PROBABILITY.

A paper by C.A.R.Hoare to be read to the Jowett Society.
Note added April 1986: at Balliol College, Oxford summer 1959

The rather large number of books which profess to give a definition of the concept of probability fall into two distinct categories : some have been written by working statisticians, and some by philosophers. The statisticians tend to fall into errors typical of the non-philosopher ; while the errors of the philosophers are typically philosophical. The fundamental source of both kinds of error may be traced to an universal lack of understanding of what exactly it is to define an abstract concept like probability, and of the methods which are applicable to such a definition. Philosophers of the so-called Oxford School justly claim to have an understanding of these things ; and, as I hope to show, their methods can be successfully applied to the elucidation not only of the concept of probability, but of a number of related concepts in statistics. Furthermore, once the correct approach to the task of defining probability has been adopted, it is not difficult to reveal the mistakes in other definitions that have been propounded ; they will be seen to bear a marked resemblance to certain well-known philosophical mistakes, which Oxford philosophers have already succeeded in revealing and refuting. Finally, an important controversy current between the proponents of rival views on the nature of probability - that concerning the probability of hypotheses, can be quite decisively resolved.

A few modern philosophers, among them Urmson, Kneale, and Toulmin, have sought to ease the task of defining probability by first analysing the usage of the adverb 'probably'. Their analyses, though interesting in themselves, do not seem to shed much light on the usage of the concept of numerical probability ; for there are many circumstances in which the adverb may correctly be used, and not the noun, and vice versa. One may say that if you toss a newly minted florin, the probability of its falling heads is one half, but one may not with propriety say that it will probably fall heads , or, for that matter, that it will probably fall tails. One may say that the theory of relativity probably gives the correct explanation of the shift in the perihelion of Mercury ; but it would be manifestly absurd to say that the probility is point nine eight seven that it does so. These first examples showsthat an analysis of the adverb 'probably' cannot be a complete account of the usage of the corresponding noun, and the second that it is unnecessary for an understanding of the usage of the noun to give an account of all the circumstances in which the adverb may be used. There are undoubtedly connections between the usage of the noun and the adverb, and it would certainly be interesting and instructive to investigate these connections ; but such an investigation can hardly begin until the concept of probability has been independently elucidated.

This paper is as free from technical terms as it can be made. There are, however, a few terms that cannot be dispensed with. The word 'trial' refers to any experiment or action whose outcome is unknown when the trial is made. To toss a coin, to cast a die, or to take a coloured ball from an urn, is to make a trial, whose outcome may turn out to be 'heads', or 'sixes', or 'a red ball' as the case may be. On a less trivial level, to mate two varieties of sweat pea, or to measure the height of a randomly chosen school-child, is to make a trial, whose outcome may be recorded, and feature in statistical reasoning. The related terms 'expectation' and 'deviation from expectation' will be defined during the course of the paper. Two other terms will constantly recur : an 'universal hypothesis' is one which explicitly or implicitly uses the concept 'all'. 'All men are mortal' is a prime example of an universal hypothesis ; as is also the more interesting statement that an unsupported body falls with an acceleration of thirty feet per second per second. A statistical hypothesis, on the other hand, is one which either explicitly or implicitly uses the concept of probability. 'The probability of heads is on half' is an obvious example ; but the statement ~~environment has no effect on intelligence~~ that intelligence is independent of environmental factors is a more interesting one. It is the statistical hypothesis which features as the subject matter of this paper.

An essential part of the answer tothe question 'In what way are statistical hypotheses used?' is given by the answer to the question 'In what way are they useful?' It is obviously very useful to know what events are going to occur in the future ; for we can then adjust our policy decisions to secure any attainable end we may desire. It is there quiteobviously useful to establish universal hypotheses which tell us quite categorically what events will and will not occur when certain conditions are repeated. But of what use is it to establish that the probability of an event's occurring is such and such? The answer is that knowledge of probabilities is also very useful when we wish to make practical decisions sensibly. A young married man with two infant sons is worried by the possibility that he will die before his sons are old enough to support their mother. If he is told that the probability of his dying so young is only one in a thousand, he may well decide that the risk is worth taking, and will refuse to take out an expensive insurance policy. If, however, he finds out that the probability is in fact nearer one in ten, he will realise that the risk of leaving his dependents destitute is too great, and may well decide to insure his life. The establishment of the truth of a statistical hypothesis is of primary concern to anyone contemplating taking out an insurance policy, or trying to decide on the merits of policies offered by rival firms. It is obviously of even greater importance to the insurance companies when they have to decide on reasonable premiums.

