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

$$P(E_1) + \ldots + P(E_n) = 1$$

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 Ω \ (E v F) be the event that something other than E or F occurs. E, F, Ω \ (E v 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, Ω\ (E v 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 $q(\mathrm{E}) > 1$, Ms A can win by choosing $S > 0$. If Mr B chooses $q(\mathrm{E}) < 0$, Ms A can win by choosing $S < 0$. Hence to be coherent, Mr B must choose $0 = q(\mathrm{E}) = 1$.

(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 $0 = q(\mathrm{E}) = 1$.

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 $E_1, E_2, \ldots, E_n, \ldots$, 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: