Philosophical Theories of Probability

Donald Gillies

ROUTLEDGE
Taylor & Francis Group

London and New York

First published 2000
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2006.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

British Library Cataloguing in Publication Data
A catalogue record for this book is available
from the British Library

Library of Congress Cataloging in Publication Data
Gillies, Donald, 1944-
Philosophical theories of probability / Donald Gillies.
p. cm. — (Philosophical issues in science)
Includes bibliographical references and index.
1. Probabilities. I. Title. II. Series
BC141 .G55 2000
121'.63—dc21   00–029113

ISBN 0-203-13224-6 Master e-book ISBN

ISBN 0-203-17945-5 (Adobe eReader Format)
ISBN 0-415-18275-1 (hbk)
ISBN 0-415-18276-X (pbk)

Contents

4 The subjective theory

So have I heard and do in part believe it.

(Shakespeare, *Hamlet*: I, i, 166)

The subjective theory of probability was discovered independently and at about the same time by Frank Ramsey in Cambridge and Bruno de Finetti in Italy. Such simultaneous discoveries are not in fact uncommon in the history of science and mathematics. Usually, however, although the independent discoverers share a common set of ideas, their treatments of the subject differ both in details and in general approach. These differences are of considerable interest, since they illustrate some of the possible variations in the theory. A detailed comparison of the views of Ramsey and De Finetti has recently been published by Galavotti (1989, 1991, 1994) in an important series of papers. In the course of expounding the subjective theory, I will discuss at various points some of these differences between Ramsey and De Finetti.

The existence of simultaneous discoveries is not perhaps so surprising. Usually there is a problem situation in the subject, and the discoverers react to this by producing somewhat similar solutions. We have seen in the previous chapter that by the mid-1920s there were many severe problems in the tradition of logical Bayesian which went back to Bayes and Laplace. Some statisticians (notably Fisher and Neyman) and some philosophers of science (notably Popper) reacted to this by rejecting Bayesianism altogether. However, another approach was to devise a new version of Bayesianism which overcame the difficulties of logical Bayesianism. This was what Ramsey and De Finetti achieved with their new subjective approach to probability.

Since Ramsey's key paper is usually referred to as Ramsey (1926) and De Finetti's earliest publications have later dates, it may appear that Ramsey is the first discoverer and that De Finetti hit on the same idea rather later. This impression is somewhat misleading, however. Ramsey's paper 'Truth and Probability' was written in 1926, and a large part of it read to the Moral Sciences Club at Cambridge, but it was not actually published until 1931. Ramsey died at the age of only 26 in 1930, having made major contributions to the foundations of mathematics, the philosophy of probability, mathematical logic and economics. His paper on probability first appeared in the collection published after his early death in 1931. De Finetti says that already by April 1928 he had written a complete exposition of the foundations of probability theory according to the subjective point of view. This may have been a little later than Ramsey, but De Finetti was the first to publish (1930a, b, c). In 1931 De Finetti (1931a) gave a full account of the philosophical aspects of the theory without formulas in his ‘Probabilism’, and provided more details about the mathematical foundations in his 1931b paper. Ramsey certainly never heard of De Finetti, and De Finetti seems not to have read Ramsey until after 1937, when his own views had been completely developed [see his new footnote (a) added in 1964 to 1937:102]. Thus, the discovery was completely independent and occurred at almost the same time.

Ramsey’s relation to the older logical tradition is very clear, since he introduces his new theory by giving detailed criticisms of Keynes’s views. De Finetti, however, does not appear to have been influenced by Keynes at the time when he devised the subjective theory. Indeed in his 1931a paper, he seems to be doubtful about what exactly Keynes’s views were, remarking in a footnote: ‘This seems to me to be Keynes’s point of view; but I cannot judge well, since I have only been able to skim his essay quickly.’ (1931a:221). Later, De Finetti expounds and criticises Keynes’s views, and remarks in a footnote: ‘I briefly saw Keynes’s book in 1929 (and I quoted it in ‘Probabilismo’ ... 1931 ...), understanding little of it, however, because of my then insufficient knowledge of English. This year I have read the German version’ (1938:362, Footnote 18). It thus seems clear that De Finetti properly studied Keynes only after his own views had been fully developed. It is also interesting to note that De Finetti’s 1938 paper is entitled ‘Cambridge Probability Theorists’; he mentions only Keynes and Jeffreys, but not Ramsey. This indicates that he probably only read Ramsey after 1938. In the light of all this, I will begin the next section with Ramsey’s criticisms of Keynes, since these follow on naturally from the previous chapter. However in the section ‘Some objections to Bayesianism’ I will give some consideration to De Finetti’s different route to subjective probability. The remaining sections will expound the subjective theory itself. ‘Subjective foundations for mathematical probability’ shows how the mathematical theory of probability can be developed on the subjective approach, and, in particular, gives a full proof of the all important Ramsey–De Finetti theorem. ‘Apparently objective probabilities in the subjective theory’ introduces the key notion of *exchangeablility,* which, as we shall see, plays a most important rôle in the theory. Both these sections are largely based on De Finetti (1937), which is my own preferred account of the theory. However, I will introduce a few changes and amplifications for the sake of clarity and will also mention some alternatives to be found in Ramsey and in De Finetti’s later work. ‘A comparison of the axiom system given here with the Kolmogorov axioms*’ and ‘The relation between independence and exchangeability*’ cover some rather mathematical points, and in another section I will present my criticism of De Finetti’s exchangeability reduction.

Ramsey's criticisms of Keynes[1]

According to Keynes there are logical relations of probability between pairs of propositions, and these can be in some sense perceived. Ramsey criticises this as follows:

> But let us now return to a more fundamental criticism of Mr. Keynes' views, which is the obvious one that there really do not seem to be any such things as the probability relations he describes. He supposes that, at any rate in certain cases, they can be perceived; but speaking for myself I feel confident that this is not true. I do not perceive them, and if I am to be persuaded that they exist it must be by argument; moreover I shrewdly suspect that others do not perceive them either, because they are able to come to so very little agreement as to which of them relates any two given propositions.
>
> (1926:161)

This is an interesting case of an argument which gains in strength from the nature of the person who proposes it. Had a less distinguished logician than Ramsey objected that he was unable to perceive any logical relations of probability, Keynes might have replied that this was merely a sign of logical incompetence, or logical blindness. Indeed Keynes does say: 'Some men – indeed it is obviously the case – may have a greater power of logical intuition than others.' (1921:18). Ramsey, however, was not just a brilliant mathematical logician but a member of the Cambridge Apostles as well. Thus Keynes could not have claimed with plausibility that Ramsey was lacking in the capacity for logical intuition or perception – and Keynes did not in fact do so.

Ramsey buttresses his basic argument by pointing out that, on the logical theory, we can apparently perceive logical relations in quite complicated cases, while being quite unable to perceive them in simple cases. Thus he says:

> All we appear to know about them [i.e. Keynes's logical relations of probability] are certain general propositions, the laws of addition and multiplication; it is as if everyone knew the laws of geometry but no one could tell whether any given object were round or square; and I find it hard to imagine how so large a body of general knowledge can be combined with so slender a stock of particular facts. It is true that about some particular cases there is agreement, but these somehow paradoxically are always immensely complicated; we all agree that the probability of a coin coming down heads is 1/2, but we can none of us say exactly what is the evidence which forms the other term for the probability relation about which we are then judging. If, on the other hand, we take the simplest possible pairs of propositions such as 'This is red' and 'That is blue' or 'This is red' and 'That is red', whose logical relations should surely be easiest to see, no one, I think, pretends to be sure what is the probability relation which connects them.
>
> (Ramsey 1926:162)

Ramsey's doubts about basing probability theory on logical intuition are reinforced by considering how logical intuition fared in the case of deductive inference, which is surely less problematic than inductive. Frege, one of the greatest logicians of all time, was led by his logical intuition to support the so-called axiom of comprehension, from which Russell's paradox follows in a few lines. Moreover, he had companions in this error as distinguished as Dedekind and Peano (for citations, see Gillies 1982: 92). Hilbert and Brouwer were two of the greatest mathematicians of the twentieth century. Yet Hilbert's logical intuition informed him that the Law of the Excluded Middle was valid in mathematics, and Brouwer's that it was not valid. All this indicates that logical intuition is not to be greatly trusted in the deductive case, and so hardly at all as regards inductive inferences.

Moreover, is so-called logical intuition anything more than a psychological illusion caused by familiarity? Perhaps it is only as a result of studying the mathematical theory of probability for several years that the axioms come to seem intuitively obvious. Maybe the basic principles of Aristotle's philosophy seemed intuitively obvious to scholars in medieval Europe, and those of Confucian philosophy to scholars in China at the same time. I conclude that logical intuition is not adequate to establish either that degrees of partial entailment exist, or that they obey the usual axioms of probability. Let us accordingly examine in the next section how these matters are dealt with in the subjective theory.

Subjective foundations for mathematical probability: the Ramsey–De Finetti theorem

In the logical interpretation, the probability of h given e is identified with the rational degree of belief which someone who had evidence e would accord to h. This rational degree of belief is considered to be the same for all rational individuals. The subjective interpretation of probability abandons the assumption of rationality leading to consensus. According to the subjective theory, different individuals (Ms A, Mr B and Master C say), although all perfectly reasonable and having the same evidence e, may yet have different degrees of belief in h. Probability is thus defined as the degree of belief of a particular individual, so that we should really not speak of *the* probability, but rather of Ms A's probability, Mr B's probability or Master C's probability.

Now the mathematical theory of probability takes probabilities to be numbers in the interval [0, 1]. So, if the subjective theory is to be an adequate interpretation of the mathematical calculus, a way must be found of measuring the degree of belief of an individual that some event (E say) will occur. Thus, we want to be able to measure, for example, Mr B's degree of belief that it will rain tomorrow in London, that a particular political party will win the next election, and so on. How can this be done?

Ramsey has an interesting discussion of this problem. His first remark on the question is that 'it is, I suppose, conceivable that degrees of belief could be measured by a psychogalvanometer or some such instrument' (1926:161). Ramsey's psychogalvanometer would perhaps be a piece of electronic apparatus something like a superior lie detector. We would attach the electrodes to Mr B's skull, and, when he read out a proposition describing the event E in question, the machine would register his degree of belief in that proposition. Needless to say, even if such a psychogalvanometer is possible at all, no such machine exists at present, and we cannot solve our problem of measuring belief in this way.

Ramsey next considers the possibility of using introspection to estimate the strength of our belief-feeling about some proposition. However, he has an interesting argument against such an approach:

> We can, in the first place, suppose that the degree of a belief is something perceptible by its owner; for instance that beliefs differ in the intensity of a feeling by which they are accompanied, which might be called a belief-feeling or feeling of conviction, and that by the degree of belief we mean the intensity of this feeling. This view would be very inconvenient, for it is not easy to ascribe numbers to the intensities of feelings; but apart from this it seems to me observably false, for the beliefs which we hold most strongly are often accompanied by practically no feeling at all; no one feels strongly about things he takes for granted.
>
> (1926:169)

Ramsey is undoubtedly correct here. When I cut a slice of bread to eat, I believe very strongly that it will nourish rather than poison me, but this belief, under normal circumstances, is not accompanied by any strong feelings, or indeed any feelings at all. Ramsey is thus led to the conclusion that: '... the degree of a belief is a causal property of it, which we can express vaguely as the extent to which we are prepared to act on it' (1926:169). I am certainly prepared to act on my belief that the bread is nourishing rather than poisonous by eating it without hesitation, even though I am not having any strong feelings at the time.

On this approach we should measure the strength of a belief by examining the character of some action to which it leads. A suitable action for measurement purposes is betting, and so Ramsey concludes: 'The old-established way of measuring a person's belief is to propose a bet, and see what are the lowest odds which he will accept. This method I regard as fundamentally sound' (1926:172). De Finetti (1930a) also introduces bets to measure degrees of belief.

Betting is of course just one kind of action to which a belief can lead. Does it therefore give a good measure of the strength of a belief as regards other sorts of actions to which a belief might lead? Ramsey defends the assumption that it does as follows:

> ... this section ... is based fundamentally on betting, but this will not seem unreasonable when it is seen that all our lives we are in a sense betting. Whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay at home.
>
> (1926:183)

My own view is that betting does give a reasonable measure of the strength of a belief in many cases, but not in all. In particular, betting cannot be used to measure the strength of someone's belief in a universal scientific law or theory (for a discussion, see Gillies 1988a:192–5). However, let us for the moment accept betting as a reasonable way of measuring degree of belief and see what this assumption leads to.

To do this, we must now present some mathematics, but, since the purpose of this book is to discuss the philosophical aspects of probability, I have tried to keep this mathematics as simple as possible, and indeed it involves no more than elementary algebra. We must first set up a hypothetical betting situation in which the rate at which Mr B is prepared to bet on E (his *betting quotient* on E) can be taken as a measure of his degree of belief in E. Then we introduce the condition of *coherence.* It will be clear that Mr B ought to choose his betting quotients in order to be coherent, and this leads to the main result (*The Ramsey–De Finetti Theorem*), which states that a set of betting quotients is coherent if and only if they satisfy the axioms of probability. I will state the axioms of probability in full and then prove the Ramsey–De Finetti theory for each one. In this way the foundations of the mathematical theory of probability will be established from the subjective point of view.

Definition of betting quotients (q)

We imagine that Ms A (a psychologist) wants to measure the degree of belief of Mr B in some event E.[2] To do so, she gets Mr B to agree to bet with her on E under the following conditions. Mr B has to choose a number q (called his *betting quotient* on E), and then Ms A chooses the stake S. Mr B pays Ms A qS in exchange for S if E occurs. S can be positive or negative, but $|S|$ must be small in relation to Mr B's wealth. Under these circumstances, q is taken to be a measure of Mr B's degree of belief in E.

A number of comments on this definition are in order. First of all it is important that Mr B does not know when choosing q whether the stake S will be positive (corresponding to his betting in favour of the event E occurring) or whether S will be negative (corresponding to his betting against E). If Mr B knew that S would be positive, it would be in his interest to choose q as low as possible. If he knew S would be negative, it would be in his interest to choose q as high as possible. In neither case would q correspond to his true degree of belief. However, if he does not know whether S is going to be positive or negative, he has to adjust q to his actual belief.