A knowledge of the relevant probabilities can be of pract-
cal use to a woman whose family has a history of haemophilia, when
ne has to decide whether she dare marry and have children. If
ne is a carrier of the disease, then the probability is one half
nat any son she has will die young as a result of inheriting the
isease ; and all her daughtersx will have to face the fearful
ilemma which is now troubling her. If, however, it transpires
nat none of her four brothers and three maternal uncles suffer
rom the disease, then it can be calculated that she herself is
ery unlikely to be a carrier, and she may marry in comparative
ecurity. If, however, she has a brother with the disease, the
robability is one half that she is a carrier, and she may well
ake the decision never to marry. A correct knowledge of the
robabilities is of the utmost importance to a woman in this un-
ortunate situation, and may well save her from condemning herself
eedlessly to a life of spinsterhood.

The paradigm case of the influence of the knowledge of probabilities on the choice of policy is in betting or gambling. If you find out that the horse 'Nonsense Rhyme' will win the next Derby, you would certainly be willing to stake a lot of money on that horse. But if you merely knew that the probability of its winning is a half, you would still bet on the horse provided that the odds you were offerered were twelve to one, six to one, three to one, or in fact any odds more favourable than evens. If you were offered odds of three to one, you would calculate your expectation of gain as follows : there is half a chance of the horse's winning, bringing you a gain of three hundred pounds ; there is half a chance of the horse's losing, entailing a loss of one hundred pounds ; one half of three hundred plus one half of minus one hundred makes plus fifty. Your expectation of gain is plus fifty pounds, and you should accept the bet. In horse-racing it it very rare to know the probabilities concerned ; but in card games and dice games the probabilities are determined by the very rules of the game, which decree that the shuffling should be fair and the dice unbiassed. The sensible player should know the probabilities involved and be able in all circumstances to determine his best policy. Historically, it was the demand of a gambler for a decision on the advisability of a certain betting policy which led to the first systematic application of the calculus of probabilities. But although gaming probabilities are historically important, and still provide a paradigm case for the use of the concept, a man who has a moral objection to gambling may still make successful use of probabilities when his opponent is Nature ; that is, when the gain which accrues to him as a result of a successful choice of policy does not entail the loss of a fellow creature. Indeed, in most non-recreational uses of probability it is Nature, so to speak, who plays the part of a oppo.ent, as the examples of the last paragraph show.

The procedure for calculating expectations is fundamental to the use of statistical hypotheses ; for it is on the basis of an expectation that policy is chosen. In general, the policy which gives the highest expectation of gain (or lowest of loss) which is to be preferred. In the horse-racing example, to reject the bet would obviously mean that the expectation of gain is zero. to accept it, that the expectation is greater than zero ; therefore the bet is accepted. To calculate an expectation is quite simple : suppose you receive alpha pounds if event A occurs, beta pounds if event B occurs, and gamma pounds if event C occurs. Suppose further that the events A, B, and C are mutually exclusive, in that not more than one of them can occur ; and exhaustive in that one of them must occur. Then the expectation is alpha times the probability of A, plus beta times the probability of B, plus gamma times the probability of C. Expectations can thus be calculated for any number of mutually expclusive events to which probabilities have been assigned ; and the technique can be extended to cover an infinity of of possible events. The concept of probability itself can be subsumed under that of expectation, for to say that the probability of A is alpha is equivalent to saying that if you were paid one unit on condition that A occurred but lost nothing if a did not occur, your expectation is alpha. But the concept expectation is wider than that of probability, for it can be used to decide policy even when the probabilities are unknown. A business man seeking a really safe investment at four percent will invest his money in Premium Bonds, knowing that his expectation is four per cent per annum, even though he is quite ignorant of the probability that he will win the thousand pound prize. The concept of expectation is used extensively in statistics, but in the rest of this paper I shall confine myself to that small section of the concept which is directly connected with probability.

