SUBJECTIVE PROBABILITY
AND ACTION GUIDANCE

MARTIN PETERSON

ABSTRACT. In the Bayesian approach to subjective probability proposed by e.g. Ramsey (1926), de Finetti (1931), and Savage (1954), probability is defined in terms of preferences over uncertain prospects. However, according to a non-Bayesian approach seldom or never discussed by philosophers, subjective probability can be defined without referring to preferences over uncertain prospects. This article argues that Bayesian theories of subjective probability are inapplicable to practical decision making, since Bayesians put the cart before the horse from the point of view of the deliberating agent: Bayesians define subjective probability and utility in terms of preferences over uncertain prospects; therefore, those probability and utility numbers cannot figure as reasons for forming new preferences over the same set of gambles. An additional aim of the article is to show that the non-Bayesian approach to subjective probability is applicable, at least in principle, to practical decision making. This aim is achieved by developing a new non-Bayesian axiomatization of subjective probability, which draws on the work of DeGroot (1970).

1. Introduction

Thomas Bayes is perhaps best known for his formula for calculating conditional probabilities.\(^1\) However, he also deserves recognition for his suggestion that probability can be defined in terms of preferences over uncertain prospects, later developed by Ramsey, DeFinetti, Savage, and others. Here is Bayes:

The probability of any event is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of the thing expected upon its happening . . . If a person has an expectation depending on the happening of an event, the probability of the event is to the probability

\(^1\)In its simplest form, Bayes’ formula tells us that the probability of a hypothesis \(H\) given a piece of evidence \(E\) is equal to the probability of \(H\) multiplied by the probability of \(E\) given \(H\), divided by the probability of \(E\).
of its failure as his loss if it fails to his gain if it happens. 
(Bayes 1763:376-77.)

It may be helpful to illustrate Bayes’ proposal in an example. Suppose that we wish to measure someone’s subjective probability that it will rain in Cambridge tomorrow. We ask the agent which of the following very generous options he prefers:

A If it rains in Cambridge tomorrow the agent wins a diamond; otherwise he wins nothing.
B If it does not rain in Cambridge tomorrow the agent wins a diamond; otherwise he wins nothing.

Bayesians believe that a rational agent should let his preference between A and B be determined by two considerations, viz. the degree to which he believes that it will rain, and how strongly he desires to win the diamond. Let us suppose that the agent desires the diamond to some degree, i.e. desires it more than no diamond. Then, if the agent prefers A over B, he thinks it is more probable that it will rain in Cambridge tomorrow than not; otherwise he would be more likely to get what he desires by choosing the other option. Furthermore, if the agent prefers B to A he thinks it is more probable that there will be no rain, for the same reason. Finally, if the agent is indifferent between A and B, the agent must consider both events to be equiprobable. This is because no other way of assigning probabilities make both options come out as equally attractive. Note that this is true irrespective of how strongly one desires the diamond, and irrespective of how desires and beliefs are aggregated into preferences.

In this article I shall defend two claims. The first is that the Bayesian theory of subjective probability is inapplicable to practical decision making. More precisely, I shall argue that Bayesians put the cart before the horse from the point of view of the deliberating agent. This theory defines subjective probability (and utility) in terms of preferences over uncertain prospects; therefore, the obtained numbers cannot figure as reasons for forming new preferences over the same set of uncertain prospects. This is not an entirely novel point. Ramsey seems to have come to a similar insight shortly after having defended the Bayesian approach in ‘Truth and Probability’ (1926). In a short note entitled ‘Probability and Partial Belief’ (1928) Ramsey observed that:

'sometimes the [probability] number is used itself in making a practical decision. How? I want to say in accordance with the law of mathematical expectation; but I
cannot do this, for we could only use that rule if we had measured goods and bads. (Ramsey 1931:256.)

Ramsey never discussed this observation any further. The argument developed in the present article might capture the idea he had in mind. However, irrespective of whether this is so, I believe my argument to be worth considering in its own right.

The second claim I defend is that non-Bayesian theories of subjective probability are not vulnerable to the above-mentioned objection. In support of this claim, I present a new non-Bayesian axiomatization of subjective probability, which draws on the work of DeGroot (1970).

