6

Subjectivistic Interpretations of Probability

The most recent and one of the most interesting interpretations of probability is the subjectivistic one: According to the subjectivistic point of view probabilities concern actual degrees of belief. Although this interpretation of probability was presented for the first time in 1926 by F. P. Ramsey, and was first claimed to be important for statistics in 1937 by Bruno de Finetti, the subjectivistic interpretation of probability had no great effects on English-speaking statisticians until the publication, in 1954, of L. J. Savage's book *The Foundations of Statistics*. Since that time the interpretation has had a considerable impact on statistical practice, though it does not by any means outweigh or replace the classical frequentist orientation of most statisticians.

The degrees of belief with which probability is concerned, according to the subjectivistic interpretation, are the actual subjective degrees of belief of people: the starting point is just the set of degrees of belief that they have.

Degrees of belief are to be interpreted behavioristically. Ramsey first proposed that degrees of belief be measured by betting odds: if one is willing to bet at odds of 1:5 on the occurrence of a three on the roll of a die, but at no higher odds, then one's degree of belief is $1/(1 + 5) = \frac{1}{6}$. As Ramsey also pointed out, there are ways of avoiding the difficulties (e.g., of decreasing marginal utility) of money bets. In fact Ramsey devised an ingenious technique by means of which both the evaluation of utilities and the evaluation of probabilities could be carried out simultaneously on the basis of a single proposition believed to the degree $\frac{1}{2}$. In a more formal style, Savage has worked out a whole system of subjective probability on the basis of the simple relation of *preference* between acts. For example, if a person prefers to stake a possible gain on the occurrence of an event *E*, rather than to stake a possible gain on the occurrence of an event *F*, then, for that person, the event *E* is more probable than the event *F*.

Although the subjectivistic theory takes as its starting point people's actual degrees of belief, it is not held that all degrees of belief in a group of propositions are equally acceptable. Some distributions of belief allow a book to be made against the holder of the beliefs. For a person to have a book against him is for him to accept a set of bets such that he cannot win, whatever happens. For example, if a person had a degree of belief equal to $\frac{4}{5}$ in the event *E*, he would, on our behavioristic interpretation, be willing to bet at odds of 4:1 in favor of the occurrence of *E*. On the other hand, if he had a degree of belief equal to $\frac{2}{5}$ in the non-occurrence of *E*, he would be willing to bet at odds of 2:3 against the occurrence of *E*. Now let him make both of these bets in dollars. Then if *E* does not occur, he wins the second bet (receiving three dollars) and loses the first bet (paying out four dollars); and if *E* does occur, he wins the first bet (receiving one dollar) but loses the second bet (paying out two dollars). In either case, then, he loses one dollar. In gambling parlance, he has had a book made against him.

Now it is possible to show that in these circumstances the person can avoid being in a position to have a book made against him by apportioning his degrees of belief in the occurrence of *E* and in the occurrence of not-*E* so that the sum of these degrees of belief is 1. A necessary and sufficient condition of not being in a position to have a book made against you on a single event *E*, is that your degree of belief in *E* and in the denial of *E*, as reflected in the least betting odds you will accept, add up to 1. In general it can be shown that a necessary and sufficient condition of not being in a position to have a book made against you is that your degrees of belief in any set of statements satisfy the axioms of the probability calculus. (This has been called the *Dutch Book Theorem* by Isaac Levi.)

A set of degrees of belief in a set of propositions (or statements, or events) is called *coherent* if and only if those degrees satisfy the axioms of the probability calculus. We may now express the Dutch Book Theorem as follows: a necessary and sufficient condition of not being in a position to have a book made against one is that one's degrees of belief be coherent.

Since not everybody's body of beliefs is coherent, the theory of probability cannot be regarded simply as a description of the way people do in fact distribute their degrees of belief. It represents an ideal of rationality: people *ought* to distribute their beliefs in accordance with the rules of the probability calculus. There is thus a close kinship between the logical view and the subjectivistic view of probability; but on the latter view there are no conditions of rationality that are to be imposed on people's degrees of belief above and beyond those embodied in the calculus of probability itself.

Savage, indeed, argues that the calculus of probability is simply a complex criterion of consistency. The person who finds himself in the position of violating theorems of the calculus is in roughly the same situation as the person whose beliefs have been shown to be logically inconsistent. He must modify his beliefs, but in neither case does the standard (the calculus of probability in the former case; the deductive calculus of logic in the latter) tell him the specific manner in which he should modify his beliefs. He must make some adjustment; but logic cannot tell him which.