We can illustrate this by a hypothetical example from the stock market. Suppose Mr B is now a jobber, and I want to find out what he thinks to be the value of a particular share (BP say). If I say to him: 'I want to sell 100 BP shares, what do you think their value is?', it will be in Mr B's interest to quote a value rather below what he thinks to be the correct one, since in this way he can hope to pick up some BP shares cheaply. Conversely, if I say to him:'I want to buy 100 BP shares, what do you think their value is?', it will be in Mr B's interest to quote a value rather above what he thinks to be the correct one, since in this way he can hope to sell some BP shares at a good profit. If, however, I ask Mr B's opinion as to the value of a BP share without saying whether I want to buy or sell, he will be forced to state his true opinion as to the value. Of course, this is only a hypothetical example to illustrate the point. In actual stock market practice, jobbers quote one price for buying and one for selling.

My next point concerns the way in which the magnitude of the stake *S* is measured, for here there is a difference between De Finetti (at least in his early papers) and Ramsey. De Finetti took the stakes to be in money, whereas Ramsey developed a theory of utility and took the stakes to be in utility as he had defined it. My own preference is for De Finetti's early approach, i.e. stakes in money, and I will now briefly discuss some of the issues involved.

If the bets are to be in money, then it is obvious that the sums used should not be too large – at least in relation to Mr B's fortune. Suppose Mr B's entire savings amount to £500. Then it would not be reasonable for Ms A to propose a bet with him on whether it will rain tomorrow with a stake of £500. On the other hand, if Mr B happens to be a billionaire, a stake of £500 might not be unreasonable, provided Ms A's research grant can cover bets of this magnitude.

Ramsey thinks that difficulties of this sort constitute a serious objection to money bets, for he writes: '... if money bets are to be used, it is evident that they should be for as small stakes as possible. But then again the measurement is spoiled by introducing the new factor of reluctance to bother about trifles.' (1926:176). It seems to me, however, that this difficulty can be overcome. Ms A has to choose a size of stake which is small enough in relation to Mr B's fortune so that the bet will not damage him financially but which is large enough to make him think seriously about the bet. I think that it would, in general, be possible to find such a level for the stakes, especially as we have to imagine Mr B as co-operating with the psychological experiment of trying to measure his degree of belief. If Mr B were totally averse to such an experiment, it would hardly be possible to carry it out.

Although there do not seem to me any major objections to money bets, I regard the introduction of a satisfactory measure of utility as a virtually impossible task. We can see some of the difficulties by giving a few quotations which illustrate Ramsey's own procedure. Ramsey writes:

> Let us call the things a person ultimately desires 'goods', and let us at first assume that they are numerically measurable and additive. That is to say that if he prefers for its own sake an hour's swimming to an hour's reading, he will prefer two hours' swimming to one hour's swimming and one hour's reading. This is of course absurd in the given case but this may only be because swimming and reading are not ultimate goods, and because we cannot imagine a second hour's swimming precisely similar to the first, owing to fatigue, etc.
>
> (1926:173–4)

I find it hard to believe that there is any satisfactory way of comparing the utility of an hour's swimming with that of an hour's reading. Both can give considerable pleasure, but the pleasures are of quite a different kind and so incomparable. Ramsey thinks that this difficulty can be overcome by introducing 'ultimate goods'. But what are these ultimate goods? No ultimate good is ever specified, and such a thing would appear to be a myth rather than a reality.

At another stage of his introduction of utility, Ramsey writes: '... we could, by offering him options, discover how he placed in order of merit all possible courses of the world. In this way all possible worlds would be put in an order of value' (1926:176). Such a procedure seems to belong to the realm of pure fantasy. Compare it with the realistic possibility of betting for a stake of £1 on whether it will rain tomorrow.

It might be objected that these arguments are directed just against Ramsey's way of introducing measurable utility, and that other more satisfactory methods might be available. Yet other methods involve similar difficulties and often lead to curious paradoxes which are difficult to resolve. Surely it is better to avoid this minefield and just consider money bets made with appropriate stakes. This latter procedure, far from belonging to the realm of fantasy can easily be carried out in practice. Indeed, De Finetti used to get his class of students to produce betting quotients on the results of Italian football games. Being of a democratic turn of mind, he invited the porter to participate as well, and the porter was nearly always the most successful. He knew more than anyone else about football.

A further objection to the betting scheme might be that it produces only very rough estimates and hardly exact numerical probabilities. De Finetti's reply to this point is that exact numerical degrees of belief are indeed something of a fiction or idealisation, but that this idealisation is a useful one in that it simplifies the mathematical calculations. Moreover, provided we do not forget that the mathematics must be understood as holding approximately, this idealisation does no harm. As De Finetti himself says:

> ... if you want to apply mathematics, you must act as though the measured magnitudes have precise values. This fiction is very fruitful, as everybody knows; the fact that it is only a fiction does not diminish its value as long as we bear in mind that the precision of the result will be what it will be.... To go, with the valid help of mathematics, from approximate premises to approximate conclusions, I must go by way of an exact algorithm, even though I consider it an artifice.
>
> (1931a:204)

My own conclusion then is that we should use the betting scheme with money bets and appropriately selected stakes, and that this does indeed give a reasonable method for measuring belief in many situations. I therefore adhere to the approach of the early De Finetti. Curiously, however, De Finetti in his later period moved in the direction of using utility, and in his last papers even abandoned the betting approach altogether. In 1957 De Finetti still hesitated to follow Savage in trying to unify probability and utility within decision theory (see quotation in Galavotti 1989:240). However, in 1964 in a new footnote to his 1937 paper he wrote: ‘Such a formulation could better, like Ramsey’s, deal with expected *utilities*’ (p. 102). In his 1970 book he used mainly decision theory to introduce subjective probabilities. He also develops a theory of utility, even though he still seems to regard this with some degree of scepticism (see De Finetti 1970:76–82). In one of his very last papers, he went as far as to repudiate the whole betting approach as inadequate, writing: ‘... betting, strictly speaking, does not pertain to probability but to the Theory of Games ... It is because of this that I invented and applied in experiments (probabilistic forecasts) the “proper scoring rules”’ (De Finetti 1981b:55). Thus, De Finetti himself moved in the direction of decision theory and utilities. However, for reasons already given, my own preference is for De Finetti’s earlier approach, and this is what I will use as the basis of the account which follows.[3]

The first problem in the subjective approach was how to measure degrees of belief. We have seen how the betting scheme offers a reasonable solution to this problem. Mr B’s degree of belief in E is measured by his betting quotient in E as elicited in the situation described above. It is worth noting that this way of introducing probabilities is in accordance with the philosophy of *operationalism.* A recent important contribution to subjective probability is Lad (1996). In this book, Lad provides a foundation for subjective probability similar to De Finetti’s but goes beyond De Finetti by showing in detail how statistics can be developed from this point of view. In the title of his book and throughout the book itself, Lad speaks of ‘operational subjective statistical methods’, which emphasises the point that subjective probability is based on operationalism. Lad writes: ‘An *operationally defined measurement* is a specified procedure of action which, when followed, yields a number.’ (1996:39). It is clear that the measurement of degrees of belief by betting quotients as just described is an operationally defined measurement in this sense. We shall return to this connection between subjective probability and operationalism from time to time in what follows.

Let us now examine a second problem which arises in the subjective approach. If the subjective theory is to provide an interpretation of the standard mathematical theory of probability, then these degrees of belief (or betting quotients) ought to satisfy the standard axioms of probability. But why should they do so? It seems easy to imagine an individual whose degrees of belief are quite arbitrary and do not satisfy any of the axioms of probability. The subjectivists solve this problem and derive the axioms of probability by using the concept of *coherence.* I will next define this concept and then comment on its significance.

Coherence

If Mr B has to bet on a number of events E_1, ..., E_n, his betting quotients are said to be *coherent* if and only if Ms A cannot choose stakes S_1, ..., S_n such that she wins whatever happens. If Ms A can choose stakes so that she wins whatever happens, she is said to have made a *Dutch book* against Mr B.

It is taken as obvious that Mr B will want his bets to be coherent, that is to say he will want to avoid the possibility of his losing whatever happens. Surprisingly, this condition is both necessary and sufficient for betting quotients to satisfy the axioms of probability. This is the content of the following theorem.

The Ramsey–De Finetti theorem

A set of betting quotients is coherent if and only if they satisfy the axioms of probability.

So far we have made a contrast between the logical theory, in which probability is degree of rational belief, and the subjective theory, in which probability is degree of belief. The concept of coherence shows that this needs a little qualification, since coherence is after all a rationality constraint, and degrees of belief in the subjective approach must be rational, at least to the extent of satisfying this constraint. De Finetti expresses this very well in the title of his 1937 paper 'Foresight: Its Logical Laws, Its Subjective Sources'. The logical laws here come from the condition of coherence. Naturally, coherence does not determine a single degree of rational belief but leaves open a wide range of choices. Thus some subjective sources for probability are also needed.

Ramsey uses the term 'consistency' for coherence, and writes that: '... the laws of probability are laws of consistency' (1926:182). The idea here is that we have to make sure that our various degrees of belief fit together and so avoid the 'contradiction' of having a Dutch book made against us. The term 'coherence' is now generally preferred, because consistency has a well-defined but different meaning in deductive logic. Even though there is an analogy, it seems better to use different terms. I will now give a detailed proof of the Ramsey–De Finetti theorem. First I will state the axioms of probability and then prove the theorem for each of them in turn.

The axioms of probability

Let E, F, ..., E_1, ... stand for events, concerning which we can have some degree of belief whether they will occur, or have occurred. Let Ω denote the certain event, which must occur. There are then three axioms of probability.

1 $0 \leq P(E) \leq 1$ for any E, and $P(\Omega) = 1$.

2 (Addition Law) If $E_1, ..., E_n$ are events which are exclusive (i.e. no two can both occur) and exhaustive (i.e. at least one must occur), then

$$P(E_1) + ... + P(E_n) = 1$$

3 (Multiplication Law) For any two events E, F

$$P(E \,\&\, F) = P(E \mid F)\, P(F)$$

The Addition Law can be stated in a different but equivalent form. For any event E, F, let E v F be the event that either E occurs or F occurs or both occur. Then we have

2' (Alternative form of the Addition Law) If E, F are any two exclusive events, then

$$P(E) + P(F) = P(E \vee F)$$

We can prove the equivalence of 2 and 2' as follows:

(a) (2 → 2') Let E, F be exclusive events, and let $\Omega \setminus (E \vee F)$ be the event that something other than E or F occurs. E, F, $\Omega \setminus (E \vee F)$ are exclusive and exhaustive events. So by Axiom 2

$$P(E) + P(F) + P(\Omega \setminus (E \vee F)) = 1$$

But E v F, $\Omega \setminus (E \vee F)$ are also exclusive and exhaustive events. So by Axiom 2

$$P(E \vee F) + P(\Omega \setminus (E \vee F)) = 1$$

Thus subtracting, we get

$P(E) + P(F) = P(E \vee F)$ i.e. Axiom 2'

(b) (2' → 2) We first prove by induction that Axiom 2' holds for any n exclusive events. The case $n = 2$ is just Axiom 2' itself. Suppose the result holds for $n - 1$, i.e. if $E_1, ..., E_{n-1}$ are any exclusive events, then

$$P(E_1) + ... + P(E_{n-1}) = P(E_1 \vee ... \vee E_{n-1})$$

Now consider n exclusive events $E_1, ..., E_n$. The events $(E_1 \vee ... \vee E_{n-1})$, E_n are also exclusive. So by Axiom 2'

$$P(E_1 \vee ... \vee E_{n-1}) + P(E_n) = P(E_1 \vee ... \vee E_n)$$

But since $E_1, ..., E_{n-1}$ are exclusive events, it follows that

$$P(E_1) + ... + P(E_n) = P(E_1 \vee ... \vee E_n)$$

But if $E_1, ..., E_n$ are exhaustive as well as exclusive, $E_1 \vee ... \vee E_n$ is the certain event with probability 1, and so Axiom 2 follows.

Proof of the Ramsey–De Finetti theorem [4]

Proof for Axiom 1

(a) Coherence → Axiom 1: Let us first consider the case of the certain event O. If Mr B chooses $q(\Omega) > 1$, Ms A can win by choosing $S > 0$. If Mr B chooses

$q(\Omega) < 1$, Ms A can win by choosing $S < 0$. Hence to be coherent, Mr B must choose $q(\Omega) = 1$. Now take any arbitrary event E. If Mr B chooses $q(E) > 1$, Ms A can win by choosing $S > 0$. If Mr B chooses $q(E) < 0$, Ms A can win by choosing $S < 0$. Hence to be coherent, Mr B must choose $0 = q(E) = 1$.

(b) Axiom 1 → coherence: If Mr B chooses $q(\Omega) = 1$, there is no way that Ms A can win, since the stake, whatever its sign, is simply passed from one to the other and then back again. For an arbitrary event E, Ms A cannot choose the sign or size of S so that she always wins if Mr B chooses $0 = q(E) = 1$.

Proof for Axiom 2

(a) Coherence → Axiom 2: Suppose Mr B chooses betting quotients $q_1, ..., q_n$, and Ms A chooses stakes $S_1, ..., S_n$. Then, if event E_i occurs, Ms A's gain G_i is given by

$$G_i = q_1S_1 + ... + q_nS_n - S_i \tag{4.1}$$

So if Ms A sets $S_1 = S_2 = ... = S_n = S$, then

$$G_i = S(q_1 + ... + q_n - 1)$$

Thus, if Mr B chooses $q_1 + ... + q_n > 1$, then Ms A can always win by setting $S > 0$. If Mr B chooses $q_1 + ... + q_n < 1$, then Ms A can always win by setting $S < 0$. Hence, to be coherent, Mr B must choose $q_1 + ... + q_n = 1$.

(b) Axiom 2 → coherence: Since Axiom 2 holds, we have $q_1 + ... + qn = 1$. Now by Equation 4.1 above, we have

$$q_iG_i = q_i(q_1S_1 + ... + q_nS_n) - q_iS_i$$

So summing over i, we get

$$q_1G_1 + q_2G_2 + ... + q_nG_n = 0 \tag{4.2}$$

Equation 4.2 shows that the Gi cannot all be positive for the following reason. The $q_i = 0$, and, since they sum to 1, at least one of them must be > 0. Hence if all the G_i were > 0, $q_1G_1 + ... + q_nG_n > 0$, which contradicts Equation 4.2. Hence, not all the Gi can be positive, which is equivalent to saying that the betting quotients are coherent. The consideration of $q_1G_1 + q_2G_2 + ... + q_nG_n$ may look like a mathematical trick, but in fact it has a simple intuitive meaning.[5] It is just Ms A's expected gain relative to the probabilities chosen by Mr B. If this expected gain is zero, Ms A cannot make a Dutch book against Mr B.

To prove the Ramsey–De Finetti theorem for Axiom 3, we need the following definition.