2. FROM PREFERENCE TO SUBJECTIVE PROBABILITY

Bayes explicitly claimed that the agent’s degree of belief can be represented by real numbers. Suppose, for instance, that you wish to measure your subjective probability that your diamond worth $10,000 will be stolen within a year. If you consider $500 to be a fair price for insuring the diamond, that is, if that amount is the highest price you are prepared to pay for a contract in which you get $10,000 if the diamond is stolen within a year, and nothing otherwise, then your subjective probability is, according to Bayes, \( \frac{500}{10,000} = 0.05 \).

Unfortunately, there are two obvious problems with this proposal. First, many people have a decreasing marginal utility for money, so the fact that they are prepared to pay $500 for insuring a diamond worth $10,000 need not show that their subjective probability is 0.05, since $10,000 is not twenty times as desirable as $500. Second, it is not clear why one should apply the principle of maximizing expected utility for aggregating beliefs and desires into preferences. Some decision maker’s may prefer some more risk averse decision rule.

The axiomatic theories of subjective probability developed by Ramsey (1926) and Savage (1954) can be thought of as attempts to fix these problems. Ramsey’s and Savage’s basic idea is to impose a set of structural requirements on preferences over gambles that restrict what combinations of preferences are legitimate. The structural requirements, or axioms, do not tell agents to evaluate gambles by calculating expected utilities. The agent is free to use what ever method he wishes for forming his preferences, as long as no combination of preferences violate the structural constraints. For example, if the agent strictly prefers option A to option B, then he must not strictly prefer B to A. Moreover, it can be proved that if all preferences satisfy the structural axioms, then the agent behaves as if he acted from a subjective probability function and a utility function that are consistent with the principle of maximizing
expected value. No claim is made about what mental or other process triggered the agents’ preferences. The theorems stated by Ramsey and Savage merely prove that a particular formal representation of those preferences is possible, which can be used for explaining and making predictions about future behaviour as long as all preferences remain unchanged.

Ramsey never worked out his axiomatization in detail. He thought that, ‘this would, I think, be rather like working out to seven places of decimals a result only valid to two’.² Savage’s work is more rigorous from a technical point of view: Let \( S = \{s, s', \ldots\} \) be a set of states with subsets \( A, B, \ldots \), and let \( X = \{x, x', \ldots\} \) be a set of consequences. Acts are conceived of as functions \( f, g, \ldots \), from \( S \) to \( X \). Let \( \succeq \) be a preference relation between pairs of acts. Now, Savage proved that if a set of six axioms (stated in Appendix A) hold true, then there exists a subjective probability function \( p \) and a real-valued function of consequences \( u \), such that \( f \succeq g \) if and only if

\[
\int_S [u(f(s)) \cdot p(s)] ds > \int_S [u(g(s)) \cdot p(s)] ds
\]

Furthermore, for every other function \( u' \) satisfying the axioms, there are numbers \( c > 0 \) and \( d \) such that \( u' = c \cdot u + d \).

A common critique of Savage’s theory is to argue that some of its axioms are questionable. Allais (1953) proposed a famous counter example to the notorious sure-thing principle (Axiom 2), which holds that states yielding identical outcomes under all acts may be disregarded by the agent. More recently, others have questioned the completeness assumption (Axiom 1), which holds that one either strictly prefers one risky alternative to the other or is indifferent between them. In order to see why this assumption may be questionable, imagine that someone asks you to state a preference between \$1000 for sure and a gamble in which you win a holiday in Hawaii if it rains in Cambridge tomorrow. If you feel that you have a clear preference between these two options, then modify the first option by either adding or withdrawing some money. Arguably, there is some point at which you no longer have a clear preference between a fixed amount of money and a chance to win a free holiday, and this need not be irrational; this indicates that the completeness axiom is too strong. The standard reply is that preference are directly revealed in choices and that you prefer whatever you choose, but this argument may draw too heavily on behavioristic ideas.

A more sophisticated response has been outlined by Joyce (1999), who suggests that the completeness axioms may be weakened: Instead of requiring that the agent’s preference ordering is complete, Joyce points

²Ramsey 1926:180.
out that we only have to assume that it can be *coherently extended* to a complete preference ordering, meaning that it must be possible to fill in the missing gaps without violating any of the axioms.

Having said all this about Savage’s work, it should be noted that theories of subjective probability can also be presented in a slightly different way. The famous Dutch Book Theorem plays a crucial role in this alternative approach, in which the agent is asked to state preferences among bets rather than gambles. (Bets are, unlike gambles, purchased. So if a bet is fair the agent has no reason to prefer one bet over another.) The Dutch Book Theorem was first touched upon by Ramsey, but in the contemporary literature it is intimately associated with the work of de Finetti.