In "The Aim of Inductive Logic," Carnap discusses the matter in more detail. He distinguishes between an *actual* credence function which is a theoretical property of an individual that, together with his utility function (also a theoretical property), provides a psychological explanation for his actions and decisions, and a *rational* credence function, which is taken to be the credence function of a perfectly rational being. A perfectly rational being will not have a book made against him; therefore his credence function at a given time will be coherent. Furthermore, if E is the observational data received by a perfectly rational being between time T_n and time T_{n+1}, then his credence functions Cr_n and Cr_{n+1} will be related by the relation that for any H,

$$\mathrm{Cr}_{n+1}(H) = \frac{\mathrm{Cr}_n(E \cap H)}{\mathrm{Cr}_n(E)} = \mathrm{Cr}_n(H|E).$$

It is but a step from this consideration to the supposition that there is a credence function for time zero, Cr_0 such that, if A represents the perfectly rational beings's total observational knowledge up to time n, his credence function at time n for any H will be $\mathrm{Cr}_0(H|A)$. A credibility function, for Carnap, is simply a generalized conditional credence function: $\mathrm{Cred}(H|A)$ is defined for an H and all consistent A, while $\mathrm{Cr}_0(H|A)$ merely represents the perfectly rational being's credence when in fact his knowledge is A. Then, as we saw in the last chapter, Carnap goes on to impose further conditions (invariance conditions, for example) of the forma counterparts of the credibility function. But the serious subjectivist will not follow him in imposing any more conditions on the credence function than are imposed by the axioms of probability themselves.

One of the most attractive things about the subjectivistic interpretation of probability is that it provides a complete rationale for induction, or, more specifically, for statistical inference. In this regard the crucial theorem of subjectivistic approach is Bayes' Theorem; indeed, among statisticians the subjectivistic theory is often called 'Bayesian'. This is slightly misleading, since even frequentists are often, in virtue of statistical background knowledge, in a position to use Bayes' Theorem.

As an example of the kind of inference that admits of a subjectivistic interpretation, but only in the most strained way a frequentist interpretation, consider the following: Let H be the statement that a certain holograph manuscript, newly come to light in the attic of an old house in Provence, is a hitherto undiscovered poem by Dante Alighieri. Let E be the statement that Dante was well known to the family that lived in that house at the time that Dante was in Provence. A person might well take the credibility of H, to be quite small, say 0.2; and similarly the probability of E might be quite small, say 0.2. The conditional credibility of E, given H, however, might quite justifiably be large: say 0.7. Thus, as subjectivists, we can calculate the conditional credibility of H, given E: it is

$$\mathrm{Cr}(H|E) = \frac{\mathrm{Cr}(H)\mathrm{Cr}(E|H)}{\mathrm{Cr}(E)} = \frac{0.2 \times 0.7}{0.2} = 0.7.$$

Now suppose we find that E is true. The subjectivist will say that we should assign a credibility of 0.7 to it. The frequentist will say that the whole analysis is nonsense, and that we might just as well admit flat out that our beliefs have changed, and not pretend to give reasons for the change.

And indeed there is a problem in the use of Bayes' Theorem, for the true subjectivist. If one coherent credence function is as good as another, a single person at two different times has as much right to two credence functions as two people at the same time. At time T_1 my credence function may be Cr_1, and my conditional credence function for the hypothesis H on the evidence E may be $\mathrm{Cr}_1(H|E) = \mathrm{Cr}_1(H \cap E)/\mathrm{Cr}_1(E)$. But suppose at T_2 I have observed E (and that that is all I have observed since T_1); the value of my new credence function for the hypothesis H will be $\mathrm{Cr}_2(H)$. The subjectivist would like to say that $\mathrm{Cr}_2(H)$ *ought* to be equal to $\mathrm{Cr}_1(H|E)$. Carnap says this, but Carnap is no true subjectivist; in the back of his mind he still entertains the belief that there is only one rational credibility function. The true subjectivist, however, can say no more than that most people feel uncomfortable about switching their credibility functions in midstream, as it were, and that his intent is to provide guidance for those who wish to remain consistent through time, as well as coherent at a given time.

