratory. The divergence between our degrees of belief in the proposition that the tray in which the culture is supposed to grow is sterile will lead us into trouble; indeed, if we have left our University positions to get rich quick in the genetic engineering boom they will cost us dearly. This cost could only be avoided if we shared degrees of belief on any matter where divergence would lead us to different conclusions concerning our experimental work. That we should be **collectively coherent** is no less important than that we should be individually coherent, and collective coherence requires shared degrees of belief.

Clearly the more apt analogy for science is not the individual laying bets according to his own beliefs, but the firm of commodity traders where the different employees must all buy and sell at the same price, lest a rival buy from the broker in the firm accepting the lowest price and sell to the one offering the highest price, guaranteeing himself a profit on every tansaction. This analogy shows the importance of collective coherence in the same way as the desire to avoid certain loss in a Dutch book gives us a reason for individual coherence, and presumably, therefore, if individual conherence is a reasonable constraint then so too is collective coherence. But the subjective theory of probability provides no basis for us to resolve our differences over the degrees of belief we ought to share - other than, of course, by conditioning on our evidence and hoping to be forced to agree before our differences force us to the wall; and this is irrelevant in the present discussion, since coherence is supposed to be maintained at all times to ensure that one must always obey the probability calculus to be rational.

In sum, then, I claim that the Dutch book argument fails due to the faults catalogued by Kennedy and Chihara, while if we set these aside, then, for two reasons, the argument in fact **undermines** the subjective theory of probability as an analysis of rationality which hopes to throw some light on scientific method. First, because the condition upon rational belief on which the argument is based cannot be knowingly met be any agent unless he has knowledge for which the subjective theory of probability cannot account. Second, because the image of individual rationality on which it is based is a fiction so far as science is concerned, while to make it more realistic we must go beyond the conception of rationality as no more than coherence, to provide a mechanism for adjudicating between various proposals concerning the degree of belief in some proposition which ought to be common among those involved in a joint enterprise.

iv. Savage´s theory.

In his [1954] Savage provides a rigorous analysis of the concept of preference between acts and shows that provided that a set of preferences meets a few plausible requirements they are related to each other as are probabilities. Thus Savage secures the axioms of the probability calculus for the subjective theory of preference, provided only that preferences must meet his constraints.

Only one of his requirements could, in my view, be **directly** challenged as a realistic and reasonable constraint on rational preference, and that is his first and fundamental principle, that preferences should be capable of being simply ordered: ie, if **x** and **y** are two acts possible for a subject then either **x** is not preferred to y or **y** is not preferred to **x**; and if a third possible act is **z** and **x** is not preferred to y and **y** is not preferred to **z**, then x is not preferred to **z**. But it may well be that preferences cannot generally be ordered in such a manner - they may not always be able to be compared, as Keynes held degrees of belief not to be in some common cases. But this would weaken Savage´s system only by restricting its scope. Whether that is significant for our analysis of induction depends upon whether preferences of the kind involved in induction are among those that can be ordered, or those that cannot. I shall suppose that they can be ordered as required, and thus not press any objections to Savage´s defence of his axioms for preference.

There are, however, problems in Savage´s approach akin to the problems we raised against the Dutch book argument. First, as Savage himself points out, the agent cannot be in a position to apply his (Savage´s) principles to bring coherence into his preferences unless he has available particular information, for example, unless the agent knows what acts are open to him.(26) But the subjective theory of probability does not account for the agent´s having such information, as I have already asserted, and will discuss further in the next section. Second, Savage constructs his theory of preference on an individualistic basis. But, as I argued above, science (indeed all knowing) is a **collective** enterprise. Now Savage does recognise the problem of providing a basis for collective decisions, and suggests that a collective could apply a version of the minimax decision strategy to reach a mutually acceptable decision.(27) This does not seem to me to resolve the problem, however, for each member of a collective who individually, or as a member of the collective, would have done better had their advice been followed, will always know this to be the case and no doubt draw it to the attention of the others. Clearly the advice to a group to forget their differences and find the decision which best fits their divergent views is to suggest to them that they pursue a decision-making policy which will build up in the group tensions liable to blow it apart.

This is fair criticism of Savage´s system, for he offers it as a normative theory, as an aid to decision-making. But so far as group decisions go, it is a bad policy, introducing instability into the group; it is bad methodology. Clearly a group would be advised to seek a joint preference ranking meeting Savage´s criterion rather than allow the point to be reached where a decision needs to be made while the members of the group still disagree concerning which course of action is to be preferred. Indeed, considering the weakness of Savage´s argument for the normative force of his principles, namely that to violate them by adopting an unsanctioned set of preferences makes one feel ´uncomfortable´ in much the same way as when some of one´s beliefs are ´logically contradictory´, it would seem that there is a stronger argument for a collective to adopt a single preference ranking than there is for an individual to ensure that his preferences abide by the principles.(28) The problem, of course, is that there is no basis in the subjective theory of probability for a group to argue out its differences on a rational basis other than by collecting evidence in the hope of coming to agreement after conditionalizing upon the information gathered, which as we said above is irrelevant to this case

In the light of these objections it is at least problematic that Savage´s proof of axioms requiring that rational preferences between acts should obey the probability calculus covers rational preference between the propositions of science. Despite the value of his work in clarifying the foundations of preference in practial decision making, therefore, it remains an open question whether the scientist ought to ensure that his degrees of belief conform to the probability calculus.

4. ON THE ADEQUACY OF THE BAYESIAN METHODOLOGY.

In this section we shall discuss problems with the methodology promoted by Bayesians of all varieties thus far identified, ie. with the claim that all there is to induction is the calculation of degrees of probability via Bayes´ theorem. I shall consider, in turn, what justification there is for the Bayesians´ claim that all learning from experience proceeds by way of Bayes´ theorem; whether Bayes theorem has the formal properties we expect of a logic of confirmation and disconfirmation; whether scientific induction requires that the conclusions of inductive inferences be accepted; and finally, relevant to subjective theories of probability alone, whether science can make do with subjective judgements of evidential support. In these discussions I accept, though in view of our analysis above and the criticisms of the Dutch book argument given by Kennedy and Chihara we are certainly not bound to do so, that degrees of belief ought to conform to the probability calculus.

a. Justifications of Temporal Credal Conditionalization.

The main topic remaining to be discussed in our analysis of the traditional Bayesian analysis of induction is the adequacy of Bayes´ theorem as a rule for inductive inference. Before taking this up, however, we must consider the possibility that Bayes´ theorem just **is** induction, and if we thought induction to have characteristics other than those of Bayes´ theorem, then we were just wrong. For this would be the situation if it were **proved** that rational changes of belief abide by Bayes´

theorem, ie. that rational changes of belief proceed by way of conditionalization upon our present beliefs, or, to use the current jargon, that changes in our beliefs over time obey temporal credal conditionalization. Are there any such proofs?

As Hacking pointed out in his attempt to make personal probability slightly more realistic, that we should **change** our beliefs in accordance with Bayes´ theorem does not follow from the Dutch book argument as discussed above, all criticisms of that argument aside.(29) For when one´s total knowledge is **K** and one is asked to bet on **h** on the hypothesis that **e**, the Dutch book argument asks that you ensure that this bet is coherent with the set of other bets made when **K** is your knowledge; however, when one comes to know that **e** is true, and one´s knowledge then becomes K´ = K+e, one is at liberty to **cancel** all former bets and begin anew, again being prompted by the Dutch book argument to ensure coherence. There is nothing in the Dutch book argument, so far as we have dealt with it above, to say that you should follow any particular rule when substituting your new set of bets for your old set. The Dutch book argument, that is, does not entail that learning from experience should go by way of Bayes´ theorem.

Since Hacking pointed this out, there has been some effort to give a justification for requiring obedience to Bayes´ theorem in the situation described, ie. for adopting temporal credal conditionalization. Teller considers various options open to the Bayesian, and I shall briefly review them here, following analysis of Hacking´s own suggestion.

i. Hacking´s argument.

Hacking´s paper on the matter was prompted at least in part by a paper by Savage discussing various problems in the theory of personal probability, among which was the problem that an agent desiring to bet on the value of a remote digit of **pi** would be required to calculate a betting quotient consistent with his degrees of belief in the relevant mathematical propositions, even if the possible reward arising from the bet was much less than the cost involved in calculating his coherent betting quotient.(30) Hacking proposed to avoid the problem by restricting the range of propositions over which one must post coherent odds to the set of propositions one explicitly accepts by direct evidence or **actually conducted** inference therefrom, rather than the deductive closure of the set of propositions one explicitly accepts. On this ´examiner´s sense´ of knowledge there would be no incoherence in posting odds on a remote digit of **pi** at variance with the odds one would post if one had deduced from what one knew just what the digit must be, for if the deduction has not been made the agent is not said to be committed to its result.

Hacking´s amendment to the theory of personal probability requires that the axioms of the theory be varied to conform to the new concept of knowledge; for example, the probability of **h** will be set at unity only if the agent has determined that –h is not possible.(31) It also entails, he claims, the dynamic assumption that $\mathbf{P_m(h|e) = P_{m+e}(h)}$ (granted, of course, the static assumption that degrees of belief should conform to the probability calculus). But since he does not explain how this comes about we are forced to guess at his reasoning. I take his argument to be that because we are not identifying knowledge with the deductive closure of the set of accepted propositions, we augment our knowledge by **e** alone on accepting **e** into our set of known propositions, and thus on the examiner's definition of knowledge there is essentially no difference between holding the set **m** and the proposition e as our knowledge and holding as knowledge only **m** but considering e as an hypothesis, at least as far as calculating the (relative) probability of **h** is concerned.

Does Hacking's proposal secure conditionalization? Not if I have correctly guessed at the reason Hacking thought it did, for while conditionalization will be obeyed when the agent accepts **e** into his knowledge, still, at later times, as he becomes aware of hitherto unnoticed consequences of **e** in conjunction with various of his previous beliefs, he may well change his beliefs in ways that are not sanctioned by conditionalization. Indeed, it is plain that Hacking's slightly more realistic personal probability cannot secure conditionalization, since slightly more realistic personal probability for the very unrealistic standard agent of Bayesian theory, who is possessed of unlimited logical abilities, is just the standard kind of probability so far as this ideal agent is concerned. Thus there remains, for such an agent, a problem with the justification of conditionalization.

Now clearly the justification of an epistemic principle cannot rest on what is contingently true about ourselves as agents, namely that at the time we accept some new propositions we do not generally see that certain consequences follow from those propositions. For there would doubtless be occasions when an agent would see that he was committed to some further proposition whose acceptance would force him to abandon the principle of temporal credal conditionalization on that occasion. The proposal for restricting conditionalization to additions to bodies of knowledge meeting the examiner´s definition does not work. But note again that this is a criticism not of an argument Hacking actually gives, but of one that we have attributed to him. That it succumbs so easily to criticism indicates that he had some other argument in mind, but I am at a loss to suggest an alternative.

ii. Teller´s argument.

Teller presents two arguments to support temporal credal conditionalization, the first an extension of the Dutch book theorem to conditional bets, devised by Lewis, the second an argument of his own, based on analysis of the conditions in which an intuitively plausible constraint on confirmation is justified. I shall not discuss the first argument since it deals only with the Dutch book **theorem**, leaving the problems of the Dutch book **argument**, ie. the identification of reasonable degrees of belief with betting quotients, unsolved.(32) The second argument, however, merits close scrutiny, for it is imaginative and rigorously presented, though ultimately, I shall argue, unsuccessful in the form given.

Teller´s argument is given in two stages, the first requiring the assumption, common to all standard Bayesian accounts of conditionalization, that the evidence upon which we conditionalize is known with certainty, ie. has probability one, the second supporting the generalization of conditionalization due to Jeffrey, which is applicable to conditionalization on uncertain evidence. (We have already noted this rule for generalized conditionalization above, in discussing Carnap´s methodology of induction.) I shall only discuss Teller´s second stage argument since, for reasons to be given below, I hold certain evidence to be a fiction, referring to the first stage only as necessary to understand the second. Also, I shall use the term ´conditionalization´ to refer to both conditionalization proper and generalized conditionalization, distinguishing between the two only where it is necessary to do so.

Before turning to the detail of Teller´s argument it should be noted that it rests upon the principle of coherence, ie. that reasonable degrees of belief conform to the probability calculus. Thus it rests upon an unsecured and controversial foundation. But his original argument for conditionalization has nothing in common with the arguments for the principle of coherence and thus does not share their difficulties. It would, therefore, be in no trouble on this account should a sound argument for the principle of coherence be found, and thus deserves discussion even though it is, since we lack a proof of the principle of coherence, at the moment unsuccessful as an argument for conditionalization.

It should also be noted that Teller´s argument is intended to apply only to circumstances where the ´set of propositions for which the agent entertains beliefs is assumed to be fixed´, thus ruling out of court objections to conditionalization such as that given by Suppes.(33) In so doing Teller prevents his theory from dealing with any induction which involves the introduction of a new idea or explanation, being concerned only with the re-apportioning of probability among old beliefs. This restricted case is still important, however, for while no actual scientific induction proceeds under any such ban on new ideas, it is certainly the case that in assessing evidence on many occasions no new idea is in fact brought into play. Teller´s object of analysis is thus sufficiently general to be of considerable interest.

In outline Teller´s argument is this. He identifies an (essentially qualitative) intuitively acceptable constraint on any acceptable confirmation function, and then proves that this principle of confirmation is equivalent to the principle of conditionalization, granted that the agent´s belief function meets certain mathematical conditions. His proof is long and detailed, and I do not propose to set it out here. (Nor do I have the mathematical background to comment authoritatively on the plausibility of the formal assumptions concerning belief which it employs.) To do justice to his paper, however, we should quote his condition on confirmation, for the power of his argument is a direct consequence of the simplicity of this principle. The condition on generalized confirmation is as follows:

$$\mathrm{GC}(\{E_i\}) \equiv_{df} (\forall i)[(P_o(E_i) = 0 \rightarrow P_n(E_i) = 0) \ \& \\ (\forall A)(\forall B)\{(A \Rightarrow E_i \ \& \ B \Rightarrow E_i \ \& \\ P_o(A) = P_o(B)) \rightarrow P_n(A) = P_n(B)\}] \qquad (34)$$

where the set $\{E_i\}$ is a set of mutually exclusive and jointly exhaustive propositions, ie. a partition of the possible outcomes of some observation.

The constraint on confirmation given by GC is essentially that, if the prior probability of an element of the partition is zero then so is its posterior probability, while if its prior probability is non-zero, and if hypotheses A and B are equally probable prior to the observation and both entail this element, then they should be equally probable following confirmation by that element of the partition.

That it follows from GC that confirmation ought to proceed by conditionalization is an exciting result, for the relative simplicity of GC makes it possible to test at an intuitive level for the conditions under which GC is an acceptable constraint upon confirmation functions. That is how Teller uses GC, as a tool for testing for conditions under which confirmation must obey conditionalization, and it leads him to defend the following 'principle of inductive logic' which sets out the conditions under which GC, and thus the principle of conditionalization, ought to be met ($\{E_i\}$ being a partition of the possible evidence as before). The principle offered as determining when changes of belief ought to obey temporal credal conditionalization, for that is what is at issue, is this:

PG (a) The agent´s initial degrees of belief are all reasonable.
(b) For any $\mathbf{E_i}$, if the agent is certain that $\mathbf{E_i}$ is false before the change, then he is certain that $\mathbf{E_i}$ is false after the change.
(c) The agent´s strength of belief in at least one of the propositions $\mathbf{E_i}$ changes, and after the change the agent´s beliefs in E_i is reasonable for each **i**.
and (d) After the change of belief, any reasons the agent might have which in fact make reasonable or justify changes in belief in any proposition $\mathbf{A} \notin \{\mathbf{E_i}\}$ are beliefs whose objects are propositions in $\{E_i\}$ or disjunctions of these propositions; or else such reasons rest indirectly on his beliefs in the E_i, or their disjunctions. (35)

Should this set of conditions be accepted as governing the conditions under which observations are incorporated into the body of scientific knowledge, then, Teller argues, GC should be met and therefore conditionalization should be the rule of inductive inference employed. But it is controversial, I think, that science follows, or ought to follow, a methodology meeting PG (a) to (d) - even if we lay aside the objection that there is as yet no sound argument for (a), the coherence requirement (and a similar comment applies to (c)). For condition (d) embodies an anti-theoretical model of science which is widely rejected. Our criticism of Teller´s argument will be restricted to arguing that (d) ought to be rejected as a serious misanalysis of the relation between theory and observation.

Teller makes it plain that the acceptability of PG (c) and (d) rests upon the acceptance of a certain account of observation, whose relevant features are essentially these conditions (36):

(i) The set of possible observation statements $\{E_i\}$ is such that the E_i are logically exhaustive and mutually exclusive, and in the course of the event observed the agent's strength of belief in at least one of the E_i changes.

(ii) If the agent's strength of belief in E_i changes then that change is caused by the environment's effects on the agent's sense organs.

(iii) Such changes in strength of belief take place without any conscious reasoning of any kind.

(iv) For any **A** which is not one of the E_i nor a disjunction of any of them, if the agent's degree of belief in **A** changes then that change in belief in **A** is not both caused by the environment's effect on the agent's sense organs and also not the result of conscious reasoning of any kind (ie., the conditions of (ii) and (iii) do not both hold for **A**).

(v) Conditions are such that following the observation the agent's beliefs in the members of $\{E_i\}$ are all reasonable.

Observations meeting these five criteria Teller calls 'G-observations', and he argues first that many observations count as G-observations, and then that for such observations the principle PG above holds (in particular, that (c) and (d) of PG hold), and thus that learning from such observations ought to take place by way of (generalized) conditionalization upon them. I shall examine his argument connecting G-observations with the satisfaction of the principle PG first, to get clear the nature of G-observations, and then argue that no observation of any interest to science counts as a G-observation.

According to Teller, it can be shown that belief ought to be amended in the light of observation as determined by (generalized) conditionalization provided that the agent's beliefs prior to the observation were stable, and the observation counts as a G-observation. The agent's beliefs are said to be stable if

> none of his beliefs constitute reasons which would make reasonable or justify changes in the degree to which he believes other propositions.(37)

I take this to entail that an agent's beliefs are stable if a change in the probability of one does not affect the probability of any other, except, of course, for changes in the probability of a theory after conditionalizing on new evidence relevant to the theory. If this interpretation is correct then we can show either that beliefs are not in general stable, or that 'evidence' must be interpreted in a liberal manner to include theories among the propositions which might count as evidence and thus prompt changes in the probabilities of further theories via conditionalization.

Suppose we restrict the concept of evidence to propositions which do not depend upon other propositions for the probability they enjoy, ie. to propositions which are directly known or at least to which a probability can be directly assigned. Consider now an inductive inference which led, via conditionalization, to the adoption of a new theory, as must surely be allowed if conditionalization is to be a credible rule of inductive inference. Stability of belief could only be maintained if the adoption of the new theory left all other beliefs unaffected, apart from such changes as were the effect of calculating the probabilities of the other theories via conditionalization upon the evidence which prompted the adoption of the new theory. Indeed, this would have to hold true not only for the relative rarity of the adoption of a new theory, which will regularly also entail the rejection of an old theory, but also for the presumably common occurrence of a change in the agent's relative strengths of belief in competing theories.

Now this will be no problem for theories whose probabilities depend only upon evidence restrictively defined as above. But it makes a mystery of how to account for the case where a theory, **A**, which has the probability it enjoys because of the probability of another theory, **B**, changes probability when B suffers a fall in probability due to the observation of **e**, even though **e** is not directly, but only through **B**, relevant to **A**. Take a very simple case, the situation where $\mathbf{T_1}$ names the painter Smith as the originator of a certain style, identified by the distinctive use of colour, T_2 asserts that Smith used a certain pigment to get the original effect for which his works are famous, $\mathbf{T_3}$ has it that the pigment was first manufactured in a certain year, which happens to be, though this is not part of the content of the theory, the time of Smith's early works, and **e** is the discovery of a record of the first production of the pigment, some time after Smith's death. Conditionalizing on e lowers the probability of T_3, and the lowered probability of $\mathbf{T_3}$ lowers the probability of $\mathbf{T_1}$ or $\mathbf{T_2}$, depending on further evidence, even though **e** by itself is irrelevant to T_1 and $\mathbf{T_2}$ and thus conditionalizing on **e** would leave the probability of these two unchanged.

To handle this situation we will have to either admit theories into the category of evidence, or admit that conditionalization upon evidence narrowly defined can **introduce instability** into a set of formerly stable beliefs. The first option is a problem for Teller's theory because the theory rests in a way we shall soon investigate on a very narrow concept of evidence, while to take the second option is to admit that change of belief is not generally given by conditionalization, even under the restrictions Teller proposes.

I am not here arguing that the shaking out of a system of beliefs following some observation relevant to one theory only through another or through a chain of other theories can never be handled by conditionalization, where each theory in turn is the evidence conditionalized upon. Rather I am arguing that some process of shaking out will **typify** bodies of knowledge or systems of belief which are anything like sciences. The relevance of this is that stability of belief requires that theories be irrelevant to each other's probabilities, while the example above shows that to be a rare special case. Thus Teller's argument for changes of belief normally following the principle PG is off to a bad start, building on the atypical case. And that appeal to the special case continues with Teller's citing G-observations as the kind of observations upon which conditionalization is singularly appropriate as a rule of inductive inference - or so I shall shortly argue. First, however, we must return to tracing out the argument which connects the definition of G-observations with the holding of PG.

Teller argues that if an observation is a G-observations then conditions (c) and (d) of PG will be met. First, (i) and (v) of the definition of G-observations entail that (c) is met, but this is no more than definitional convenience, for the term ´reasonable´ is given no analysis sufficient to allow us to identify reasonable beliefs. We are told only that coherence is a condition of reasonableness, and that certain kinds of environmental impact on one´s senses do not lead to reasonable belief, an example being that while suffering a blow to the head might lead one to believe that money grows on trees such a belief would not be reasonable.(38) With this we will not doubt agree, but the matter cannot be settled in so simplistic a manner, for we will want to say, or rather, to maintain his a-theoretical analysis of observation, Teller will need us to be able to say, that an agent might arrive at the reasonable belief that a rock emits invisible rays which burn but are not hot, after receiving burns while handling a radio-active material - and this is not obviously reasonable, especially if asserted by (say) a twelth century alchemist. Teller´s concept of observation will clearly have trouble supporting the distinction between reasonable and unreasonable beliefs, for it is plain that the reasonableness of a belief cannot be a product of the causal antecedents of belief alone, the epistemic context also having a role to play.