Put in Ramsey’s words, the Dutch Book Theorem proves that ‘If anyone’s mental condition violated these laws [the probability axioms] . . . [h]e could have a book made against him by a cunning better and would then stand to lose in any event.’ That is, if the agent is prepared to place fair bets on a sufficiently large number of events and if there is no bet that will yield a sure loss no matter what happens, then the agent’s betting dispositions can be represented by a real-valued function that obeys the axioms of the probability calculus. A common interpretation of this result is that coherent betting dispositions are identical with subjective probabilities.

The argument to be spelled out in the next section is valid no matter whether Savage’s or DeFinetti’s version of the Bayesian framework is preferred. The central assumption both theories have in common, which will be attacked here, is that an action guiding probability function can be defined in terms of preferences over some sort of uncertain prospects.

3. Do Bayesians put the cart before the horse?

The Bayesian approach to subjective probability is not action guiding. Briefly put, the point is that Bayesians put the cart before the horse from the point of view of the deliberating decision maker: An agent who is able to state preferences over a set of uncertain prospects already knows what to do. Therefore, a Bayesian agent does not get any new, action-guiding information from the theory. In what follows, I shall articulate what I take to be the most forceful version of this argument. Similar, but less detailed, concerns have been raised by others. In

---

3Ramsey (1926:182).

In order to spell out the argument in more detail, it is helpful to imagine two agents, A and B, who are exactly parallel with respect to their psychological make-up. They hold the same beliefs about past, present, and future events, and they like and dislike the same movies, the same food, the same wines, etc. For instance, they both believe, to the same degree, that it will rain tomorrow, and they also dislike this to the same degree. There is, however, one important difference. Agent A is able to express her preferences over gambles in the way required by Bayesians. Therefore, for a very large set of gambles agent A knows whether she prefers one gamble to another or is indifferent between them; and of course, her preferences conform to the rest of the axioms employed in this approach as well.

Agent B is more like the rest of us, so in most cases she does not know if she prefers one gamble to another. However, B’s inability to express preferences over gambles is not due to any odd structure of her tastes and beliefs; rather, it is just a matter of an insufficient capacity to process large amounts of information (so her preferences conform to the axioms proposed by advocates of the Bayesian approach in an implicit sense). Since B’s beliefs (and tastes) are exactly parallel to those of A, it follows that B’s gambling-behavior ought to be exactly parallel to A’s.

Agent A is designed to be the kind of highly idealized, rational agent described by Bayesians. That is, A has all of the properties these theorists wish we all had. Now suppose you are A and ask yourself: What advice might I receive from a Bayesian theory of subjective probability? Do I get any ‘new’ information that can be used for calculating what acts or gambles to choose, that is, does the theory provide me with any action-guidance? The subjective probability function returned by the theory, which can be used for describing you as if you were maximizing subjective expected utility, is obtained by reasoning backwards. Let us imagine that you preferred a gamble in which you receive ten dollars if a republican wins the next election (and nothing otherwise), to a gamble in which you receive the same amount if a democrat wins (and nothing otherwise). Then the theory ‘reveals’ that your subjective degree of belief in the first event is higher than your belief in the latter. Suppose that by offering you a number of additional gambles we find that your gambling behavior conforms to the expected utility principle,

---

4See, for example, Savage (1954:7).
given that we ascribe a subjective probability of .6 to the event \( R = \text{the next president is a republican} \) and a subjective probability of .4 to the event \( D = \text{the next president is a democrat} \), as well as utilities that are linear in dollar amounts. Can the probabilities established in this way be used for making choices among gambles on events \( R \) and \( D \)? Suppose that you are to decide between the gamble in which you win ten dollars if \( R \) takes place (and nothing otherwise) and the gamble in which you win the same amount if \( D \) takes place (and nothing otherwise). Since the possible outcomes of both gambles are identical, your subjective probabilities will determine your choice. However, it would of course be ridiculous to decide between the two gambles by using the subjective probabilities derived above, since those probabilities were established by asking you to directly state a preference between the same gambles.

The moral of this example is simple. If your preferences over gambles are used as input data in the construction of a subjective probability function, then those probabilities merely reflect what was already known about your preferences. Thus, there is little point in using subjective probabilities derived from preferences over gambles for guiding your preferences over gambles which you have already made up your mind about.