We now know of what sort of use a statistical hypothesis can be once it has been established ; but the question immediately arises 'How do we establish a statistical hypothesis?'. The task of the gambler is simple enough ; he knows that certain statements of probability are true, because he knows that his opponents are not card-sharpers or cheats ; but no-one could calculate from the rules of some game what the chance is of his dying before the age of fifty. To obtain factual knowledge of this sort we have to use the only method which can ever yield factual knowledge of any sort - that is the inductive method. First we observe closely certain events that have occurred in the past ; these may suggest to us a hypothesis. From the hypothesis deductions may be drawn, and then further observations may be made, or experiments performed, and if consequences deduced from the hypothesis fail to occur, the hypothesis is rejected. This is in effect the procedure used by statisticians. A doctor, perhaps formulates a hypothesis that in England a T.B. patient is equally likely to be male as female.. The statistician can draw a sample of , say, four hundred patients to test this hypothesis. The expected number of men in the sample, according to the hypothesis, is two hundred, for if you were offered one unit if a patient turned out to be male, your expectation on each patient would be one half, and your expectation on all four hundred patients would be four hundred times one half, that is two hundred. Now suppose in fact when the sample is inspected, it turns out that there are only one hundred and ~~sx~~ sixty men, forty less than the expectation. ~~This meansxthatxif~~ then there is said to have occurred a deviation of forty from expectation. This means that if you had followed the advice of the hypothesis, and bet evens on each patient that he would be a man, you would have lost forty units ; if you had bet evens on women, you would, of course, have won forty units ; but if you had been offered a bet on women at odds just slightly less favorable than evens, the hypothesis would have told you to reject the bet, and you would have missed the oportunity of winning , say, thkirty nine units. Thus the occurrence of a deviation is an unfavourable sign, for it means that you have exposed yourself to the chance of losing money, and to the chance of missing the opportunity of winning it. If your original hypothesis had been that the chance of a patient's being a man is ~~six~~ four tenths, then expectation would have been one hundred and sixty men, the same as the observed value, and there would have been no deviation from expectation. In that case you would have rejected the bet of evens on men, and avoided losing forty units ; and you would have accepted the bet on women, and thus avoided failure to win thirty nine units. Thus a hypothesis according to which the deviation from expectation is zero or quite small leads to policies that would have turned out in a real sense to be better or at least as good in every case as policies which are based on a hypothesis according to which the deviation is large. ~~Thusxaxhypothesis~~

nsu a large deviation indicates that a hypothesis is not a very ood one to guide our decisions, for it has proved to be dangrous and liable to lead to large losses ; and therefore large eviations incline us to reject a hypothesis as false.

The problem immediately arises, how can the largeness of a eviation be measured? How can one deviation be compared with nother with respect to size? Clearly the absolute size of a deiation is no guide, for the larger the number of trials, the arger the deviation tends to be; a deviation of forty in four hunred is sufficient to lead us to reject the hypothesis concerned, ut a deviation of forty one in four thousand would excited no coment.o Relative size is of no more use than absolute size, for a eviation of one in two trials, which would occur if the first two osses of a coin resulted in heads, is quite normal, even though t is a deviation of one hundred percent ; whereas if ninety heads ccurred in the first hundred tosses we should certainly suspect he fairness of the coin, though in this case the relative deviation s only eighty per cent. Statisticians have elaborated a measure f the size of deviation which is intuitively quite appealing ; it s based on the so-called significance of a deviation, that is, n the probability, calculated according to the very hypothesis, f obtaining a deviation as great or greater than that actually bserved. In the case cited, the statistician first assumes the octor's hypothesis of equiprobability is true. He can then calulate the probability that exactly three or twenty three patients ut of four hundred are men, and by adding, can find the probbility that less than one hundred and sixty or more than two hundred and forty patients are male. This is obviously the same as the probability that the number of male patients should differ from expectation, two hundred, by forty or more. This probability is in fact very low, less than one thousandth, which is said to be a very high level of significance, and would generally be considered quite high enough to justify a rejection of the hypothesis in question. In general, when a deviation is observed that is very improbable, it is said to be highly siginificant, and the hypothesis according to which the deviation and the probability have been calculated is to be rejected. This procedure can be interpreted in a more intuitively acceptible way. The doctor who propounded the original hypothesis could himself calculate the probability of a deviation as great or greater than forty in a sample of four hundred patients, and on the basis of his calculation accept odds of a thousand to one against such an occurrence. But if he had, on the basis of the hypothesis, accepted such odds, he would have lost not forty, but a thousand units ; furthermore, even if his bet had been successful, he only stood to win one unit. So large a loss when considered in compariso with so nugatory a possible gain, would be quite unacceptible to most people ; and therefore most people will reject the hypothesis that leads to, to the acceptance of so disastrous a bet.