Just how thoroughly a-theoretical is his account of observation, and the reasons for his adopting this implausible view, can best be made clear if we quickly set out the reasoning by which Teller asserts that an observation´s meeting (iv) of the criteria for G-observations entails, provided the set of beliefs is stable, that clause (d) of PG will be satisfied.(39) The argument proceeds as follows.

First, define the relation 'stems from ' as follows: proposition **X** stems from the set Y if **X** is a member of **Y** or a disjunction of members of **Y**. Now recast (d) of PG to say that after the observation causing a change of belief in E_i, any reasons an agent might have which make reasonable a change of belief in a proposition **A** which is not a member of $\{E_i\}$ either stem from $\{E_i\}$ or are beliefs at which the agent has arrived by a chain of reasoning whose original premises are propositions which stem from $\{E_i\}$. Now since the agent's beliefs are assumed to be stable, any reason for changing the degree of belief in A must be a new reason. Teller's case that an observation's meeting (iv) of the definition of G-observations justifies (d) of PG thus rests upon proving that if (iv) holds then if the agent's reason for changing belief in **A** does not stem from $\{E_i\}$ then it is a belief which is supported by reasons which ultimately stem from $\{E_i\}$. That is quickly done. Let R be the agent's reason for changing belief in **A**. R is either caused or reasoned, for if it is neither caused nor reasoned it is unreasonable. Say R is caused and not reasoned. Then it follows from (iv) that R stems from $\{E_i\}$ after all, for (iv) says that belief in all propositions which are not members of $\{E_i\}$ is not purely caused, so R must be a member of $\{E_i\}$. Now let R be reasoned, and let R' be the reason the agent has for R. We can repeat the analysis just conducted, and therefore conclude that the chain of reasons must either be infinite (which would presumably be unreasonable) or lead back to $\{E_i\}$. Thus it is shown that (d) of PG is secured by (iv) of the definition of G-observations, and thus, all other conditions being met, conditionalization must be obeyed for observations which are G-observations. But are there any such observations? That is the question we must now address.

Note first that if an observation is to count as a G-observation it must be completely free of any theoretical presupposition. This is made clear by the argument we have just rehearsed, if it is not plain from the definition of G-observations itself. But all observations are theory-laden, as Popper showed long ago.(40) And speaking of the **causes** of belief rather than the **reasons** for belief does not evade the point at all. Take, for example, the observation Teller offers as a clear case of being caused to know a proposition, namely seeing that the sun is shining (in this following Jeffrey).(41) This involves reasoning, **and that reasoning depends upon theory.** For one does not **see** that the sun is shining, one sees a bright light above and feels a sensation as though the light source were hot, and one infers from background theory that there is no **likely** cause of such sensations other than that the sun is shining. Of course there are other **possible** causes of such sensations, eg. having in one´s sleepy state rubbed liniment on one´s face rather than moisturizing cream, accounting for the heat, and a nearby sportsfield turning on their lighting towers, accounting for the bright light. Or, to take another example, one is certainly not caused to see that the evening star is shining, as this author is well aware, having been brought up short when the evening star which he had just seen, or so he thought, and was now waxing lyrical about, rapidly sank towards the horizon and then landed at a distant airport.

An observation which one can **get wrong** is not purely caused, for the effect of a cause cannot be anything other than the effect of that cause; if you **saw** something, **it was there**, thus if you were caused to see something, it was there. It makes no sense to say that one was **caused to see** something that was not there; rather one came to **think** one saw something which was not there, because one was **caused to have sensations** which one associated with the thing being there and one did not consider any further explanations of one's sensations. Thus one is not caused to see the stick bend in the water, for it does not bend in the water; rather one is caused to have sensations which one would normally associate with the stick having bent, but one has learned not to allow the association by coming to know that it rests upon the theoretical assumption that light travels in straight lines, which here does not hold true.

But all observations can be got wrong; the sensation caused by liniment can be mistaken for that caused by a heat source, and consequently one is not caused to observe a heat source, one reasons that one observed a heat source after discounting other possible explanations of one's sensation. It matters not at all that one does not often **actually consider** explanations of one's sensations other than that which one unreflectingly accepts, for while one does not consciously reason in this case that is only because one accepts **without argument** a claim which requires the support of argument if it is to be justified, a truism which can become apparent in embarrassing ways such as when the object of one's display of astronomical knowledge for the benefit of a new love lands safely and on time not long after sunset.

Rejecting Teller's causal account of observation leads us to reject the claim that some observations count as G-observations. But if there are no G-observations then the conditions under which Teller provides a convincing defence of his principle PG are never met. Thus conditionalization is not secured as a reasonable rule of induction even in the restricted case Teller examines (requiring stability of belief, reasonableness of initial and final beliefs, and a fixed set of propositions for belief to be divided among), for while conditionalization is proved equal to the weaker GC, the defence of GC itself comes unstuck, due to the impossibility of any observation meeting the criteria for showing that PG is justified as a principle governing inductive inference.

Perhaps Teller would object to my criticism by pointing out that if we insist that all observations are based upon reasoning from caused sensations then the set of propositions among which belief is shared will not be fixed, for the reasoning from sensations might and indeed commonly will, when such reasoning is consciously and deliberately conducted, include the consideration and discounting of hitherto unconsidered propositions. But this would not weaken my criticism, for the point then applies to this earlier restriction; Teller argues his case for conditionalization under restrictions which exclude all cases of scientific induction, since all observation statements are theory-laden.

Nor is it likely that Teller's argument could be repaired, for once we allow the possibility that a new concept or theory will be introduced, either as an hypothesis for evaluation as an explanatory theory or as a low level piece of background knowledge in the justification of an observation statement, we allow the possibility that the agent will completely review his theoretical commitments and be led to changes of belief more comprehensive than would be occasioned by conditionalization and unable to be generated by that rule of inference. Such changes provide ´cogent arguments against conditionalization´ as Teller admitted had been shown by Suppes.(42) At most Teller´s argument could be used as an after the event check on the reasonableness of the manner in which we had conducted an inductive inference, for if in fact our observation had involved us in considering no new hypotheses, nor prompted such consideration, then for that case the conditions of G-observations (doubts over the meaning of ´reasonable belief´ aside) would have been met, and there would be a basis for urging that GC should also have been met. But even this would be problematic, since it could always be argued that while we did not in fact consider new hypotheses in representing our caused sensations by a reasoned observation statement, we ought to have done so (perhaps on the grounds of simplicity), and thus that our inductive inference was not, after all, reasonable.

Despite these criticisms, however, Teller´s paper does provide a useful backing for conditionalization just by proving (under restrictions that have not been challenged here) that conditionalization is equivalent to the weaker principle of confirmation GC, for GC has considerable intuitive appeal.(43) Should we find, however, as we shall, that taking conditionalization as a rule of scientific induction leads to grave difficulties in the typical case, then the intuitions which defend GC will be countered by those which defend any features of current scientific practice found to be incompatible with conditionalization, and in this contest, it should surely be CC that gives way.(44)

iii. Skyrms' argument

Another argument for conditionalization can be constructed from material in a number of recent papers discussing the relationships between conditionalization, Jeffrey's modified rule of conditioning, and the maximum entropy rule of inference taken from Jaynes' work on prior probability distributions. The fullest account of the matter, so far as the justification of the principles of inference involved is concerned, is given by Skyrms. His argument is based upon the idea of levels of probability - where level one contains the agent's probabilities for various events and level two contains level one plus probabilities of these level one probabilities - and it purports to show that if we define levels of probability in a certain way then conditionalization at level one is equivalent to applying Jeffrey's probability kinematics at level two; and that this, subject to some restrictions, is a special case of applying the maximum entropy rule of inference. Thus, if there were some justification for the maximum entropy rule, or for Jeffrey's rule, this would be passed on to conditionalization in the specified circumstances. His argument, however, is in fact designed to run the other way, for he also shows that conditionalization at level two is equivalent to probability kinematics at level one, and he does this in order for the justification he believes to be given to conditionalization from Lewis'

Dutch book argument to be passed on to probability kinematics.(45).

Perhaps some might, however, find the principle of maximum entropy as a basis for inductive inference to be so intuitively appealing that they will take the connections Skyrms traces out to provide a rationale for conditionalization. That would be, I think, an error, for while the informal notions of maximizing entropy, or minimizing information, are intuitively appealing bases for inference, the actual rules of inference are a formal treatment which, while based on a clear argument, are not obviously unobjectionable. Indeed, as we shall see in the discussion of the neo-Bayesians below, there are objections to these rules of inference, in particular that - contrary to what Skyrms claims to have shown - they are inconsistent with conditionalization.

iv. Shafer´s analysis of conditionalization and subjective probability.

The final argument for conditionalization to be considered arises from a version of one of Bayes´ arguments, applied by Shafer to subjective probability. Shafer claims that Bayes´ argument for his proposition 3, which Shafer translates into modern notation as the standard definition of conditional probability, can in fact be reconstructed as a justification of conditionalization under certain circumstances. Shafer´s case proceeds by proving, along the lines of Bayes´ proof, the proposition

(*) $P_p(B) = P_n(A \cap B)/ P_n(A)$

the right hand side of which we know as the standard definition of conditional probability, while the left hand side is the probability we are to assign B at time **p** immediately after coming to know **A**, Bayes effectively proving, therefore, that the probability assigned **B** by conditioning on A at time **n** is the one we ought to assign it at p after learning A as expected.(46)

The proof requires, however, that a certain kind of relation exist between the events A and **B**, and the main purpose of Shafer´s exposition is to make clear just what relationship this is, to which end he provides a mathematical formalism for charting out the various paths which lead from one event to other events. From this it is clear that for the probability of B to be given by conditioning on **A** it must be the case both that **A** lies above B (thus preceding it in time) and that from the time at which we calculate the probability of **B** if **A** happens (by conditioning on A) there is only one path to A. Finally, by $P(B|A)$, we must mean the probability of B immediately after **A** happens. As Shafer puts it:

> what we need to assume in order to assure uniqueness for the probability of **B** immediately after **A** has happened is that there is only one way A can happen - ie. only one possibility for what other events have happened at the point where **A** has just happened.(47)

Conditionalization will be justified, therefore, only when this assumption is justified. When will this be? Before we consider Shafer´s answer let us reflect on the nature of the condition.

In fact the condition for the applicability of conditionalization which Shafer specifies, though he does not use this term, is that we have a (stochastic) model for the phenomena under study such that it follows from the model that the probability of B, given the addition of the event A to the model, is given by the probability of B conditional on A. For only on the basis of such a model can we rule out the possibility that something other than A which is relevant to the probability of B will occur prior to A, making the probability of B at the moment immediately after A´s occurrence not given by P(B|A). Thus Shafer´s condition on the applicability of conditionalization is that it can only be applied when we are sure we have been able to

> guard not only against the possibility not only that B itself may or may not have happened by the time A happens but also against the possibility that other events affecting the probability of B may or may not have happened by then.(48)

Now this, I think, is very interesting in the light of the development of the analysis of conditionalization since the time that Hacking pointed out the difficulty with it for subjective probability. For, while he did not say so, there is no difficulty with conditionalization for frequency theories of probability, for these theories define probabilities always relative to a stochastic model, say random sampling from an urn containing various coloured balls, and in the terms of the model the probability that a ball B will be drawn second given the fact that ball A is drawn first cannot be anything other than the conditional probability P(B|A). To adopt any other probability abandons the model.

With this in mind we can recast Suppes´ point that conditionalization only applies where there is no conceptual change as saying that we can only use conditionalization when we stick by our model. Abandoning the model requires us to abandon conditionalization. To this Shafer adds that for the determination of subjective probabilities to be bound by conditionalization it would be required not only that we stick by our model, but that **the model must also work**, for otherwise an event can occur between the first and second drawing - say some balls escaping through a hole in the urn - which will affect the probability of **B** immediately after **A**´s occurrence so that its probability at that time is not given by conditionalizing on **A**.

This last requirement applies to subjective probability, but not to frequency theories of probability, because subjective probabilities have **an existence independent of the model** - they are our degrees of belief - and thus they will respond to events which occur even if they are not provided for by the model; whereas probabilities on the frequency theory are, strictly speaking, merely ratios of classes defined by the model, and have no existence apart from the model. But of course agents use probabilities derived from frequency accounts as bases for their degrees of belief, and when they do so they must make the same assumption that is required, as Shafer shows, for the subjectivists to justify using conditionalization, namely that the **model is accurate**, or as he puts it, that events should be ´determined´ as expected, or that the possibility for the occurrence of events should be as we believe it to be.

But on what basis can we be assured that events will follow the model we have adopted, and thus that we can be justified in using conditionalization? Shafer suggests first that we would be so assured if the model, or the ´step by step determination of events´, was ´entirely objective´, it being ´nature that moves down the rooted tree´ which represents the temporal pattern of possible occurrences. Or it might be that our knowledge keeps pace with nature, and thus that the order of events is the order of our learning imposed upon us by the order of actual occurrences. Or, finally,

> A third approach is to interpret the step-by-step determination of events in an entirely subjective way. This means requiring that the events represented by our rooted tree all be events **that we learn some given fact**. And it means that the tree itself must be part of our knowledge: at the initial node we assign probabilities not just to future possibilities as to what facts we might learn, but also to the various orders in which we might learn these facts. In this context, conditioning on an exact event means conditioning on all we have learned. And the justification for this conditioning is based on our assigning probabilities beforehand to the possibilities for how our knowledge might develop.(49)

To clarify this, note that if we asked an agent who based his degrees of belief on a frequency model for probabilities why he intended to adopt the conditional probability **P(B|A)** as his probability for B after learning **A** ie, why he would change his degrees of belief in accordance with conditionalization, he would (presumably) reply that according to his frequency model there was, other than **A**, no possible event occurring prior to B which was relevant to his probability for B. An agent following Shafer´s advice would reply to the same question by noting that according to his model for the probabilities of learning various facts at various times there was now a zero probability of learning anything relevant to the probability of B prior to its occurrence, other than A. If the frequentist's model breaks down, ie. if an event relevant to the probability of B but not provided for in the model occurs, conditionalization is abandoned. Similarly if the subjectivist's model breaks down, ie. if he learns something relevant to the probability of B other than A prior to the occurrence of B, then conditionalization is abandoned.

There are two things to note about this analysis of conditionalization. First, it provides, under the restrictions given, not a justification for abiding by conditionalization, but rather a specification of the conditions under which it is justifiable. From all that has been said so far, the agent is free not to abide by conditionalization even if learning proceeds as planned. It is obvious, however, that all plausible problems with conditionalizing have been cleared away, and that for an agent not to abide by conditionalization when the conditions are met is to act in a way which is at the very least unmotivated. Indeed in later work Shafer goes further and argues, I think convincingly, that an agent meeting the given requirements for learning who did not change his beliefs by conditionalization would be incoherent in the sense that he would be violating his own expectations.(50)

The second thing to note about what thus becomes Shafer's justification of conditionalization is that we are only going to be committed to conditionalization by a **good** model for our future beliefs - or, to use terminology more appropriate to the subjectivist version, by a good anticipated history of our learning - ie. if our learning proceeds according to a step-by-step plan undisturbed by any unexpected information. I suspect that this restriction would lead to conditionalization being mandated as a procedure for learning from experience only in very artificial situations and very rarely indeed in actual scientific practice. Certainly the problems with the likelihood principle, discussed below, lead us to conclude that it is not generally the case that conditionalization is an adequate inductive logic. But we ought at least to grant Shafer that he has shown that if our learning has proceeded according to our anticipated history, then we ought to abide by conditionalization. If we accept the probability calculus as a constraint upon rational belief, therefore, we need to provide a reason for any case in which we reject conditionalization as a constraint on rational learning from experience.

b. Formal Problems with Bayes´ Theorem as a Logic of Confirmation.

Criticism of Bayes´ theorem as a logic of confirmation has focused on two questions: whether the evidential import of an experiment is exhausted by the likelihood function, which must be the case if Bayes´ theorem is to be a complete logic of confirmation; and the problems of taking account of conceptual innovation or theory change within a Bayesian framework. I shall discuss these in turn.

i. The Likelihood principle.

If we suppose that conditionalization by application of Bayes' theorem to the results of experiments constitutes a complete and adequate inductive logic, and Bayesians such as de Finetti, Savage, and also Jeffreys do claim this, as we have already noted, it follows immediately that the entire evidential import of our experiments is given by the likelihood function $p(\mathbf{e}|\mathbf{h})$, where $\mathbf{h}$ is the hypothesis and $\mathbf{e}$ is the outcome of the experiment under consideration. For Bayes' theorem states that the posterior probability of $\mathbf{h}$ given $\mathbf{e}$ is given by the likelihood function and the prior probabilities of $\mathbf{h}$ and $\mathbf{e}$, ie.

$$p(h|e) = p(e|h).p(h)/p(e)$$

Thus that the input of the experiment is given entirely by $p(\mathbf{e}|\mathbf{h})$, the likelihood function, is entailed by the assertion that Bayes' theorem is a complete inductive logic. One possible line of criticism against the Bayesian theory of induction, therefore, is that the evidential import of an experiment is not generally given entirely by the likelihood function, for it would follow immediately from this that Bayes' theorem is not a complete inductive logic. Let us call the identification of the evidential content of an experiment with the likelihood function associated with it the 'likelihood principle'. Our immediate task is thus to consider what arguments have been offered for the likelihood principle, and to offer our own arguments on the matter. I shall argue that the principle is by no means obviously acceptable, and that in certain cases it is dubious indeed.

Let us first get clear just what the likelihood principle states. This, and the fundamental importance of the likelihood principle to the Bayesian account of induction, was made plain by Savage, who wrote:

> One of the most obvious, ubiquitous, and valuable consequences of the Bayesian position I know is what I call the likelihood principle.

and

> The likelihood principle says this: ... given the likelihood function in which an experiment has resulted, **everything** else about the experiment - what its plan was, what different data might have resulted from it, the conditional distributions of statistics under given parameter values, and so on - is irrelevant.(51)

Among Bayesians the likelihood principle is generally accepted as a consequence of Bayes´ theorem being a complete inductive logic, rather than given independent backing. I can find no defence of the principle in Jeffreys, nor in de Finetti, nor in Jeffrey, while Savage defends some of its consequences for the way we must evaluate statistical data. Indeed, the only purported proof of the principle I have found mentioned in the literature is Birnbaum´s attempt to derive it from his sufficiency and conditionality principles, but that has been found to be deficient because it assumes a version of the sufficiency principle no less in need of support than the likelihood principle which it is supposed to support, as Hacking pointed out.(52) And in a later paper Birnbaum goes on to argue that the likelihood principle is incompatible with other and more secure concepts of statistical evidence, so Birnbaum provides no stout defence of the principle.(53) Edwards wrote a book recommending it, but this defence was conducted largely by example rather than analysis and thus is susceptible to critique by counter-example such as Hacking provided in his review.(54) Finally Fisher from time to time commented favourably on the likelihood principle, and Barnard showed that measuring support by likelihood has some appealing features, but none of this work constitutes a proof of the principle.(55)

This dearth of attempts to prove the likelihood principle cannot be accounted for by its not being controversial, for it entails the rejection of familiar principles of statistical practice which set out rules to follow in collecting evidence for statistical inference, eg. that the experimental set-up should include randomization to prevent unconsidered bias affecting the data, or that experimenters must decide before collecting the data how much data is to be collected, and certainly cannot stop collecting data just when they think they have enough, lest this assessment be consciously or unconsciously based upon how the data are looking at that point in the planned sample. But these rules cannot be supported on the likelihood principle, for such facts about the data do not influence the likelihood function. However, there are arguments against such rules which do not proceed from the assumption of the likelihood principle, such as that there surely cannot be any difference between the evidential import of two identical samples if one was got from an experiment planned to yield that quantity of data while the other was intended to give a larger sample but could not due to a cut in funding; or that if a sample was derived from an experimental arrangement obviously liable to give misleading results, say because all the plants treated with one fertilizer were on the northern slope while all those treated with another were on the southern, it will not change our reasoned assessment of the evidential value of the sample to learn that the experimental arrangement was randomly chosen.

Such arguments may have been taken to be adequate to prevent the likelihood principle being undermined by these controversial consequences. But I think that some of the points of this kind must have had an impact, since Savage, for example, implicitly accepts that the likelihood function cannot express the full evidentiary significance of a body of data in agreeing that scientists should publish as much detail as possible about their experiments. In taking this view Savage does not admit that conditionalization does not constitute a complete and adequate inductive logic, but in urging scientists to publish their data, as well as the inferences they have drawn therefrom, he opens the door to the other scientists employing whatever rules of confirmation they find attractive. This proposal, which Savage surely did not intend to endorse, obviously endangers the whole project of laying down criteria for reasonable inductive inference, and threatens to legitimate the subjectivism inherent in trusting the scientist's good sense to draw appropriate conclusions from the data, a proposal which some authors in the field have advocated.(56)

We conclude that while the likelihood principle is not proved, it is plausible, and thus attacks on the principle are not likely to refute the suggestion that conditionalization is a complete logic of induction. However our discussion should make us wary of holding to conditionalization as a complete logic of induction if we are tempted to bolster that position with **ad hoc** supplementary rules of confirmation whose promulgation is fostered by lingering doubts about the acceptability of the likelihood principle.

ii. Conditionalization and theoretical innovation.

There is one possible interpretation of the likelihood principle under which it is surely false, namely the notion that the likelihood function published by one scientist contains all of the information which another might find in the evidence which the function is intended to represent. The reason for this is that the likelihood function as published by one scientist will not cover **all possible hypotheses** - for, as Hacking notes, that would be impossible - but will only cover the set of hypotheses thought relevant by the scientist publishing the likelihood function.(57) It is always possible, therefore, that another scientist might find in the complete evidence support for an hypothesis not covered by the published likelihood function, and thus in this sense the likelihood function is plainly not the evidential equal of the evidence itself. The likelihood function at best, therefore, can be taken as the equal of the evidence in the context of the examination of a given set of hypotheses, conditionalization upon the evidence leading to a new probability distribution over these hypotheses. But how does conditionalization, that is to say the Bayesian theory of induction, deal with the problem of assessing support for hypotheses which are not members of this privileged set, say hypotheses which only occur to the scientists once the evidence starts to come in and it is seen that some initial assumption is dubious given the data now to hand?