Let us now consider what happens if you apply the probabilities established above for choosing among some other gambles. For example, imagine that you are offered a choice between a gamble in which you win twenty dollars if event \( R \) occurs and a fair coin lands heads (and nothing otherwise), and a gamble in which you win five dollars if event \( D \) occurs (and nothing otherwise). In this case it would be pointless to decide between the new gambles based on your subjective probabilities for \( R \) and \( D \). In order to see why, remember that the probabilities for \( R \) and \( D \) were derived by not only considering your preference between the two initial gambles. You were also offered a number of additional gambles. As carefully explained by Savage, the set of gambles you have to state preferences over includes all possible gambles in the world under consideration. More precisely, Savage’s completeness axiom (Axiom 1) requires that the agent’s preferences over all possible gambles is a weak order.\(^5\) Therefore, your preference between the new gambles was already known to you when you established your subjective probability for \( R \) and \( D \). The conclusion is that advocates of the Bayesian approach take too much for granted. What they use as input data to their theory is what decision theorists would like to obtain as

output. In that sense, Bayesians put the cart before the horse from the point of view of the deliberating agent.

Now consider the non-ideal agent B. For B the situation is slightly different. It is of course true that B ought to behave as A would have behaved, but since A’s preferences are not accessible for B, B cannot simply use A as a tool for figuring out what to do. However, despite this, it is commonly assumed that a representation theorem is normatively relevant for a non-ideal agent in more or less indirect ways. For example, Savage claimed that, ‘According to the [Bayesian approach], the role of the mathematical theory of probability is to enable the person using it to detect inconsistencies in his own real or envisaged behavior.’\(^6\) Let us see how this may work.

First, suppose that B has access to some of her preferences among gambles, but not all, and also assume that she has partial information about her probability and utility function. Then, the Bayesian representation theorems can be put to work for ‘filling the missing gaps’ of the preference ordering, by using the initially incomplete information for reasoning back and forth, thereby making the preference ordering less incomplete. This is essentially what Joyce (1999) has in mind when he speaks about incomplete preference that can be ‘coherently extended’ to a complete ordering; but see also Hansson (1988:143-4). In this extension process, some preferences might be found to be incoherent with the initial preference ordering, and for this reason be ruled out as illegitimate.

It should be admitted that for B, the Bayesian approach is to some extent action guiding. Since the Bayesian axioms prevent B from forming whatever new preferences he likes, he gets some ‘negative’ action guidance from the theory. However, two important problems remain. First, when forming new preferences, B will frequently face situations in which the existing preferences and the Bayesian axioms do not uniquely determine the new preferences. Hence, the new preferences acquired by B are to a certain extent arbitrary. From an action guiding perspective this is obviously problematic. Second, even if the initial information happens to be sufficiently rich for filling the gaps, this manoeuvre offers no justification of the preference ordering. How were the agent’s preferences obtained, and why not revise all of them, no matter whether they are inconsistent or not?\(^7\)

\(^6\)Savage (1954:57).
\(^7\)Hansson (1988:143-4) makes the same point.
In summary, you are either an ideal agent, or a non-ideal one. In the first case, you do not get any new information out of the Bayesian approach, because the relevant information (your preferences over gambles) actually served as input data to the theory. In the latter case you merely get partial ‘negative’ information out of the theory. Of course, these problems do not show that every insight learned from the Bayesian approach to subjective probability is mistaken, but it seems fair to say that it legitimizes the discussion of an alternative approach.

4. A non-Bayesian alternative

A number of mathematicians and statisticians have pointed out that degrees of belief can be measured without considering preferences over uncertain prospects, and without making use of desires or any other evaluative concepts. The most prominent example of such a non-Bayesian approach is DeGroot (1970). His theory starts from the assumption that agents can make qualitative comparisons of whether one event is more likely to occur than another. DeGroot thereafter shows that if the agent’s qualitative judgements are sufficiently fine-grained and satisfy a number of structural conditions, then there exists a unique probability function that assigns real numbers between 0 and 1 to all events, such that one event is judged to be more likely than another if and only if it is assigned a higher number. So in DeGroot’s theory, probabilities are obtained by fine-tuning qualitative data, thereby making them quantitative. The probabilistic information is present from the beginning, but after putting qualitative information to work the theory becomes quantitative.