Definition of conditional betting quotient

$q(\mathrm{E} \mid \mathrm{F})$, the conditional betting quotient for E given F, is the betting quotient which Mr B would give for E on the understanding that the bet is called off and all stakes returned if F does not occur.

Ramsey remarks that 'Such conditional bets were often made in the eighteenth century.' (1926:180).

Proof for Axiom 3

In all parts of the proof, we shall use the following notation

$$q = q(\mathrm{E}\ \&\ \mathrm{F})$$
$$q' = q(\mathrm{E} \mid \mathrm{F})$$
$$q'' = q(\mathrm{F})$$

(a) Coherence → Axiom 3, using determinants: Suppose Mr B chooses betting quotients q, q', q'' as above, and Ms A chooses corresponding stakes S, S', S''. Three possible cases can occur, and we shall calculate Ms A's gain in each case.

1 E and F both occur

$$G_1 = (q - 1)\,S + (q' - 1)S' + (q'' - 1)\,S''$$

2 E does not occur, but F occurs

$$G_2 = qS + q'S' + (q'' - 1)\,S''$$

3 F does not occur

$$G_3 = qS + \quad + q''S''$$

For fixed G_1, G_2, $G_3 > 0$, these are three linear equations in three unknowns, S, S', S''. Thus, they always have a solution, unless the determinant vanishes. So, for coherence, we must have

$$\begin{vmatrix} q-1 & q'-1 & q''-1 \\ q & q' & q''-1 \\ q & 0 & q'' \end{vmatrix} = 0$$

Subtracting the bottom row from the top two rows, and then the middle row from the top row gives

$$\begin{vmatrix} -1 & -1 & 0 \\ 0 & q' & -1 \\ q & 0 & q'' \end{vmatrix} = 0$$

Then expanding by the first row, we get

$$- q'q'' + q = 0$$
So $q = q'q''$ as required.

For those unfamiliar with the theory of determinants, the following gives a proof of the same result without using determinants.

(b) Coherence $\rightarrow$ Axiom 3, without using determinants: Suppose Ms A chooses $S = +1$, $S' = -1$, $S'' = -q'$, we then have

$$G_1 = (q - 1) + (1 - q') + q' - q'q'' = q - q'q''$$
$$G_2 = q - q' - q'q'' + q' = q - q'q''$$
$$G_3 = q - q'q''$$

So all Ms A's gains are positive, unless $q \leq q'q''$.

Similarly, if Ms A chooses $S = -1$, $S' = +1$, $S'' = q'$, all her gains are positive unless $q \geq q'q''$. So, to be coherent, Mr B must choose $q = q'q''$, as required.

(c) Axiom 3 $\rightarrow$ coherence: We have to show that if $q = q'q''$, the betting quotients are coherent, i.e. Ms A's gains G_1, G_2, G_3 cannot all be positive. Using the method employed for Axiom 2, we need to consider Ms A's expected gain given the probabilities chosen by Mr B, and then show that it is zero. Ms A's expected gain is in fact $\lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3$ where

$$\lambda_1 = q'q'',\ \lambda_2 = (1 - q')q'',\ \lambda_3 = 1 - q''.\ \text{Since } 0 \leq q', q'' \leq 1, \text{ each } \lambda_i \geq 0.$$

Now

$$\lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3 = \alpha S + \beta S' + \gamma S'',$$

where

$$\alpha = q'q''(q - 1) + (1 - q')q''q + (1 - q'')\, q$$
$$= q''(q'q - q' + q - qq' + (1 - q'')q'), \text{ since } q = q'q''$$
$$= q''(q'q - q' + q'q'' - qq' + q' - q'q'')$$
$$= 0$$

$$\beta = q'q''(q' - 1) + (1 - q')q''q' = 0$$

$$\gamma = q'q''(q'' - 1) + (1 - q')q''(q'' - 1) + (1 - q'')q'' = 0$$

Hence $\lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3 = 0$.

But now at least one of the $\lambda_i > 0$, for either $q'' \neq 1$, when $\lambda_3 > 0$, or $q'' = 1$, when $\lambda_1 = q'$, $\lambda_2 = 1 - q'$. In this case, either $q' \neq 1$, when $\lambda_2 > 0$, or $q' = 1$, when $\lambda_1 > 0$. It follows that not all the G_i can be positive, and so Mr B's betting quotients are coherent, as required.

The Ramsey–De Finetti theorem is a remarkable achievement, and clearly demonstrates the superiority of the subjective to the logical theory. Whereas in the logical theory the axioms of probability could only be justified by a vague and unsatisfactory appeal to intuition, in the subjective theory they can be proved rigorously from the eminently plausible condition of coherence. Indeed, given the Ramsey–De Finetti theorem, it is difficult to deny that the subjective theory provides a valid interpretation of the mathematical calculus of probability – though it is of course possible to hold that there are other valid interpretations of this calculus. In addition, the subjective theory solves the paradoxes of the Principle of Indifference by, in effect, making this principle unnecessary, or at most a heuristic device. In the logical theory, the principle was necessary to obtain the supposedly unique a priori degrees of rational belief, but, according to the subjective theory, there are no unique a priori probabilities. Different individuals can choose their a priori probabilities in different ways, and, provided they are coherent, there need be nothing wrong with these different choices. Thus, if the Principle of Indifference is used as a heuristic device, and suggests two different possibilities for the a priori probabilities, there is no contradiction. Mr B might choose one of these possibilities as his subjective valuation, and Ms D might choose the other. Ramsey is well aware of the superiority of the subjective to the logical theory in these respects and states them as follows:

> In the first place it gives us a clear justification for the axioms of the calculus, which on such a system as Mr Keynes' is entirely wanting. For now it is easily seen that if partial beliefs are consistent they will obey these axioms, but it is utterly obscure why Mr Keynes' mysterious logical relations should obey them. We should be so curiously ignorant of the instances of these relations, and so curiously knowledgeable about their general laws.
>
> Secondly, the Principle of Indifference can now be altogether dispensed with; ... To be able to turn the Principle of Indifference out of formal logic is a great advantage; for it is fairly clearly impossible to lay down purely logical conditions for its validity, as is attempted by Mr Keynes.
>
> (Ramsey 1926:188–9)

There remain, however, some problems connected with the subjective theory, and in particular the question of how probabilities which appear to be objective, such as the probability of a particular isotope of uranium disintegrating in a year, can be explained on this approach. De Finetti tackles this problem by introducing the concept of *exchangeability,* and I will give an account of this below (pp. 69– 83). Before going on to this, however, there is a matter which may well be of interest to mathematicians. Nearly all advanced treatments of mathematical theory of probability are today based on the Kolmogorov axioms (see Kolmogorov 1933). Now the axioms given above are of course similar to the Kolmogorov axioms, but do nonetheless differ on one or two points. It certainly seems worth examining these divergences from standard mathematical practice to see what significance they have. In general, in this book my aim is to discuss the philosophical side of probability using as little mathematics as possible, indeed no more than quite elementary algebra. Sometimes, as here, however, it will be useful to discuss issues which require a knowledge of more advanced mathematical approaches to probability (random variables, measure theory, analysis, etc.). My plan is to place such discussions in sections marked with an asterisk and to arrange them so that they can be read by mathematicians but omitted by non-mathematicians without losing the general thread of the argument.

A comparison of the axiom system given here with the Kolmogorov axioms*

De Finetti assigns probabilities to events E, F, ..., including the certain event which we have denoted by Ω. In Kolmogorov's mathematical approach, probabilities are assigned to the subsets of a set Ω. This difference does not seem to me an important one, since it would be fairly easy to map De Finetti's treatment into set-theoretic language. A more significant divergence comes with the treatment of conditional probabilities. Kolmogorov introduces these by definition (see Kolmogorov 1933:6), so that

$$P(E \mid F) = \text{def}\, \frac{P(E \,\&\, F)}{P(F)} \quad \text{for } P(F) \neq 0$$

The case P(F) = 0 is dealt with by Kolmogorov later in his monograph (1933:Chapter V). Thus, in Kolmogorov's treatment an equality is established by definition which in the treatment we have just given is a substantial axiom (Axiom 3) requiring an elaborate proof, and is indeed the multiplication law of probability.

In fact, this is not the only instance in mathematics where a substantial assumption appears in the form of a definition, but the practice does not seem to me a good one. I would argue that it is better to state important assumptions as axioms (or derive them as theorems) and try to keep definitions as far as possible as mere abbreviations. This inclines me to prefer De Finetti's treatment to Kolmogorov's on this point. This would amount to taking P(E | F) as a primitive (undefined) term in the axiom system and characterising it by an axiom, rather than introducing it by an explicit definition.

It is clear that De Finetti's approach is more natural for the subjective theory, since conditional probabilities can be introduced as conditional betting quotients defined within a particular betting scheme. It is then by no means obvious that these conditional betting quotients obey our Axiom 3; indeed the proof is quite long. Moreover, similar considerations apply in the other interpretations of probability. We have seen in Chapter 3 that the notion of the conditional probability of h given e is a primitive and fundamental notion within the logical theory. It thus seems natural to take it as a primitive notion in an axiom system, as Keynes does. As we shall see in Chapters 5 and 6, the notion of conditional probability is also primitive in the frequency and propensity interpretations. On this point I side with De Finetti rather than Kolmogorov, and I favour the introduction of conditional probabilities by an axiom rather than a definition. This, moreover, leads to a rather elegant symmetry in the axiomatic treatment between the addition and multiplication laws of probability.

The next important difference between De Finetti and Kolmogorov concerns the issue of finite versus countable additivity. De Finetti's Axiom 2 (the Addition Law) can, as we have seen, be stated in the equivalent form: if $E_1, ..., E_n$ are events which are exclusive,

$$P(E_1 \vee ... \vee E_n) = P(E_1) + ... + P(E_n).$$

The question now arises whether we can extend the Addition Law from the finite case to the countably infinite case, that is to say whether we can legitimately go from finite additivity to countable additivity. This would involve adopting as an axiom the following stronger form of the Addition Law.

> Addition law for countable additivity: If $E_1, ..., E_n, ...$ is a countably infinite sequence of exclusive events, then
>
> $$P(E_1 \vee ... \vee E_n \vee ...) = P(E_1) + ... + P(E_n) + ...$$

Kolmogorov's treatment of this question is interesting. In the first chapter of his monograph he allows only finite additivity. Then in the second chapter he adds to his five previous axioms a sixth axiom (the axiom of continuity) which is equivalent to the Addition Law for countable additivity as just stated. Kolmogorov does, however, appear to have some reservations about his axiom, for he says:

> Since the new axiom is essential for infinite fields of probability only, it is almost impossible to elucidate its empirical meaning, as has been done, for example, in the case of Axioms I – V in §2 of the first chapter. For, in describing any observable random process we can obtain only finite fields of probability. Infinite fields of probability occur only as idealised models of real random processes. *We limit ourselves, arbitrarily, to only those models which satisfy Axiom VI.* This limitation has been found expedient in researches of the most diverse sort.

Kolmogorov here argues that countable additivity goes beyond what can be checked empirically, but that its adoption is nonetheless justified because of its usefulness in a whole range of research.

De Finetti shares Kolmogorov's doubts about countable additivity, but he regards them as a reason for limiting oneself to finite additivity.[6] Thus he says that:

> [The assumption of countable additivity] is the one most commonly accepted at present; it had, if not its origin, its systematization in Kolmogorov's axioms (1933). Its success owes much to the mathematical convenience of making the calculus of probability merely a translation of modern measure theory.... No-one has given a real justification of countable additivity (other than just taking it as a 'natural extension' of finite additivity).
>
> (1970:vol. 1, 119)

De Finetti, however, thinks that one should not introduce new axioms simply on the grounds of mathematical convenience, unless these axioms can be justified in terms of the meaning of probability. Now in the subjective theory, probabilities are given by an individual's betting quotients. A given individual will always bet on a finite number of events, and it is difficult to imagine bets on an infinite number of events. Thus the subjective theory would seem to justify finite, but not countable, additivity. De Finetti gives a number of other arguments in favour of finite additivity and against countable additivity. We shall here consider one more of these.

If we adopt countable additivity, then it becomes impossible to have a uniform distribution over a countable set, such as the positive integers $\{1, 2, ..., n, ...\}$. For suppose we put $P(i) = p$ for all i. If $p > 0$, then $P(1) + P(2) + ... + P(n) + ...$ becomes infinite, whereas by the axioms of probability it should be $= 1$. If we put $P(i) = 0$ for all i, then by countable additivity $P(\{1, 2, ... , n, ...\}) = P(1) + P(2) + ... + P(n) + ... = 0$, whereas, by Axiom 1, $P(\{1, 2, ..., n, ...\}) = P(\Omega) = 1$. However, if we adopt only finite additivity, then the second half of the argument is blocked, so that it becomes possible to have a uniform distribution over the positive integers. De Finetti regards it as a counterintuitive feature of the axiom of countable additivity that it prevents us from having such uniform distributions. After all, for any finite n, however large, we can introduce a uniform distribution over the positive integers $1, 2, ..., n$ by setting $P(i) = 1/n$, $i = 1, ..., n$. However, if we postulate countable additivity over the infinite collection of positive integers $1, 2, ..., n, ...$, we can only have what he terms 'extremely unbalanced partitions' (1970:Vol. 1, 122). He explains his meaning here more fully later on when he says that countable additivity: 'forces me to choose some finite subset of them [i.e. the countable class in question, e.g. the positive integers] to which I attribute a total probability of at least 99% (leaving 1% for the remainder; and I could have said 99.999% with 0.001% remaining, or something even more extreme).' (1970:Vol. 2, 351) This argument does not perhaps go very well with the previous argument which suggests that on the subjective approach one should always limit oneself to finite collections of events and not consider probability distributions over countable sets at all.

Not all probabilists agree with De Finetti's attitude to countable additivity within the subjective theory. Adams (1964) presented a proof that countable additivity does follow from the assumptions of the subjective approach. This proof has been considerably simplified by Williamson (1999), which also discusses the philosophical problems involved. Williamson devises a betting situation in which it would seem quite reasonable to bet on a countable number of events. Suppose Ms A tells Mr B that in a sealed parcel in the next room there is the computer print-out of a positive integer, and asks him to give a betting quotient on this number being n for all n. Now of course Mr B would realise that the practicalities of technology must impose some upper bound on the value which the hidden number could take. However, this upper bound is hard to determine, and the problem is a very open-ended one. Rather than fix on a particular upper bound, it would be easier for Mr B to produce an infinite sequence of betting quotients. Actually, the infinite is often brought into applied mathematics for exactly this kind of reason.