Hacking follows a line on this problem which was raised by Bartlett in the discussion at the conference on Bayesian statistics recorded in Savage (ed) [1962]. Bartlett pointed out that an unexpected hypothesis cannot ever get a non-zero probability via conditionalization if its initial probability was zero, while zero initial probabilities will surely be common since, as Barnard had been insisting, it is not possible to enumerate all of the hypotheses which might possibly explain some set of data. To save conditionalization in such circumstances Savage suggested that one's prior distribution should not allocate the total probability to the set of hypotheses actually considered, but should leave a little bit of prior probability for 'something else', ie. the unexpected hypothesis.(58) But this **ad hoc** strategy will not work, for at least three reasons.

First, as Hacking insisted, it is just not plausible that if you ever come to attribute a non-zero probability to some hypothesis then you always attributed a non-zero probability to it; indeed, this must be false, since it is possible that we might come to believe an hypothesis of which we had formerly thought it not logically possible that it should be true, for example when we initially take it to be logically necessary that every state should have a cause, and then come to accept a physical theory which allows that some events are not caused. For such changes of belief, which are clearly in need of inductive support, conditionalization does not provide a reasonable analysis.

Second, the set of hypotheses which have to be incorporated under the umbrella of ´something else´ will be infinite in number, and since the prior of each cannot be infinitesimal if significant confirmation of any of these hypotheses is not to take an infinity of favourable observations, the prior distribution will be improper, which is at least undesirable.

Thirdly, and this is the main point Barnard was making in the discussion of Savage´s paper, you cannot compute the likelihood function on the hypothesis ´something else´, but only on some definite hypothesis. It will not do, therefore, to add ´something else´ to the list of hypotheses over which the likelihood function is defined, since it will go indeterminate for the argument ´something else´ (from which Barnard inferred that probabilities got from Bayes´ theorem may not be directly comparable one with another, normalization being destroyed by discontinuity of the likelihood function, and that therefore the major advantage of assessing evidential support by posterior probabilities rather than likelihoods was chimerical).

One possible response to this set of difficulties for conditionalization is to suppose that Bayesian learning can proceed **without hypotheses**, by adopting a prior distribution over the set of possible outcomes and allowing conditionalization to determine a posterior distribution as the evidence comes in. For example, if the task concerns learning the grammar of a simple language, we would begin not by considering various hypotheses about the grammar, lest some new hypotheses be suggested by the data, but rather by enumerating the possible sentence structures, both grammatical and ungrammatical, then we could group these together in all possible subsets, and use conditionalization to infer, from data concerning the admissibility of various sentences, a posterior distribution over the various sets of sentences. Since each set of sentences could be identified with a possible grammar for the language, a posterior distribution for the sets would be a posterior distribution for the various possible grammars.

As Suppes shows, however, humans simply do not have the computing power to tackle the enormous combinatorial problems presented by such a simple learning strategy.(59) Rather we must reduce the number of alternatives we have to consider by introducing some structure into the field of possible grammars, which we do by employing various concepts to group the possible grammars together. Once this is allowed, however, we run the risk of adopting a conceptual framework which is not conducive to effective learning, and which may even mislead us, by failing to separate, for example, two possible patterns of words which, when we review the data, ought evidently to be distinguished. The concepts which structure the field of possibilities, that is, incorporate hypotheses about relevant differences, as they must if any structure is to be given to the field.

The consequence of this is that while different sets of concepts will generally lead us, via conditionalization, to distinct posterior distributions, there is no way that conditionalization can get us from one set of concepts to another.

Putting together the argument of Suppes with those of Bartlett and Barnard, we conclude that conditionalization cannot generally provide an adequate account of induction since, if induction in any ordinary case is to be humanly possible, the set of possible outcomes of an observed state must be restricted or structured in some way. This is done by adopting some particular set of conceptualizations of the possible outcomes - or set of hypotheses describing what might be the possible outcomes - but having adopted some such structure there is no way that we can, without abandoning temporal credal conditionalization, allow observations to modify the structure.

We conclude, therefore, that conditionalization, because it entails the likelihood principle, defines a methodology for induction which is far from problem free. We shall come back to this problem in relation to Levi´s theory of induction to see if he has been able to solve it. We shall not need to consider it in relation to the neo-Bayesians since they offer no new methodological proposals apart from the new rules of inference offered in place of Bayes´ theorem.

c. Acceptance vs. Partial **Belief as a Basis for Induction.**

The Bayesian theory of induction does not provide for a proposition to be accepted on the basis of an inductive inference, rather the outcome of the inference is merely a change to the proposition's probability. Now there are a number of objections to the notion that science does not require inductive inference to justify acceptance, but I do not intend to review that debate here.(60) I shall take up just one problem, the problem of the acceptance of evidence, which seems to me to be fundamental and has not received sufficient attention in the literature.

In the previous Chapter, in the course of discussing Carnap's Bayesian methodology of induction, we noted a problem concerning the premises of inductive inferences, namely that for conditionalization to apply it must be the case that the probability of the evidence is one. Now on any conception of science which recognises the complex structure of experiment it is implausible that the evidence upon which inductive inferences are based can generally (if indeed ever) be assigned a probability of one (if we accept, as is traditional, that a proposition's having maximal probability entails its incorrigibility). Some solution to this problem must therefore be found if the Bayesian theory of induction is not to be jettisoned along with naive empiricism, and not surprisingly, therefore, Bayesians have given the matter considerable attention, Carnap placing his trust in Jeffrey's generalization of conditionalization which allows for conditioning upon uncertain evidence. Having noted this in the previous Chapter, I went on to claim that Jeffrey's generalized conditionalization rule avoided the problem but at the cost of creating a new one, since in typical cases the probability of the evidence would be so low that the probability being derived from (generalized) conditionalization would also be low, and the hypothesis seeking support would not be well confirmed by the evidence. Later in the Chapter I argued that Kyburg's theory of induction faces a similar problem due to its allowing accepted claims to remain less than maximally probable. I must now make good my claim concerning Jeffrey conditionalization.

Jeffrey's rule for conditioning on uncertain evidence is

$$\text{JC.} \quad PROB(A) = prob(A|B)\ PROB(B) + prob(A/-B)\ PROB(-B) \qquad (61)$$

where 'prob' is the agent's initial probability function and 'PROB' his final probability function.

Suppose that A is the (test) hypothesis that a certain wire is surrounded by a magnetic field, **E** is the event that a compass needle brought near the wire moves, BK_1 is the background knowledge, or better, experimental hypothesis, that the compass needle moves if and only if it is placed in a magnetic field (ie is moved across lines of magnetic flux) and $\mathbf{BK_2}$ is the experimental hypothesis that there is no background field near the wire. Can we use JC to determine the probability of **A**? That depends upon how we interpret it, but on what is plainly the interpretation it ought to be given, the answer is no.

One interpretation of JC would allow us to write

$$(1) \quad \mathrm{PROB}_{BK}(A) = \mathrm{prob}_{BK}(A|E).\mathrm{PROB}_{BK}(E) + \mathrm{prob}_{BK}(A|-E).\mathrm{P}_{BK}(-E)$$

where BK denotes the conjunction of $\mathbf{BK_1}$ and $\mathbf{BK_2}$. This would give a value for PROB(A), but it illegitimately assumes that **BK is certain,** whereas the elements of BK, themselves being hypotheses confirmed by earlier experiments, are, on the Bayesian account of scientific induction, only probable. Thus what is wanted instead of (1) is

$$(2) \quad \mathrm{PROB}_{BK}(A) = \mathrm{prob}(A|E.BK).\mathrm{PROB}(E.BK) + \mathrm{prob}(A|-(E.BK)).\mathrm{PROB}(-(E.BK))$$

But (2) cannot be computed, for prob(A|-(E.BK)) is not well defined, being the probability that there is a magnetic field around the wire given that either the compass needle did not move, or that it might move or fail to move whether a field is present or not, or that there is a background field around the wire.

Since, however, probabilities cannot be negative, from (2) we can infer that $\mathrm{PROB}_{BK}(A)$ lies in the interval bounded above by 1, and below by

$$(3) \quad \mathrm{PROB}_{BK,L}(A) = \mathrm{prob}(A|E.BK).\mathrm{PROB}(E.BK)$$

and we might hope that JC in this form will provide a means by which **A** can come to have a high posterior probability on the basis of uncertain evidence. In typical cases of evidence gained from scientific experiments, however, that will not be the case, as the following shows.

Suppose, as is typical in non-statistical contexts, that the conjunction of E , the evidence, and BK, the set of hypotheses stating the conditions under which the experiment can be considered reliable, entails A, the hypothesis. Then we have

(4) $\quad PROB_{BK,L}(A) = PROB(E.BK)$

Should the typical case not obtain, (4) will **overestimate** $PROB_{BK,L}(A)$. Thus assuming (4) will lead to no loss of generality for our argument.

Now in the case considered above BK is the conjunction of two hypotheses concerning experimental controls, which are independent of each other and E, and thus

(5) $\quad PROB_{BK,L}(A) = PROB(E)\ PROB(BK_1)\ PROB(BK_2)$

From the form of (5) it is apparent that if an experiment relies on a great number of independent experimental controls each of which has a high but not maximal probability of obtaining at the time of the experiment, then $PROB_{BK,L}(A)$ will be too low to account for us having great confidence in the hypotheses confirmed by such complex experiments. The same would be true of an experiment which required a relatively small number of controls, each of which, however, could only be tested by a further controlled experiment. And the same point would apply to even a simple experiment which took place against a background of knowledge assumed to be unproblematic, for if we accept the Bayesian analysis of induction then we must accept that many, if not most, of the elements of this background field are themselves only probable; and thus the probability that the field as a whole does not contain a falsehood which would prejudice the experiment in question, if this probability can be defined, is surely very low indeed.

Our conclusion here need be no more, however, than that the Bayesian account of induction, whether based on standard or Jeffrey conditionalization, stands or falls with an account of science which assumes that observations are relatively theory free. Our next task must be to consider whether some other of the recent sophistications of the Bayesian programme avoids that criticism. In this regard we shall need to consider only Levi´s theory of induction, most recently set out in his [1980], since the other recent development within the Bayesian paradigm, the neo-Bayesian theories inspired by Jaynes´ theory of maximum entropy as a basis for determining prior probability distributions, treats constraints on probability distributions simply as the given. In this respect these recent theories constitute, therefore, no amendment to the traditional theory, save by allowing constraints to include probability distributions for relevant parameters, as Jeffrey does, and we have already seen that this will not solve the problem.

d. Science and Subjectivism.

One of the controversial features of the Bayesian programme has been the adoption by some of its members, those I have identified as the subjectivists, of a subjective theory of probability. Before leaving the discussion of the traditional Bayesian accounts of induction we should briefly consider whether scientific induction could be reasonably identified with subjective judgement.

It must first be admitted that so called objective statistics are not free of subjective elements in the sense that in applying objective methods one is forced to make decisions or accept principles of inference which are not objectively justified, as we have been at pains to argue. Indeed, particularly in relation to Neyman-Pearson statistics, Savage was fond of arguing that his subjective theory of statistical inference was in fact more objective since it brought the subjective decisions out into the open for public consideration.(62) One can grant Savage's point, however, and yet conclude that subjective statistics, while not inferior to so called objective statistics, is still inadequate, since what is wanted is a thoroughly objective statistics, or theory of induction. The point requiring analysis, then, is whether we do require, for our science to be rationally based, an objective basis for inductive inference.

The main plank in the subjectivists' rejection of the supposed need for objective probabilities is the claim that typically differences of opinion will be overwhelmed by the evidence, and thus observers with different subjective prior probabilities will commonly be brought by the evidence to agree.(63) Even when carefully stated, however, the claim is false. For in the typical case two observers will surely differ in their **background knowledge** if they differ appreciably in the degree of belief they allocate to the occurrence of some event, and in that case they will not necessarily draw the same conclusion from whatever evidence they come to share. The subjectivists´ point, that is, assumes that the evidence is given, but it is not; rather it is constructed on the foundation of a field of background knowledge, as I have briefly sketched above. Thus the evidence gathered by two observers of the same experiment will only propel them towards agreement if they share some background knowledge which requires them to draw certain similar inferences from the observations. Again we see that Bayesianism rests upon a naive empiricist account of observation, for although the subjectivists´ argument can be applied to elements of the background knowledge as well, the day of convergence receds unhopefully into the future once we realize what a complex set of beliefs may be held by one observer and not shared with another, and how evidence can only get to bear on any one of these beliefs through the mediation of others.

It is not enough, therefore, to found rational scientific practice, for one scientist to be able to say to another, ´If you disagree with me let´s get some more evidence´, as Savage suggested, for the two may agree no more closely when the new evidence is in, drawing divergent conclusions from it, and this pattern can be repeated again and again. In order that the agreement between scientists which is both required and observed might be obtained, it is necessary for one scientist to be able to prove to another that there are good grounds for giving up his former beliefs in favour of some alternative account of the matter under investigation. That will require two fundamental shifts away from the subjective Bayesian account of scientific inference, for it entails that some changes of belief do not abide by conditionalization, but rather proceed by extensive revision of one´s present set of probabilities; and it also entails that there is the possibility of an objective basis for changing one´s probabilities, since to prove to the satisfaction of an opponent who does not share some relevant beliefs that they have been making an error which they ought now repudiate, we require an objective foundation for argument. Insisting upon such changes leads us to Levi´s theory, for among his other innovations are fundamental departures from the traditional Bayesian account of induction in relation to the points we have just raised.

e. Concluding Comments on the Standard Bayesian Theories.

Before leaving the standard Bayesian accounts of induction, however, let us sum up the criticisms we have made. First, neither the subjective nor the objective theory provides a sound case for requiring that a rational agent´s degrees of belief should conform to the probability calculus. Given that there are prices to be paid for accepting the axioms of the calculus as a constraint upon belief, particularly in regard to the possibility of having the results of complex experiments given a sufficiently high probability to allow the experimental evidence to confer significant degrees of confirmation upon the hypotheses tested, this failure is a significant defect. In addition to this the objective theory has yet to solve the problem of finding an adequate basis on which to specify prior probabilities which are legislative for rational belief, while the subjective theory has not given a satisfactory defence of its requirement that induction proceed by the rule of conditionalization, both Teller's and Shafer's proofs not applying to typical scientific situations. Moreover there are arguments applicable to the objective and subjective theories alike which cast doubt upon the likelihood principle upon which the whole Bayesian theory of induction rests. Finally, the subjective theory is unable to provide a basis for rational scientific practice, this requiring, so I have argued, an objective basis for settling disputes.

Our procedure now will be to consider the fresh approaches to a Bayesian account of induction given by Levi and the neo-Bayesians, and determine, should they be free of other problems, whether they make any impression on the list just set out, where those problems are relevant to the new theories. We shall find, however, that neither of the new theories is free of problems of their own, and thus conclude that neither is capable of supplying a rational foundation for a version of induction in the Bayesian tradition.

5. LEVI´S THEORY OF INDUCTION.

I shall examine Levi´s theory of inductive inference as developed in his [1980]. This work extended the approach to learning put forward in his [1967] - with changes of varying degrees of importance, none of which will be discussed here - building in particular on the attempt to construe scientific inference as a process of decision-making in the Bayesian mould, governed by distinctively epistemic utilities. Our task here is to give an outline of the major elements of Levi´s decision-theoretic account of scientific inference, and then review the critical literature to which Levi´s theory has given rise, selecting what I take to be the most important of the criticisms levelled at the theory for discussion and elaboration.

a. An Outline of Levi´s Epistemology of Scientific Inference.

Standard Bayesian theory provides an account of how an ideally rational agent ought to make decisions concerning various courses of action open to him; he should consider the utility of each act under each of the various states of nature that might obtain, and pick the act which has the highest expected utility, where the expected utility of an act is calculated by multiplying the utility of the act in each state of nature by the probability of that state and summing over all states. Now scientific inference in this scheme of things can be no more than a process by which one comes to fix one´s probabilities for the various states of nature, or one´s credal probability function, as we shall call it. What would happen, however, if we classed among the possible acts open to an agent that he should accept a certain hypothesis? This would seem to open the way for a Bayesian account of how we come to accept, and not merely allocate degrees of belief to, hypotheses concerning the states of the world.

Now to a strict Bayesian, this is a confusion; what need have we of acceptance when we already have degree of belief, and degree of belief is all we require for practical deliberation - indeed, what does ´acceptance´ mean in this context?(64) We have seen, however, that there is a good reason to want to provide for the acceptance of hypotheses, for only then, so we have argued, will we be able to account for the reliability accorded the results of complex experiments. Taking up this option will require, however, a clearer sense to be given to the notion of acceptance, as well as the development of a theory of epistemic utilities - the utilities relevant to the decision to accept or reject an hypothesis - and the specification of a rule of acceptance, if the rule of maximizing expected utility is not appropriate. Levi´s responses to all of these challenges, plus his theory of credal probability, are sketched below.

i. Acceptance into **K**, infallibility and corrigibility.

According to Levi to accept some claim **h** is to adopt **h** as part of one´s standard for serious possibility, or, more precisely, to incorporate it into one´s (deductively closed) corpus K which functions as one´s standard for assessing serious possibility. The notion of serious possibility Levi takes as primitive, giving nothing that 'amounts to an explication of serious possibility' nor 'a definition in other terms'. He insists only that serious possibility is relative to an agent at a particular time, or relative to a certain corpus; that a proposition cannot have a positive credal probability unless it is a serious possibility - although a proposition can be a serious possibility and yet have zero credal probability; and finally, and most definitively, that 'h is a serious possibility according to **X** at t if and only if **h** is consistent with his corpus of knowledge at t.'(65) More revealing than these elements of a definition, however, is the example of serious possibility given just before the more formal presentation: suppose we are about to toss a coin and wish to consider the various possible outcomes of the toss; we shall surely not consider the possibility that the coin will fly to Alpha Centauri, nor that the Earth will explode on impact with the coin, for while both of these outcomes are logically possible they are not seriously possible, if one 'takes for granted even the crudest folklore of modern physics' - presumably because they are not consistent with accepted laws of physics. Rather, the seriously possible outcomes are heads, tails, and perhaps that the coin will come to rest on its edge, for it is only such outcomes which are consistent with the hypotheses forming the corpus. Thus what is involved in accepting a statement is explained by noting that those statements we accept are used to rule out of contention logically possible states of affairs whose occurrence would entail the falsity of an accepted claim.

Note, for later reference, that a corpus which does not already contain universal laws, such as the laws of physics alluded to in the coin example, will function as only a very ineffectual standard of serious possibility, for many implausible propositions will be consistent with such a restricted corpus. Hence, should a strong corpus prove to have a role in legitimating induction, Levi's theory will face a serious problem in showing how this precondition of reasonable induction can be met. This will prove to be of particular concern in the light of Levi's commitment to statements of objective chance as the necessary basis for direct inference.

Further light is thrown on acceptance by noting that since X's corpus is his standard for serious possibility, if **h** is a member of **K** then **-h** is inconsistent with K and thus if **h** is a member of **K** its falsity is not a serious possibility; it implies, therefore, if X accepts a claim into his corpus, that he takes the accepted claim to be true, and infallibly so, since its negation is not a serious possibility. Thus Levi embraces an unfashionable infallibilism, on the strength of the argument that to accept fallibilism entails rejecting his account of how 'knowledge functions as a resource for enquiry and deliberation', viz, as a standard for serious possibility.(66)

Additional support for his infallibilism is sought in a criticism of Peirce's 'categorical fallibilism' - the thesis that all logical possibilities should be considered serious possibilities at all times - Levi pointing out that this doctrine is consistent with his infallibilism if one's corpus is restricted to logical truths; but that the two jointly entail, therefore, the **incorrigibility** of knowledge.(67) Now the history of science cannot be understood if we hold to incorrigibility, at least for contingent claims, and Levi implies that we can only avoid incorrigibilism by adopting infallibilism, thus allowing contingent claims into the corpus and consquently holding them to be infallibly true even though conceding that they might one day be removed. But the reader here is hard pressed to avoid great confusion, since one is entitled to infer, it would seem, from holding that h is infallibly true, that h will never turn out to be false, ie. never be rejected from the corpus, and thus that the corpus is incorrigible after all. However, while we may reasonably suspect that Levi´s theory will have trouble accounting for the rejection of h once it is declared infallibly true, one should decline the easy knock-down argument suggested; it simply puts too much weight on the concepts of fallibility and corrigibility, while the incongruities Levi is led to defend show that the informal concepts are not well suited to the role they are called upon to play. Serious criticism of Levi´s epistemology, therefore, must and will be delayed until we can deal with his formal proposals.

ii. Epistemic utility.

Levi asserts uncontroversially that the aim of enquiry is error free information, from which it plausibly follows that the epistemic value of a possible addition to one´s corpus ought to a be a function of its content and probability. Levi supplies such a function after formalizing the idea of information.

The set of possible answers to some question, relative to **K**, the agent X's corpus at time **t**, is represented by the ultimate partition U, a set of hypotheses h_i such that K entails that at least and at most one of the h_i is true and each of the h_i is consistent with K. A possible answer is represented by the rejection of some or all of the h_i, X suspending judgement between all the unrejected h_i, or, equivalently, accepting the disjunction of the unrejected h_i. Now if one potential answer involves the rejection of a proper subset of the h_i rejected by another answer, then the former is less informative than the latter.(68) Suppose that we then assign each of the h_i an information measure such that the sum of the information in all of the h_i is 1, and stipulate that the information given by a potential answer g is the sum of the information of the h_i rejected by that answer. The assignment of informational values to the elements of **U** is then given by a probability function, **M(g)**, defined on the possible answers, where if $M(h_i)$ is the informational value of h_i, **M(g)** is the informational value of **rejecting g**, ie. of accepting the disjunction of the h_i which are not disjuncts of g (together with their deductive consequences).(69)

Using this definition of informational value, and other considerations given in his [1969], Levi suggests representing the epistemic utility of rejecting all the h_i which are elements of **g**, ie. of accepting all h_i which are not elements of **g**, by **(1 - a)M(g)** when g is true and it is rejected erroneously, and by **a + (1 - a)M(g)** when **g** is false and rejecting it avoids error.(70) Thus the expected epistemic utility of rejecting g would be given by

$$EU(-g) = Q(-g)[a + (1 - a)M(g)] + Q(g)(1 - a)M(g)$$

where **Q(g)** is the agent´s credal probability for g, and **a** takes values from .5 to 1, larger values representing increased desire to avoid error. We shall return to the role of this shortly.

iii. Credal probability and inductive logic.