In what follows I shall develop what I take to be an improved version of DeGroot’s theory. Most of the technical framework will remain the same, but the set of objects among which the agent is supposed to compare is drastically reduced: In the improved theory the agent merely has to consider beliefs about his own choices. The basic set-up of the new theory is as follows. Consider a hypothetical ‘horse lottery’, in which you win $100 dollars if and only if you bet on the winning horse:

<table>
<thead>
<tr>
<th></th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lottery 1</td>
<td>$100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lottery 2</td>
<td>0</td>
<td>$100</td>
<td>0</td>
</tr>
<tr>
<td>Lottery 3</td>
<td>0</td>
<td>0</td>
<td>$100</td>
</tr>
</tbody>
</table>

TABLE 1

8See also Koopman (1940) and Good (1950).
Suppose that the agent’s utility for money is state-independent, i.e. that $100 is worth equally much no matter which state occurs. In this set-up, the agent’s choice will presumably be governed exclusively by which state he thinks is most likely to occur: Lottery 1 will be chosen over lottery 2 if and only if the agent considers state 1 to be more likely to occur than state 2, and so on. Now, in order to totally eliminate any reference to a desire for money the following step will be taken: Suppose that the agent is not making the choice for himself, but is rather advising a friend about which horse lottery he should enter. For example, the agent may say that, ‘If you prefer $100 over $0, then I recommend you to choose lottery 1 over lottery 2’. The truth of this statement does not depend on the strength of the speaker’s desires. Despite this, the statement obviously carries information about the speaker’s subjective probability. He thinks that state 1 is more likely to occur than state 2. Moreover, in order to avoid the conditional form of the recommendation, one may prefer to consider conjunctions such as, ‘I believe that you prefer $100 over $0, and I therefore recommend you to choose lottery 1 over lottery 2’. This is a purely factual report about the speaker’s own beliefs, i.e. the statement expresses his subjective probability, but is says nothing about the speaker’s desires.

From a behavioristic point of view it can be argued that subjective probability ought to be linked to observable choices, rather than merely linguistic behaviour. Let us assume, for the sake of the argument, that this objection is valid. Then, in order to meet the behavioristic requirement, one may assume that the agent in not only advising his friend, but is in fact also making the decision on his behalf. This is to say that the agent decides which lottery his friend will enter based on (i) his own beliefs about the world, and (ii) his beliefs about his friend’s desires. Of course, the agent’s choice will be influenced by his desire to help his friend, but this desire is constant and will not affect his choice behaviour. Therefore, the agent’s choice behaviour can be interpreted as a direct report of his subjective probability.

It is important to keep in mind that we only consider horse lotteries, i.e. lotteries in which the agent’s friend will either win some fixed amount of money or nothing. This means that the agent does not have to make any assumption about the relative strength of his friend’s desire for money—it is sufficient to assume that $100 is worth more than $0. Compare this set-up with the analogous theory based on

---

9As pointed out by Schervish et al (1990), this assumption does not always hold true: If the agent lives in Japan and therefore needs to exchange dollars for yen, and the three states are different exchange-rates, this assumption does not hold true.
Savage’s axioms. In Savage’s set-up it would make little sense to ask
the agent to make choices on behalf of a friend. This is because there
is no way in which one could know the strength of the friend’s desire
for various amounts of money and other goods. The utility functions
of others are in general unaccessible, so it would make no sense to try
to separate beliefs from desires in the way proposed above.

The theory outlined here can easily be formalized. Let $S = \{s_1, s_2, \ldots\}$
denote a set of states to which the agent seeks to assign subjective
probabilities, and let $L$ be a set of horse lotteries corresponding to the
elements of $S$.\footnote{We assume that $S$ meets the following requirements: (i) $s \in S$ (ii) If $s \in S$, then $s^c \in S$ (iii) If $S_1, S_2, \ldots$ is an infinite sequence of sets from $S$, then $\bigcup_{i=1}^{\infty} S_i \in S$.} $L$ is constructed from the elements of $S$ as described
in Table 1. Statisticians sometimes prefer to assign probabilities to
events rather than states. However, note that states can be thought
of as events that occur, and vice versa. As usual, $s_1 \cup s_2$ denotes the
union of $s_1$ and $s_2$, i.e. the complex state in which at least one of
the two states $s_1$ or $s_2$ obtains. $s_1 \cap s_2$ is the intersection of the two
states, i.e. the state consisting of the elements that the two states
have in common.\footnote{An example might help to explain this. If $s_1$ is the state in which the die lands
showing either 1, 2 or 3, and $s_2$ is the state in which the die lands showing either
3, 4 or 5, the intersection is the state in which the die lands showing 3.} A pair of states is mutually exclusive if and only if
$s_1 \cap s_2 = \emptyset$.