The above account gives the logic of a rejection rule for statistical hypotheses. It is similar to the rejection rule for universal hypotheses, in that it involves first the temporary assumption of the hypothesis, and then the calculation of the logical consequences of the hypothesis, and then the observation of the relevant facts to test whether the consequences in fact occur. But there is one important and obvious dixtinction between the two types of rejection rule ; for an universal hypothesis is refuted absolutely and irreversibly by a single unfavourable observation, and can never again afterwards be reaffirmed. When the first black swan was discovered the hypothesis that all swans are white had to be abandoned, and however many white swans were subsequently observed, this particular hypothesis could never be revived. But this is not so with statistical hypotheses. We may start with the hypothesis that a particular coin is unbiassed. After tossing it four hundred times we may find that it has fallen heads only one undred and sixty times. A deviation of forty in a sequence of four hundred trials, is, as we have already seen, very improbable, and therefore we may well reject our original hypothesis. Then perhaps the coin is tossed another four thousand times, and in the whole series of four thousand four hundred tosses the number of heads is two thousand two hundred and seven. Then we should be inclined to believe that we were wrong in rejecting the hypothesis, and we may, without logical impropriety, reaffirm it. It is this irrefutibility of statistical hypotheses which make them impossible to define in terms of the simply verifiable and falsifiable statements of observation. But it is exactly this which makes statistical hypotheses so useful to the practical scientist working is such fields as agriculture, forestry, genetics, or psychology ; for in these disciplines practically any universal hypothesis is refuted and becomes unusable after the first ten or twenty observations have been made. In fact, progress in these fields has only been made possible in the last half century by the discovery and development of more refined statistical methods.

The rule of use and the rule of rejection ordain the relationships between statistical hypotheses and the verifiable predictions or semi-predictions which may be based on them, and between statistical hypotheses and the verifiable statements of observation that may be brought forward to disestablish them. But in addition to these relationships, there is a relationship which statistical hypotheses may bear to one another : this relationship is one of logical implication. One universal hypothesis may logically imply another ; and just as the logic of quantification explores these deductive relationships between universal hypotheses, so does the logic of probability examine deductive relationships between statistical hypotheses. The logic of probability is used extensively in the calculation of expectations, deviations, and significance levels, and it is therefore essential to the ~~proper understanding~~ application both of the rule of use and of the rule of rejection. The axioms of the logic of probability are less intuitively obvious than those of ordinary logic, but it can be shown that no other alternative logic could possibly serve the purpose of helping us to decide our betting policies; for a contravention of the traditional rules must lead to immediate disadvantage, as was shown by Frank Ramsay in his interesting article in The Foundations of Mathematics. The law of negation for example says that the probability that an event does not occur is equal to one minus the probability that it does occur ; or, rephrased as a logical implication, the statement that the probability of A is p logically implies the statement that the probability of not-A is one minus p. Let us consider the case of a man who contravenes this rule, and simultaneously asserts S1 The

S1. The probability of A is two thirds

S2. The probability of A not occurring is two thirds.

According to S1, such a man must accept a bet which gives him one pound if A occurs, and lose him two pounds if A does not occur. Similarly, according to S2 he must accept a bet which gives him one pound if A does not occur, and lose him two pounds if A does occur. So he must accept both bets. Then if A occurs he wins one pound on the first bet, and loses two pounds on the second : total loss one pound. But if A does not occur he loses two pounds on the first bet, and wins one pound on the second : total loss again one pound. Thus whether A occurs or not, our friend will lose a pound. Such is the penalty of contravening the law of negation. The rule of addition states that if A and B are mutually exclusive events, the probability of either A or B occurring is equal to the probability of A plus the probability of B (minus the probability of both A and B occurring). That is to assert a logical relationship between three statistical hypotheses :

P1. The probability of A is x

P2. The probability of B is y

P3. The probability of A or B is z (where z equals x plus y)

the rule states that any two of these assertions logically imply the third (provided that A and B are exclusive events.) Now suppose a man contravenes this rule, and simultaneously asserts

S1 The probability of A is one third

S2 The probability of B is one third.

S3. The probability of A or B is one half.

Such a man would have to accept the following three bets : Bet I gives him four pounds of A occurs, and loses two pounds otherwise. B II gives him four pounds of B occurs and loses him two otherwise. B III gives him three pounds if neither A nor B occurs, and loses him three pounds if either of them occur. Then there are three possibilities P1. That A occurs, and that the man wins four pounds on the first bet, loses two pounds on the second, and three on the third: total losss one pound. The same loss is suffered if B occurs. If neither A nor B occurs, the man wins three pounds on bet three, but loses two pounds on each of the other bets. So it transpires that in any event the wretched man is condemned to losing one pound, a loss which might convince him of the validity of the traditional rule of the addition of probabilities.