A noteworthy feature of this example is that a uniform distribution is highly implausible. On the contrary, we would expect small numbers to be more probable than very large ones. In general, in any betting situation in which we approximate the large open-ended finite by the infinite, the unbalanced distributions described by De Finetti, far from being counterintuitive, are just what we would expect.

Williamson's other point is that, once we have introduced a betting scheme for a countably infinite number of events, it only requires one extra condition to derive the axiom of countable additivity by exactly the same Dutch book argument which De Finetti uses for finite additivity. This extra condition is that only a finite amount of money should change hands. Assuming this, let us see how the proof of Axiom 2 must be modified if we have, instead of a finite number of events $E_1, ..., E_n$, a countably infinite number $E_1, ..., E_n, ...$. Because only a finite amount of money should change hands, Ms A's gains G_i must all be finite, which means in turn that the series $q_1S_1 + ... + q_nS_n + ...$ must converge. Moreover, from Axiom 1, it follows that $q_1 + ... + q_n + ... \leq 1$. If in the proof of Axiom 2 given above, we replace the finite sums by infinite series, then, using the above results, all the series converge, and the proof goes through just as before. So, if we allow bets over a countable infinity of events (as seems eminently reasonable in the kind of situation described above), and if we specify that only a finite amount of money should change hands (which can hardly be avoided), then the axiom of countable additivity does follow rigorously from exactly the same Dutch book argument which De Finetti uses to establish finite additivity. This argument of Williamson's seems to me to show that countable additivity is completely justified within the subjective theory, and that De Finetti was wrong to deny it.

This result seems to me to strengthen rather than weaken the subjective theory. On De Finetti's approach, mathematicians who adopted the subjective theory of probability would have to use a mathematical theory somewhat different from the standard one. Many would surely regard this as an argument against becoming a subjectivist. Williamson's argument shows that such doubts are quite unnecessary, and that it is perfectly possible both to be a subjectivist and to use the standard mathematical theory. Moreover, as Williamson points out, countable additivity strengthens the subjective theory as against the logical theory. Suppose we were betting on a countably infinite sequence of events E_1, E_2,..., E_n, ..., and suppose we had no reason to prefer E_i to E_j for all *i, j,* then the logical theory with its Principle of Indifference would seem to require a uniform distribution. Countable additivity forces a skew distribution on us, thus preventing a logical interpretation and introducing a subjective element. So, ironically, De Finetti's defence of a uniform distribution in this context is more of a defence of the logical view than of his own subjective approach.

Apparently objective probabilities in the subjective theory: exchangeability

So far the subjective theory has had considerable success. Starting from the analysis of probability as the degree of belief of an individual, it has shown how such degrees of belief can be measured, and how from the simple and plausible condition of coherence the standard mathematical axioms of probability can be derived. All this establishes beyond doubt that subjective probabilities are at least one of the valid interpretations of the mathematical calculus. Moreover, there are a number of situations where the subjective analysis of probability looks highly plausible. Examples would be the probability of it raining tomorrow, the probability that a particular party will win the next election or the probability of a particular horse winning a race. Such probabilities can plausibly be said to be subjective, or at least to involve a considerable subjective component. Yet there are other probabilities which do seem, at first sight at least, to be completely objective. Suppose we have a die which is shown by careful tests to be perfectly balanced mechanically, and which in a series of trials has given approximately the same frequency for each of its faces. Surely for such a die $P(5) = {}^1/_6$, and this is an objective fact, not a matter of subjective opinion. Then again consider the probability of a particular isotope of uranium disintegrating in a year. This is surely not a matter of opinion, but something which can be calculated from quantities specified in textbooks of physics. Such a probability looks every bit as objective as, for example, the mass of the isotope. How is a supporter of the subjective theory of probability to deal with cases of this sort?

Actually there are two possible approaches. First of all, it could be admitted that the examples we have cited, and others like them, are indeed objective, and consequently that there are at least two different concepts of probability which apply in different circumstances. This was the position which Ramsey (1926) adopted, and I will discuss it in Chapter 8. Second, however, it could be claimed that all probabilities are subjective, and that even apparently objective probabilities, such as the ones just described, can be explicated in terms of degree of subjective belief. This was the line adopted by De Finetti, and I will next consider his argument in detail.

De Finetti states the problem as follows:

> It would not be difficult to admit that the subjectivistic explication is the only one applicable in the case of practical predictions (sporting results, meteorological facts, political events, etc.) which are not ordinarily placed in the framework of the theory of probability, even in its broadest interpretation. On the other hand it will be more difficult to agree that this same explanation actually supplies rationale for the more scientific and profound value that is attributed to the notion of probability in certain classical domains, ...
>
> (1937:152)

Nonetheless, De Finetti does think that the subjective account of probability is adequate even in these 'classical domains', for he continues:

> Our point of view remains in all cases the same: *to show that there are rather profound psychological reasons which make the exact or approximate agreement that is observed between the opinions of different individuals very natural, but that there are no reasons, rational, positive, or metaphysical, that can give this fact any meaning beyond that of a simple agreement of subjective opinions.*
>
> (1937:152)

Let us now see how De Finetti works out this view by taking a simple example. Suppose we have a coin which is known to be biased, but for which the extent of the bias is not known. An objectivist would say that there is a true, but unknown, probability p of heads, and that we can measure p roughly by making n tosses (for large n), observing the number r of heads and setting $p \approx r/n$. The exact relation between p and r/n will depend on the particular objective theory adopted.

How then does a subjectivist like De Finetti deal with this case? The first step is to consider a sequence of tosses of the coin which we suppose gives results: E_1, ..., E_n, ..., where each E_i is either heads (H_i) or tails (T_i). So, in particular, H_{n+1} = Heads occurs on the n + 1th toss. Further, let e be a complete specification of the results of the first n tosses, that is a sequence n places long, at the ith place of which we have either H_i or T_i. Suppose that heads occurs r times on the first n tosses. The subjectivist's method is to calculate $P(H_{n+1} \mid e)$, and to show that under some general conditions which will be specified later $P(H_{n+1} \mid e)$ tends to r/n for large n. This shows that whatever value is assigned to the prior probability $P(H_{n+1})$, the posterior probability $P(H_{n+1} \mid e)$ will tend to the observed frequency for large n. Thus, different individuals who may hold widely differing opinions initially will, if they change their probabilities by Bayesian conditionalisation, come to agree on their posterior probabilities. The objectivist wrongly interprets this as showing that there is an objective probability, but, according to De Finetti, 'objective probability' is a metaphysical concept devoid of meaning. All that is happening is that, in the light of evidence, different individuals are coming to agree on their subjective probabilities. Such is the argument. Let us now give, in our simple case, the mathematical proof which underpins it.

Suppose that $P(E_i) \neq 0$ for all *i*, so that also $P(e) \neq 0$. We then have by Axiom 3

$$P(H_{n+1} \mid e) = \frac{P(H_{n+1} \,\&\, e)}{P(e)} \tag{4.3}$$

To proceed further we introduce the condition of *exchangeability.* Suppose Mr B is making an a priori bet that a particular *n*-tuple of results ($E_{i1}\ E_{i2}\ \ldots\ E_{in}$ say) occurs. Suppose further that heads occurs *r* times in this *n*-tuple. Mr B's betting quotients are said to be *exchangeable* if he assigns the same betting quotient to any other particular *n*-tuple of results in which heads occurs *r* times, where both *n* and *r* can be chosen to have any finite integral non-negative values with $r \leq n$. Let us write his prior probability (or betting quotient) that there will be *r* heads in *n* tosses as $\omega_r^{(n)}$. There are ${}^nC_r$ different ways in which *r* heads can occur in *n* tosses, where, as usual, ${}^nC_r = n!/(n - r)!\ r! = n(n - 1) \ldots (n - r + 1)/r(r - 1) \ldots 1$. Each of the corresponding *n*-tuples must, by exchangeability, be assigned the same probability, which is therefore $\omega_r^{(n)}/{}^nC_r$. Thus

$$P(E_{i1}E_{i2}\ldots E_{in}) = \frac{\omega_r^n}{{}^nC_r} \tag{4.4}$$

Now e, by definition, is just a particular *n*-tuple of results in which heads occurs *r* times. Thus, by exchangeability,

$$P(e) = P(E_{i1}E_{i2}\ldots E_{in}) = \frac{\omega_r^n}{{}^nC_r} \tag{4.5}$$

Now H_{n+1} & e is an $(n + 1)$-tuple of results in which heads occurs $r + 1$ times. Thus, by the same argument,

$$P(H_{n+1} \,\&\, e) = \frac{\omega_{r+1}^{(n+1)}}{{}^{n+1}C_{r+1}} \tag{4.6}$$

And so, substituting in Equation 4.3, we get

$$\begin{aligned} P(H_{n+1} \mid e) &= \frac{{}^nC_r}{{}^{n+1}C_{r+1}} \frac{\omega_{r+1}^{(n+1)}}{\omega_r^{(n)}} \\ &= \frac{n!}{(n-r)!\,r!} \frac{(r+1)!\,(n-r)!}{(n+1)!} \frac{\omega_{r+1}^{(n+1)}}{\omega_r^{(n)}} \\ &= \frac{r+1}{n+1} \frac{\omega_{r+1}^{(n+1)}}{\omega_r^{(n)}} \end{aligned} \tag{4.7}$$

Equation 4.7 gives us the result we want. Provided only $\omega_{r+1}^{(n+1)}/\omega_r^{(n)} \rightarrow 1$ as $n \rightarrow \infty$ (a very plausible requirement), we may choose our prior probabilities $?_r^{(n)}$ in any way we please, and still get that as $n \rightarrow \infty$, $P(H_{n+1} \mid e) \rightarrow r/n$ (the observed frequency), as required.

To sum up then: according to the objectivist, there is a real objective probability p of heads, and the observed frequency r/n gives an increasingly better estimate of p as $n \rightarrow \infty$.

According to the subjectivist, the 'real objective probability p' is a metaphysical delusion. Different people may, subject only to coherence, have different prior probabilities $P(H_{n+1})$. However, coherence + exchangeability + one other plausible assumption ($\omega_{r+1}^{(n+1)}/\omega_r^{(n)} \rightarrow 1$, as $n \rightarrow \infty$) ensure that $P(H_{n+1} \mid e) \rightarrow r/n$ as $n \rightarrow \infty$. Thus, as the evidence piles up, the people who disagree a priori will come to agree a posteriori. This 'exact or approximate agreement between the opinions of different individuals for rather profound psychological reasons' is what gives rise to the illusion of objective probabilities.

In n tosses, we can have either 0, 1, 2, ..., or n heads. So, by coherence,

$$\omega_0^{(n)} + \omega_1^{(n)} + \omega_2^{(n)} + \ldots + \omega_r^{(n)} + \ldots + \omega_n^{(n)} = 1 \tag{4.8}$$

In the subjective theory, we can choose the $\omega_r^{(n)}$ (the prior probabilities) in any way we choose subject only to Equation 4.8. However, we can also, though this is not compulsory, make the 'Principle of Indifference' choice of making them all equal so that

$$\omega_o^{(n)} = \omega_1^{(n)} = \omega_2^{(n)} = \ldots = \omega_{\mathrm{r}}^{(n)} = \ldots = \omega_n^{(n)} = 1/(n+1) \tag{4.9}$$

Substituting this in Equation 4.7, we get

$$P(H_{n+1} \mid e) = \frac{r+1}{n+2} \tag{4.10}$$

This is a classical result – Laplace's Rule of Succession.

The Rule of Succession has been used to try to solve Hume's problem of induction. Suppose, having read Hume, we are worried about whether the Sun will rise tomorrow. Now recorded history goes back at least 5,000 years, and the Sun has been observed (in the appropriate latitudes) to rise every single morning during all that time. At least, if the Sun had failed to rise one morning, it is a reasonable presumption that this fact would have been recorded. So our evidence is that the Sun has risen each morning for 1,826,250 days. To calculate the probability of its rising tomorrow, we use Equation 4.10 with $r = n = 1{,}826{,}250$. This gives the probability of the Sun's rising tomorrow as approximately 0.9999994. If this reasoning is correct, then we should no longer be troubled by Humean doubts, but should be able to look forward with very great confidence to the Sun rising tomorrow!

But not everyone is convinced by the argument, and the Rule of Succession has been subjected to quite a number of harsh criticisms. I will here describe one based on an example due to Popper.[7] Suppose the inhabitants of London wake up one summer morning to find that although according to their clocks it should be day, it is in fact still night outside. They switch on their radios and televisions and learn that something quite extraordinary has happened. The Earth appears to have stopped rotating. It is still night in London, while on the opposite side of the globe, the Sun is staying fixed at one position in the sky. Of course this quite contradicts all the known laws of physics. Moreover, apart from the strange change in the apparent movements of the Sun, everything else seems to be continuing just as before, a situation which again contradicts all the known laws of physics. Scientists the world over confess that they are baffled and cannot understand what is happening. Copies of the philosophical works of Hume are selling well.

Given this bizarre, but at least imaginable, situation, what would be the probability of the Sun's rising again as usual the next morning? It is easy to calculate according to the Rule of Succession. In Equation 4.10 above, we now have $r = n - 1$, and $n = 1{,}826{,}251$. So the probability of the Sun's rising the next day is 0.9999989. In other words, if we stick to the Rule of Succession, the quite extraordinary events just described would reduce the probability of the Sun's rising the next day by 0.0000005, i.e. 5×10^{-7}. Obviously this is quite wrong. There would be such a state of confusion that no one would have the least idea of whether the Sun would rise the next day or not. Certainly no one would assign a probability of 0.9999989 to its doing so. This example shows that, although the Rule of Succession may give reasonable answers in some cases, it gives absurd answers in others and so cannot be considered valid in general. On the other hand, it is not clear what exactly is wrong with the rather convincing chain of reasoning which was presented above and which led to the Rule of Succession. Rather than pursue this problem immediately, I will first present a general criticism of De Finetti's analysis of apparent objectivity in terms of exchangeability. This criticism casts light on why the Rule of Succession fails so dramatically in some cases, as I will then show.