Let B be an agent´s set of permissible degrees of belief such that each member of B is a degree of belief in some hypothesis in the agent´s corpus K relative to some evidence in **K**. If **Q(h;e)** is a member of **B** Levi stipulates that Q(**h;e**) is a normalized probability measure, thus adopting, without argument, the principle of **credal coherence** which has the effect of constraining degrees of belief to abide by the probability calculus.(71) That this principle is not defended is of course a problem for Levi´s theory, since, as we have seen, its proof is far from a trivial matter. However, we shall not labour that point any further here.

Levi adopts credal coherence as a principle of inductive logic, that is to say, as a principle which specifies membership conditions for **B**, the set of permissible degrees of belief or, to use his term, credal functions **Q(h;e)** for **h** and **e** in **K** . He adds to this first principle two of a different kind, which do not specify membership conditions for **B** but concern the number of the members of **B**, and set down relationships between them, namely **credal consistency**, which specifies that B is non-empty if and only if **K** is consistent, and **credal convexity**, which specifies that the convex combination of any two members of **B** is also a member of B.(72) Credal convexity is a distinctive feature of Levi's system, for he offers it in place of the stronger standard requirement that B contain a single Q-function, ie. that the agent have a unique (though perhaps interval valued) degree of belief for given pairs **(h;e)** in **K**. Levi's reasons for opting for credal convexity can best be brought out by pursuing the next part of his system.

Given a corpus **K** at some specified time an agent will have a set of credal commitments **B** which can be thought to be determined by a function defined on K and taking values in **B**, ie. the agent has a function **C** which represents his **confirmational commitment** at a given time. That C is defined on K, ie. that **C(K) = B**, constitutes a **total knowledge requirement**, since each Q-function in B is required to be informed by K, the whole of the agent's knowledge.

Now define B′ as the **conditionalization** of B with respect to K and K′, where **K′** is K augmented by **e**, if and only if for every Q in B there is a Q′ in B′, and vice versa, such that if **f** is consistent with K then Q′**(h;f) = Q(h;f&e)**. Levi can now define the principle of **confirmational conditionalization** as specifying that if K′ is a consistent expansion of **K**, **C(K′)** is the conditionalization of C**(K)** with respect to K and K′.(73)

While we hold a single C-function, that is a single confirmational commitment, holding to confirmational conditionalization will require us to abide by temporal credal conditionalization, ie. to changing our Q-functions in accordance with Bayes´ theorem, but adopting the principle of confirmational conditionalization does not require us to hold to a single C-function and Levi does not recommend this approach, which he dubs ´confirmational tenacity´. But what reasons could one have for changing one´s confirmational commitment? Levi´s answer to this question brings us to the heart of his conception of inductive logic.

According to strict Bayesians such as de Finetti and Savage, inductive logic is just the requirement of credal coherence. Levi calls those who share this view ´coherentists´. Levi himself believes that the requirement of coherence must be supplemented by a principle of direct inference, which specifies how one is to arrive at credal probabilities given knowledge of **chances**, calling those who share his view ´objectivists´.(74) Now the principles of inductive logic, so far as the coherentists recognise them, leave the agent free to choose his Q-functions subject only to the requirement of coherence, and for particular **h** and **e** this does not rule out any particular Q-function provided appropriate choices are made elsewhere; and while adopting an objectivist inductive logic narrows the range of choice of Q-functions open to the agent it still leaves a wide field of choice. Levi opines that some authors, and he aims his point particularly at those in the strict Bayesian tradition, would have an agent arbitrarily choose a

Q-function from among those permissible for him, calling such theorists ´intemperate personalists´. Levi, however, objects to choice when there is no rational or objective basis for choice, and as there is no such basis for choice between coherent sets of Q-functions on the coherentist´s model of inductive logic, he rejects the intemperate personalist´s position and requires that if the agent has no rational basis for choice between Q-functions then the agent should suspend judgement between the Q-functions permissible for him.(75) That is the basis for his adoption of credal convexity in place of credal uniqueness.

If we adopt (a) the view that the principles of inductive logic are the sole rational basis for declaring some Q-functions to be impermissible for an agent (with a certain corpus and at a certain time), and (b) the principle that one should not rule out a Q-function without a rational basis for so doing, we seem to be forced to accept a great degree of latitude in specifying permissible Q-functions. One group who hoped that this would not be so were the ´necessitarians´, as Levi calls them, led by Jeffreys and Carnap, who hoped to find principles of inductive logic sufficiently strong to make only one Q-function permissible for the agent given **h** and e. But Levi argues (in later chapters) that this project has failed, and therefore, since he rejects the intemperate personalists´ suggestion for narrowing the range of permissible Q-functions, which involves rejecting (b), and yet wishes to narrow the range to ensure that Q-functions place significant constraints upon the set of degree of belief permissible for an agent, Levi proposes that we should reject (a). This involves letting non-formal, or non-logical, or contextual factors, influence permissible confirmational commitments. (76)

Now, that contextual factors can be allowed to influence choice of confirmational commitment is controversial, and Levi´s theory of what factors can be allowed to influence the choice and how their influence is to be exercised has been found by his reviewers to be unclear. We shall therefore list the theory of contextualism, as we shall call it, for later discussion in our critique of Levi´s epistemology and proceed with our description of its main features.

iv. Expansion of the corpus: observation and induction.

An agent in expanding his corpus risks importing error, but will do so in order to gain more information. The balancing of the desire for information against the desire to avoid error is the essence of Levi´s account of inductive inference, as is already plain in his definition of epistemic utility. Since we can expand our corpus in two ways, by routine and inferential expansion, we shall have to examine how that balance is struck in the two cases. We shall deal first with Levi´s account of routine expansion.

The idea of routine expansion is that the agent settles on a programme for gaining information, and then executes this programme regardless of the results obtained. The important point, and the point which distinguishes routine from inferential expansion, is that each addition to the corpus is not subject to any test, rather the programme as a whole was subjected to evaluation before being adopted - presumably by inferential expansion of the methodological commitments of one´s corpus, though Levi does not make this clear - and thus the warrant of each routine addition is the decision to follow the programme rather than any feature of the actual element added.(77) Levi makes this plain by comparing routine expansion with the statistical tests promoted by the Neyman-Pearson school, which are distinguished by the requirement that **prior to conducting the test** it should be decided what outcomes of the experiment involved will lead to the rejection of the null hypothesis, and this policy is then carried through no matter what experimental result is actually achieved, a feature of the tests which drew criticism from Hacking since the actual result could fall in the rejection region and yet strongly indicate that the alternative hypothesis is false or even be inconsistent with it. Routine expansion shares this feature, for it can lead to the adoption of a claim which is inconsistent with the agent´s corpus at that time. Now according to the agent such a claim is certainly false, yet it is accepted because the expansion strategy was adopted, and this in itself constitutes a warrant for the acceptance of the outcomes of the routine process.(78) It follows from this that Levi´s account of routine expansion must be supplemented with an account of how we can contract the corpus to regain consistency, but we shall leave that question for the moment.

Levi´s account of inferential expansion is based upon the idea that the agent should decide to add some hypothesis to his corpus only if certain conditions are met, these conditions being designed to ensure that expansion leads neither to the passing up of an option that would, for some permissible Q-function and utility function, maximize epistemic utility, nor requires the agent to chose arbitrarily between admissible options. The main elements of Levi´s proposal are these.

Consider a partition U of hypotheses h_i such that the disjunction of all the h_i is logically true and the conjunction of any two is logically false; suppose that the agent has determined that at least one of the h_i is true, and the problem is to decide which one to accept into K as the true hypothesis. The agent has various options open to him, namely to accept any one of the h_i or to accept the disjunction of the members of some subset of the h_i. The agent therefore requires some principle for selecting between these alternative expansion strategies or cognitive options.

Levi begins with the Bayesian decision rule that one should chose that option which maximizes expected utility. But if, as is generally the case according to Levi´s theory, the agent does not have a unique Q-function, nor a unique utility, or U-, function, then the rule of maximizing expected utility will not usually suffice to determine the optimal strategy for the agent to pursue, since different options for expanding one´s corpus will yield maximum expected utility on different Q-functions. Call any any option which has maximum expected epistemic utility, relative to given Q & U-functions, ´E-admissible´. Levi´s first step away from the Bayesian decision theory is to replace the rule of chosing to maximize expected epistemic utility with chosing between E-admissible options.(79)

To provide a basis for choice between E-admissible options Levi turns to comparing the relative strength of the various options for expansion open to the agent. Define an option as accepting the disjunction of all of the h_i unrejected by the decision rule adopted; one option is said to be weaker than another if the h_i rejected by the former are a proper subset of the h_i rejected by the latter. The stronger option rejects all of the h_i rejected by the weaker, and at least one more besides; or, equivalently, the disjunction accepted on the stronger strategy entails that accepted on the weaker strategy. Levi now proposes, following the rule for ties of his earlier work, that if there is a unique weakest E-admissible option then this should be adopted as the appropriate expansion of the corpus, departing from the principle of maximizing expected utility, since this generalized rule for ties entails that the Bayesian principle is not sufficient to yield a decision. (Note that adopting the weakest E-admissible strategy leaves the way open for further expansions of the corpus with the elimination of h_i belonging to the weakest disjunction, while adopting a stronger disjunction would rule out some h_i without reason.)(80)

Having argued that an option's maximizing expected epistemic utility is not a sufficient basis on which to select it, Levi now considers whether it is necessary that an option meet this criterion for it to be selected as the optimal expansion strategy. He concludes that it is not, since as he shows with an example, the agent may have Q-functions such that the strongest option which expresses suspension of belief between the options which yield maximum expected utility on his various Q-functions, ie.

suspension of belief between all E-admissible options, may itself not be E-admissible. This option ought to be admissible, but none weaker can reasonably be. (Note that adopting a disjunction of the h_i which is weaker than the strongest which expresses suspension of judgement between the E-admissible options would be to accept some h_i which is rejected by all the E-admissible options.)(81)

While the rule of choice based upon Levi´s generalization of maximizing epistemic utility, namely that one should choose an E-admissible option, turns out to be neither necessary nor sufficient to determine the optimal expansion strategy, the Bayesian´s motivation for choice remains the basis of Levi´s proposal. For the strongest disjunction of the h_i which are unrejected by at least one of the E-admissible strategies - which is the option recommended by the new rule for ties - will be E-undominated, ie. there will be no other option which has higher epistemic utility for all Q & U-functions; and E-admissibility is still fundamental since it is violated only to suspend choice between E-admissible options.(82)

This development of Levi´s theory leads to a simple rule of rejection for the h_i:

> Reject h_j in U if and only if for every Q-function in B, $Q(h_j) < q.M(h_j)$.

Levi does not rest content with this for a moment, however, noting that the agent may have differing values of q, the index of caution (a function of the parameter, a, introduced above), and have no unique

M-function, which requires that the simple rule be complicated to provide for the rejection of all h_j rejected by the simple rule for all permissible Q & M-functions when the greatest lower bound upon the agent's index of caution, q, is employed. As Levi sums up this development,

> The idea behind this procedure is that an element of the ultimate partition U is to be rejected if and only if it is rejected according to all permissible rankings of cognitive options with respect to epistemic utility...(83)

These developments of Levi's theory lead to great latitude in the choice of expansion strategies. Levi tries to specify more clearly the range of options open to the agent by describing the conditions just set out as restricting choice to the set of P-admissible options, where an option is P-admissible if and only if it is (a) no stronger than the weakest E-admissible option, (b) no weaker than any other option meeting condition (a), and (c) E-undominated.

Now while P-admissibility is a more restrictive criterion for acceptance than E-admissibility it will not necessarily leave us with any usefully definite expansion strategy. First, because if we have any significant latitude in our choice of Q or M-functions, while there may be a uniquely P-admissible option this option will be very weak, that is to say it will entail the rejection of only a few of the hypotheses in U under consideration; restricting the permissible Q & M-functions is thus a serious problem for Levi's theory. Second, because if the options are not comparable for strength, which will be the case if no ultimate partition can be given to encompass them all or if certain other deviations from

Levi´s account of induction are permitted, or if the problem at hand is not one of inductive expansion, then all E-admissible options will be P-admissible.(84) In these cases we will require a further basis for selection of options.

We shall not consider this new difficulty in any detail, however, since such cases do not arise in inductive expansion, at least if we grant Levi´s account of inductive expansion as the selection of an option from among several defined by the rejection of some subset of members of an ultimate partition. Thus we shall not detail Levi´s proposal to discriminate between P-admissible options, merely noting that the further principle of selection, called S-admissibility, is selected from a number of models of decision making under uncertainty proposed in the literature, Levi favouring maximin, though he is also prepared to endorse the leximin strategy.(85) Should we reject Levi´s account of induction, therefore, his theory would reduce to using maximin to select between all those options which maximize expected utility for some Q & M-function.

Before we leave Levi´s theory of discrimination between options for inductive expansion, let us pause to reflect upon the status of the theory. For the theory we have just summarized is, as the reader no doubt appreciates, a complex and sophisticated account of expansion, so far removed from simple accounts of induction that we must deliberately step back from the detail of the theory from time to time in order to maintain our bearings on the problem with which we, and Levi, are concerned, namely, finding an account of induction which enables us to found scientific inference in reason. It is important, therefore, not to become bogged down in the detail of Levi´s theory and miss the significance of what he is proposing. For though he does not hide it, it is worth highlighting the fact that in trying to find criteria to constrain the field of options for expansion open to the agent we are in fact giving a model of the factors which we think ought to influence scientific investigators in their choice of an inductive logic. There may be good reason for the model Levi has incorporated in his definitions, for example for his insistence that where several options are E-admissible we ought to suspend judgement between them, but such reasons must be given; there is nothing inevitable about the model Levi proposes. Now Levi himself notes the problem, and after admitting that he ´has no proof´ of the superiority or adequacy of the model of induction he proposes, asserts that ´the best that can be done´ is to explore the further development of his theory to see if its consequences are ´acceptable in an account of good scientific theory´.(86) Given the originality and complexity of Levi´s account, we ought at this stage to list its adequacy as a rational reconstruction of scientific practice as an open question, for which Levi´s comparisions of his theory with certain rivals provides useful material. And we should also note, of course, that since the theory aims to be normative we will be unable to justify it in purely descriptive or, in the absence of some proof that some system of induction is adequate, comparative terms. Levi´s theory, therefore, at least at present, stands in need of justification and thus cannot itself be taken to provide an adequate justification for scientific inference.

Notwithstanding this problem, however, Levi´s theory has the considerable virtues of broad scope, detailed construction and rigour. If we wish, therefore, to criticize it, it is incumbent upon us to find problems with it rather than rest content with pointing out weaknesses in its justification thus far. We shall turn to that task after considering the most controversial aspect of Levi´s theory, his account of the process by which we can come to replace a theory held in our corpus by a rival presently considered certainly false, after finding evidence inconsistent with our old theory.

v. Rejection and replacement of items in the corpus.

On the face of it Levi´s theory will have grave problems attempting to account for the rejection of an element of the agent´s corpus, and for the replacement of one element by some other claim which is a rival to the one now accepted. Levi puts the view of the critic in this matter very clearly:

> For X to contract his corpus is for him to surrender error-free information. Replacement involves not only the abandonment of error-free information, but also the substitution of information that, from X´s point of view at t, is certainly and infallibly false. If X does take all items in his corpus to be infallibly true and seeks error-free information, it appears to be counter-productive for him to contract or to replace certainties with hypotheses he is certain are false. Counter to what I have claimed, infallibilism presupposes incorrigibilism.(87)

The solutions he offers to these related difficulties are as follows.

In respect of contraction, Levi points out that while an agent, in surrendering an element of his corpus, is giving up what is to him is error-free information, he will have sufficient reason so to do in the circumstances which prompt contraction. There are two cases: the discovery by the agent that his corpus is inconsistent; and the desire of the agent to set some accepted hypothesis aside in order to allow the consideration of some hypothesis inconsistent with it. If the agent, via routine expansion or error of computation, imports into his corpus a claim inconsistent with some element of the corpus (recall that the corpus is deductively closed, so all conjunctions of elements are elements), he will have a sufficiently good reason to contract the corpus to remove the inconsistency, for an inconsistent corpus cannot serve as a standard for serious possibility, and to be able to use one´s corpus as such a standard outweighs the desire to hold onto information.(88) In the second case, if the agent becomes aware that there is a rival to some element of his corpus, say h´ inconsistent with **h** in K, then the agent may be prepared to surrender h in order to give **h´** a hearing, if on grounds of informational content **h´** is preferable to h. Contraction, therefore, could and would be motivated either by the desire to retain a consistent corpus or by the desire to increase the informational content of the corpus, and in neither case would the agent´s view of his corpus as incorrigible mandate against contraction - or so Levi argues; we shall consider the adequacy of his argument below.

But suppose the agent decided to contract his corpus for one of the two reasons discussed. How would he arrive at a strategy for so doing?; for there would ordinarily, if not always, be a great number of ways of removing the contradiction or preventing contradiction arising from the adoption of the rival hypothesis **h´**. Levi suggests that the agent should be guided by the desire to avoid unnecessary loss of information, and thus should opt for that contraction strategy which minimizes information loss, grading the members of the corpus with respect to their corrigibility - even though all are incorrigible - according to their informational value. But this has problems; for example, it would seem to follow that we should reject whatever logical principle we hold that bans accepting contradictions, since presumably logical principles have no content. Such criticisms will be taken up below.

We come now to the final aspect of Levi´s theory listed for discussion, his theory of replacement. Replacement occurs when the agent shifts from a corpus K to K´ inconsistent with K, which events Levi identifies with scientific ´revolutions´. Here he has particular trouble to account for the agent´s willingness to replace his present corpus with another, for to do so seems ´utterly counterproductive from his point of view´. But Levi suggests the following rationale: suppose the agent, rather than substituting K´ for K in one step, which would be inexplicable on Levi´s theory, first contracts K to form a corpus consistent with K´, and then applies the rules for evaluating alternative expansion strategies to determine whether to move back to K or to adopt K´, quite likely expanding the neutral contracted K first with additional evidence relevant to the selection of an option for expansion. While replacement could not be justified in one step on Levi´s theory, for the agent would see that the substitution would lead to the adoption of beliefs which are certainly false by his present lights, there is some hope for the two stage process, for the agent should be **myopic** with respect to the further consequences which might follow from contraction; but this too will require further analysis in our critical discussion of Levi´s theory below.

This completes our summary of Levi´s theory, and we turn now to the major criticisms which have been raised against it in the literature.

b. Criticisms of Levi´s Theory.

Levi´s [1980] has been reviewed at length by many of the leading writers on induction and the foundations of statistics, and his reviewers, while respectful of the scope and rigour of Levi´s account of learning from experience, have typically subjected to the book to searching criticisms.(89) I do not intend here to review this critical literature in detail, but rather to pick out, for discussion and development, what seem to me to be the points which go to the heart of Levi´s theory. I shall also concentrate on philosophical rather than technical features of Levi´s theory. For criticisms of Levi´s formal system the reader is referred to the reviews of Harper and Spielman particularly.

I shall take up four issues: Levi´s attempt to square infallibilism and corrigibilism, which is basic to his theory of contraction and replacement; the rejection of pedigree epistemology; the possibility of his theory's explaining how we can come by a corpus sufficiently rich to justify inductive inferences; and Levi's rejection of the traditional ban upon allowing contextual factors to enter into the logic of theory appraisal.

i. Corrigibilism and contraction.

One of Cohen's criticisms of Levi's theory attempts to show that its account of contraction is unsatisfactory, and that to improve it we would have to surrender the idea that the corpus is the agent's basis for determining serious possibility.(90) Cohen notes that Levi bases the relative assessment of the informational content of hypotheses upon how many elements of the relevant ultimate partition they conflict with. Thus, for example, how many facts two rival hypotheses explain is not relevant to measuring the content of the rivals - and he argues further that Levi could not incorporate such a factor into his assessment of the content of rival hypotheses, for while one is in the corpus the other is taken to be certainly false and thus any evidence for it - such as that it explains certain facts - must be discounted. But suppose that T which is in the corpus does explain more facts than T'; then, Cohen claims, T' does not merit a hearing, regardless of its informational content. While if T' explains no fewer facts than T, it surely ought to be considered a serious possibility. Thus, Cohen claims, we must give up the idea that the corpus is a standard for serious possibility in the sense that any claim inconsistent with a member of the corpus is surely false.

I do not think that Cohen´s criticism is successful. (But when we see why it fails we see that a related and more serious criticism can be sustained against Levi´s theory.) For since T and T´ are members of the same ultimate partition, any virtues of the rivals relevant to their truth or probability, such as their relative abilities to explain the relevant facts, would have been incorporated into our assessment of their relative merits at the time of the original expansion (errors of computation aside, as always). Thus even if we had a case where T was a member of the corpus and some rival, T´, which is inconsistent with T, was therefore considered not to be seriously possible, the fact that T´ explained more facts than T would not upset this determination since this factor would have been taken account of in the decision to accept T rather than T´.

While this explanation deals with Cohen´s point, it is not clear, however, that it is available to Levi, since he rejects the notion of pedigree epistemology, and if this entails his accepting the thesis that a corpus has no history, then there will be no memory in the corpus of having rejected T´ to accept T and thus T´ will be able to insist upon a hearing despite being formerly rejected.(91) Then the difficulty Cohen raises will bite, because either T´ will explain more than T and thus deserve to be considered seriously possible, or it will explain less and thus not deserve a hearing, regardless of its informational content as measured by Levi. And worse, if the corpus has no history, T´ can insist upon a hearing no matter how many times it has been considered and rejected in favour of T. Clearly, we must allow the corpus a memory or the agent will be sentenced to endlessly considering the same choice between T and T´, and Levi´s epistemology will have no relevance to real problems.