I have now given, with almost inexcusable oversimplification, an outline of the analysis of the usage of the word 'probability'. I would be very surprised if such an analysis were to satisfy a working statistician or scientist. His obvious objection might be expressed : 'Yes, yes ; but we know all that already. After all, we have been using the concept of probability for many years, and we ought to know how to use it correctly by now. What we want from a philosopher is not a repetition of what we know, but a true definition of the meaning of the words we use daily ; for this you have not given us ; and we shall not be satisfied until we get it.' In order to meet such an objection it would be necessary to give a short course of instruction in the discoveries od modern philosophy. I would point out that there are many ways of defining a concept ; and the way chosen to define a particular concept depends on the nature of the concept in question,; for

for certain kinds of concept only admit of certain appropriate kinds of definition ; and to force the wrong sort of definition on a concept can be proved to be mistaken, by showing that the wrong definition leads to a misrepresentation of the usage of the concept. A simple property may be defined ostensively ; a biological species may be defined per genus et differentiam ; but it would be absurd to maintain that the only legitimate form of definition was the latter form, and attempt to define the colour green per genus et differentiam. The field of mathematics provides apposite analogies. The concept of infinity may be given a contextual definition, the ancestor relation may be given an impredicative definition, and the word 'theorem' itself may be given a recursive definition. But if a man objects that these forms of definitions do not truly define, then he is faced with the fact that the concepts in question are strictly indefinable by any other means whatsoever. Brogress in the field of logic and the foundations of mathematics has only been made possible by the discovery and acceptance of new kinds of definition ; in philosophy also, progress has only resulted from the realisation that abstract concepts like 'all' or 'good' cannot be given non-circular definitions of the traditional kind ; but that there meaning could only be adequately explained by analysing their usage ; and that in fact, to know what a word means is nothing more than to be able touse it correctly. If the objector remained unconvinced, I would challenge him to say exactly which of the many kinds of definition of probability he would consider adequate ; and then I would try to prove that such a definition is in principle impossible. I do not know whether such a line of reasoning would be acceptable to a scientist ; I only hope it is acceptable to the philosopher.

The more philosophical of the authors who have published their works on probability have adhered to some sort of 'degree of belief' view of the nature of probability. Proponents of such a view in its simplest form maintain that to assign a probability to an event is to report on an act of introspection, by which your intensity of belief in the proposition that the event will occur has been measured. If I say that the probability of heads is one half, I am supposed to mean nothing more than that I entertain towards the proposition that the coin will fall heads an intensity of belief that is measured as one half along a standard scale which stretches from nought to one. But this is quite simply not true. In common with most people, my feelings of belief do not have degrees ; I either believe a proposition, disbelieve it, or else remain in doubt about its truth. Even if a more subtle intuition would reveal that my feelings of belief could be graded, perhaps into as many as five categories, I must remain utterly incapable of assigning to this feeling, or to any other, a numerical measure of intensity. But let us suppose even that it is possible in principle to measure the intensity of my feelings of belief ; surely I would not then assess the probability of an event by the intensity of my feeling ; surely I would rather attempt to adjust the intensity of my feeling in accordance with my knowledge of the probabilities concerned, just as I try to adjust the intensity of my feelings of moral indignation or approval to the degree of moral wickedness or goodness which I believe to be inherent in the acts I am considering. Philosophers will see a great similarity between the degree of belief theory in its most naive form and other subjectivist definitions in other branches of philosophy ; and will therefore be able to adapt the familiar arguments against subjectivism to this case.

More recent proponents of the degree of belief view adopt a behaviourist approach. They have taken into account the arguments advanced by Professor Ryle, and realise that belief is not a feeling at all, and therefore that a measure of its intensity by introspection is quite irrelevant. Their analysis of the concept of degree of belief is very similar to the one I have given of the use to which statistical hypotheses may be put. They give a procedure for determining the numerical value of our own and other people's degrees of belief by finding out what odds they are willing to accept. If a man is willing to accept any odds more favourable than evens when betting that heads will accur on the toss of a particular coin, and is also willing to accept any odds more favourable than evens when betting that tails will occur, then we may with reason assert that his degree of belief in the prediction that heads will occur is to be measured as one half. A similar analysis is to be applied to discover one's own degrees of belief. The error of this analysis is that it makes the assertion of a statistical hypothesis tantamount either to a simple prediction of ones own behaviour when faced with a betting situation, or else to an offer of betting odds, of the form 'Iwill bet anyone who cares to accept such and such odds that so and so will occur'. Such an utterance is in effect performative, insofar as it commits the utterer to a certain line of action in certain situations. Now it is true that if a man asserts a statistical hypothesis, and then refuses to accept odds calculated as acceptible on the basis of this own hypothesis, he is thereby convicted of hypocrisy ; but then so is a man who asserts that State Schools are in every way as good as Public Schools, if , at great expense, he sends his own sons to Public School. But this does not prove that the assertions in question are merely performative ; for they obviously have other connotations. A hypocrite is not necessarily a liar.