To explain my criticism of De Finetti's exchangeability argument, I will begin by quoting an important passage in which he describes some general features of the argument. It is precisely these features which I will then criticise. The passage runs as follows:

> Whatever be the influence of observation on predictions of the future, it never implies and never signifies that we *correct* the primitive evaluation of the probability $\mathbf{P}(E_{n+1})$ after it has been *disproved* by experience and substitute for it another $\mathbf{P}^*(E_{n+1})$ which *conforms* to that experience and is therefore probably *closer to the real probability*; on the contrary, it manifests itself solely in the sense that when experience teaches us the result A on the first n trials, our judgment will be expressed by the probability $\mathbf{P}(E_{n+1})$ no longer, but by the probability $\mathbf{P}(E_{n+1} \mid A)$, i.e. that which our initial opinion would already attribute to the event E_{n+1} considered as conditioned on the outcome A. Nothing of this initial opinion is repudiated or corrected; it is not the

> function **P** which has been modified (replaced by another P*), but rather the argument E_{n+1} which has been replaced by $E_{n+1} | A$, and this is just to remain faithful to our original opinion (as manifested in the choice of the function P) and coherent in our judgment that our predictions vary when a change takes place in the known circumstances.
>
> In the same way, someone who has the number 2374 in a lottery with 10,000 tickets will attribute at first a probability of 1/10,000 to winning the first prize, but will evaluate the probability successively as 1/1000, 1/100, 1/10, 0, when he witnesses the extraction of the successive chips which give, for example, the number 2379. At each instant his judgment is perfectly coherent, and he has no reason to say at each drawing that the preceding evaluation of probability was not right (at the time when it was made).
>
> (De Finetti 1937:146–7)

This passage puts very clearly the difference between De Finetti's position and that of an objectivist – particularly an objectivist with Popperian tendencies. For such an objectivist, any evaluation ***P*** of a probability function is just a conjecture as to the values of the real objective probabilities, and, like any conjecture it should be severely tested. If these tests show that it is inadequate in anyway, it should be replaced by a new conjecture ***P****** which fits the facts better. In De Finetti's scheme, we do not try to test or refute our prior probabilities $\mathbf{P}(E_{n+1})$, we simply change them into posterior probabilities P(En + 1 | A) by Bayesian conditionalisation. Different people may start with different prior probabilities, but, as the evidence mounts up, their posterior probabilities will tend in many circumstances to converge producing the illusion of the existence of an objective probability.

My argument against De Finetti can be stated in general terms as follows. The prior probability function ***P*** will in all cases be based on general assumptions about the nature of the situation under study. Now if these assumptions are broadly correct, then De Finetti's scheme of modifying ***P*** by Bayesian conditionalisation will yield reasonable results. If, however, the initial assumptions are seriously wrong in some respects, then not only will the prior probability function be inappropriate, but all the conditional probabilities generated from it in the light of evidence will also be inappropriate. To obtain reasonable probabilities in such circumstances, it will be necessary to change ***P*** in a much more drastic fashion than De Finetti allows, and, in effect, introduce a new probability function ***P****. This line of thought could be summarised as follows. De Finetti's scheme of allowing changes only by Bayesian conditionalisation is too conservative. Sometimes, in order to make progress, much more drastic changes in ***P*** are needed than those which he allows. I will give an example of such a situation in a moment. However, to explain the general character of this example, it will be desirable to examine the relation between the concepts of independence and exchangeability. As this involves some technicalities I will discuss the matter in the next section. I will then give an informal summary of the main points of this section before giving my example of a situation in which De Finetti's method of changing prior probabilities only by Bayesian conditionalisation proves to be inadequate.

The relation between independence and exchangeability*

In a certain sense the concept of exchangeability is the equivalent within the subjective theory of the objectivist's notion of independence. This does not mean that the concept of independence does not apply in the subjective theory. Two events E, F are defined to be independent, if P(E & F) = P(E) P(F). This definition can of course be applied when the probabilities involved are given a subjective meaning. The trouble is that while in objective approaches the assumption of independence is a very important one which applies in many cases, independence in the subjective sense turns out to be an assumption which can rarely, if ever, be made. If we make the mathematical assumption of independence, giving the probabilities an epistemological meaning, this turns out to give a case in which no learning from experience can occur. We can see this in the context of the subjective theory by exploring what happens if we change the assumption of exchangeability to that of independence. This amounts to assuming that

$$P(E_{i1} \ \& \ E_{i2} \ \& \ \ldots \ \& \ E_{in}) = P(E_{i1})\, P(E_{i2}) \ldots P(E_{in})$$

It follows in particular that $P(H_{n+1} \ \& \ e) = P(H_{n+1})\, P(e)$. Substituting this into Equation 4.3 above, we get

$$P(H_{n+1} \mid e) = P(H_{n+1})$$

So within the Bayesian framework no learning from experience can occur. De Finetti must have realised this very early on in his development of the subjective theory for he writes:

> If the outcome of the preceding trials can modify my opinion, it is for me *dependent* and *not* independent.... If I admit the possibility of modifying my probability judgment in response to observation of frequencies; it means that – by definition – my judgment of the probability of one trial is not independent of the outcomes of the others ....
>
> (1931a:212)

In general, an individual such as our Mr B will want to modify his probability judgements in response to observation of frequencies, and so it follows that the assumption of independence will rarely, if ever, be made within the subjective theory. At first sight this may seem rather a severe blow to the subjective approach, since objectivists frequently and successfully make assumptions of independence. This was no doubt one factor which stimulated De Finetti to invent his new concept of exchangeability. Roughly speaking where an objectivist assumes independence, a subjectivist will assume exchangeability. De Finetti proved a general theorem showing how the two concepts are linked, I will next state his result.

Let us first define exchangeability for a sequence of random variables (or random quantities as De Finetti prefers to call them) $X_1, \ldots, X_n, \ldots$. These are exchangeable if, for any fixed n, X_{i1}, X_{i2}, ..., X_{in} have the same joint distribution no matter how $i1$, ..., in are chosen. Now let Y_n be the average of any n of the random quantities X_i, i.e. $Y_n = (1/n)(X_{i1} + X_{i2} + \ldots + X_{in})$, since we are dealing with exchangeable random quantities it does not matter which $i1$, $i2$,..., in are chosen. De Finetti first shows (1937: 126) that the distribution $\Phi_n(\xi) = \mathrm{P}(Y_n \leq \xi)$ tends to a limit $\Phi(\xi)$ as $n \to \infty$, except perhaps for points of discontinuity. He goes on to say:

> Indeed, let $\mathbf{P}_\xi(\mathrm{E})$ be the probability attributed to the generic event E when the events E_1, E_2, ..., E_n, ... are considered independent and equally probable with probability ?; the probability $\mathbf{P}(\mathrm{E})$ of the same generic event, the E_i being exchangeable events with the limiting distribution $\Phi(\xi)$, is
>
> $$\mathbf{P}(\mathrm{E}) = \int_0^1 \mathbf{P}_\xi(\mathrm{E})\,\mathrm{d}\Phi(\xi)$$
>
> This fact can be expressed by saying that the probability distributions $\mathbf{P}$ corresponding to the case of exchangeable events are linear combinations of the distributions $\mathbf{P}_\xi$ corresponding to the case of independent equiprobable events, the weights in the linear combination being expressed by $\Phi(\xi)$.
>
> (De Finetti 1937:128–9)

This general result can be illustrated by taking a couple of special cases. Suppose that we are dealing with a coin-tossing example and the generic event E is that heads occurs r times in n tosses. Then

$$\mathbf{P}_\xi(\mathrm{E}) = {}^nC_r\,\xi^r(1-\xi)^{n-r}$$

So

$$\mathbf{P}(\mathrm{E}) = \omega_r^{(n)} = {}^n\mathrm{C}_r \int_0^1 \xi^r(1-\xi)^{n-r}\,\mathrm{d}\boldsymbol{\Phi}(\xi)$$

If, in particular, F(?) is the uniform distribution, we have

$$\omega_r^{(n)} = {}^n\mathrm{C}_r \int_0^1 \xi^r(1-\xi)^{n-r}\,\mathrm{d}\boldsymbol{\Phi}(\xi)$$

$= {}^{n}C_{r}\, B(r + 1, n - r + 1)$, where B is the beta function

$= 1/(n + 1)$ (cf. Equation 4.9 above)

Comparing these results with our earlier calculations involving exchangeability, we can see how exchangeability and independence are related.

De Finetti interprets these mathematical results as showing that we can *eliminate* the notions of objective probability and independence (which in his view are metaphysical in character) in favour of those of subjective probability and exchangeability. Alternatively, we could speak of his results as a *reduction* of objective probability and independence to subjective probability and exchangeability. The idea is that when an objectivist assumes independence, and formulates corresponding mathematical equations, a subjectivist can simply reinterpret these equations as being about subjective probabilities and exchangeability. This interpretation eliminates the objectivist's metaphysical notions and gives the real empirical meaning of the equations. I will call this argument *De Finetti's exchangeability reduction* and will criticise it in the next section.

Criticism of De Finetti's exchangeability reduction

In the previous section, it has been shown that exchangeability is in a sense the subjective equivalent of objective independence. De Finetti takes this to mean that we can eliminate the objectivist's notion of independence in favour of exchangeability. From the objectivist's point of view, however, the relation can be read, so to speak, in the opposite direction as showing that we can only apply exchangeability *when the situation is objectively one of independence.* However, not all sequences of events are independent. On the contrary, there are many situations in which the outcome of a particular event is very strongly dependent on the outcomes of previous events. In such situations we would expect that the use of exchangeability, and the calculations with it explained above, would give completely erroneous results. This is indeed the case, as I will illustrate in a moment by means of an example. My conclusion is that far from our being able to reduce the notion of objective independence to that of exchangeability, the concept of exchangeability is actually parasitic on that of objective independence and so redundant. In order to use exchangeability in a way which does not lead to erroneous and misleading results, we have first to be sure that the situation is objectively one of independence. We can only acquire such a conviction by conjecturing that the situation is one of independence and testing this assumption rigorously. If our conjecture passes these tests, then we can use the exchangeability calculation without going far wrong, but there is no need to do so, since we handle the problem in the standard way, using independence and objective probabilities. In this case then, exchangeability is unnecessary. If, on the other hand, our tests show that the situation is not one of independence, then the use of exchangeability will give misleading results and should be avoided. In neither case therefore is there any reason for using exchangeability.

To illustrate this argument, it would be possible to use any sequence of events which are dependent rather than independent. I have chosen one very simple and at the same time striking example of dependence. This is the game of red or blue.[8] At each go of the game there is a number s which is determined by the previous results. A fair coin is tossed. If the result is heads, we change s to $s' = s + 1$, and if the result is tails, we change s to $s' = s - 1$. If $s' \geq 0$, the result of the go is said to be blue, whereas if $s' < 0$, the result of the go is said to be red. So, although the game is based on coin tossing, the results are a sequence of red and blue instead of a sequence of heads and tails. Moreover, although the sequence of heads and tails is independent, the sequence of red and blue is highly dependent. We would expect much longer runs which are all blue than runs in coin tossing which are all heads. If we start the game with $s = 0$, then there is a slight bias in favour of blue, which is the initial position. However, it is easy to eliminate this by deciding the initial value of s by a coin toss. If the toss gives heads we set the initial value of s at 0, and if the toss gives tails we set it at -1. This makes red and blue exactly symmetrical, so that the limiting frequency of blue must equal that of red and be $^1/_2$. It is therefore surprising that over even an enormously large number of repetitions of the game, there is high probability of one of the colours appearing much more often than the other. Feller (1950:82–3) gives a number of examples of these curious features of the game. Suppose for example that the game is played once a second for a year, i.e. repeated 31,536,000 times. There is a probability of 70 per cent that the more frequent colour will appear for a total of 265.35 days, or about 73 per cent of the time, whereas the less frequent colour will appear for only 99.65 days, or about 27 per cent of the time.

Let us next suppose that two probabilists – an objectivist (Ms A) and a subjectivist (Mr B) – are asked to analyse a sequence of events, each member of which can have one of two values. Unknown to them, this sequence is in fact generated by the game of red or blue. Possibly the sequence might be produced by a man-made device which flashes either 0 (corresponding to red) or 1 (corresponding to blue) on to a screen at regular intervals. However, it is not impossible that the sequence might be one occurring in the world of nature. Consider for example a sequence of days, each of which is classified as 'rainy' if some rain falls, or dry otherwise. In a study of rainfall at Tel Aviv during the rainy season of December, January and February, it was found that the sequence of days could be modelled successfully as a sequence of dependent events. The particular kind of dependence used was what is known as a *Markov chain,* that is to say the probability of a day being rainy was postulated to depend on the weather of the previous day, but not on the weather of days further back in the sequence. In fact, the probabilities found empirically were probability of a dry day given that the previous day was dry = 0.75, and probability of a rainy day given that the previous day was rainy = 0.66. (For further details see Cox and Miller 1965:78–9.) It is clear that this kind of dependence will give longer runs of either rainy or dry days than would be expected on the assumption of independence. It is thus not impossible that the sequence of rainy and dry days at some place and season might be represented quite well by the game of red or blue.

Let us return to our two probabilists and consider first the objectivist (Ms A). Knowing that the sequence has a random character, she will begin by making the simplest and most familiar conjecture that the events are independent. However, being a good Popperian, she will test this conjecture rigorously with a series of statistical tests for independence. It will not be long before she has rejected her initial conjecture, and she will then start exploring other hypotheses involving various kinds of dependence among the events. If she is a talented scientist, she may soon hit on the red or blue mechanism and be able to confirm that it is correct by another series of statistical tests.

Let us now consider the subjectivist Mr B. Corresponding to Ms A's initial conjecture of independence, he will naturally begin with an assumption of exchangeability. Let us also assume that he gives a uniform distribution a priori to the $\omega_r^{(n)}$ (see Equation 4.9 above) so that Laplace's Rule of Succession holds (Equation 4.10). This is just for convenience of calculation. The counterintuitive results would appear for any other coherent choice of the $\omega_r^{(n)}$. Suppose that we have a run of 700 blues followed by two reds. Mr B would calculate the probability of getting blue on the next go using Equation 4.10 with $n = 702$ and $r = 700$. This gives the probability of blue as $^{701}/_{704} = 0.996$ to three significant figures. Knowing the mechanism of the game, we can calculate the true probability of blue on the next go, which is very different. Go 700 gave blue, and go 701 gave red. This is only possible if s on go 700 was 0, the result of the toss was tails and s became -1 on go 701. The next toss must also have yielded tails or there would have been blue again on go 702. Thus s at the start of go 703 must be -2, and this implies that the probability of blue on that go is zero. Then again let us consider one of Feller's massive sessions of 31,536,000 goes. Suppose the result is that the most frequently occurring colour appears 73 per cent of the time (as pointed out above there is a probability of 70 per cent of this result, which is thus not an unlikely outcome). Mr B will naturally be estimating the probability of this colour at about 0.73 and so much higher than that of the other colour. Yet in the real underlying game, the two colours are exactly symmetrical.