Now Levi did wish, I think, to deny the corpus a memory, in line with his idea that it does not matter how a claim came to be admitted to the corpus, rather what matters is whether the claim can hang onto its place in the face of challengers. But it does matter how the claim got in if the challengers now faced are the very ones just defeated. Thus it is false that

> In general, neither the origins of X´s having h in his corpus nor the grounds on which he justified adding h to his corpus in the first place will be relevant to deciding whether to remove h unless considerations of origin can be shown to have a bearing on whether elimination of h will improve X´s corpus.(92)

No great change to Levi´s outlook is required to avoid the proble we have raised, however; indeed, it might be thought to be already covered by the rather vague caveat Levi attaches to his rejection of pedigree epistemology in the passage quoted. It will suffice that the corpus contain a record of the circumstances in which T was admitted, and that this record be used to determine not whether T may be relied upon at some future time, for that has already been settled in the affirmative when T was adopted and the decision will stand while T remains a part of the corpus, but rather be used to determine whether a rival T´ is **sufficiently novel** to be accepted as a challenger to T.

But the problem we have been discussing is not the most serious of the difficulties Levi´s epistemology faces in this area. For as the reply to Cohen I have sketched out makes plain, once T is ensconced in K, it is possible that facts could come to be accepted which, had they been known when T was accepted, would have led to T´s acceptance, and yet their acceptance does not provide grounds for rejecting T to give T´ a further hearing, for they bear not on the informational content of the rivals, but rather on their probabilities, and once T is accepted its probability is frozen at 1 and its rivals´ frozen at 0. For example, say T asserts that a certain coin is biased towards heads such that the chance of heads on a single toss is greater than 0.6, while T´ has it that the chance of tails is greater than 0.6, and we originally conduct a trial of 100 tosses, leading to T being accepted, but our research assistant continues testing, unknown to us, and later reports that a further trial of 100 tosses so heavily favoured tails that on the whole set of 200 tosses T´ would be accepted. Having already accepted T there appears little we can do, since the further result is not inconsistent with T, nor does T´ have any informational virtue which could prompt contraction by rejecting T to give T´ a hearing. Though the evidence now goes against it, we are stuck with T.

Short of accepting some fact inconsistent with T, then, T will keep its place in the corpus no matter how far its probability would sink if floated free to be judged by the evidence at hand at the time, and thus no matter that if the contest that T won to get its place in K were re-held, T would now be unsuccessful. Levi´s epistemology is thus shown to recommend an unacceptable dogmatism.

There is a further line of criticism advanced against Levi´s theory of contraction by a number of authors, most forcefully by Kaplan. Kaplan argues that an agent who has T in his corpus would never contract K by surrendering T in order to give T´ a hearing, for he would see that one possible outcome of surrendering T is eventual acceptance of T´, and from the agent´s present point of view, T´ is certainly false. Anxious to avoid importing error into his corpus, the agent will refuse to give up T so as not to allow the situation to arise where he could be led to accept a false hypothesis, like the member of Alcoholics Anonymous who will not accept a first drink in order avoid the situation where he will be led on to a thorough binge. Thus, Kaplan argues, despite Levi´s construing replacement as a two stage process, infallibilism does entail incorrigibilism.(93)

Now Levi had anticipated the kind of response Kaplan makes, and had put up a proposal to deal with it, namely that the agent should be **myopic** in relation to the possibility that one of the consequences of contraction will be the importation of error into K. Levi claimed that when surrendering T to give T´ a hearing the agent should look no further than the first stage and not consider the further consequences, for this would, as Kaplan claims, give the agent sufficient reason to hang on to T.(94) But this ploy will not work, for it is completely **ad hoc**, and thus it gives us no reason for being myopic with respect to the future possibilities of importing error while still looking forward to the possibility of gaining further information, which is needed to give us a basis for contraction, as Spielman rightly points out.(95). Levi, that is, needs the agent to be forward looking with respect to the possibility of gaining information as a consequence of contraction - for only then does the agent have a **reason for** contraction - but myopic when it comes to considering the possibility that contraction will have have the eventual consequence of importing error into the corpus - to avoid the agent having a **reason against** contraction. But he cannot have it both ways, and thus his theory of replacement is in serious trouble.

ii. Levi´s rejection of pedigree epistemology.

We have already seen that Levi´s rejection of pedigree epistemology must be modified in a small respect to allow the corpus to retain the memory of the circumstances under which every element was admitted. A more drastic revision of the thesis is necessary, however, for Levi´s attack on epistemological pedigrees in fact undermines his attempt to circumscribe the ways in which an agent can reasonably amend his corpus.

Suppose agent X with corpus K at t´ meets agent Y with corpus K´ at t´, and after comparision of the results they have both got for various problems of mutual interest, X is disposed to conclude that K´ is superior to K, say because while X cannot answer various problems which have been bothering his discipline for some years, Y´s corpus supplies what seem to be adequate answers. X would simply like to throw up K for K´, but he is a follower of Levi, and can find no way of getting from K to K´ while abiding by Levi´s strictures on admissible changes to

K. Now X notes that he could have got to K´ if only his corpus at t, some time past, had been Ky rather than Kx, and he muses that since Levi rejects pedigree epistemology he would not object to X´s having begun his systematic expansion of his knowledge with Ky rather than Kx, there being - as Levi wittily puts it - ´no immaculate preconceptions´, all our beliefs having ´dark origins´, all ´born on the wrong side of the blanket´.(96) But now X will surely become agitated, for a choice made unthinkingly some time ago, which according to his mentor was of no consequence so far as the rationality of his future beliefs were concerned, is now preventing him from adopting the corpus he thinks to be superior to his own. Should X be compelled to endlessly suffer the consequences of some youthful folly? He will surely find this requirement too much to take, and simply throw up K for K´ on the basis that, having now seen the consequences of having had as his corpus at t Kx rather than Ky, he has changed his mind about what ought to have been his beliefs before he began his systematic expansion of his knowledge. But if X can do this once he can do it as often and whenever he likes; by suitable choice of antecedents, no such choice requiring justification according to the anti-pedigree epistemologist, X can attain whatever corpus he desires at the time without contravening any of Levi´s constraints on rational change of corpus.

Levi´s epistemology, that is, is susceptible to the same criticism as is commonly levelled at subjective Bayesians; the agent is simply not constrained by Bayes´ theorem at all, since there is nothing to prevent him working backwards to provide himself with whatever initial beliefs are required to lead to whatever it is he wants to believe at the present time. Now the reason this criticism is successful against the subjective Bayesian is that his theory of induction is anti-pedigree; our criticism of Levi works for the same reason.(97) We conclude that Levi's epistemology must either take account of pedigrees, or, less pejoritavely put, require that initial knowledge states, and not just changes of state, be rationally justified, on pain of complete collapse of rational constraints upon final knowledge states.

iii. Chance and inductive inference.

As we noted in passing in our brief exposition of Levi's epistemology, statements of chance form the basis of his theory of direct inference. They are required in particular to generate the inferences to the likelihoods required for conditionalization. My interest, following Kyburg's suggestion, is to get clear just what is entailed by a chance statement and then to determine whether Levi has or can provide an account of inference to chance statements. If he cannot, then even if we should agree to put aside our other criticisms, his epistemology would be in deep trouble since it would be unable to account for the satisfaction of a main precondition of its application, viz the adoption of statements concerning chances.

In fact Levi's account of chance is not completely clear, at least to me, but it is possible to determine to what an agent is committed upon accepting a chance statement, and that will suffice for our analysis.

Suppose X accepts that the chance of coin a landing heads when tossed in a certain manner at time t, which defines a trial of type **S**, equals the chance of it landing tails, equals 0.5. X is then committed to its being seriously possible that the coin should land heads, and also that it should land tails; and X is committed to the **direct inference** that if the particular toss in question is also known to be a trial of kind T, and it is known that this extra information is stochastically irrelevant to the events in the sample space, then all Q-functions in **K** at t ought to assign equal credal probabilities to the two possible outcomes of the toss. Most importantly, then, via direct inference, knowledge of chances is the foundation for the agent´s credal probability distribution. Note that for the direct inference to go through we must have in K that all information about a trial, other than that which defines it as a trial of the kind for which a chance distribution is known, does not define a new kind of trial with a different chance distribution.(98)

Chances, then, are the foundation for direct inference. And direct inference is a necessary pre-requisite for conditionalization, for the likelihoods employed in Bayes´ theorem, and the priors as well in some cases, are derived from direct inference. Thus securing knowledge of chances is fundamental to Levi´s epistemology. But it is also problematic, for direct inference cannot go through unless we have the knowledge in K to enable us to make the necessary judgements of stochastic irrelevance, and that requires considerable prior knowedge of chances. Levi faces the difficulty head on:

> one cannot justify knowledge of chances by deriving it from the testimony of the senses and the records of memory without a background of theoretical assumptions. (99)

However, this does not lead Levi to amend his approach, for he declares epistemologies which try to get by without presuming significant background knowledge ´bankrupt´, declares Hume´s problem to be ´insoluble´, and styles himself as getting on with ´positive efforts to construct an account of the revision of knowledge´, a task in which it is ´in keeping with what we already know that in order to test statistical hypotheses, we need to have some prior knowledge of chances´.

But despite Levi´s strong words on the matter, his epistemology is surely in trouble on the point of justifying knowledge of chances. As Kyburg observes:

> It seems to follow from Levi´s view that we cannot [obtain knowledge of chances - MR]... rationally: there is no rational procedure for revising a body of knowledge that contains no knowledge of chances so as to arrive at a corpus of knowledge that does contain knowledge of chances. (100)

This surely spells disaster for Levi´s epistemology, unless we severely restrict its domain of application. For Levi´s theory is rather like an investment guide directed to helping agents make their **second** million which provides no guidance on how to get to the point where the problem arises. Of course there is no reason to dismiss Levi´s theory just for this reason, any more than there is to declare useless a money making guide for the very rich. But just as a strategy for getting **richer** does not aid the poor, so Levi´s epistemology does not help those who have not yet been able to justify adopting claims concerning chances. Rather Levi´s epistemology applies only after we have solved the problem of getting knowledge of chances; rather than solving Hume´s problem, therefore, it presupposes that it has been solved, or, as Levi would doubtless prefer to put it, it puts Hume´s problem to one side in order to get on with some interesting problems we face once some basic knowledge of chances is assured.

iv. Levi and the ´curse of Frege´.

There are several points in Levi´s logic of learning from experience where the context of the enquiry plays a part in determining the appropriate decision, most significantly in the choice of questions to be considered, the specification of a set of possible answers to the question selected, which is choice of an ultimate partition, and the choice of value for q, representing the relative importance of a gain in information over avoiding acceptance of false propositions. It is possible to argue that allowing the context of enquiry to influence such decisions allows elements which are beyond rational control to undermine the effort at rational control of the growth of knowledge, but I accept Levi´s contention that one cannot simply accept that this is so. The case against contextualism must be made out, it cannot be simply assumed. But have we not assumed that contextualism is not a permissible approach to induction, and thus begged the question against Levi, in our discussion of the acceptance of chance statements? I do not believe so.

In our discussion of Levi´s analysis of chance statements we have only posed the problem of how we are to account for our possessing such statements, we have not challenged the view that **having got them** they may then be relied upon in further enquiries or, along with other elements of our corpus, be employed as elements of the background knowledge influencing formal factors in our logic of enquiry. To use the more familiar terminology of Levi´s [1967], we have not challenged the project of constructing a logic of **local** rather than **global** induction, we have merely insisted, as we surely have a right to do, that the logic of local induction be able to account for the existence of the background knowledge. We have concluded that it cannot.

Of course it might be said that the background which is needed to get a local enquiry off to a sound start is the result of a prior local enquiry, but on pain of infinite regress we shall have to think of some other answer in the end. Perhaps it will then be said that the background required for a local investigation is in part merely hypothesis rather than part of the agent´s corpus of knowledge. And that might be the case, but to opt for that line of argument is to supplant Levi´s carefully developed epistemology by an element foreign to it, an element perhaps of Popper´s epistemology, which would be a strange addition indeed to Levi´s infallibilism.

The problem which, following Kyburg, we have identified with the knowledge conditions requiring to be met before one´s corpus can support inductive inference according to Levi's theory, is not, therefore, to be sidestepped by reference to the unwarranted rejection of the propriety of the influence of epistemic and practical contexts upon inductive logic, that is by the embracing of contextualism, or the local approach to induction.

Thus the problem with the inference to chances, together with the other criticisms we have discussed, lead us to reject as an adequate basis for induction Levi's attempt to provide a Bayesian analysis of induction as decision-making controlled by distinctively epistemic values. We turn now to the final theory of induction to be considered, the neo-Bayesians.

6. NEO-BAYESIAN THEORIES OF INDUCTION.

The most recent theory of induction to be defended in the literature is the neo-Bayesian theory, a theory which first arose from Jaynes´ work on objective prior probabilities and then, on the basis of his employment of the concept of entropy as a measure of information in this work, led to a new rule of inference. Others have now refined the rule of inductive inference suggested by Jaynes´ work, and a burst of very recent papers is investigating the relationships between the new rule and the established Bayesian rules of inference given by conditionalization and Jeffrey´s probability kinematics. I argue, however, that this recent work, although promising, so far as the mathematically unsophisticated reader is able to grasp where it promises to lead us, has failed to deal with earlier claims of incompatibility between classical and neo- Bayesian inference, and in consequence it is not possible to predict that the new school will be able to make good its present promise of providing an intuitively acceptable basis for inductive inference.

a. Jaynes´ Theory of Maximum Entropy Inference.

The most important recent defender of the objective Bayesian theory is the physicist E T Jaynes, who employs the concept of entropy as a measure of information to give a basis for induction. We shall briefly describe and criticize his proposals, following Seidenfeld´s critique.(101)

The central and appealing proposition underlying Jaynes´ theory is that we can find an objective basis for induction if, for a given problem, we can identify a probability distribution which incorporates all of the information we have on the problem but apart from those constraints leaves our expectations, and thus the distribution, maximally indeterminate. We can formalize this proposition if we can find a formal measure of informational content for probability distributions such that informational content is determined by the degree of uncertainty of the distribution. Such a measure was given by Shannon's definition of entropy as a measure of information content. For discrete distributions, the measure of H, the entropy of the distribution, is given by

$$(*) \qquad H = -\sum_m P_m \log P_m \qquad (102)$$

where P_m is the probability of the mth possible outcome of the experiment involved in the problem at hand - eg. the prediction of the results of a set of drawings of balls from an urn. Thus the objective basis of inductive, or statistical, inference, as Jaynes sees it, is the maximization of H subject to whatever constraints are given by the knowledge we have of the problem.

Now Jaynes' original proposal was to use (*), or a generalization of it for continuous parameters, to find objective prior distributions. But (*) itself incorporates no requirement that the probabilities involved be prior rather than posterior probabilities, and while the generalization of (*) does require a prior it also, like (*), can be used with any amount of information. Thus (*) provides a means of conducting inductive inferences in its own right, quite apart from any use it has in finding priors for feeding into Bayes´ theorem.(103) It is of great interest, therefore, to consider whether this new basis for inductive inference can be shown to be a reasonable basis for such inferences.

b. Criticism of the Neo-Bayesian Approach to Inductive Inference.

Criticism of Jaynes´ theory has come from Friedman and Shimony, and Seidenfeld, both critiques claiming to show that maximum entropy inference, ie. inference that proceeds according to the rule ´choose the distribution that maximizes (*) subject to any constraints given´ is inconsistent with conditionalization.

If I have understood their paper correctly, Friedman and Shimony show that for a simple example of inductive inference (though mathematical technicalities make the analysis of the case complex, and I shall not set out their calculations here) it is entailed by Jaynes´ solution to the inference problem that the random variable involved should take a particular value. That is, Friedman and Shimony present an example of inductive inference which, if we apply the rule (*) to determine a posterior for the expected value of the random variable, concentrates the posterior probability on **just one** of the possible values, giving us, if we accept the inference, certain knowledge **a priori**.(104) This is given as a counter-example to Jaynes´ claim that the rule of maximum entropy provides us with a basis to ´honestly describe what we know´, and it serves also as a counter-example to the claim that the rule provides us with an objective and rationally compelling basis for induction.

Seidenfeld´s argument is similar, presenting a case where employing maximum entropy inference, ie. using (*) as a rule of inductive inference, rather than just as a rule for determining a prior for feeding into Bayes´ theorem, leads to a posterior distribution which contains more information than that at which we arrive by conditionalizing upon the information given, even though the same prior is employed for both inference paths.(105) Specifically, the inference by the maximum entropy rule leads to a posterior which is normally distributed, whereas the inference via conditionalization leads to Student´s t-distribution, the former but not the latter allowing the variance of the population to be calculated. Seidenfeld´s argument thus provides further reason to take the identification of entropy and information, in the sense in which information is given by the data or conditionalization upon it, to be problematic. Thus, at the present time at least, we cannot take (*) to be an acceptable rule of induction, since it leads to posteriors containing information to which we have no right, so far as our right to information is determined by what is given by the data or conditionalization upon it.

But if we are taking (*) as an autonomous rule of inference and not merely as a means of generating prior probabilities, why should we be impressed by the problems raised by Friedman and Shimony, and Seidenfeld, for these rely on assuming that inductive inference ought to be conducted by Bayes´ theorem, or at least that the information to which we are entitled is that which we can get from conditionalization upon the data. This point was made by Williams in connection with his principle of minimizing relative information as a basis for inductive inference quite independent of conditionalization.(106) And the point has some force, for although the probability calculus is employed in the derivation of (*), in the generalization of it to continuous parameters, and in the justification of Williams´ principle by Hobson in his derivation of the measure of relative information employed, these arguments do not presuppose that learning ought to abide by conditionalization.(107)

A problem looms, however, for as we have seen above there is - in rather restricted circumstances admittedly, but those restrictions are not relevant here - a strong argument for employing conditionalization as a rule for learning from experience which rests upon no more than the concept of reasonable belief which is required to get the probability calculus itself as a constraint on rational belief. That argument was given by Shafer, and in our discussion above we judged it compelling. But even if that judgement be over-hasty, we certainly will be faced faced with an inconsistency if we can find neither a basis for rejecting Shafer´s claim that an agent who accepts that his degrees of belief ought to be constrained by the probability calculus is also committed to conditionalization, provided only that his learning proceeds according to plan, nor a basis for rejecting the arguments of Friedman and Shimony, and Seidenfeld, that maximum entropy inference is inconsistent with conditionalization. For we shall then have it that maximum entropy inference is inconsistent with the probability calculus itself, and since the calculus is presupposed in the definition of entropy, that the theory of maximum entropy inference is self-inconsistent. Clearly this matter requires investigation.(108)

A full analysis of the problem, however, is beyond the scope of this thesis, for it would be a major project in itself, requiring both a high level of mathematical sophistication to enable a clear analysis of the issues involved to be given, and a sound sense of philosophical significance to ensure that the epistemological problems did not come to be dominated by the mathematical problems and promises of interesting developments. Much as I would like to read such an analysis, I am in no position to write it. Thus I content myself with noting the claims that have already been advanced, some of which will have to be proved ill-founded if consistency is to be regained.

First we have Shafer´s claim that the basis on which an agent can be committed to having his beliefs obey the probability calculus also requires that he accept conditionalization as the appropriate rule for learning from experience, at least in certain kinds of situations. The same result is accepted by some, Skyrms for example, on the basis of extending the Dutch book argument to conditional bets. Second (in what strikes me as the logical order - it is not the temporal order of the publication of results), there is Skyrms´ purported proof that conditionalization on second order of beliefs (which include the agent´s beliefs about his (first order) beliefs about the world) is equivalent (given certain conditions which we need not mention) to Jeffrey´s probability kinematics on first order beliefs. Thus, according to Skyrms, kinematics is entailed by conditionalization. Then, as he later claims to show, kinematics is entailed by Jaynes´ maximum entropy rule, and kinematics entails Jaynes´ rule (under restrictions which are not relevant here).(109)

Thus we have it claimed that accepting the probability calculus as a constraint on reasonable belief entails accepting conditionalization which entails the reasonableness of probability kinematics which entails Jaynes´ maximum entropy rule. But we also have from Friedman and Shimony and from Seidenfeld that conditionalization is inconsistent with Jaynes´ maximum entropy rule. At least one of these claims must be given up, or alternatively be so restricted that under those restrictive conditions some other of the set of claims breaks down.

It might be thought that this problem will be resolved by some relatively minor consideration such as close attention to the conditions under which the various results have been proved, and thus that it is just a problem with the formalisms involved. But I think that the problem likely goes deeper than that, for such a resolution would proceed by splitting off maximum entropy from probability kinematics or kinematics from conditionalization, leaving the deeper relationship intact. The deeper relationship in question is that the probability calculus as a constraint on reasonable belief is employed in the construction of the formalism for entropy that leads to Jaynes´ rule of inference, and to Hobson´s variation, as already noted, and thus any proof that Jaynes´ rule and conditionalization give incompatible results will lead to trouble in the presence of justifications of conditionalization which proceed from only the constraints on rational belief which are required to get the probability calculus. To escape the trouble, therefore, conditionalization must be shown to be not justifiable in the same manner as coherence, or the claims of Friedman and Shimony and Seidenfeld must be disproved.

It may be that Friedman and Shimony, and Seidenfeld, can be shown to have erred in their attempts to prove an inconsistency between Jaynes' rule and conditionalization, but there has been no attempt to do so in the literature with which I am familiar. And by the nature of their arguments it would seem that the only plausible source of error would be a mathematical miscalculation, since their logic is simple and direct. It is implausible, then, that this is the way out of the problem.