Philosophers have tried to express the other connotations of an assertion which assigns a probability, by defining that term as a rational degree of belief. The addition of the word rational does supply some of the deficiencies of the original formulation. The word rational has two aspects : firstly it is prescriptive, in that it advises others also to adopt certain betting policies. To say that one half is a rational degree of belief is to advise others to accept any odds more favourable than evens when they have to decide on a policy whose outcome is dependent on the occurrence of the event whose probability is said to be one half. The scientist who establishes intricate probability formulae is not merely saying what odds he himself is willing to accept : he is informing others of what odds it is sensible or rational to accept. The second aspect of the meaning of the word rational is that it presupposes some general rule or principle which dictates the procedure for establishing or refuting a claim to rationality. Proponents of the degree of belief theory make no attempt to elucidate what this rule of procedure can be ; they suggest that intuition is the basis of a claim to rationality, a suggestion which philosophers now realise to be unacceptible. The rule of rejection for statistical hypotheses seems to provide the requisite procedure at least for defeating a claim of rationality ; and this is what is necessary to complete the analysis of the concept rat. deg. bel.

The chief defect of all forms of the degree of belief theory s that they represent a statistical hypothesis as arising in the ind of the asserter either as a result of an act of introspection, r as a more or less arbitrary decision of policy, or simply as prediction of the asserters behaviour.. They leave entirely out f consideration the laborious experimentation, the meticulous d rawing of samples, and the intricate statistical calculations hat are in fact made before a statistical hypothesis is asserted r rejected. It is this aspect of the usage of statistical hyp- theses which is most emphasized by the rivals of the degree of elief theory, who hold what is known as the relative frequency iew of the nature of probability. This view also gives a pro- edure for the determination of probabilities. Consider a type f trial which can be repeated an indefinitely large number of imes ; that is to say, however many times it has already been erformed, it can always be performed once again. When each trial has been performed, there may occur one of two outcomes, A or not-A. Then a sequence of trials is started, and their outcomes are noted. Thus a sequence of A's and not-A's is obtained, which may be made indefinitiely long by continuing the trials. After each trial the number of time A has occurred so far is counted, and the number is divided by the number of trials performed to date, thus arriv- ing at the relative frequency of the occurrence of As among the first n trials. These relative frequencies then form a sequence of fractions as long as the original sequence of trials from which they are cal- culated. This sequence of fractions may also therefore be extended beyond all limit, and may be considered to behave mathematically like and infinite sequence. Under certain conditions, therefore, it may be said to converge to a rela, number, and if it does so, this number is said to be the probability that the trial has outcome A.

This sort of analysis as it stands is quite unacceptible, for it fails to give any practical method for establishing or disestablishing a statistical hypothesis. According to this view, in order to ~~establish~~ discover the value of a probability we first need to perform an infinite number of trials ; then make and infinite number of calculations of relative frequency ; and then make an infinite number of comparisons to establish convergence. This is, of course, quite impracticable. But when the concept of convergence is analysed in the usual way, the frequency view is not as absurd as it seems ; indeed it expresses a very importand aspect of the rule of rejection. According to this analysis, to state that, for instance, the probability of heads is one half is to claim that for every epsilon greater than zero there is a whole number N which has the following property : if the trial is performed N times and more, every relative frequency caleulated after the Nth. toss is between one half plus epsilon and one half minus epsilon. This statement describes its own rule of rejection, for the person who asserts it must be prepared to specify for every value of epsilon a value of N which he thinks will have the required property.; and if it turns out not to have this property, that is , if it turns out that after more than N trials the relative frequency is greater than half plus epsilon or less than half minus epsilon, then the asserter must be prepared to abandon his original assertion ; for such is the rule of rejection

This rule of rejection is in effect very similar to the one already described in this paper. To choose and N for a given epsilon is equivalent to choosing a significance level, which, if attained by the observations, will induce you to reject your hypothesis. Once epsilon has been given, you can calculate from the hypothesis the probability that a relative deviation greater than epsilon should occur after the Nth. xxxx trial ; and therefore you can choose your N to make this probability say less than one in a hundred : and then if it turns out that after the Nth trial the deviation is greater than e, an event will have odcurred that you would have been willing to bet ninety nine to one against : and by choosing that particular N you have in effect chosen to reject the hypothesis which might have led you to make a bet of this degree of unfortunateness. In practice, however, an adoption of the relative frequency rejection rule would lead to a rather different type of calculation of significance from the type in general use today. Nevertheless, the relative frequency formulation of the rule of rejection is a distinct improvement on that which I previously described. In the first place it lays just emphasis on the requirement that the sequence of trials must be indefinitely extendible, since in order to choose an N for every epsilon it is necessary to be able to choose it as large as we please ; and therefore to test the hypothesis we must be able to perform at least N trials. Furthermore the relative frequency interpretation is applicable to hypotheses which use the concept of expectation, and which therefore do not allow of a calculation of significance level. Thirdly, this formulation seems to be more adequate for the theory of sequential sampling, which may eventually supersede the more traditional forms of statistical procedure. Therefore in the rest of this paper the phrase 'rule of rejection ' should be taken to refer to the relative frequency formulation of that rule.