We see that Mr B's calculations using exchangeability will give results at complete variance with the true situation. Moreover, he would probably soon notice that there were too many long runs of one colour or the other for his assumption of exchangeability to be plausible. He might therefore think it desirable to change his assumption of exchangeability into some other assumption. Unfortunately, however, he would not be allowed to do so according to De Finetti, for, to quote again a section of the key passage given above:

> ... when experience teaches us the result A on the first n trials, our judgment will be expressed by the probability $\mathbf{P}(E_{n+1})$ no longer, but by the probability $\mathbf{P}(E_{n+1} | A)$, i.e. that which our initial opinion would already attribute to the event E_{n+1} considered as conditioned on the outcome A. Nothing of this initial opinion is repudiated or corrected; it is not the function **P** which has

> been modified (replaced by another **P***), but rather the argument E_{n+1} which has been replaced by $E_{n+1} \mid A$, and this is just to remain faithful to our original opinion (as manifested in the choice of the function **P**) ...
>
> (1937:146)

Yet if we assume exchangeability a priori when the sequence of events is in reality dependent, no amount of modifying our prior probabilities $P(E_{n+1})$ to posterior probabilities $P(E_{n+1} \mid A)$ by Bayesian conditionalisation will produce probabilities which accord with the real situation. De Finetti's exchangeability analysis only looked plausible in the first place because it was applied to coin tossing, and we know from long experience that tosses of a coin can validly be considered to be objectively independent. Unless we know that the events are objectively independent, we have no guarantee that the use of exchangeability will lead to reasonable results.

This point explains why the Rule of Succession leads to such erroneous results in the case in which the Sun mysteriously fails to rise one morning. Of course our background knowledge tells us that successive risings of the Sun are not independent events, but are highly dependent. This explanation of the situation can be reinforced by considering a case in some respects like the example of the Sun rising, but in which we do know that the events are independent. In such a case, as we shall see, the Rule of Succession gives perfectly reasonable results.

Suppose we have a large number of balls in a container. The container is thoroughly shaken, a ball is drawn, its colour is noted and it is then replaced. We can suppose that, as part of our background knowledge, we have a detailed acquaintance with all the mechanisms involved so that we can be sure that the drawings are independent. We do not, however, know the number of balls in the container or their colour. In fact, there are 1,000,000 balls of which 999,999 are yellow (corresponding to the Sun rising), and one is black (corresponding to its failing to rise). Suppose a yellow ball is drawn 737,856 times, and then a black ball is drawn. The Rule of Succession gives 737,856/737,858 = 0.9999972 to seven significant figures for the probability of drawing a yellow ball on the next occasion. This is actually not unreasonable in the circumstances. The results so far indicate that there must be an overwhelming preponderance of yellow balls in the container. So that, even if there are a few black balls, we are still much more likely to get a yellow ball on the next draw, provided the container is shaken very thoroughly (independence assumption). The Rule of Succession gives a reasonable result in this case of drawing balls from a container, but an absurd result in the case of the Sun failing to rise. This is because we know that independence applies in the case of drawing the balls, and that it doesn't apply in the case of the Sun either rising or failing to rise. This reinforces our conclusion that we can only apply exchangeability if we are sure on the basis of our background knowledge that the events concerned are objectively independent.

This concludes my criticism. Let us now see how a supporter of De Finetti might try to answer it. De Finetti himself does say one or two things which are relevant to the problem. Having shown that exchangeable events are the subjective equivalent of the objectivist's independent and equiprobable events, he observes that one could introduce subjective equivalents of various forms of dependent events, and, in particular, of Markov chains. As he says:

> One could in the first place consider the case of classes of events which can be grouped into Markov "chains" of order 1,2, ..., m, ..., in the same way in which classes of exchangeable events can be related to classes of equiprobable and independent events.
>
> (De Finetti 1937: Footnote 4, 146)

We could call such classes of events *Markov exchangeable.* De Finetti argues that they would constitute a complication and extension of his theory without causing any fundamental problem:

> One cannot exclude completely *a priori* the influence of the order of events.... There would then be a number of degrees of freedom and much more complication, but nothing would be changed in the setting up and the conception of the problem ..., before we restricted our demonstration to the case of exchangeable events ...
>
> (1937:145)

Perhaps De Finetti has in mind something like the following. Instead of just assuming exchangeability, we consider not just exchangeability but various forms of Markov exchangeability. To each of these possibilities we give a prior probability. No doubt exchangeability will have the highest prior probability. If the case is a standard one, like the biased coin, this high prior probability will be reinforced, and the result will come out moreover less like that obtained by just assuming exchangeability. If, however, the case is an unusual one, then the posterior probability of exchangeability will gradually decline, and that of one of the other possibilities will increase until it becomes much more probable than exchangeability. Does a scheme of this sort resolve the problems which have been raised? I will now argue that it does not.

The main problem with the approach just sketched is that it is unworkably complicated, and moreover these complications are quite unnecessary since they can be eliminated completely on the objective approach. I will deal with these points in turn. What leads to so much complication is that on this approach *it is necessary to consider all the possibilities which might arise at the very beginning of the investigation.* In order to set up his prior probabilities, Mr B has to consider every possible kind of dependence which might arise in the sequence of events, and assign each a prior probability. Now there is a very large number of different forms of dependence. De Finetti mentions Markov chains of different orders, but there are non-Markovian forms of dependence as well. Even if Mr B listed all the forms of dependence which have been so far explicitly defined and studied by mathematicians, he could still miss the one which applies to the sequence of events he is considering because this might be of a hitherto unstudied form. Yet for Mr B to list and assign prior probabilities to all forms of dependence known at present would be a task of such complexity so as to exceed most human powers. It is a testimony to the difficulty of this task that no one has, to my knowledge, carried it out in detail. Moreover, and this is my second point, all this complication is eliminated completely by adopting the objective approach. Our objectivist Ms A, when considering a sequence of events of a hitherto unstudied type, need only consider a single possibility to begin with. She could start with the conjecture that the events are independent with constant probabilities for the various outcomes. She does not need to bother a priori with other hypotheses of dependence, variable probabilities, or whatever, because, being a good Popperian, *she will subject her initial conjecture to a series of rigorous statistical tests.* Perhaps these tests will corroborate her initial conjecture in which case an elaborate a priori consideration of other possibilities would have been a waste of time and trouble. Perhaps, however, the test will refute her conjecture, in which case she will, at that stage and in the light of the results obtained, attempt to devise some new hypothesis. By approaching the problem in this step-by-step fashion, it is rendered tractable, whereas the Bayesian attempt to consider all possibilities a priori is quite unworkable.

Let us now consider another way in which the criticism we have made might be answered. A subjectivist might argue that De Finetti's requirement that prior probabilities should be changed only by Bayesian conditionalisation, i.e. from $\mathbf{P}(E_{n+1})$ to $\mathbf{P}(E_{n+1} | A)$ is too strong. Maybe prior probabilities should generally be altered in this fashion, but perhaps if exceptional results appear, as in the game of red or blue, prior probabilities could be altered in some quite different fashion to take account of the new circumstances. This solution of the difficulty certainly appeals to common sense, and would, I am sure, be adopted in practice. Unfortunately, however, it destroys the basis of De Finetti's exchangeability reduction, and even of Bayesianism in general. The exchangeability reduction works by arguing that whatever prior probabilities a set of different people adopt, their posterior probabilities will converge towards the same value. However, this argument is only valid on the assumption that all members of the set are changing their prior probabilities to posterior probabilities by Bayesian conditionalisation. If they are allowed at any time to change their priors in some quite different fashion (as on the present suggestion), there is no guarantee that their posterior probabilities will become at all similar. After 500 events, Mr B might suddenly decide to change to some form of Markov exchangeability, while Ms C continues to use exchangeability. After 700 events their posterior probabilities could be completely different. Moreover, it is one of the most attractive features of Bayesianism that it offers a simple mathematical formula for the way in which a rational person should change his or her beliefs in the light of evidence. If we now say: 'well, sometimes rational people should use this mathematical formula to change their beliefs, but, of course, it is quite open to them whenever they feel like it to change their beliefs in a completely different way', then surely we have lost that very feature which made Bayesianism an appealing theory.

I conclude that De Finetti's exchangeability reduction does not work, and it will be obvious that my arguments against this reduction can be used against Bayesianism in general. I will consider this matter briefly in the next section.

Some objections to Bayesianism

Most Bayesian statisticians use Bayesianism in something like the following form. They suppose that, in a given problem, there is a set of possible hypotheses to be considered. This set can be written $\{H_\theta\}$ where $\theta \varepsilon$ I, for some set I, usually an interval of the real line. The parameter θ is given a prior distribution $\mu(\theta)$ say, and this is changed to a posterior distribution $\mu(\theta \mid e)$. These distributions are in effect over the set of hypotheses under consideration. So let us set $P(H_\theta) = \mu(\theta)$ and $P(H_\theta \mid e) = \mu(\theta \mid e)$.

We can test this approach using the following simple 'black box' model. Mr B is confronted with a black box which flashes a figure (either 0 or 1) on to a screen at regular intervals $t = 0, 1, 2, ..., n, ...$ Let the sequence of figures be $x_0, x_1, x_2, ...,$ $x_n, ...$ It is generated by some process unknown to Mr B. Mr B has to assign probabilities of the form $P(x_n \mid x_1, x_2, ..., x_{n-1})$ when he knows the value of $x_0, x_1, x_2,$ $..., x_{n-1}$ but not that of x_n. These probabilities are taken as his betting quotients in the usual gambling game played with Ms A on the value of the nth figure. Mr B tackles this problem by using the standard approach of a Bayesian statistician described in the first paragraph of this section. If e states the observed values of $x_0,$ $x_1, x_2, ..., x_{n-1}$, he uses $P(H_\theta \mid e)$ to calculate $P(x_n \mid e)$.

In this framework, we can restate the objection, based on the game of red or blue, and given previously (p. 79). Suppose Mr B chooses H_θ = the sequence is independent with Prob(1) = θ, $0 \le \theta \le 1$. Suppose further that the sequence is in reality generated by the game of red or blue with red = 0, blue = 1. Arguing as in the previous section, we can show that Mr B's systematic use of Bayesian conditionalisation as his means of learning will produce a sequence of probabilities at complete variance with reality. Bayesian conditionalisation will not therefore be a very effective learning strategy.

The obvious reply which a Bayesian might make to this argument is that Mr B has considered too narrow a class of hypotheses and a broader class should have been introduced. Albert has, however, shown that there is a serious difficulty with this reply.[9] Albert asks us to suppose that the 0s and 1s flashing on the screen of the black box are generated by what he calls a *Chaotic Clock.* This device is illustrated in Figure 4.1. There is one pointer that can point to all real numbers in the interval I = [0, 1], where the vertically upward position is zero and the vertically downward position is $^1/_2$. Initially, the pointer deviates by an angle $\omega = 2\theta\pi$ from the vertically upward position, thus pointing at the real number θ. At $t = 1, 2, ..., n, ...,$ the pointer moves by doubling the angle ω.

In terms of the chaotic clock, Mr B can form hypotheses as to how the sequence of 0s and 1s is generated. H_θ might be that θ is the initial position of the pointer and that if the pointer comes to rest in the left hand side of the dial, the screen of the black box shows 0, while otherwise it shows 1. For technical reasons, Albert (1999) considers a slight modification of this chaotic clock set of hypotheses, and Suppose Mr B adopts any learning strategy whatever, i.e. he chooses his

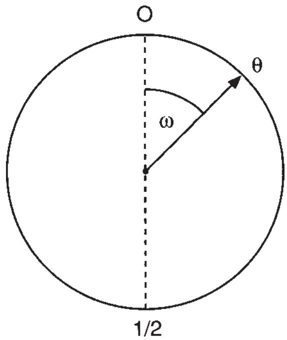

Figure 4.1 A chaotic clock

sequence of $P(x_n \mid e)$ in any arbitrary way. There then exists a prior probability distribution μ over the set of modified chaotic clock hypotheses such that Mr B's probabilities are produced by Bayesian conditioning of μ.

Albert's result is very striking indeed. His chaotic clock hypotheses are by no means absurd. After all, chaos theory is used in both physics and in economics. Indeed, hypotheses involving chaos are quite plausible as a means of explaining, for example, stock market fluctuations. If Mr B were really faced with a bizarre sequence of 0s and 1s, why should he not consider a hypothesis based on chaos theory? His imaginary situation is not so very different from the real situation of traders in financial markets who sit glued to their computer screens and make bets on what will appear shortly. Yet if Mr B is allowed to consider the chaotic clock set of hypotheses, then any learning strategy he adopts becomes a Bayesian strategy for a suitable choice of priors. In effect, Bayesianism has become empty.

It follows that a Bayesian of the type we are considering in this section (Mr B say) is caught on the horns of a dilemma. Mr B may adopt a rather limited set of hypotheses to perform his Bayesian conditionalisation, but then, as the example of the game of red or blue shows, if his set excludes the true hypothesis his Bayesian learning strategy may never bring him close to grasping what the real situation is. This is the first, or 'red or blue', horn of the dilemma. If Mr B responds by saying he is prepared to consider a wide and comprehensive set of hypotheses, these will surely include hypotheses from chaos theory and thus anything he does will become Bayesian, making the whole approach empty. This is the second, or 'chaotic clock', horn of the dilemma.

These difficulties with Bayesianism and, more specifically, with De Finetti's exchangeability reduction do indicate that there may be a need for objective probabilities and a methodology for statistics based on testing. This is therefore a good point at which to begin considering the principal objective theories of probability which will be dealt with in the next three chapters. I will, however, conclude the present chapter by considering in the last section the historical background to De Finetti's introduction of the subjective theory.

De Finetti's route to subjective probability

Earlier (pp. 52–3) I showed how Ramsey arrived at the subjective theory of probability through a criticism of Keynes's logical theory. This was not the way that De Finetti came to the theory, however, since, as I pointed out earlier, he only studied Keynes's views on probability carefully after he had already formulated the subjective theory. But what then was De Finetti's route to subjective probability?