The other possible path back to consistency, however, would be to reject conditionalization, though this would be a rather drastic proposal, requiring the rejection of the classical Bayesian theory of inference as well as the arguments of Shafer, and Lewis' modified Dutch book argument. Perhaps that is the correct approach, but it would place a great burden on the neo-Bayesian analysis of inductive inference, for far from receiving support from their relationship with the classical school their proposed rules of inference would have to stand on their own merits. It could not be confidently predicted that the explication of the concept of information which underlies the neo-Bayesian analysis would emerge unscathed from the intense scrutiny to which it would then be subjected.

For if it was accepted that one must choose between conditionalization and probability kinematics, and the arguments which support these rules, on the one hand, and a neo-Bayesian rule on the other; or alternatively, splitting probability kinematics off from conditionalization, and thus requiring choice between Bayes´ theorem and its support, or kinematics plus the neo-Bayesian rule; then it would be seen that the neo-Bayesian theory is based on an abstract and really only sketchily developed analysis of the concept of information, and that the solid philosophical foundation for probability kinematics is its ties with conditionalization, not with the neo-Bayesian rule.

Whatever is the outcome of the present rapid development of the analyses of rules of inference within the neo-Bayesian camp, therefore, at this time no rationally acceptable proposal has been given, for the analysis has yet to deal with an inconsistency in its very foundations.

We conclude, therefore, that, relative to the policy set out at the beginning of this Chapter, this following our earlier analysis of what is required to rebut Hume´s inductive scepticism, the Bayesian account of induction has provided us with no solution to Hume´s problem.

CHAPTER 8: CONCLUSION.

Following our analysis of Hume´s discussion of the rational foundations of inductive inference we adopted the view that Hume´s inductive scepticism has been defeated only if a form of inductive inference is presented which meets two criteria. First, it must not allow the critic who accepts the premises of an inductive inference to reject the conclusion of the inference (or refuse to attach a specified probability to the conclusion of the inference) without transgressing against a principle or principles of rationality to which the critic can be committed. Second, the model of inductive inference given must require as premises only information which does not require inductive support.

Sticking to this policy the conclusion to this thesis must be that among the theories of inductive inference we have examined, and we have dealt with all of the major accounts of inductive inference based upon the notion that inductive inference is a form of probable inference (which was identified in Chapter 2 as the dominant suggestion for dealing with Hume´s problem), there is no proposal which defeats Hume´s inductive scepticism. We are forced to this conclusion because we have found no system of inductive inference which the critic cannot reject, for all have more or less significant problems, as discussed in detail above. But there are, despite our being forced to this conclusion, some lines open for further work on the problem of induction, in addition to the further development of the systems of inductive and statistical inference we have examined.

First, it is important to see that while none of the systems we have examined successfully **defeats** Hume's inductive scepticism, it is not the case that the criticisms we have made of the versions of inductive inference examined show that inductive scepticism can **never** be defeated. In so saying I take it that Hume's inductive scepticism **is defeated** only if the conditions set out in the first paragraph of this chapter are met, while his scepticism is shown to be **indefeasible** only if it is shown that no system of induction can meet these conditions. Showing Hume's inductive scepticism to be indefeasible would thus require it to be shown that a necessary step in meeting one of the two conditions for defeating his inductive scepticism cannot be made good for **any** system of induction, and we have not shown this. We have shown, that is, that none of the dominant systems of induction presently on offer meets the two criteria for defeating Hume's inductive scepticism, but we have not shown that no **possible** system meets these criteria.

The introduction of probability as a basis for induction has thus achieved a subtle but important advance in the struggle with Hume. For Hume's sceptical critique of induction, which the reader will recall did not consider the possibility of reasonable probable inference, is indefeasible with respect to inductive inferences which aim to be provably truth-preserving. For Hume´s showed that **necessarily** no inductive inference can be provably truth-preserving, since an inductive inference´s being provably truth-preserving would entail the inconsistency of accepting the premises of a (correct) inductive inference while denying the conclusion, whereas there can be no such inconsistency - granted, as we have accepted that Hume proved, that the principle of the uniformity of nature can not be included as a premiss of inductive inference. In contradistinction to this we have shown that no system of probable inference among those examined does **in fact** justify its conclusion, but not that this failure is **necessary**. The possibility that some author will yet defeat Hume´s inductive scepticism by providing a system of induction based on probability and meeting the two criteria set out above thus remains open.

Second, in order to keep clear the nature of the problem which we face while Hume´s inductive scepticism remains undefeated, recall that neither Hume´s argument, nor any criticisms of inductive inference here or in the literature which are akin to Hume´s, provide any warrant for **rejecting** the conclusion of any inductive inference. That would be to commit the fallacy of misconceived refutation. Inductive scepticism does not mandate or even recommend rejecting the conclusion of any inductive inference. Rather it consists in the awareness that every version of induction we have thus far been able to construct allows the critic to refrain from adopting the conclusion. But from this it follows, and this is where inductive scepticism bites, that when one accepts the conclusion of some induction one´s doing so will not, should the critic press for an explanation of this change of belief, be justifiable by reason alone.

Third, given that we find, as we surely should, that it is intolerable that inductive scepticism should remain undefeated, we could consider the possibility of abandoning the conception of justification embodied in our criteria for determining whether inductive scepticism has been defeated. We would doubtless be loath to take this course, since the conception of justification we have adopted is well motivated. In particular, it is universalist, in the sense that the justification of some inference as reasonable to perform applies to all agents in all circumstances. Should we drop either the insistence that the standards of rationality by which an inductive inference is intended to be justified must be applicable for all possible agents rather than being particular to those that share a certain interest in a given enquiry, or the requirement that justifiable inductive inference not require factual assumptions which could be justified only by prior inductive inferences, we appear to open the door to relativism in order to shut the door on inductive scepticism.

Fourth, we could return to the other responses to Hume's problem of induction which I have not considered in this thesis - Popper's fallibilism, attempts at inductive justifications of induction and arguments along the lines of Reichenbach's attempted vindication - to consider whether, contrary to my assessment (based mainly on Salmon's critiques of these alternatives), some successful attack on Hume can be launched from one of these positions.

It would be beyond the scope of this thesis to consider any of these alternative proposals for refuting Hume´s inductive scepticism. It is my own view, however, that the third offers some hope for establishing the rationality of making at least certain inductive inferences for any agent who adopts an appropriate set of background assumptions concerning matters of fact, and whose inductive enquiry is prompted by a particular interest which legitimates the adoption of the specified rule of inductive inference. It is also my view that this can be done while controlling the opportunity for shielding beliefs from criticism which is incorporated in allowing the agent to assume a body of background knowledge. My initial work on this line of investigation has convinced me that it is promising but problematic. This is not the place to launch a new investigation, however, even though the desire to do so, thereby to avoid the dismal conclusion which the present enquiry has forced upon us, is strong.

NOTES TO CHAPTER 1.

1. Stove's analysis is given in his [1966] and especially his [1973]. The influence of this work on my own view of Hume will be apparent to the reader familiar with these texts, though I disagree with Stove on some points concerning the logic of Hume's argument, as will be made clear below. I have also profited from Hacking's [1975], though I shall not discuss the issues he raises here, since my interest in Hume centres on the impact of his work on contemporary philosophy not on its historical antecedents and effects.
2. Stove's presentation of the main points of Hume's argument is given in his [1973]:Part II, with handy summaries on pp. 45 and 51. On Hume's probabilism Stove's [1966] is of independent interest. See also Hacking [1975]:180f.
3. Stove [1973]:34 and 50.
4. Hume [1978]:86f.
5. Hume [1978]:650.
6. Stove [1973]:33. For evidence of Hume's conception of inference as proceeding on the basis of psychological exigencies see passages such as that at the foot of pp.46 of his [1977], and for a detailed study see Passmore [1952]:18f.
7. Stove [1973]:50
8. Outside of mathematics Hume had scant regard for formal rules of reasoning, as Passmore shows. See also Hume [1978]:175.
9. Hume [1975]:30.
10. Hume [1978]:88f.
11. Stove [1973]:42f.
12. Hume [1978]:127. Note that Hume thought that chances must all be equal, taking the more probable event to have more of the equal chances leading to its occurrence.
13. Hume [1978]:136.
14. Hume [1975]:34.
15. Hume [1978]:652.
16. The quotations are, respectively, from Hume [1978]:91, 92, 652 and [1975]:32.
17. Hume [1975]:38.
18. To be fair to Stove he does recognize the difference in Hume's attitude to the **a priori** and predictive-inductive inference (Stove [1973]:53), but he locates this difference in Hume's assessment of the psychological force of the inference, notes that the two forms of inference are **logically** on a par, and takes Hume's sceptical conclusion to be just the logical part of his assessment. But we ought to translate Hume's psychological claims into philosophical claims and thus make sense of them, not exclude them from the analysis; thus we ought to retain Hume's distinction between the force of **a priori** and inductive inferences. Therefore, since the **a priori** inference has no force, it is unreasonable, while the inductive inference, since it is compelling, is not unreasonable, but rather is lacking rational justification. For further discussion of this point see below, Chapter 3 section (d).

NOTES TO CHAPTER 2.

1. I shall discuss most of the dissolutionists' arguments, but not all of them, since in my view those in Moore [1952]; Ryle [1957] and (following him), Toulmin [1958]; and Will [1947]; have been thoroughly rebutted by, respectively, Brodbeck [1952]; Achinstein [1960]; Alexander [1958]; and Stopes-Roe [1960] and Wang [1950].
2. Will [1942]:507. It should be noted here that Will does not address his argument to Hume specifically, but there is no doubt that Hume was his prime target.
3. In addition to Will, the following authors have made essentially the same point: Ambrose [1947]:254; Black [1949]:61; Edwards in Swinburne (ed) [1974]:36 (originally published in 1949); Pap [1949]:186; and Strawson [1952]:250.
4. Edwards [1974]:38.
5. Other authors to have made this point are Black [1949]:75; Strawson [1952]:250; and Will [1942]:507f.
6. This view is shared by Black [1949]:74; Pap [1949]:186; and Strawson [1952]:250. A close relative of (3), which I will discuss below, is defended by Ayer [1952]:49f; and another by Goodman [1979]:62.
7. Strawson [1952]:257.
8. Ambrose [1947]:254.
9. Black [1949]:61.
10. Edwards [1974]:36.
11. Strawson [1952]:250.
12. Black [1949]:61.
13. Edwards [1974]:36.
14. It may be objected here that we do not require that the premises of an inductive inference make belief in the conclusion rational, but merely that they make rational the belief that the conclusion is probable. Adopting the suggestion, however, would not affect my point, and it would introduce into the present argument a complication which is unnecessary. Therefore the objection will be put aside till later chapters, where the suggestion can be examined in detail.
15. This point is put forcefully by Stove in the course of his critique of the ordinary language solution of Hume's problem of induction. See Stove [1973]:109f.
16. See Edwards [1974]:36; Strawson [1952]:250; and Black [1949]:66.
17. Goodman [1979]:62.
18. Ayer [1952]:50.
19. Strawson [1952]:257.
20. See Salmon [1965]:277, and Urmson [1974]:passim.
21. My comments here draw heavily on Salmon [1965]:278.
22. Goodman [1979]:64.
23. Carnap [1968]:226.
24. Haack [1976]:118.

NOTES TO CHAPTER 3.

1. The clearest account, to my mind, is in Goosens [1979]:82.
2. Stove [1966]:188 and 209f.
3. Stove [1973]:129.
4. Stove [1973]:57f.
5. Stove [1966]:198.
6. Stove [1973]:69.
7. Stove [1973]:72.
8. Stove [1973]:76.
9. See Stove [1973]:Chapter 1 and especially pp.9, and [1977]; and Bradley [1977].
10. Stove [1973]:57.
11. Of course further support for Stove´s interpretation of Hume´s conclusion as the statement of logical probability he offers comes from his interpreting Hume´s claim as ´Inductive inferences are unreasonable´. But in both this Chapter and the first I have given reasons why that suggestion ought to be rejected in favour of ´Inductive inferences are not rationally justified´, and this interpretation does not lend itself to being interpreted as a statement of logical probability.
12. Stove [1973]:57.
13. Stove [1973]:59.

NOTES TO CHAPTER 4.

1. For Giere´s account of foundationism, and his alternative, see his [1975].
2. For references to the historical literature on probability, and a philosophical analysis of the three concepts of probability mentioned, see Weatherford [1982].
3. For references to the historical literature on statistics see Giere [1979].
4. Giere [1977]:21.
5. Giere ([1977]:67n.8) recognises the difficulty of characterizing Fisher´s position, and notes that part of the difficulty stems from Fisher´s having revised his earlier works in line with his developing, and perhaps significantly changing, views. Thus the matter of his position in the conceptual framework is complicated by his historical position. Coming before the development of the testing paradigm by Neyman and Pearson, and the elaboration of what I will call the ´confirmation paradigm´, Fisher may with justice be seen as the fount from which both sprang; but I think that he, on seeing where one was led by the testing paradigm, opted for the confirmation paradigm.
6. For an overview of Fisher´s contribution to statistics see Savage [1976].
7. Seidenfeld [1978]:709. Seidenfeld notes that the frequentist position does not require one to adopt a frequency theory of probability.
8. Hacking [1980].
9. Toulmin [1950]:31f. Kyburg [1970]:5 provides a brief critique of the modal theory.

NOTES TO CHAPTER 5.

1. Reichenbach [1949]:360f.
2. Giere [1975].
3. Hacking [1980]:141. See also Giere [1975]:217, and note that Giere reports that Pearson is reputed not to share Neyman's enthusiasm for the behaviouristic view of statistics, and thus I refer to Neyman-Pearson statistical methods but Neyman's behaviourism.
4. Neyman and Pearson [1933]:141f. Note that I use the symbol '=>' to mean greater than or equal to, and '<=' for less than or equal to.
5. Hacking [1980]:153.
6. Neyman [1941]:379.
7. My counterposing conclusions and decisions draws on Tukey [1960]. I do not make the distinction in the same manner, however, for what seems to make the difference for Tukey is the degree of surety of the judgement (to use that as a term to cover both conclusions and decisions), whereas it seems to me that a decision (say, to go to the beach tomorrow), can be more certain than an inference (say, that the weather will be fine tomorrow). Nor do I think it is helpful to follow another suggestion and limit decisions to the judgements which issue in actions, while conclusions lead to beliefs, since there is nothing wrong with the notion that after seeing the evidence we **decide** to accept, or believe, a certain claim. Rather, what makes the difference between conclusions and decisions, in my view, is that conclusions are reached by a process of judgement which is abstract with respect to the individual, and is thus compelling for all agents with the same information, whereas in reaching a decision one can properly take account of the personal consequences of the various acts under consideration. Conclusions, then, are objective, whereas decisions include an element of subjective assessment. Thus judgements reached by weighing the evidence plus objectively determined public utilities, such as Levi's epistemic utilities, I will count as conclusions (though, where it is not important to keep the distinction clear, I will employ the common terminology and call such inferences 'decisions').
8. Neyman [1957]:7.
9. Neyman [1957]:8 - first emphasis added.
10. Hacking [1980]:143 n.1 notes the connection between Neyman's theory of confidence intervals and aspects of Peirce's philosophy of induction, but as I understand it that connection was not known to Neyman or to other statisticians at the time. Neyman's major paper is his [1937], but his [1977] should also be consulted, for it concentrates on the logical foundations of the theory.
11. There are many presentations of the theory of hypothesis testing both in the original literature and the many text-books in which the Neyman-Pearson account of statistical methods is the received theory of statistics. My description of hypothesis testing is based mainly on Neyman's introductory text (Neyman [1950], especially Chapter 5), which although mathematically uncomplicated, is especially strong on the logical foundations of the tests.
12. Each such hypothesis is called a simple hypothesis, for it suffices with M to determine a distribution for **x**. An hypothesis which

assigned p to a sub-interval of [0,1] would be a disjunction of simple hypotheses, or a complex hypothesis. I shall be concerned primarily with the problems of testing simple hypotheses, since tests of complex hypotheses involve complications which, while important and duly noted below, detract from the fundamental problems of Neyman-Pearson statistics.
13. Neyman [1950]:272.
14. Fisher [1955]:73.
15. Giere [1975]:329.
16. See, for example, Lindgren [1976]:278f.
17. See, for example, Neyman and Pearson [1933]:146.
18. See Neyman [1950]:270, and [1977]:106.
19. Neyman [1950]:274.
20. Churchman [1948]:218.
21. Lindgren [1976]:295.
22. Neyman and Pearson [1933]:141f.
23. Hacking [1965]:105. See also Jeffreys [1963]:376.
24. Hacking returns to this criticism in his [1980] defence of Neyman-Pearson statistics, but he does not provide a solution, suggesting that we need to introduce the notion of chance into the reliability programme, in order to treat error frequencies as chances applicable to the single case of the employment of a test. But this will undermine the reliability programme just as surely as the notion of probability which Bernoulli's theorem requires; for once we admit chances, inferences to chances (presumably from observed relative frequencies), and inferences from chances to future relative frequencies, it is hard to see what inferences are left to be supported by the Neyman-Pearson rationale.
25. On the first two points mentioned see Hacking [1965]: Chapter 7, and on the third, Pratt [1961].
26. Giere [1977]:32.
27. Wald's original [1942] suffers from this defect, but it remains an exciting analysis of the possibilities. Giere's [1977] brings philosophical considerations to the fore, and tries to find a basis for determining epistemic costs in order to make the decision-theoretic account a more realistic account of science - an attempt which I find promising. But even if Giere should be successful, other problems would remain for Neyman-Pearson statistics.

NOTES TO CHAPTER 6.

1. Fisher [1936]:3/451. Note that this paper was originally published in 1936 and was reprinted in Volume 3 of Fisher's Collected Papers, the page reference applying to the latter volume. This form of citation will be used for all of Fisher's texts included in the Collected Papers, ie. original publication date followed by the number of the volume of the Collected Papers in which the paper is reprinted, followed by the page reference to that Volume.
2. Fisher adopted a modified version of the frequency theory of probability. For a discussion of his concept see Seidenfeld [1979]:71f.
3. Fisher [1935a]:7.
4. Fisher [1925a]:10, 13th. ed. 1958.
5. Fisher [1935a]:7.
6. Fisher [1951]:5/188.
7. Fisher [1935a]:16.
8. Fisher [1935a]:17.
9. See, eg. Anscombe [1963]:84.
10. Fisher [1956]:42.
11. As Hacking noted in his [1965]:81. Spielman [1974]:215 credits Bertrand with the point.
12. Savage [1976]:473.
13. Fisher [1935a]:15.
14. Seidenfeld [1979]:79.
15. Fisher [1956]:42.
16. Fisher [1935a]:13.
17. Giere [1977]:25.
18. Fisher [1955]:5/344.
19. Fisher [1956]:46.
20. Spielman [1974]:217.
21. Carlson [1976] misses the point in reply to Spielman, arguing that in the situations for which significance tests are suited such extra information would not be available. This leads only to the conclusion that when tests of significance are appropriate they are unjustified, and when they are justified they are inappropriate.
22. Spielman [1974]:225.
23. Hacking [1965]:79.
24. The phrase is not mine, but I use it even though I cannot recall its source, since it neatly captures the unreasonableness of hypothetico-deductive inference as form of inference upon which confirmation, rather than disconfirmation (for which it is more suited), may be based. Salmon [1966]:154 devotes some attention to the inference, commenting that apart from its being fallacious it is also susceptible to Hume's sceptical critique. I shall not discuss it further.
25. Fisher [1925b]:2/16.
26. The empirical presupposition of estimation, the fixing of a family of distributions, is similar to the empirical presupposition of an hypothesis test in the Neyman-Pearson system, since it in effect fixes a class of admissible hypotheses (distinguished by the different values given to the free parameters of the family). Though I will not pursue the point, it

being wide of our mark, this suggests that it is not Fisher's tests of significance which are the rivals of Neyman-Pearson tests of hypotheses, but his theory of estimation, tests of significance being a distinct and more primitive statistical procedure. There is some justification for this view in Anscombe [1963]:81, and Fisher [1935a]:17.
27. Hacking [1965]:173; Seidenfeld [1979]:91.
28. Fisher [1922]:1/282; and for a further example see his [1934]:3/145.
29. Fisher [1956]:148.
30. Fisher [1935b]:3/247.
31. Fisher [1956]:151. As Seidenfeld [1979]:223 notes, Fisher here exposes the weakness of proposed pragmatic vindications of induction of the kind proposed by Reichenbach.
33. Seidenfeld [1979]:106.
34. Fisher [1945]:4/508.
35. Neyman [1941]:388.
36. Fisher [1930]:2/431.
37. Fisher [1956]:34.
38. Hacking has changed his views somewhat since his [1965], mainly in relation to his book's assertion of the primacy of the concept of likelihood for statistical inference. When I refer to a view of Hacking's without a date attached, I am referring to his 1965 position. For his later thought see his [1980]. On the influence of Hacking's work on professional statisticians, see Barnard [1972].
39. Hacking [1965]:122.
40. Hacking [1965]:125.
41. Hacking [1965]:122.
42. Hacking [1965]:90. The probability expressions here refer to long run relative frequencies or chances, Hacking's explication of probability$_2$. His account of support, the concept Carnap explicated with his probability$_1$, is discussed below.
43. Hacking [1972]:135.
44. Hacking [1965]:40.
45. Hacking [1965]:41 and 45.
46. Hacking [1965]:55.
47. Hacking [1965]:63f.
48. Hacking [1965]:109.
49. Hacking [1965]:1.
50. Kyburg [1974]:Chapter 6.
51. Hacking [1972]:136.
52. Hacking [1965]:178.
53. Hacking [1965]:184 and Barnard [1972]:131.
54. Hacking [1965]:139.
55. On this point see Kyburg [1974]:66
56. The main point I make about fiducial inference is not the problem which has received most attention in the literature, nor the point which deserves most attention. That point, which has been aired at length, is that we can, by using different pivotals, get different probability distributions for an hypotheses seeking support from a fiducial inference. This problem, and Fisher's attempts to ensure uniqueness, are discussed in great detail by Seidenfeld ([1979:130), and he concludes that the problem is fatal to fiducial inference. I accept Seidenfeld's authority on this point, and go on to develop a further problem with Hacking's account of

fiducial inference.
57. Hacking [1965]:135
58. Hacking notes the relationship between his ´support´ and Carnap´s ´confirmation´ but he does not follow it up. See his [1965]:32.
59. The basis for this analysis is given in the Introduction to Jeffrey (ed) [1980].
60. In the LFP Carnap´s discussion reads as I have claimed, but in his [1963] he denies that in the discussion in the LFP he was trying to refute Hume´s inductive scepticism on the basis that inductive inferences based on $probability_1$ may be rationally justified, claiming that the discussion in the LFP was intended merely to show that if we adopt a c-function of the kind there recommended, then we can show that using this c-function as a basis for our decision-making will mean that long run success is a reasonable expectation, ie. that on the basis of a chosen c-function, it is highly confirmed that using the c-function will lead to success in the long run. In this later work Carnap shifts the burden of justification onto the a priori justifications given for his axioms. We shall deal with the LFP as already indicated, however, and consider the a priori justification below.
61. Much attention has been given to the claims for induction we interpret Carnap as making, these being given prominence by Burks [1963], Hay [1952], Nagel [1963] and Salmon [1966]:210. The focus of our discussion will not be, however, to give arguments against Carnap´s claims, as these authors have, but to show that what Carnap offers as reasons for his view do not stand scrutiny.
62. Note that if we accept Carnap´s 1963 claims about his discussion of the presuppositions of induction in the LFP then the sentence from which we have just quoted constitutes his major analysis of the justification of induction in the LFP, for by the end of the sentence $probability_1$ is accepted as a basis for rational inductive inference, and according to Carnap in 1963, the discussion following does not aim to justify this acceptance.
63. LFP:178. All quotations in the remainder of this section are taken from LFP:178-182 and will not be separately referenced.
64. Apart from the main problems with the line of Carnap´s argument, dealt with in the body of the text, there are two problems with this claim which deserve mention. First, X would not be ´clearly justified in following the inductive method´ if he knew the truth of (3), for while (3) might be true as a claim about the long run it might also be the case that for X´s finite lifetime as a decision-maker in need of a rational inductive method, he would be more successful if he followed some method inconsistent with induction as Carnap defines it. Carnap´s argument, that is, is a long run justification, and it shares all of the problems of that species of justification which we discussed in our earlier analysis of the reliability paradigm. Second, whether (3) is true or false will depend on what evidence is available to X; if X chooses to use as evidence of tomorrow´s weather the point on the compass to which his cat heads on being let into the house after dinner, then (3) may well turn out to be false. Carnap´s response to such criticisms is of course that X must base his inductions on the total relevant evidence, but the same problem could arise with the totality of X´s evidence. Moreover, the total evidence requirement results in X not knowing whether his inductions are justified

by the truth of (3), since the justification of an induction requires that X know that he has taken account of the total relevant evidence, and it is doubtful that in any realistic situation X would be in a position to know this.