The relative frequency interpretaion of a statistical hypothesis, and its rule of rejection, have all the characteristics which we have found to belong to statistical reasoning. Firstly, the rule of rejection gives only a sufficient reason for rejecting the hypothesis, and not a sufficient condition of its falsity. For it may happen that for a given epsilon the N we choose does not have the required property ; but it may turn out later that another larger N does have the property, and then we may with propriety reassert the hypothesis ; a thing which we could not do if the failure of the N first chosen proved that the hypothesis was false. Secondly, the rule of rejection does not give any reason for accepting a hypothesis. But this again is characteristic of all hypotheses employed in the application of the hypothetico-deductive method. A hypothesis can only be supported by the failure of many honest attempts to refute it ; and this applies as much to statistical as to universal hypotheses. Thus there are neither sufficient nor necessary conditions for the truth of a statistical hypothesis ; and this is evident from the very formulation of the hypothesis according to the relative frequency view. The hypothesis contains mixed quantifiers, - for all epsilon there is an N such that for all numbers greater than N etc. - and it is a commonplace that such statements can be analysed neither in terms of necessary nor of sufficient conditions. It is in an attempt to give such necessary and sufficient conditions that the concept of convergence in a completed infinite sequence is required. Philosophers may recall an idea of early Wittgenstein that an universal hypothesis should be defined as an infinite conjuction of statements ascribing to particulars the property which the universal hypothesis ascribes to all particulars. The crude form of the relative frequency view defines a statistical hypothesis as the disjunction of all those infinite conjunctions of statements describing the outcome of particular trials which satisfy the convergence condition. Working statisticians, who form the majority of proponents of the relative frequency view, undoubtedly find it useful in their calculations to adopt such a definition ; and there is no reason why the philosopher should not allow them this useful computational device, just as they may allow the logician to treat universally quantified statements as if they were infinite conjunctions. Nevertheless a philosopher is quite right in refusing to allow as a legitimate form of definition one which uses the concept of infinity : for this must violate the basic principle of all definition, that you must not define obscurum per obscurius.

The analysis of the concept of probability that I have given applies exclusively to the probability of events, or by a slight transference, to the probability of the statement describing the event being true. Some philosophers maintain that not only can there be a probability that such a verifiable statement should turn out true, but that there are also probabilities that hypotheses themselves are true. The desire to establish that hypotheses can have a probability of being true is a natural one, for if a hypothesis can be shown to have a very high probability, then we should be justified in believing, asserting, betting, and otherwise relying on the truth of the hypothesis, just as we are justified in predicting and relying on the occurrence of an event which we know has a very high probability of occurring. This would immediately confer that lon-sought prize, a justification of induction. Unfortunately, like all such justifications, it turns out to be circular ; for in order to establish that any assignation of probabilies is correct we have to collect a number of instances and then make a step of inductive inference ; and therefore induction is necessary to justify any step of induction, and we might just as well not have embarked on this kind of justification. Furthermore, if hypotheses could be assigned probabilities without the use of induction we would not need to use hypotheses at all, and we would have no need of induction whatsoever. For the rules of probability would enable us to transfer the probability of the hypothesis onto the predictions made on the basis of the hypothesis, and therefore we should know the exact probability of every future event, and that without using any hypothesis, statistical or otherwise, to support the assignation of probabilities. These would all be assigned by the rules of probability logic ; and therefore must logically be correctly assigned.