Note that each state corresponds to a horse lottery in $L$. For in-
stance, $s_1 \cup s_2$ corresponds to the lottery $l_1 \cup l_2$, i.e. the lottery in which
you win if at least one of the two states $s_1$ or $s_2$ materializes. Hence,
to perform a set-theoretic operation on horse lotteries is equivalent to
performing the corresponding operation on the underlying states.

The relation ‘is chosen by the agent rather than’ is a non-evaluative
binary relation between lotteries in $L$. $l_1 \succ l_2$ means that $l_1$ is chosen
by the agent rather than $l_2$, and $l_1 \sim l_2$ means that both lotteries are
judged to be equally choice-worthy by the agent. Furthermore, $l_1 \succeq l_2$
is an abbreviation of ‘either $l_1 \succ l_2$ or $l_1 \sim l_2$, but not both’. Let us
now consider the following five axioms, which are supposed to hold for
all elements $l_x$ and $l_y$ in $L$:

\textbf{Axiom 1.} $L \succ \emptyset$ and $l_x \succeq \emptyset$.  

The leftmost part of Axiom 1 articulates the plausible assumption that
the union of all lotteries in $L$ would be chosen over no horse lottery at
all, whereas the rightmost part holds that each individual lottery in $L$
is at least as choice-worthy as no horse lottery at all.
Axiom 2. For any two horse lotteries $l_x$ and $l_y$, exactly one of the following three relations hold: $l_x \succ l_y, l_y \succ l_x, l_x \sim l_y$.

Axiom 2 is a completeness axiom. It is not entirely uncontroversial. For example, if the set of states contains extremely disparate elements, such as ‘rain here toady’ and ‘all humans go extinct soon’, some agents will perhaps argue that it is impossible to make a choice among the corresponding horse lotteries on behalf of a friend. However, in response to this, note that the axioms spelled out here are supposed to be normative requirements for rational agents. If you are facing a decision in which the probabilities for very different events are relevant, it is not unreasonable to require that an ideal agent should be able to compare them. Of course, Axiom 2 resembles the ordering axiom in the Bayesian approach, according to which rational agents must be able to rank any set of alternative risky acts without (explicitly) knowing the probabilities and utilities associated with the potential outcomes. However, note that Axiom 2 is less demanding than its Bayesian counterpart, since it merely holds for horse lotteries.

Axiom 3. If $l_x, l_y$ and $m_x, m_y$ are four horse lotteries such that $l_x \cap l_y = m_x \cap m_y = \emptyset$ and $m_i \succeq l_i$ for $i = 1, 2$, then $m_x \cup m_y \succeq l_x \cup l_y$. If, in addition, either $m_x \succ l_x$ or $m_y \succ l_y$, then $m_x \cup m_y \succ l_x \cup l_y$.

In order to explain Axiom 3, suppose that some states can obtain in either of two mutually exclusive ways, for example (1) ‘the coin lands heads and it rains’ and (2) ‘the coin lands tails and it rains’; and (3) ‘the coin lands heads and the sun shines’ and (4) ‘the coin lands tails and the sun shines’. Then, if the agent considers (3) to be more likely than (1) and (4) to be more likely than (2), he will conclude that it is more likely that the sun is going to shine, and will of course therefore make his recommendations for his friend accordingly.

Axiom 4. If $l_{x_1} \supset l_{x_2} \supset \ldots$ and $m$ is some horse lottery such that $l_i \succeq m$ for $i = 1, 2, \ldots$, then $\bigcap_{i=1}^{\infty} l_i \succeq m$.

Axiom 4 guarantees that the probability function is countably additive. It follows from Axiom 4 that $m \sim \emptyset$, because no matter how unlikely $m$ is to yield a prize ($m \succ \emptyset$), it is impossible that for every $n$ between one and infinity, the intersection $\bigcap_{i=1}^{\infty} l_i$ is at least as likely as $m$ to yield a prize, given that each $l_n$ is more likely to yield a prize than $l_{n+1}$. For an intuitive interpretation, suppose that each $l_n$ corresponds to a state $s_n$ that denotes an interval on the real line between $n$ and infinity.