65. **Cont**:12.

66. LFP:285. To avoid seeming triviality we should read ´probability$_1$´ as ´degree of confirmation´.

67. LFP:165f.

68. Carnap [1963]:975.

69. **Cont**:14.

70. **Cont**:14. A Q-predicate is a complex term consisting of, for each primitive predicate, either the predicate or its negation. Thus for **n** primitive predicates there are 2 to the n Q-predicates. For example, if the 2 primitive predicates of a language are ´round´ and ´hard´, the Q-predicates are: Q_1 = round and hard; Q_2 = round and not hard; Q_3 = not round, and hard; Q_4 = not round, and not hard. If an object was hard it could be described by the term $Q_1 \vee Q_2$, and completely described by one of Q_1 or Q_2.

71. Nagel [1963]:798, and Salmon [1967]:763.

72. **Cont**:26

73. In the Appendix to **Cont** Carnap acknowledges that he failed to take account of the minimax solution to the problem of estimating a relative frequency, and there subjects it to analysis. He rejects the minimax method on the grounds that it violates several of his requirements and does not meet other requirements it would be desirable for a c-function to meet. But Carnap´s criticism may well beg the question against the minimax method, for the idea behind minimax may be taken to be as basic as those intuitions which motivate Carnap´s adequacy requirements for c-functions.

74. **Cont**:48 and 54.

75. **Cont**:42. Note that rejecting very small values of L is not the only way to respond to the problem of strong confirmation from small samples; it would also be possible to develop a theory of the weight of evidence, as Keynes suggested. And in view of the problem Kyburg ([1974]:123f) identifies with larger values of L, that might turn out to be the better option.

76. **Cont**:38.

77. **Cont**:55.

78. **Cont**:59f.

79. **Cont**:56.

80. **Cont**:78. e´ denotes a method of estimation not below the lower bound on L, and e_o is the straight rule, which is below the bound. Indeed, that there is such a bound above e_o is Carnap´s main result here, since it provides a basis for criticism of the common statistical practice of using e_o as an estimator, establishing that the requirements of unbiasedness, and that an estimator have minimum mean squared error, are incompatible. (**Cont**:79)

81. I shall refer to these texts as **IL&II**, **Replies** and **Basic** respectively, distinguishing between the two halves of **Basic**, when necessary, by calling the first **Basic** I and the second **Basic** II. **IL&II** was published in Lakatos (ed) [1968], **Replies** in Schilpp [1963], **Basic** I in Carnap and Jeffrey (eds) [1971] and **Basic** II in Jeffrey

(ed) [1980].
82. I use 'C' to denote the confirmation functions of the Basic System to distinguish these from the c-functions of the first system, for the two functions are formally distinct. Hilpinen [1973] gives an accessible and interesting review of the innovations introduced in **Basic**.
83. **Basic** I:122.
84. **Basic** I:161.
85. Kemeny [1963]:732.
86. On some occasions Carnap invoked inductive intuition to support instantial relevance and other principles which I have not given details of here. By this intuition he did not mean **a priori** reason, but rather the considered judgement of the philosophical community, as Jeffrey [1973]:306 makes clear. But, in contra-distinction to the present case, Carnap required considered opinions on many problems which had not been formulated prior to his work, and thus he had no body of thought to turn to for justification. Perhaps wider debate would have agreed with his judgement, but this is not easy to predict.
87. **Cont**:14.
88. Nagel [1963]:798.
89. Nagel's criticism of the assumption of the symmetry of Q-predicates ran along similar lines: it is a matter for experiment to determine whether there are any differences in the real to be recorded by distinctions between predicates - eg. to determine whether a thick wire is the evidential equal of a thin wire in a test of electrical resistance.
90. Carnap [1963]:992.
91. **LFP**:485.
92. I think that this is the point Howson was wanting to make in the penultimate paragraph of his [1975]. I shall consider the matter at greater length in connection with Kyburg's conception of epistemological randomness which, like Carnap's adoption of the principle of symmetry of individuals, is founded upon the assumption of the irrelevance of any information we do not possess.
93. Spielman [1976] offers a proof of the result that the assumption of the symmetry of individuals amounts to the **a priori** assumption that either the individuals are all identical, or that they form a collective in von Mises' sense. Either assumption is unjustified. Spielman's proof is suggestive, for quite apart from its formal success, it leads one to conclude that any useful application of a language which could draw no distinction between two individuals would require either that the two individuals be identical, or that their differences be irrelevant to the investigation for which the language has been adopted. As Spielman rightly points out, either assumption, and thus the adoption of a language which forces one of these assumptions upon one, is unwarranted **a priori**.
94. **Basic** I:120.
95. **Basic** II:20.
96. **Basic** II:30.
97. **Basic** II:34 and 42.
98. **Basic** II:44 and 93. Note that the meaning of the parameter **L** is not exactly what it was in **Cont**. **L** now characterizes only C-functions which are applicable to families all of whose attributes are at an equal distance one from another. Such families are referred to as having n-equality. Families which do not have n-equality require confirmation

functions outside the continuum indexed by **L**.
99. Basic II:34 and 106.
100. **Basic** II:11.
101. Basic I:51f.
102. Hilpinen [1973]:327. Notable among the philosophers Hilpinen refers to would be Lakatos, in view of his [1968]:363, where he argues, as we have here, that Carnap´s inductive logic presupposes the choice of a language, and, following Nagel and Putnam (in Schilpp [1963]:804 and 779, respectively) that what evidence confirms which hypothesis depends in part on the language or framework adopted. The present discusion, following Hilpinen, extends that critique to Carnap´s second system of inductive logic and shows just how and why C(h|e) depends on framework choice. The two discussions come at the point from different angles, however: Putnam and Nagel argued that in actual scientific practice confirmation depends upon a theoretical background, that this ought to be reflected in our confirmation theory, and thus that there is more to confirmation than degree of confirmation as Carnap explicates it; while the present analysis shows that in fact **on Carnap´s model** there is more to confirmation, or empirical support, than degrees of confirmation, viz choice of a theoretical framework, and if Carnap´s theory is to be adopted as a complete logic of scientific inference it would need to be supplemented by a theory of framework choice.
103. I do not wish to give the impression that Carnap would definitely have objected to this characterization of his theory of C-functions as only one half of the logic of empirical support. Indeed, the dependence of degrees of confirmation on the theoretical background may have been in part a response to the earlier criticism of his first system by Putnam and Nagel. Certainly one at times gains the impression that workers in Carnap´s programme were of the view that his confirmation functions were restricted to measuring support within a given framework, and that framework choice was the business of another research programme in the philosophy of science, Popper´s theory being seen as a candidate for this role. Thus Bar-Hillel in Lakatos(ed)[1968]:69 refers to Carnap´s theory as synchronic, and Popper´s as diachronic.
104. Ayer [1968]:104.
105. I will refer to the first mentioned paper as **ILRD**. It was published in Carnap and Jeffrey (eds) [1971].
106. **ILRD**:7
107. **ILRD**:29.
108. A discussion which covers what seem to me to be the main issues and includes a contribution by Carnap is to be found in the paper by Kyburg, and the comments upon it, in Lakatos (ed) [1968].
109. **ILRD**:29
110. Lakatos(ed) [1968]:146ff.
111. Spielman [1981]:57.
112. Kyburg [1979]:432. Material in brackets added.
113. Kyburg [1982a]:13.
114. Kyburg´s theory of probability has undergone some development since he first set it out, but the main lines have remained constant. My summary is based upon his definition given in a recent essay introducing the concept which was published as Chapter 9 of his [1983].
115. Kyburg [1977b]:205. Apparently this paper was originally intended as

a joke and the new theory of the acceptance of universal generalizations was put forward tongue in cheek. Kyburg now claims, however, that he thinks that his new theory is substantially correct. On this change of heart see his Self-Profile and Bibliography in Bogdan (ed) [1982].

116. Kyburg [1977c]:85

117. Kyburg [1977c]:86.

118. Kyburg [1977c]:92. I have simplified Kyburg's example somewhat to facilitate its presentation and analysis.

119. Kyburg [1977c]:95.

120. See Kyburg [1977c] for the origin of the description introduced.

121. Kyburg's theory of inverse inference is set out in his [1974] Chap.14, and summarized in non-technical terms in his [1979]:422. A critique of the inference is given by Seidenfeld in his [1978] and [1979].

122. The debate was conducted in Kyburg [1977a] and [1980], and Levi [1977], [1978] and [1981].

123. Kyburg [1974]:286.

124. Kyburg [1974]:287.

125. Kyburg [1974]:290 and 295.

126. Levi [1977]:20.

127. Kyburg [1977a]:517f.

128. Levi [1979]:732f.

129. Kyburg [1980]:114.

130. Levi [1981]:549.

131. Kyburg [1974]:295. See also Kyburg [1980]:103, second paragraph, the comment on MT 11.3 of his [1974].

132. Levi [1977]:10 and 23f, respectively. Kyburg clarifies the second requirement in his [1977a]:507.

133. Levi [1977]:9.

134. Of course, since probabilities are interval valued, according to Kyburg, the completely indeterminate probability [0,1] is a reasonable probability to be assigned on the basis of ignorance. For simplicity I ignore this, and thus use 'probability' to mean 'determinate probability', ie. the probability identified with a subinterval of [0,1].

135. Kyburg [1982b]:138f. In this brief examination I consider only one of the three complications that Kyburg considers, for we need only to consider one example to make our point. Thus I ignore Kyburg's rule for dealing with the case in which we must choose between taking as the reference set the set of all possible outcomes, or the product set defined by the cross product of the sets of sub-outcomes which lead to the final outcome, eg. choice of urn and choice of ball from urn; and also his rule for handling the possibility that we can identify a subset of the outcome, x.

136. In addition to the passage quoted above from Levi on the reasonableness of direct inference see Kyburg [1977a]:507. Note that the criterion given covers only subsets of the suggested class as rival reference classes. Other kinds of rivals would require other rules to eliminate.

137. Levi [1978]:736.

138. Kyburg [1980]:114. See also Kyburg [1982a]:26.

139. Kyburg [1965], reprinted in Swinburne (ed) [1974]:65. Kyburg's paper was a contribution to a symposium with Barker and Salmon, reprinted in Swinburne. Salmon responded in detail to Kyburg's paper; here I consider

only one argument of Kyburg´s, which proposes just the kind of solution to the problem of induction left open in the preceding discussion of Kyburg´s system of probability and induction.

NOTES TO CHAPTER 7.

1. Jeffreys [1961]:15.
2. Jeffreys [1961]:20.
3. Jeffreys [1973]:24.
4. The passage containing this claim is placed at the head of the first chapter of Jeffreys [1961].
5. On these points see Cohen [1977], the discussion of Jeffrey conditionalization below, and Kyburg [1968b]:59f respectively.
6. Jeffreys [1961]:28, and [1973]:32.
7. Hacking [1967]:314.
8. Jeffreys [1961]:118.
9. Huzurbazar [1976]:31 and 46.
10. Huzurbazar [1976]:46.
11. Hacking [1965]:203f.
12. Seidenfeld [1979b]:422. In reply, Rosenkrantz (Rosenkrantz [1979]:449) argues that a pair of experiments such as Seidenfeld describes should be treated as a composite experiment, and we should choose a prior suitable for the composite rather than either of its parts. However, when one experiment is performed the other may not even be thought of, since the data which support some hypothesis may be derived from diverse sources over long periods of time. Rosenkrantz´s suggestions would thus entail withdrawing one´s probability for an hypothesis when data from some new relevant experiment became available, and recalculating one´s posterior based on a composite experiment formed from all of the experiments which support the hypothesis in question. Such an approach to induction is inconsistent with temporal credal conditionalization, and also with confirmational conditionalization. Moreover, the approach recommended would be hard to put into practice, and certainly is at odds with actual scientific practice. Rather than pursue Rosenkrantz´s line of argument, therefore, it would be better to concede that Seidenfeld has shown that on Jeffreys´ theory one´s prior probability will depend upon which of the pieces of evidence supporting some hypothesis was the first to be considered, since this (perhaps composite) experiment was used to set the initial prior probability distribution.
13. Jaynes [1971].
14. Savage [1967a]:602.
15. De Finetti [1972]:183.
16. De Finetti [1972]:147.
17. Jeffrey [1984]:85f.
18. Ramsey [1964]:80.
19. Ramsey [1964]:78.
20. Evidence for believing that Ramsey´s text, as printed in both his [1950] and in Kyburg and Smokler (ed) [1964], contains an error is found in other commentaries. Von Wright [1962]:332 gives Ramsey´s definition as follows: if the agent is indifferent between the conditioned options G1 if p, G2 if not, and G3 if p, G4 if not, then the agent´s degree of belief in p is given by (writing V1 for the value of G1, etc) (V4 - V2)/(V1 - V3 + V4 - V2), which, if we put V3 = V4 = a, making the second option a for certain, and substitute **b** for V1, and **g** for V2, is

identical to the formula I suggest. Agreement with Jeffrey's reporting of Ramsey is noted in the text below. Kyburg [1968a]:55, however, gives (g - a)/(b - a), of which I can make no sense as a betting quotient, unless he too intended what he wrote as a ratio to be a quotient, and vice versa (and intended to express the agent's payout if he loses the bet as a negative gain).

21. Ramsey [1964]:81.
22. On his Dutch book style proof of the axiom of total probability, and his assumption of the remaining axioms for the probability calculus see Jeffrey [1965]:41;49:69;and 84.
23. De Finetti [1964]:100.
24. De Finetti [1964]:102.
25. For details of the argument see Jackson and Pargetter [1976] and Kennedy and Chihara [1979]. On the reasons for the modifications to the Dutch book argument suggested by Jackson and Pargetter see also Baillie [1973].
26. Savage [1967b]:309.
27. Savage [1954]:172f.
28. Savage [1954]:21.
29. Hacking [1967]:313f.
30. Savage [1967b].
31. Hacking [1967]:320.
32. Teller [1973]:222. Teller also reports that he has considered the possibility of extending Savage's justification for the axioms of the probability calculus to cover temporal credal conditionalization, and found it not to be possible. The details are given in his [1976]:215.
33. Teller [1973]:218.
34. Teller [1973]:247. The bracket after '$P_o(B)$' is missing in Teller's text.
35. Teller [1973]:250.
36. Teller [1973]:253.
37. Teller [1973]:240.
38. Teller [1973]:254.
39. What follows is a gloss on Teller [1973]:241f and 255f.
40. Popper [1959]:94.
41. Jeffrey [1965]:171.
42. Suppes [1966] passim.
43. Though I admit the intuitive appeal of GC I do not mean to imply that it is beyond challenge, as Kyburg shows in his commentary on Teller's [1976].
44. To be fair to Teller we should note that he does not assume that conditionalization, if justified, will be justified as a general rule. Indeed, his aim is to find out what restrictions must be placed on (generalized) conditionalization if it is to be valid, ie. to identify a restricted domain in which it is valid. My argument below, then, will be that the features of scientific practice which are incompatible with conditionalization are so pervasive that the domain of validity of conditionalization is so restricted, if it is not empty, that conditionalization does not count as a justifiably reasonable rule of induction.
45. Skyrms [1980]:122.
46. Shafer [1982]:1083.

47. Shafer [1982]:1078.
48. Shafer [1982]:1078.
49. Shafer [1982]:1086.
50. Shafer [1983]:454.
51. Savage [1964]:184.
52. Birnbaum [1962]:284, and Hacking [1965]:220.
53. Birnbaum [1969]:127.
54. Edwards [1972]; Hacking [1972].
55. Hacking [1965]:65 gives some history of the likelihood concept in statistical inference.
56. Savage [1962]:99. On subjectivism, consider the conclusion reached by Lykken after thoughtfully criticizing the use of statistics in psychological research (Lykken[1968]:158): ´The value of any research can be determined, not from statistical results, but only by skilled, subjective evaluation of the coherence and reasonableness of the theory, the degree of experimental control employed, the sophistication of the measuring techniques, the scientific or practical importance of the phenomena studied, and so on.´ He does not consider the problems involved in agreeing on a definition of ´skilled´ in this context if we cannot agree on objective criteria to evaluate the factors mentioned.
57. Hacking [1965]:221.
58. Savage (ed) [1962]:79-84.
59. Suppes [1966]:43f.
60. For a more general discussion of problems with a theory of induction which does not provide for acceptance, see Kyburg [1968b] and the discussion of that paper. For another view of the relation between evidence and acceptance/degree of belief see the papers by Levi and Jeffrey in Swain (ed) [1970].
61. Jeffrey [1965]:158. The rule quoted is for the special case where the observation concerns an event which can be described simply as either occurring (the event B) or not occurring (the event -B); Jeffrey also gives a generalization to the case where the observation concerns a more complex situation, say the determination of the colour of a piece of cloth, when at least three colours are given non-zero prior probabilities. We need not consider this more complex case.
62. Savage [1977]:15.
63. Savage [1954]:68.
64. Jeffrey [1968]:321f.
65. Levi [1980]:5.
66. Levi [1980]:13.
67. Levi [1980]:14f.
68. Levi [1980]:46.
69. Levi [1980]:48.
70. Levi [1980]:50.
71. Levi [1980]:77. Levi´s attention in his discussion of credal coherence is focused not on justifying the requirement itself, but rather on giving arguments for some special features of the probability axioms he employs. I shall not be discussing these formal details of his theory, since none of my criticisms of his account of induction bear upon them. For details see Levi [1980]:77, and the references given there.
72. Levi [1980]:78.
73. Levi [1980]:81.

74. Levi [1980]:87. Note that just how we get the knowledge of chances which Levi takes to be the foundation of direct inference is a problem for Levi´s epistemology, to which we shall return below.
75. Levi [1980]:89.
76. Levi [1980]:91.
77. Levi [1980]:34f.
78. Levi [1980]:39.
79. Levi [1980]:96.
80. Levi [1980]:133.
81. Levi [1980]:134.
82. Levi [1980]:136.
83. Levi [1980]:137.
84. Levi [1980]:139 and 144. See also his essay in Bogdan ed [1982]:203.
85. Levi [1980]:148. Levi´s discussion of these decision rules contains a powerful criticism of those who argue that all such decisions are closet Bayesian decisions on the basis that for any such decision there is a credal state which would yield the decision as the option with maximal epistemic utility. But, as Levi points out, the credal state has already been taken account of in relation to the epistemic utility of the various options, and the current step is to provide for choice between options not eliminated by the earlier generalized Bayesian stage.
86. Levi [1980]:141.
87. Levi [1980]:58.
88. For Levi´s theory of contraction see his [1980]: Chap. 3.
89. The reviews which I have in mind are Backman [1983], Cohen [1982], Harper [1983], Kaplan [1983a] and [1983b], Kyburg [1984], and Spielman [1983]. For an unflattering evaluation of Levi´s efforts see Fetzer [1982].
90. Cohen [1982]:299.
91. The rejection of pedigree epistemology amounts to the claim that it does not matter how one came by the corpus one in fact has; rather, what matters is how one can improve it - as is plain from the text to n.97 below.
92. Levi [1980]:2.
93. Kaplan [1983b]:315.
94. Levi [1980]:70
95. Spielman [1983]:202.
96. Levi [1980]:1.
97. This point was suggested to me by a related criticism of Levi made by Kyburg [1984]:353.
98. Levi [1980] 251-254.
99. Levi [1982]:212.
100. Kyburg [1984]:354.
101. Seidenfeld´s paper was written as a critique of Rosenkrantz´s theory. However, Rosenkrantz responded (Rosenkrantz [1979]:441) by denying that he adopted Jaynes´ theory of maximum entropy inference as a basis for induction in his [1977], and thus he claimed that Seidenfeld´s critique was ´peripheral´ to his theory. It seems to me, however, particularly in view of Rosenkrantz [1977]:59, that Seidenfeld´s assessment of Rosenkrantz´s theory was well based. Be that as it may, however, it is certainly the case that if one accepts that Rosenkrantz did not adopt maximum entropy as a basis for inductive inference, then his theory is a

version of the traditional Bayesian account and thus, despite his innovations in respect of simplicity, which he emphasizes in his reply to Seidenfeld, it does not require separate discussion here.