De Finetti (1995) gives some reminiscences about when he first concluded that probability was subjective. As far as he could remember, the adoption of this philosophical position occurred very early in his intellectual career, and in fact:

> When I was a student, probably two years before graduating, while I was studying a book of Czuber's, *Wahrscheinlichkeitsrechnung* ... In that book there was a brief account of the various conceptions of probability, presented very sketchily in the first few paragraphs. Now I don't remember well the contents of the book either in general or regarding the various conceptions of probability. It seems to me that he mentioned De Morgan as representative of the subjective point of view.... Comparing the various positions it seemed to me that all the other definitions were meaningless. In particular the definition which is based on the so-called "equally probable cases" seemed to me unacceptable.
>
> (De Finetti 1995:111)

Czuber's book on probability was published in 1903, with a second enlarged and revised edition appearing in 1908–10. It was an important work in the early decades of the twentieth century and is referred to extensively by Keynes. It is worth noting that Keynes states that Czuber gives one of the best accounts of the paradoxes of geometrical probability (Keynes 1921:47), but that nonetheless Czuber thought that some form of the Principle of Non-sufficient Reason was indispensable.

In De Finetti's (1931a) first systematic account of the philosophy of probability, there are, however, no references to either Czuber or De Morgan. Instead, he cites mainly the writings of the French school of probabilists: Bertrand, Borel, Lévy and Poincaré. These writers were of course steeped in the Laplacean tradition, and their writings (particularly those of Bertrand and Borel) contained detailed discussions of the paradoxes of the Principle of Indifference. Thus, although De Finetti's reading must have been considerably different from Ramsey's, he was faced with the same problem situation – namely the difficulties for the traditional Laplacean kind of Bayesianism created by paradoxes of the Principle of Indifference. These paradoxes arose because of the perceived need to generate a single correct probability by some kind of logical process. They are thus resolved by the subjective move which allows different people to have different prior probabilities without this creating a contradiction.

However, De Finetti does not focus narrowly on the problems generated by the Principle of Indifference, but he rejects the whole Laplacean outlook, both

Laplace's determinism and his acceptance of the enlightenment value of rationality. Regarding determinism, De Finetti says:

> Certainly, we cannot accept determinism; we cannot accept the "*existence*", in that famous alleged realm of darkness and mystery, of immutable and necessary "*laws*" which rule the universe, and we cannot accept it as true simply because, in the light of our logic, it lacks all meaning....
>
> Nature will not appear ... as a monstrous and incorrigibly exact clockwork mechanism where everything that happens is what must happen because it could not but happen, and where all is foreseeable if one knows how the mechanism works.
>
> (1931a:169–70)

De Finetti returns often in his writings to this criticism of determinism and to a consideration of what should replace it. He also (De Finetti 1931a) explicitly rejects enlightenment rationalism in favour of a relativistic, and even irrational, mentality. Thus he says:

> ... the subjective theory of probability ... [is] ... an example of the application of the relativistic mentality to such an increasingly important branch of modern mathematics as the probability calculus, and as an essential part of the new vision of science which we want to give in an irrationalist, and, as we shall say, probabilist form.
>
> (1931a:172)

As we observed at the end of Chapter 2, these anti-enlightenment themes are very characteristic of the twentieth century, and perhaps especially of the 1930s when De Finetti was writing.

Although De Finetti refers to all the French authors mentioned above, his most frequent reference is to Poincaré's chapter on the calculus of probabilities in *Science and Hypothesis* (1902: Chapter XI, 183–210). Here Poincaré does indeed introduce subjective probability, which he says is the appropriate concept when a gambler is trying a single *coup* (1902:187–8). However, Poincaré goes on to argue that there is objective probability which manifests itself in a long sequence of repetitions. It looks as if De Finetti accepted Poincaré's notion of subjective probability but did not see any need for having objective probability as well. However, Poincaré has an argument for objective probability based on the insurance business. How could insurance companies make regular profits, he asks, if there was not some objective reality corresponding to their probability calculations? This argument obviously puzzled De Finetti, because he comments on it as follows:

> It seems strange that from a subjective concept there follow rules of action that fit practice. And Poincaré keeps explaining why the subjective explanation seems insufficient to him, mentioning practical applications in the field of insurance. "There are many insurance companies that apply the

> rules of the probability calculus, and they distribute to their shareholders dividends, whose objective reality is incontestable."
>
> (De Finetti 1931a:194)

Poincaré's example might be criticised in the light what happened at Lloyd's of London. This insurance company not only failed to distribute dividends, but even brought financial disaster to many of its 'names'. Is this an argument for the subjective approach to probability? Did the managers of Lloyd's formulate subjective probabilities for various events, which, although perfectly coherent, were rather unlucky? Or were they a bunch of incompetents who failed to apply the probability calculus correctly? Unfortunately, the whole matter is surrounded by great obscurity and allegations of fraud and corruption. So it is difficult to draw any definite conclusion.

We can now consider another important difference between Ramsey and De Finetti. It is to De Finetti rather than Ramsey that we should attribute the concept of exchangeability. This remark needs a little qualification since one of Ramsey's manuscript notes, published for the first time in 1991, does contain a derivation of Laplace's Rule of Succession in the special case $r = n$ using an argument quite similar to the one given above (pp. 70–3). Ramsey make the derivation under the condition: 'Suppose chance a priori of μ out of n + 1 being A is f(μ), all permutations equally probable.' (1991:278). The condition of all permutations being equally probable is equivalent in this context to De Finetti's exchangeability. Galavotti, who was the first to publish this passage, suggests that Ramsey took this condition 'from his teacher Johnson, who had introduced a 'permutation postulate' (1994:333).[10] However, we have here only a short unpublished note dealing with a very special case. This does not compare with De Finetti (1930b:121), who defined the concept explicitly,[11] and then went on to develop the mathematical theory of exchangeable random quantities in a series of important papers which culminated in his 1937. Since De Finetti wanted to eliminate objective probabilities completely in favour of subjective probabilities, he had more of a stimulus for developing the theory of exchangeability than had Ramsey, who, in his 1926 book at least, advocated, like Poincaré, a two-concept view of probability with both objective and subjective probabilities. I will return to Ramsey's two-concept view in Chapter 8, after I have given a detailed account of the two principal objective theories of probability in Chapters 5, 6 and 7.

4 The subjective theory

1 A good discussion of these criticisms of Ramsey's is to be found in Cottrell (1993: 30–2).

2 The heroine and hero of this betting scenario are named after the principal characters in Samuel Richardson's novel of 1740 *Pamela; or, Virtue Rewarded.* Pamela Andrews (Ms A) is a servant girl in the home of Mr B (always referred to thus in the novel). Mr B, who is very rich, attempts to seduce Pamela, but she virtuously refuses his advances, and eventually he decides to marry her. The novel was a best seller at the time of its publication and exerted an enormous influence

on the development of European literature. Presumably in Richardson's fictional setting, it must have been important for Ms A to ascertain Mr B's degrees of belief in various propositions.

3 For an interesting discussion of the money versus utility problem which is more sympathetic to the views of Ramsey, see Sahlin (1990: 41–3).

4 The proof which follows is based on De Finetti 1937, but expanded to fill in the details. A shorter but mathematically more sophisticated proof is to be found in Paris (1994: 19–23).

5 This was pointed out to me by Ladislav Kvasz.

6 A full account of his views on the question of finite versus countable additivity can be found in De Finetti (1970), and I discuss these views in my review of the book (Gillies 1972b: 142–5). In that article I give references to the original Italian edition of De Finetti's book, but in what follows here my references will be to the English translation which appeared in 1974.

7 I learnt of this example from a typescript version of Popper's (1957a), which was circulating in LSE when I was a graduate student there in 1966–8. Popper considers a situation in which the Sun has risen 1,000,000 times in succession but then fails to rise for 10 days. He uses this to criticise the subjective theory of learning in general terms for giving too much authority to past experience, and making a revision of our ideas practically impossible. Although nearly all of the typescript is reprinted in Popper (1957a, 1983), this example is rather curiously omitted. A possible reason is that the example is not effective against all versions of the subjective theory of learning. As Howson and Urbach point out (1989: 81), Bayesianism implies falsificationism in the sense that refuted hypotheses acquire probability 0. Let us consider then a version of subjective Bayesianism which is concerned with the learning of general laws in the sense of trying to assign probabilities to such laws in the light of evidence. Such an approach would have assigned a probability to the universal law that the Sun rises every morning in the light of the 1,000,000 sunrises in succession. However, this probability would drop to zero after the first failure. Thus Popper's example is not a good argument against all versions of the subjective theory of learning, but it does yield a very strong argument against the Rule of Succession as I will show in what follows.

8 The game of red or blue is described in Feller (1950: 67–95), which contains an interesting mathematical analysis of its curious properties. Popper read of the game in Feller, and he had the idea of using it to argue against various theories of induction. Popper (1957a: 358–60) (reprinted 1983: 301–5) uses the game to criticise what he calls 'the simple inductive rule', while later (Popper 1957a: 366–7, reprinted in 1983: 323–4) he uses the game to try to prove the impossibility of an inductive logic. The first of these arguments seems to me valid, and I have adapted it to produce the criticism of De Finetti's exchangeability reduction given here. The second of Popper's arguments seems to me less convincing, since it is perfectly possible that an inductive logic could be devised which could accommodate cases like the game of red or blue. Indeed I give arguments in favour of the possibility of an inductive logic (Gillies 1996: 98–112).

9 The mathematical part of Albert's argument is to be found in Albert (1999), where Theorem 1 is what is here called the *Anything Goes Theorem.* The more philosophical part of the argument will be published soon. I am most grateful to Max Albert for sending me an unpublished typescript with a full discussion of both the mathematical and philosophical sides of the argument, as well as for some helpful discussions of the question and its relation to the argument involving the game of red or blue.

10 An interesting account of W. E. Johnson's contribution to this question is to be found in Zabell (1989).

11 De Finetti initially used the term 'equivalent' (in Italian *equivalente*), but the term 'exchangeable' has now become standard.

References

In general, works are cited by their date of first publication, but the exact edition from which quotations are taken is also specified, and its date given if different from the first edition. Occasionally two places are given where a particular work can be found. In this case the quotations in the text are from the first cited place, but the second is added as it might be more accessible to the reader. English translations of works in foreign languages are cited whenever possible.

Adams, E. (1964) 'On Rational Betting Systems', *Archiv für Mathematische Logik und Grundlagenforschung* 6: 7–29, 112–28.

Albert, M. (1992) 'Die Falsifikation Statistischer Hypothesen', *Journal for General Philosophy of Science* 23: 1–32.

—— (1999) 'Bayesian Learning when Chaos looms large', *Economics Letters* 65: 1–7.

Ayer, A. J. (1963) 'Two Notes on Probability', in *The Concept of a Person and other Essays,* Macmillan, 188–208.

Bacon, F. (1620) *Novum Organum,* English translation in R. L. Ellis and J. Spedding (eds), *The Philosophical Works of Francis Bacon,* Routledge, 1905, 212–387.

Bateman, B. W. (1988) 'G. E. Moore and J. M. Keynes: A Missing Chapter in the History of the Expected Utility Model', *American Economic Review* 78: 1098–106.

—— (1996) *Keynes's Uncertain Revolution,* University of Michigan Press.

Bayes, T. and Price, R. (1763) 'An Essay towards Solving a Problem in the Doctrine of Chances', reprinted in E. S. Pearson and M. G. Kendall (eds.) *Studies in the History of Statistics and Probability,* Griffin, 1970, 134–53.

Born, M. (1935) *Atomic Physics,* Blackie, 7th Edition, 1966.

Cantelli, F. P. (1935) 'Consideration sur la convergence dans le calcul des probabilités', *Annales de l'institut Henri Poincaré* V: 1–50.

Carnap, R. (1950) *Logical Foundations of Probability,* 2nd Edition, University of Chicago, 1963.

Childers, T. (1996) 'On the Relation between the Normative and the Empirical in the Philosophy of Science', unpublished PhD thesis, University of London.

Church, A. (1936) 'An unsolvable problem of elementary number theory', *American Journal of Mathematics* 58: 345–63.

—— (1940) 'On the concept of a random sequence', *Bulletin of the American Mathematical Society* 46: 130–5.

Cottrell, A. (1993) 'Keynes's Theory of Probability and its Relevance to his Economics. Three Theses', *Economics and Philosophy* 9: 25–51.

Cox, D. R. and Miller, H. D. (1965) *The Theory of Stochastic Processes,* Methuen.

Cramér, H. (1946) *Mathematical Methods of Statistics,* Princeton University Press, 1961.

Czuber, E. (1903/1908–10) *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlersausgleichung, Statistik und Lebensversicherung.* Teubner.

Daston, L. (1988) *Classical Probability in the Enlightenment,* Princeton University Press.

David, F. N. (1962) *Games, Gods and Gambling. The Origins and History of Probability and Statistical Ideas from the Earliest Times to the Newtonian Era,* Harper.

Davis, J. B. (1994) *Keynes's Philosophical Development,* Cambridge University Press.

De Finetti, B. (1930a) 'Fondamenti Logici del Ragionamento Probabilistico', *Bollettino dell' Unione Matematica Italiana* 5: 1–3. Reprinted in De Finetti, 1981a, 261–3.

—— (1930b) 'Funzione Caratteristica di un Fenomeno Aleatorio', *Memorie della Reale Accademia dei Lincei* IV, 5: 86–133. Reprinted in De Finetti, 1981a, 267–315,

—— (1930c) 'Problemi Determinati e Indeterminati nel Calcolo della Probabilità', *Rendiconti della Reale Accademia Nazionale dei Lincei* XII, 9: 367–73. Reprinted in De Finetti, 1981a, 327–33.

—— (1931a) 'Probabilism', English translation in *Erkenntnis* 1989, 31: 169–223.

—— (1931b) 'On the Subjective Meaning of Probability', English translation in De Finetti, 1993, 291–321.

—— (1936) 'Statistics and Probability in R. Von Mises' Conception', English translation in De Finetti, 1993, 353–64.

—— (1937) 'Foresight: Its Logical Laws, Its Subjective Sources', English translation in H. E. Kyburg and H. E. Smokler (eds.), *Studies in Subjective Probability,* Wiley, 1964, 93–158.

—— (1938) 'Cambridge Probability Theorists', English translation in *The Manchester School of Economic and Social Studies* 1985, 53: 348–63.

—— (1970) *Theory of Probability.* vols 1 and 2, English translation. Wiley. 1974.

—— (1981a) *Scritti (1926–1930),* Cedam: Padua.

—— (1981b) 'The Role of "Dutch Books" and "Proper Scoring Rules"', *British Journal for the Philosophy of Science* 32: 55–6.

—— (1991) *Scritti (1931–1936),* Pitagora: Bologna.

—— (1993) *Induction and Probability,* Clueb: Bologna.

—— (1995) *Filosofia della Probabilità,* Il Saggiatore.

Dostoyevsky, F. M. (1866) *The Gambler,* English translation in Penguin Classics, 1976.