Unfortunately, our analysis of the meaning of the concept of probability shows that the assignation of a probability to a hypothesis is not only meaningless in the sense that it is strictly useless ; but meaningless in the sense that there are not, nor never can be, any considerations whatsoever that could lead us to reject such an assignation. A probability assigned to the outcome of a future event will help us to decide whether to accept a bet which will bring us profit if the event does occur, and loss if the event does not occur. At some point in the future we shall know whether the event has occurred or not, and we shall know who has to pay whom. It would be stupid to bet on an event if we shall never know whether it has occurred or not ; that is, it is foolish to bet on the truth of a proposition if you will never know whether it is true or not. But this is exactly the case if you bet on a hypothesis ; for it is a characteristic of a hypothesis that however many observations supporting it have been made, it may still be, and even turn out to be, false. If I bet at any odds whatsoever that all swans are either black or white, I should never be in a position to claim the money, even if the hypothesis is true. Of course if a red swan happens to be observed then I shall at once lose my bet :: but this only makes the bet doubly foolish ; for I can never win it, although I may lose. If the hypothesis has mixed quantifiers, then I can never either win or lose. The same considerations apply when your betting opponent is not a fellow- man, but Nature herself. for nature can never offer us a reward conditional upon the truth of a hypothesis, or a penalty conditional upon its falsity. If in fact she could do so, we would not need to assign probabilities to hypotheses, or even use the laborious process of induction to establish them. We would only have to make the bet with nature on the truth of the hypothesis ; and if we win the reward that will prove once and for all that the hypothesis is true ; and if the penalty, we shall know immediately that the hypothesis is false. The fact that this is not the procedure adopted by scientists proves that there is no such reward, and therefore that a knowledge of the probabilities of hypotheses is of no use to us ; and indicates that the assignation of a probability to a hypothesis is to this extent meaningless.

But even if it were useful to us to establish the numerical probability of hypotheses, we could never in fact do so ; since there is no evidence of any kind that could be brought forward to support or reject such an assignation of probability. Evidence can, of course be brought forward to support or refute the hypothesis itself, but never to support the metahypothesis that the probability of the hypothesis being true is such and such. For in order to reason about the probability of the statement A we must be able to observe a trial which may or may not have the outcome described by the statement A : in fact we must be able to observe as long a sequence of such trials as we wish. But there is no sort or kind of repeatable trial which has as outcome the truth or falsity of a hypothesis. It is patently absurd to say that in four hundred trials the hypothesis has been true two hundred and twenty times. and this absurdity is final proof of the meaninglessness of an assignation of a probability to a hypothesis.

4th page from end.

There has been suggested a procedure for assigning proba-ilities to hypotheses which depends first on assigning to a ypothesis an a priori probability, that is a probability it has efore any evidence has been collected, and then calculate from he laws of probability what the probability of the hypothesis is iven any set of evidence which may be at hadd. Naturally no ev-idence can be brought forward to support the assignatioh of an priori probability, and the procedure actually adopted may be escribed and supported as follows. Consider a statement P, of hose truth or falsity we are in ignorance. Someone offers to et us at odds just more favourable than to us than evens that P s true ; and someone else offers to bet, again at favourable dds, that P is false. Our best policy is to accept both bets, for we are then sure to gain ; but suppose that we are only allowed to accept one of them ; which should we accept ? Our best policy is to toss a coin which we know to be unbiassed, and decide which way to bet according to which way the coin falls; for then we know that our expectation of gain is greater than zero whether P is true or not . For if P is true, there is half a chance that we would have bet on its truth and havewon say ten pounds, and half a chance that we would have bet on its falsity and lost nine pounds : expectation of gain is therefore ten shillings if P is true. But a similar calculation shows that our expectation of gain is ten shillings if P is false. Thus in any case we stand to win, provided that we have chosen which way to bet by the toss of a coin. Now this policy can be wrongly interpreted as follows : If the coin falls heads we are willing to bet any odds more fav-ourable than evens that P is true ; that is we are willing to ad-

adopt that very policy that is recommended by the hypothesis that the probability of P's being true is one half. But if the coin falls tails, we are again willing to adopt the same policies as are recommended by ~~xk~~ an assignation of the probability one half ot the falsity, and therefore, by the laws of logic, to the truth of P. Thus whether the coin falls heads or tails we act on the assumption that the probability of P's truth is one half ; therefore one half is a reasonable probability to assign a priori to P. But this line of reasoning is quite invalid, as is obvious when it is stated in these bald terms. The theory of games shows that there are circumstances in which our best policy is to make our practical decisions at random, but to use a random device to decide what line of action we are to adopt is not the same as assigning a probability to any statement whatsoever. The application of the theory ~~xm~~ of games to statistical inference is a very interesting development, well worth exploring by philosophers ; but the subject is unforunately very technical, and I have not yet been able to do any work on it. Nevertheless ~~xm~~ I do know that there is nothing in the theory which ascribes to hypotheses any form of probability, a priori or otherwise, and on the basis of the arguments here put forward I would venture to assert categorically that no mathematical techmique, however refined, will ever provide the basis for assertions of the probility of hypotheses.