Axiom 5. The agent believes that there exists a random variable which has a uniform distribution on the interval $[0, 1]$. 
Axiom 5 is different from the other axioms. Of course, what agents believe cannot be directly observed, so in order to verify this axiom one has to observe what recommendations are stated by the agent. In order to understand what work is carried out by Axiom 5, suppose that an agent wishes to determine her subjective probability for the two states 'rain here within an hour' and 'no rain here within an hour'. Then, since the set of states $S$ only contains two elements, it is not possible to obtain a quantitative probability function by only comparing lotteries corresponding to those two events. The set of states (and hence the set of lotteries) has to be extended in some way. Axiom 5 is the key to this extension. In a uniform probability distribution all elements (values) are equally likely to occur. As an example, think of a roulette wheel in which the original numbers have been replaced with an infinite number of points in the interval $[0,1]$. Then, by applying Axiom 5 the set of states can be extended to the union of the two original states and the infinite set of states 'the wheel stops at $x$ ($0 \leq x \leq 1$'), etc. Now consider the following theorem.

**Theorem 1.** Axioms 1-5 are jointly sufficient and necessary for the existence of a unique probability function $p$ that assigns a real number in the interval $[0,1]$ to all elements in a set of states $\{s_x, s_y, \ldots\}$, such that $l_x \succeq l_y$ if and only if $p(s_x) \geq p(s_y)$.

**Proof.** A convenient way to prove Theorem 1 is to start from DeGroot's (1970:79-81) analogous theorem. This means that we shall first show that Axioms 1 - 5 imply the existence of a unique probability function, and thereafter verify that the function in question satisfies the axioms of the probability calculus.

Part (1): The subjective probability function $p$ is constructed by first applying Axiom 5, according to which there exists a random variable which has a uniform distribution on the interval $[0,1]$. Let $G[a,b]$ denote the event that the random variable $x$ lies in the interval $[a,b]$. Then consider the following lemma: If $l_x$ is any element in $L$, then there exists a unique number $a^*$ ($1 \geq a^* \geq 0$) such that $x \sim G[0,a^*]$. (For a proof, see DeGroot 1970:77-78.) Now, if $l_x$ is any element in $L$ we can apply the lemma and let $p(l_x)$ be defined as the number $a^*$. Hence, $l_x \sim G[0,p(l_x)]$. It follows that $l_x \succeq l_y$ if and only if $p(l_x) \geq p(l_y)$, since $l_x \succeq l_y$ if and only if $G[0,p(l_x)] \succeq G[0,p(l_y)]$.

---

12Note that the axiom does not require that the random variable in question really exists. It is sufficient that the agent believes that a random variable exists. The non-Bayesian approach makes no assumption about the nature of the external world; all that matters is the structure of internal subjective believes. Axiom 5 is thus consistent with the fact that the world might be deterministic.
Part (2): A probability function $p$ is characterized by the following conditions: (i) $p(l_x) \geq 0$ for all $l_x$, (ii) $p(L) = 1$, (iii) $p(\bigcup_{i=1}^n l_i) = \sum_{i=1}^n p(l_i)$. In order to show that these conditions are fulfilled, note that it follows from the definition of $p$ above that $p(l_x) \geq 0$ for every $l_x$. This verifies (i). Moreover, $L = G[0, 1]$, which entails that $p(L) = 1$; this verifies (ii). In order to verify (iii), we have to show that $p(\bigcup_{i=1}^n l_i) = \sum_{i=1}^n p(l_i)$. To start with, consider the binary case with only two elements $l_{x1}$ and $l_{x2}$; that is, we want to show that if $l_{x1} \cup l_{x2} = \emptyset$, then $p(l_{x1} \cup l_{x2}) = p(l_{x1}) + p(l_{x2})$. In the first part of the proof we showed that $l_{x1} \sim G[0, p(l_{x1})]$. Hence, $l_{x1} \cup l_{x2} \sim G[0, p(l_{x1} \cup l_{x2})]$. According to a lemma proved by DeGroot (1970:79) it also holds that $L' \sim G[p(l_{x1}), p(l_{x1} \cup x_2)]$. Now, note that $G[p(l_{x1}), p(l_{x1} \cup l_{x2})] \sim G[0, p(l_{x1} \cup l_{x2}) - p(l_{x1})]$. Also note that by definition $L' \sim [0, p(L')]$. Hence, $p(l_{x1} \cup l_{x2}) - p(l_{x1}) = p(l_{x2})$. By induction this result can be generalized to hold for any finite number of disjoint elements. □