102. Jaynes [1968]:229. For Shannon´s original derivation of (*) see Shannon and Weaver [1949] Appendix 2.

103. The independence of (*) as a rule of inference is plain from Jaynes [1968]:229, though he continues to speak of using (*) as a way of finding a prior even when clearly involved in finding a distribution constrained by experimental results, ie. a posterior probability distribution.

104. Friedman and Shimony [1971]:383.

105. Seidenfeld [1979b]:430f.

106. Williams [1980]:134 n2.

107. Hobson [1971]:36f.

108. The inconsistency we are considering here will arise both for Jaynes´ original rule of maximum entropy inference, and Williams´ principle of minimum relative information, as bases for inductive inference, for as Hobson makes plain in his development of the concept of relative information, Shannon´s formula for entropy, which is the basis of Jaynes´ theory, is derivable from the formula for relative information. See Hobson [1971]:43, and Seidenfeld [1979]:n.24. Thus I will not give any separate consideration to minimum relative information inference in this regard, and in my discussion I will refer to both the maximum entropy, and minimum information versions of the neo-Bayesian principle of inference as the ´neo-Bayesian rule´.

109. Skyrms [1980]:124 and 134. Related results, employing a more abstract mathematical formalism, but a less precise philosophical analysis, are given in Domotor et al [1980], while Domotor [1980] purports to show that in general Jeffrey and Field conditionals cannot be reduced to, and thus presumably cannot be justified by, conditionalization.

BIBLIOGRAPHY.

Achinstein P [1960] From Success to Truth. **Analysis** Vol 21 6-9.

Alexander H G [1958] General Statements as Rules of Inference. Feigl et al (eds) [1958].

Ambrose A [1947] The Problem of Justifiying Inductive Inference. **Journal of Philosophy** Vol 44 253-271.

Ansombe F J [1963] Tests of Goodness of Fit. **Journal of the Royal Statistical Society** Series B, Vol 25 81-94.

Asquith P and Kyburg H (eds) [1979] **Current Research in Philosophy of Science** Philosophy of Science Association.

Ayer A J [1952] **Language, Truth and Logic** New York, Dover Publications.

[1968] Induction and the Calculus of Probabilities. **Logique et Analyse** Vol 11 95-144.

Backman W [1983] Practical and Scientific Rationality: A Difficulty for Levi´s Epistemology. **Synthese** Vol 57 269-276.

Baillie P [1973] Confirmation and the Dutch Book Argument. **British Journal for the Philosophy of Science** Vol 24 393-397.

Bar-Hillel Y [1968] Inductive Logic as ´the´ Guide in Life. Lakatos (ed) [1968].

Barker S F [1965] Is There a Problem of Induction? **American Philosophical Quarterly** Vol 2 271-273.

Barnard G A [1972] Review of Hacking [1965]. **British Journal for the Philosophy of Science** Vol 23 123-132.

Bennett J H (ed) [1971-1974] **Collected Papers of R A Fisher** in 5 Vols, Adelaide, University of Adelaide Press.

Birnbaum A [1962] On the Foundations of Statistical Inference. **Journal of the American Statistical Association** Vol 57 269-306.

[1969] Concepts of Statistical Evidence. Morgenbesser et al (eds) [1969].

Black M [1949] **Language and Philosophy** New York, Cornell University Press.

Black M (ed) [1950] **Philosophical Analysis** Englewood Cliffs N J, Prentice Hall.

Bogdan R (ed) [1982] **Henry E Kyburg Jr and Isaac Levi** Dordrecht, D Reidel.

Bradley M [1977] Stove on Hume. **Australasian Journal of Philosophy** Vol 55 69-73.

Brodbeck M [1952] An Analytic Principle of Induction? **Journal of Philosophy** Vol 49 747-750.

Buck R and Cohen M (eds) [1971] **Boston Studies in the Philosophy of Science.** New York, Humanities Press.

Burks A [1963] On the Significance of Carnap´s System of Inductive Logic for the Philosophy of Induction. Schilpp (ed) [1963].

Carlson R [1976] The Logic of Tests of Significance. **Philosophy of Science** Vol 43 116-128.

Carnap R [1950] **Logical Foundations of Probability** Chicago, University of Chicago Press (2nd ed 1962).

[1952] **The Continuum of Inductive Methods.** Chicago, University of Chicago Press.

[1962] The Aim of Inductive Logic. Nagel et al (ed) [1962].

[1963] Replies and Systematic Expositions. Schilpp (ed) [1963].

[1968] Inductive Logic and Inductive Intuition. Lakatos (ed) [1969].

[1971] Inductive Logic and Rational Decisions. Carnap and Jeffrey (eds) [1971].

[1971-1980] **A Basic system of Inductive Logic**, in 2 parts: Part 1, Carnap and Jeffrey (eds) [1971] Part 2, Jeffrey (ed) [1980].

Carnap R and Jeffrey R (eds) [1971] **Studies in Inductive Logic and Probability** Vol 1 Los Angeles, University of California Press.

Chappell V C (ed) [1966] **Hume** London, Macmillan.

Churchman C W [1948] **Theory of Experimental Inference** New York, Macmillan.

Cohen L J [1977] **The Probable and the Provable** Oxford, The Clarendon Press.

[1982] Review of Levi [1980] **Mind** Vol 91 297-301.

Colodney R G (ed) [1966] **Mind and Cosmos** Pittsburg, University of Pittsburg Press.

[1977] **Logic, Laws and Life** Pittsburg, University of Pittsburgh Press.

Domotor Z [1980] Probability Kinematics and Representation of Belief Change. **Philosophy of Science** Vol 47 384-403.

Domotor Z, Zanotti M and Graves H [1980] Probability Kinematics. **Synthese** Vol 44 421-442.

Edwards, A W F [1972] **Likelihood** Cambridge, Cambridge University Press.

Edwards P [1974] Russell´s Doubts About Induction. Swinburne (ed) [1974]. Original publication, 1949.

Feigl H [1950] De Principiis Non Disputandum... ? Black (ed) [1950].

Feigl H, Scriven M and Maxwell G (eds) [1958] **Minnesota Studies in the Philosophy of Science** Vol 2, Minneapolis, University of Minnesota Press.

Fetzer J H [1980] Review of Levi [1980] **Philosophical Books** Vol 23 35-37.

Finetti B de [1964] Foresight: Its Logical Laws, Its Subjective Sources. Kyburg and Smokler (eds) [1964]. Original publication (in French), 1937.

[1972] **Probability, Induction and Statistics** London, Wiley and Sons.

Fisher R A [1922] On the Mathematical Foundations of Theoretical Statistics. Bennett ed [1971-1974] Vol 1.

[1925a] **Statistical Methods for Research Workers** Edinburgh, Oliver and Boyd, 13th ed 1958.

[1925b] Theory of Statistical Estimation. Bennett (ed) [1971-1974] Vol 2.

[1930] Inverse Probability. Bennett (ed) [1971-1974] Vol 2.

[1934] Probability, Likelihood and Quantity of Information in the Logic of Uncertain Inference. Bennett (ed) [1971-1974] Vol 3.

[1935a] **The Design of Experiments** New York, Hafner Publishing Company, 8th ed 1971.

[1936] Uncertain Inference. Bennett (ed) [1971-1974] Vol 3.

[1945] The Logical Inversion of the Notion of the Random Variable. Bennett (ed) [1971-1974] Vol 4.

[1951] Statistics. Bennett (ed) [1971-1975] Vol 5.

[1955] Statistical Methods and Scientific Inference. Bennett (ed) [1971-1974] Vol 5.

[1956] **Statistical Methods and Scientific Inference** London, Macmillan, 3rd ed 1973.

Friedman K and Shimony A [1971] Jaynes´ Maximum Entropy Prescription and Probability. **Journal of Statistical Physics** Vol 3 381-384.

Giere R N [1975] The Epistemological Roots of Scientific Knowledge. Maxwell and Anderson (eds) 1975.

[1977] Testing Versus Information Models of Scientific Inference. Colodney (ed) [1977].

[1979] Foundations of Probability and Statistical Inference. Asquith and Kyburg (eds) 1979.

Godambe V P and Sprott D A [1971] **Foundations of Statistical Inference** Toronto, Holt, Rinehart and Winston.

Goodman N [1979] **Fact, Fiction and Forecast** 3rd ed Cambridge Mass., Harvard University Press. Original publication, 1955.

Goosens W K	[1979]	Stove and Inductive Scepticism. **Australasian Journal of Philosophy** Vol 57 79-84.
Haack S	[1976]	The Justification of Deduction. **Mind** Vol 85 112-119.
Hacking I	[1965]	**Logic of Statistical Inference** Cambridge, Cambridge University Press.
	[1967]	Slightly More Realistic Personal Probability. **Philosophy of Science** Vol 34 311-325.
	[1972]	Likelihood. **British Journal for the Philosophy of Science** Vol 23 132-137.
	[1973]	Propensities, Statistics and Inductive Logic. Suppes et al (eds) [1973].
	[1975]	**The Emergence of Probability** Cambridge, Cambridge University Press.
	[1980]	The Theory of Probable Inferenece: Neyman, Peirce and Braithwaite. Mellor (ed) [1980b].
Harper W L	[1983]	Review of Levi [1980] **Journal of Philosophy** Vol 80 367-376.
Harper W L and Hooker C A (eds)	[1976]	**Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science** 3 Vols, Dordrecht, D Reidel Pubishing Co.
Hay W H	[1952]	Professor Carnap and Probability. **Philosophy of Science** Vol 19 170-177.
Hilpinen R	[1973]	Carnap´s New System of Inductive Logic. Synthese Vol 25 307-333.
Hintikka J and Suppes P (eds)	[1966]	**Aspects of Inductive Logic** Amsterdam, North Holland Publ. Co.
Hobson A	[1971]	**Concepts in Statistical Mechanics** New York, Gordon and Breach.
Howson C	[1975]	The End of the Road for Inductive Logic? **British Journal for the Philosohy of Science** Vol 26 143-149.
Hume D	[1975]	**An Enquiry Concerning Human Understanding** Oxford, Clarendon Press. Original publication, 1777.
	[1978]	**A Treatise of Human Nature** Oxford, Clarendon Press. Original publication, 1739.
Huzurbazar V S	[1976]	**Sufficient Statistics** New York, Marcel Dekker.
Jackson F and Pargetter R	[1976]	A Modified Dutch Book Argument. **Philosophical Studies** Vol 299 403-407.
Jaynes E T	[1968]	Prior Probabilities. **IEEE Transactions on Systems Science and Cybernetics** Vol SSC-4, 227-241.
	[1971]	The Well-Posed Problem. Godambe and Sprott (eds) [1971].

Jeffrey R [1965] **The Logic of Decision** New York, McGraw-Hill.

[1968] Review of I. Levi´s **Gambing With Truth** **Journal of Philosophy** Vol 65 313-322.

[1970] Dracula Meets Woolfman: Acceptance Vs Partial Belief. Swain (ed) [1970].

[1973] Carnap´s Inductive Logic. **Synthese** Vol 25 229-306.

[1984] De Finetti´s Probabilism. **Synthese** Vol 60, 73-90.

Jeffrey R (ed) [1980] **Studies in Inductive Logic and Probability** Vol 2. Los Angeles, University of California Press.

Jeffreys H [1961] **Theory of Probability**, 3rd ed, Oxford, Clarendon Press. Originally published in 1939.

[1973] **Scientific Inference**, 3rd ed, Cambridge, Cambridge University Press. Originally published in 1931.

Kaplan M [1983a] Practical and Scientific Rationality. A Bayesian Perspective on Levi´s Difficulty. **Synthese** Vol 57 277-282.

[1983b] Review of Levi [1980] **The Philosophical Review** Vol 22 310-316.

Kemeny J G [1963] Carnap´s Theory of Probability and Induction. Schilpp (ed) [1963].

Kennedy R and Chihara C [1979] The Dutch Book Argument: Its Logical Flaws, Its Subjective Sources. **Philosophical Studies** Vol 36 19-33.

Korner S [1957] **Observation and Interpretation** London, Butterworths.

Kuhn T [1962] **The Structure of Scientific Revolutions** Chicago, University of Chicago Press,

Kyburg H E Jr [1965] Comments on Salmon´s ´Inductive Evidence´. **American Philosophical Quarterly** Vol 2 274-276.

[1968a] Bets and Beliefs. **American Philosophical Quarterly** Vol 5 54-63.

[1968b] The Rule of Detatchment in Inductive Logic. Lakatos [1968].

[1970] **Probability and Inductive Logic** Toronto, Macmillan.

[1974] **Logical Foundations of Statistical Inference.** Dordretch, Reidel.

[1977a] Randomness and the Right Reference Class. **Journal of Philosophy** Vol 74 501-521.

[1977b] All Acceptable Generalizations are Analytic. **American Philosophical Quarterly** Vol 14 201-210.

[1977c] A Defense of Conventionalism. **Nous** Vol 11 75-95.

[1978] Subjective Probability: Criticisms, Reflections and Problems. **Journal of Philosophical Logic** Vol 7 157-180.

[1979] Tyche and Athena. **Synthese** Vol 40 415-438.

[1980] Conditionalization. **Journal of Philosophy** Vol 77 98-114.

[1982a] Self-Profile. Bogdan (ed) [1982].

[1982b] Reply to Spielman and Harper. Bogdan (ed) [1982].

[1983] **Epistemology and Inference** Minneapolis. University of Minnesota Press.

[1984] Review of Levi [1980] Nous Vol 18 247-354.

Kybyrg H E and Smokler H E (eds) [1964] **Studies in Subjective Probability** New York, Wiley and Sons.

Lakatos I [1968] Changes in the Problem of Inductive Logic. Lakatos (ed) [1968].

Lakatos (ed) [1968] **The Problem of Inductive Logic** Amsterdam, North Holland.

Levi I [1967] **Gambling With Truth** London, Routledge and Kegan Paul.

[1970] Probability and Evidence. Swain (ed) [1970].

[1977] Direct Inference. **Journal of Philosophy** Vol 74 5-29.

[1978] Confirmational Conditionalization. **Journal of Philosophy** Vol 75 730-737.

[1980] **The Enterprise of Knowledge** Cambridge, Massachusetts, The MIT Press.

[1981] Direct Inference and Confirmational Conditionalization. **Philosophy of Science** Vol 48 532-552.

[1982] Self-Profile in Bogdan (ed) [1982].

Lindgren B [1976] **Statistical Theory** 3rd ed New York, Macmillan.

Lykken D T [1968] Statistical Significance in Psychological Research. **Psychological Bulletin** Vol 70 151-159.

Maxwell G and Anderson R (eds) [1975] **Induction, Probability and Confirmation** Minneapolis, University of Minneasota Press.

Mellor H (ed) [1980] **Prospects for Pragmatism** Cambridge, Cambridge University Press.

Morgenbesser S, Suppes P and White M [1969] **Philosophy, Science and Method** New York, St Martin´s Press.

Moore A [1952] The Principle of Induction. **Journal of Philosophy** 49 741-747 and 750-758.

Nagel E [1963] Carnap´s Theory of Induction. Schilpp (ed) [1963].

Nagel E, Suppes P and Tarski A (eds) [1962] **Logic, Methodology and the Philosophy of Science** Stanford, Stanford University Press.

Neyman J [1937] Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. Neyman [1967].

[1941] Fiducial Argument and the Theory of Confidence Intervals. Neyman [1967].

[1950] **A First Course in Probability and Statistics** New York, Holt, Rinehart and Winston.

[1957] ´Inductive Behaviour´ as a Basic Concept of Philosophy of Science. **Bulletin of the International Statistical Institute** Vol 25 7-22.

[1967] **A Selection of Early Statistical Papers of J Neyman** Berkley, University of California Press.

[1977] Frequentist Probability and Frequentist Statistics. **Synthese** Vol 36 97-131.

Neyman J and Pearson E S [1933] On the Problem of the Most Efficient Tests of Hypotheses. Neyman and Pearson [1967].

[1967] **Joint Statistical Papers** Berkley, University of California Press.

Pap A [1949] **Elements of Analytic Philosophy** New York, Macmillan.

Popper K [1959] **The Logic of Scientific Discovery** London, Hutchinson.

Pratt J W [1961] Review of Lehmann´s **Testing Statistical Hypotheses. Journal of the American Statistical Association** Vol 56 163-167.

Putnam H [1963] ´Degree of Confirmation´ and Inductive Logic. Schilpp (ed) [1963].

Ramsey F P [1950] **The Foundations of Mathematics and Other Essays** ed R B Braithwaite, New York, Humanities Press.

[1964] Truth and Probability. Kyburg and Smokler (eds) [1964]. The essay was written in 1926.

Reichenbach H [1949] **Theory of Probability**, 2nd ed Berkley, University of California Press.

Rogers B [1971] Material Conditions on Tests of Hypotheses. Buck and Cohen (eds) [1971].

Rosenkrantz R [1977] **Inference, Method and Decision** Dordrecht, D. Reidel Publishing Company.

[1979] Bayesian Theory Appraisal. **Theory and Decision** Vol 11 441-451.

Ryle G [1957] Predicting and Inferring. Korner S (ed) [1957].

Salmon W [1957] Should We Attempt to Justify Induction? **Philosophical Studies** Vol 8 33-48.

[1965] The Concept of Inductive Evidence (with a rejoinder to Barker and Kyburg). **American Philosophical Quarterly** Vol 2 265-270 and 277-280.

[1966] The Foundations of Scientific Inference. Colodney (ed) [1966].

[1978] Unfinished Business: the Problem of Induction. **Philosophical Studies** Vol 31 1-19.

Savage L J [1954] **The Foundations of Statistics** New York, Wiley and Sons.

[1962] Subjective Probability and Statistical Practice. Savage (ed) [1962].

[1964] The Foundations of Statistics Reconsidered. Kyburg and Smokler (eds) [1964]. Originally published in 1961.

[1967a] Implications of Personal Probability for Induction. **Journal of Philosophy** Vol 64 593-607.

[1967b] Difficulties in the Theory of Personal Probability. **Philosophy of Science** Vol 34 305-310.

[1976] On Re-reading R A Fisher. **The Annals of Statistics** Vol 4 441-500.

[1977] The Shifting Foundations of Statistics. Colodney (ed) [1977].

Savage (ed) [1962] **The Foundations of Statistical Inference** London, Methuen and Co.

Schilpp P A (ed) [1963] **The Philosophy of Rudolf Carnap** La Salle, Open Court.

Seidenfeld T [1978] Direct Inference and Inverse Inference. **Journal of Philosophy** Vol 75 709-730.

[1979a] **Philosophical Problems of Statistical Inference** Dordrecht and Boston, D Reidel Publishing Company.

[1979b] Why I am not an Objective Bayesian. **Theory and Decision** Vol 11 413-440.

Shafer G [1982] Bayes´s Two Arguments for the Rule of Conditioning. **The Annals of Statistics** Vol 10 1075-1089.

[1983] A Subjective Interpretation of Conditional Probability. **Journal of Philosophical Logic** Vol 12 453-466.

Shannon C E and Weaver, W [1949] **The Mathematical Theory of Communication** Urbana, University of Illinois Press.

Skyrms B [1980] Higher Order Degrees of Belief. Mellor (ed) [1980].

Spielman S [1974] The Logic of Tests of Significance. **Philosophy of Science** Vol 41 211-226.

[1976] Carnap's Robot and Inductive Logic. **Journal of Philosophical Logic** Vol 15 407-414.

[1983] Review of Levi [1980] **Theory and Decision** Vol 15 199-210.

Stopes-Roe H [1960] Recipies and Induction. **Analysis** Vol 21 115-120.

Stove D C [1966] Hume, Probability and Induction. V C Chappell (ed) [1966].

[1973] **Probability and Hume's Inductive Scepticism** Oxford, Clarendon Press.

[1977] Hume, Kemp Smith and Carnap. **Australasian Journal of Philosophy** Vol 55 189-200.

Strawson P F [1952] **Introduction to Logical Theory** London, Methuen.

Suppes P [1966] Concept Formation and Bayesian Decisions. Hintikka and Suppes (eds) [1966].

Suppes P, Henkin L Joja A and Mosil G (eds) [1973] **Logic, Methodology and Philosophy of Science IV** Amsterdam, Noth-Holland Publishing Company.

Swain M (ed) [1970] **Induction, Acceptance and Rational Belief** Dordrecht, D Reidel Publishing Company.

Swinburne R (ed) [1974] **The Justification of Induction** Oxford, Oxford University Press.

Teller P [1973] Conditionalization and Observation. **Synthese** Vol 26 218-253.

[1976] Conditionalization, Observation and Change of Preference. Harper and Hooker (eds) Vol 1 205-259.

Toulmin S [1950] Probability. **Proceedings of the Aristotelian Society** Vol 24 27-62.

[1958] **The Uses of Argument** Cambridge, Cambridge University Press.

Tukey J W [1960] Conclusions versus Decisions. **Technometrics** Vol 2 423-33.

Urmson J O [1974] Some Questions Concerning Validity. Swinburne (ed) [1974]. Original publication 1953.

Von Wright G H [1962] Remarks on the Epistemology of Subjective Probability. Nagel, Suppes and Tarski (eds) [1962].

Weatherford R [1982] **Philosophical Foundations of Probability** London, Routledge and Kegan Paul.

Wald A [1942] **On The Principles of Statistical Inference** South Bend Indiana, University of Notre Dame Press.

Will F L [1942] Is There a Problem of Induction? **Journal of Philosophy** Vol 39 505-512.

[1947] Will the Future be Like the Past? **Mind** Vol 56 332-347.

Williams P M [1980] Bayesian Conditionalization and the Principle of Minimum Information. **British Journal for the Philosophy of Science** Vol 31 131-144.