Duhem, P. (1904–5) *The Aim and Structure of Physical Theory,* English translation by Philip P. Wiener of the second French edition of 1914, Atheneum, 1962.

Earman, J. and Salmon, W. C. (1992) 'The Confirmation of Scientific Hypotheses', in M. H. Salmon *et al.* (eds), *Introduction to Philosophy of Science.* Prentice Hall, Chapter 2, 42–103.

Edwards, A. W. F. (1978) 'Commentary on the Arguments of Thomas Bayes', *Scandinavian Journal of Statistics* 5: 116–18.

Farjoun, E. and Machover, M. (1983) *Laws of Chaos,* Verso.

Feller, W. (1950) *Introduction to Probability Theory and Its Applications,* Third edition, 1971, Wiley.

Fetzer, J. H. (1981) *Scientific Knowledge: Causation, Explanation, and Corroboration,* Boston Studies in the Philosophy of Science, Vol. 69, D. Reidel, Dordrecht.

—— (1982) 'Probabilistic Explanations', *PSA* 2: 194–207.

—— (1988) 'Probabilistic Metaphysics', in J. H. Fetzer (ed.) *Probability and Causality,* D. Reidel, 109–132.

—— (1991) 'Critical Notice: Philip Kitcher and Wesley C. Salmon, (eds.), *Scientific Explanation*; and Wesley C. Salmon, *Four Decades of Scientific Explanation. Philosophy of Science', Philosophy of Science* 58: 288–306.
—— (1993) 'Peirce and Propensities', in Edward C. Moore (ed.) *Charles S. Peirce and the Philosophy of Science,* University of Alabama Press, 61–71.
Fisher, R. A. (1925) *Statistical Methods for Research Workers,* Oliver and Boyd.
—— (1935) *The Design of Experiments,* Oliver and Boyd.
Fleck, L. (1935) *Genesis and Development of a Scientific Fact,* English translation, University of Chicago Press, 1981.
Fraser, R. G. F. (1931) *Molecular Rays,* Cambridge University Press.
Fry, T. C. (1928) *Probability and its Engineering Uses,* Van Nostrand.
Galavotti, M. C. (1989) 'Anti-Realism in the Philosophy of Probability: Bruno de Finetti's Subjectivism', *Erkenntnis* 31: 239–61.
—— (1991) 'The notion of subjective probability in the work of Ramsey and de Finetti', *Theoria* LXII(3): 239–59.
—— (1994) 'F. P. Ramsey and the Notion of "Chance"', *Proceedings of the 17th International Wittgenstein-Symposium (Kirchberg am Wechsel, 14-21 August 1994)* 330–40.
—— (1995) 'Operationism, Probability and Quantum Mechanics', *Foundations of Science* 1: 99–118.
—— (1999) 'Some Remarks on Objective Chance (F. P. Ramsey, K. R. Popper and N. R. Campbell)', in M. L. Dalla Chiara *et al.* (eds.) *Language, Quantum, Music,* Kluwer, 73–82.
Gibbon, E. (1776–88) *Decline and Fall of the Roman Empire,* John Murray. 1862.
Gillies, D. A. (1971) 'A Falsifying Rule for Probability Statements', *British Journal for the Philosophy of Science* 22: 231–61.
—— (1972a) 'Operationalism', *Synthese* 25: 1–24.
—— (1972b) 'The Subjective Theory of Probability', *British Journal for the Philosophy of Science* 29: 138–57.
—— (1973) *An Objective Theory of Probability,* Methuen.
—— (1982) *Frege, Dedekind, and Peano on the Foundations of Arithmetic,* Van Gorcum.
—— (1987) 'Was Bayes a Bayesian?', *Historia Mathematica* 14: 325–46.
—— (1988a) 'Induction and Probability', in G. H. R. Parkinson (ed.), *An Encyclopaedia of Philosophy,* Chapter 9, 179–204.
—— (1988b) 'Keynes as a Methodologist', *British Journal for the Philosophy of Science* 39: 117–29.
—— (1990) 'Bayesianism versus Falsificationism. Review of Howson and Urbach 1989', *Ratio* New Series III(1): 82–98.
—— (1991) 'Intersubjective Probability and Confirmation Theory', *British Journal for the Philosophy of Science* 42: 513–33.
—— (1996) *Artificial Intelligence and Scientific Method,* Oxford University Press.
Gillies, D. A. and Ietto-Gillies, G. (1991) 'Intersubjective Probability and Economics', *Review of Political Economy* 3(4): 393–417.
Gnedenko, B. V. (1950) *The Theory of Probability,* English translation, Chelsea, 1962.
Hacking, I. (1965) *Logic of Statistical Inference,* Cambridge University Press, 1965.
—— (1975) *The Emergence of Probability. A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference,* Cambridge University Press, Paperback edition, 1984.
Hicks, J. (1979) *Causality in Economics,* Blackwell.

Hilbert, D. (1899) *Foundations of Geometry,* English translation, Open Court, 1971.

Howson, C. and Urbach, P. (1989) *Scientific Reasoning. The Bayesian Approach,* Open Court.

Humphreys, P. (1985) 'Why Propensities cannot be Probabilities', *The Philosophical Review* 94: 557–70.

Iversen, G. R., Longcor, W. H., Mosteller, F., Gilbert, J. P., and Youtz, C. (1971) 'Bias and Runs in Dice Throwing and Recording: a Few Million Throws', *Psychometrika* 36(1): 1–19.

Jaynes, E. T. (1973) 'The Well-Posed Problem', *Foundations of Physics* 4(3): 477–92.

Jeffreys, H. (1939) *Theory of Probability,* Oxford University Press.

Kendall, M. G. and Babington Smith, B. (1938) 'Randomness and Random Sampling Numbers', *Journal of the Royal Statistical Society* 101: 147.

Kendall, M. G. and Babington Smith, B. (1939a) 'Randomness and Random Sampling Numbers', *Journal of the Royal Statistical Society (Supplement)* 6: 57.

Kendall, M. G. and Babington Smith, B. (1939b) *Tables of Random Sampling Numbers,* Cambridge University Press.

Keynes, J. M. (1921) *A Treatise on Probability,* Macmillan, 1963.

—— (1936) *The General Theory of Employment, Interest and Money.* The Collected Writings of John Maynard Keynes, Vol. VII, Macmillan, Cambridge University Press for the Royal Economic Society, 1993.

—— (1938) 'My Early Beliefs', in The Collected Writings of John Maynard Keynes, Vol. X, *Essays in Biography,* Macmillan, Cambridge University Press for the Royal Economic Society, 1985.

Kolmogorov, A. N. (1933) *Foundations of the Theory of Probability,* Second English Edition, Chelsea, 1956.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions,* University of Chicago Press.

Lad, F. (1983) 'The Construction of Probability Theory', *Science and Society* 47: 285–99.

Lad, F. (1996) *Operational Subjective Statistical Methods,* Wiley.

Lakatos, I. (1968) 'Changes in the Problem of Inductive Logic', in *Philosophical Papers,* Vol. 2, Cambridge University Press, Paperback edition, 1983, Chapter 8, 128–200.

Laplace, P. S. (1814) *A Philosophical Essay on Probabilities,* English Translation of the 6th French Edition, Dover, 1951.

Locke, J. (1690) *An Essay concerning Human Understanding,* Everyman Library, J. M. Dent and Sons, 1961.

McCurdy, C. S. I. (1996) 'Humphreys's Paradox and the Interpretation of Inverse Conditional Propensities', *Synthese* 108: 105–25.

Mach, E. (1883) *The Science of Mechanics: A Critical and Historical Account of its Development,* 6th American Edition, Open Court, 1960.

Mellor, D. H. (1971) *The Matter of Chance,* Cambridge University Press.

Miller, D. W. (1994) *Critical Rationalism. A Restatement and Defence,* Open Court.

—— (1996) 'Propensities and Indeterminism', in A. O'Hear (ed.) *Karl Popper: Philosophy and Problems,* Cambridge University Press, 121–47.

Milne, P. (1986) 'Can there be a Realist Single-Case Interpretation of Probability?', *Erkenntnis* 25: 129–32.

Monk. R. (1990) *Ludwig Wittgenstein. The Duty of Genius,* Jonathan Cape.

—— (1996) *Bertrand Russell. The Spirit of Solitude,* Jonathan Cape.

Newton, I. (1687) *Philosophiae Naturalis Principia Mathematica,* Andrew Motte's English translation of 1729, revised by Florian Cajori, University of California Press, 1960.

Paris, J. B. (1994) *The Uncertain Reasoner's Companion. A Mathematical Perspective,* Cambridge University Press.
Pascal, B. (1670) *Pensées,* English translation, Penguin, 1972.
Pearson, K. (1900) 'On the Criterion that a given System of Deviations from the Probable in the case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling', Reprinted in *Karl Pearson's Early Statistical Papers,* Cambridge University Press, 1956, 339–57.
Peirce, C. S. (1910) 'Notes on the Doctrine of Chances', Reprinted in *Essays in the Philosophy of Science,* The American Heritage Series, Bobbs-Merrill, 1957, 74–84.
Poincaré, H. (1902) *Science and Hypothesis,* English translation, Dover, 1952.
Popper, K. R. (1934) *The Logic of Scientific Discovery,* 6th revised impression of the 1959 English translation, Hutchinson, 1972.
—— (1957a) 'Probability Magic or Knowledge out of Ignorance', *Dialectica* 11(3/4): 354–74.
—— (1957b) 'The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory', in S. Körner (ed.) *Observation and Interpretation, Proceedings of the Ninth Symposium of the Colston Research Society, University of Bristol,* 65–70 and 88–9.
—— (1957c) 'The Aim of Science', *Ratio* 1(1): 24–35.
—— (1959a) *New Appendices to the Logic of Scientific Discovery,* in 6th revised impression of the 1959 English translation, Hutchinson, 1972, 307–464.
—— (1959b) 'The Propensity Interpretation of Probability', *British Journal for the Philosophy of Science* 10: 25–42.
—— (1963) *Conjectures and Refutations,* Routledge and Kegan Paul.
—— (1983) *Realism and the Aim of Science,* Hutchinson.
—— (1990) *A World of Propensities,* Thoemmes.
Ramsey, F. P. (1926) 'Truth and Probability', in Ramsey, 1931, 156–98. Reprinted in H. E. Kyburg and H. E. Smokler (eds), *Studies in Subjective Probability,* Wiley, 1964, 61–92.
—— (1928) 'Further Considerations', in Ramsey, 1931, 199–211.
—— (1929) 'Last Papers. C. Probability and Partial Belief', in Ramsey, 1931, 256–7.
—— (1931) *The Foundations of Mathematics and other Logical Essays,* Routledge and Kegan Paul.
—— (1991) *Notes on Philosophy, Probability and Mathematics,* Bibliopolis.
Reichenbach, H. (1949) *The Theory of Probability,* University of California Press.
Richardson, S. (1740) *Pamela; or Virtue Rewarded,* Penguin Classics, 1985.
Runde, J. (1994) 'Keynes After Ramsey: In Defence of *A Treatise on Probability', Studies in History and Philosophy of Science* 25(1): 97–121.
—— (1996) 'On Popper, Probabilities and Propensities', *Review of Social Economy,* LIV(4): 465–85.
Russell, B. (1905) 'On Denoting', in R. C. Marsh (ed.) *Bertrand Russell. Logic and Knowledge. Essays 1901–50,* Allen and Unwin, 1956, 41–56.
—— (1912) *The Problems of Philosophy,* Williams and Norgate, n.d.
—— (1914) *Our Knowledge of the External World,* Allen and Unwin, 1961.
—— (1959) *My Philosophical Development,* Allen and Unwin, 1969.
—— (1967) *The Autobiography of Bertrand Russell,* Vol. 1. *1872–1914,* Allen and Unwin.
Rutherford, D. E. (1951) *Classical Mechanics,* 2nd Edition, Oliver and Boyd, 1957.

Ryder, J. M. (1981) 'Consequences of a Simple Extension of the Dutch Book Argument', *British Journal for the Philosophy of Science* 32: 164–7.
Sahlin, N. -E. (1990) *The Philosophy of F. P. Ramsey,* Cambridge University Press.
Salmon, W. C. (1979) 'Propensities: a Discussion Review of D. H. Mellor *The Matter of Chance', Erkenntnis* 14: 183–216.
Sambursky, S. (1954) *The Physical World of the Greeks,* English paperback edition, Routledge and Kegan Paul, 1963.
Skidelsky, R. (1983) *John Maynard Keynes.* Vol. 1. *Hopes Betrayed 1883–1920,* Macmillan.
—— (1992) *John Maynard Keynes.* Vol. 2. *The Economist as Saviour 1920–1937,* Macmillan.
Soros, G. (1987/94) *The Alchemy of Finance. Reading the Mind of the Market,* 2nd Edition, 1994, John Wiley.
Soros, G. (1992) Introduction to W. Newton-Smith and J. Tianji (eds.) *Popper in China,* Routledge.
Student (Gosset, W. S.) (1908) 'The probable error of a mean', *Biometrika* 6: 1–25.
Sucar, L. E., Gillies, D. F., and Gillies, D. A. (1993) 'Objective Probabilities in Expert Systems', *Artificial Intelligence* 61: 187–208.
Todhunter, I. (1865) *A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace,* Chelsea, 1965.
Von Mises, R. (1919) 'Grundlagen der Wahrscheinlichkeitsrechnung', Reprinted in Von Mises, 1964b, 57–106.
—— (1928) *Probability, Statistics and Truth,* 2nd revised English edition, Allen and Unwin, 1961.
—— (1938) 'Ernst Mach und die empiristische Wissenschaftsauffassung', Reprinted in Von Mises, 1964b, 495–523.
—— (1940) 'Scientific Conception of the World: On a New Textbook of Positivism', Reprinted in Von Mises, 1964b, 524–9.
—— (1964a) *Mathematical Theory of Probability and Statistics,* Academic Press.
—— (1964b) *Selecta II,* American Mathematical Society.
Wald, A. (1937) 'Die Widerspruchsfreiheit des Kollektivsbegriffes', *Ergebnisse eines Mathematischen Kolloquiums* 8: 38–72.
—— (1938) 'Die Widerspruchsfreiheit des Kollektivbegriffes', in *Wald: Selected Papers in Statistics and Probability,* MacGraw-Hill, 1955, 25–41.
Williamson, J. O. D. (1999) 'Countable Additivity and Subjective Probability', *British Journal for the Philosophy of Science,* 50: 401–416.
Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus,* Routledge and Kegan Paul. 1963.
Zabell, S. L. (1989) 'The Rule of Succession', *Erkenntnis* 31: 283–321.