5. Subjectivism and the meaning of probability

Bayesians and non-Bayesians disagree about the meaning of probability. Bayesians define probability by making certain claims about preferences over uncertain prospects. However, the relation between partial degrees of belief and preferences may not be as tight as they think. A preference invariably requires a desire and a belief. If I tell the waiter that I prefer tuna over salmon, then I believe that making this statement will lead to an outcome that I desire more than the alternative outcome. However, the converse does not hold. An agent can certainly believe something to a certain degree without having any preferences. The independence of beliefs and desires was stressed already by Hume, and few contemporary philosophers are willing to entirely reject his idea. Therefore, it seems natural to ask: Why on Earth should a theory of subjective probability involve assumptions about preferences, given that preferences and beliefs are separate entities? Contrary to what is claimed by Ramsey, Savage and DeFinetti, emotionally inert decision makers failing to muster any preference at all (i.e. people that have no desires) could certainly hold partial beliefs.

Here is a slightly different way of putting this argument. If the Bayesians are right, a decision maker cannot hold a partial belief unless he also has a number of desires, all of which are manifested as preferences over uncertain prospects. However, contrary to what is claimed by Ramsey, Savage and DeFinetti, emotionally inert agents could also
hold partial beliefs. The assumption that there is a necessary link between probabilities and desires is therefore dubious, and this shows that the meaning of probability is not captured by the Bayesian analysis.

At this point Ramsey, Savage and DeFinetti may of course object that they have never attempted to analyse the meaning of probability. All they seek to do is to offer a procedure for explaining and predicting human behaviour. However, if this line of reasoning is accepted we must look elsewhere for a definition of probability. The non-Bayesian definition of probability is silent about preferences over uncertain prospects, which makes it philosophically more attractive. In addition, the new theory can help indecisive agents to guide their decisions.

Appendix A: Savage’s axioms

Let $S = \{ s, s', \ldots \}$ be a set of states with subsets $A, B, \ldots$, and let $X = \{ x, x', \ldots \}$ be a set of consequences. Acts are conceived of as functions $f, g, \ldots$, from $S$ to $X$. Let $\geq$ be a preference relation between pairs of acts. We stipulate that $f$ and $g$ agree with each other in $B$ just in case $f(s) = g(s)$ for all $s \in B$. Furthermore, $f \succeq g$ given $B$, if and only if, if it were known that $B$ does not obtain, then $f$ is weakly preferred to $g$; and $B$ is null, if and only if, $f \succeq g$ given $B$ for every $f, g$. Finally, $A$ is not more probable than $B$ (abbreviated $A \leq B$) if and only if $f_A \succeq f_B$ or $x \succeq x'$, for every $f_A, f_B, x, x'$ such that: $f_A(s) = x$ for $s \in A, f_B(s) = x'$ for $s \in \neg A, f_B(s) = x$, for $s \in B, f_B(s) = x'$, for $s \in \neg B$. Now, Savage proposes the following axioms:

SAV 1. $\geq$ is complete and transitive.

SAV 2. If $f, g,$ and $f', g'$ are such that:
   
   (1): in $\neg B$, $f$ agrees with $g$, and $f'$ agrees with $g'$,
   (2): in $B$, $f$ agrees with $f'$, and $g$ agrees with $g'$,
   (3): $f \succeq g$;

then $f' \succeq g'$

SAV 3. If $f(s) = x, f'(s) = x'$ for every $s \in B$, and $B$ is not null, then $f \succeq f'$ given $B$, if and only if, $x \succeq x'$.

SAV 4. For every $A$ and $B$: $A \geq B$ or $B \geq A$.

SAV 5. It is false that, for every $x, x'$: $x \succeq x'$.

SAV 6. Suppose it false that $f \succeq g$; then, for every $x$, there is a (finite) partition of $S$ such that, if $g'$ agrees with $g$ and $f'$ agrees with $f$ except on an arbitrary element of the partition, $g'$ and $f'$ being equal to $x(B)$ there, then it will be false that $f' \succeq g$ or $f \succeq g'$.
REFERENCES


*E-mail address: mbp24@cam.ac.uk*