Richard Jeffrey, in *The Logic of Decision*, has offered a more general approach to the problem of changing beliefs. One of the problems with the standard subjectivistic (and also the logical) approach is that experience is taken to render evidence statements completely certain: $\mathrm{Cr}_2(E)$ is taken to be $\mathrm{Cr}_1(E|E) = 1$. But often we should take account of evidence which is uncertain. When $\mathrm{Cr}_2(E)$ merely differs from $\mathrm{Cr}_1(E)$, Jeffrey suggests we compute our new credence function as follows:

$$\mathrm{Cr}_2(H) = \mathrm{Cr}_1(H|E)\,\mathrm{Cr}_2(E) + \mathrm{Cr}_1(H|\bar{E})\,\mathrm{Cr}_2(\bar{E}).$$

Again, the conditional probabilities are supposed not to change. It is hard to see what rules that possibility out. The difference between $Cr_1(E)$ and $Cr_2(E)$ is supposed to be a result of experience. But it might also result from a change in blood sugar. And it is perfectly possible that certain physiological changes—or perhaps mere reflection—should change the value of the credence function for H—say, a general hypothesis—instead of its value for an 'observation' such as E. It is one thing to take the Dutch Book Theorem as an argument for having a credence function which is coherent at any given time; it is quite another thing to use either Bayes' Theorem of Jeffrey's Theorem to change beliefs.

Much of the appeal of Bayesian statistics to statisticians stems from the fact that *a priori* probabilities or distributions that are fed into Bayes' Theorem, which represent the most subjective element of the theory, often turn out to be less and less important as the empirical evidence provided by observed frequencies or observed distributions increases. The subjective element is out in the open, they say, but the actual results of the statistical inference often do not depend in any quantitatively significant way on that subjective element.

This phenomenon—the swamping of the influence of prior probabilities by that of empirical observations—has not gone unnoticed. Even an extreme frequency theorist, von Mises, makes a point of remarking on it, in order to argue that Bayes' Theorem is useful in statistics. He makes much of the fact, which is central to all of statistical inference from the subjectivist's point of view, that often whatever probability assessments or estimates you start with are very quickly overwhelmed by the importance of empirical evidence. Reichenbach, too, makes important use of Bayes' Theorem. But most frequency-oriented statisticians in recent years have eschewed the use of Bayes' Theorem, except where there are available known prior probabilities that can be interpreted empirically, primarily on the ground that lacking this sort of prior information one could only assign prior probabilities subjectively, and that this subjectivity would therefore infect all of their results. (Here is a particularly clear example of the utter lack of communication between members of opposing camps, even when they use the same words.)

Three kinds of objections have been brought against the subjectivistic interpretation of probability; in view of the fact that there are those who accept the subjectivistic interpretation, it may be inferred that these objections are not decisive. The first two objections apply as well to the logical theories of the preceding chapter; the third applies only to the subjectivistic theory.

One thing which has been held to be implausible about either the subjectivistic theory or the logical program for a theory is that degrees of belief are regarded as being completely precise, like betting ratios. The logical theory has the difficulty that it legislates (or intends to legislate) in complete detail about degrees of belief in the absence of any evidence at all. The subjectivistic theory has the difficulty that any such degree of belief is permissible, so long as it does not conflict with other degrees of belief. On either theory, we must suppose that the probability of any event—such as that of finding furry animals on Mars, or that of experiencing rain in Detroit on Tuesday—can be evaluated to any desired number of decimal places. On neither theory is there any intrinsic distinction between probabilities that are based on the careful assessment of extensive evidence and probabilities that are (logically or psychologically) *a priori*. I take the probability of heads on the first toss of a coin that you casually pull out of your pocket to be one-half; and after performing extensive mechanical tests, static and dynamic, and after running a long series of trials, and subjecting them to extensive statistical tests, I conclude that the probability of heads on the next toss of that coin is one-half. On either the logical or the subjectivistic theory, this means that the net effect of all that work is zero; the probability is unchanged. A probability of one-half is just a probability of one-half.

This objection has to some extent been met by recent developments in the foundations of the theory (formalism plus interpretation) of probability. I. J. Good and C. A. B. Smith have both proposed generalizations of the subjectivistic theory in which one speaks of upper and lower probabilities.

A second objection, also applicable both to the logical and to the subjectivistic interpretation, is that both interpretations allow the derivation, on the basis of purely *a priori* probabilities, of very high probabilities concerning long sequences of future events. For example, let us consider a sequence of balls drawn with replacement from an urn. Let the probability that the first ball is purple be 0.01, on either logical or subjectivistic grounds, and similarly, let the probability that the second ball is purple, given that the first one is purple, be 0.02. Given the usual conditions regarding draws from urns, it is then possible to show that the probability is 0.99 that an arbitrarily long sequence of draws will produce less than 10% purple balls, and to show this, furthermore, on *a priori* grounds alone. The probability is 0.9996 that less than half the balls will be purple. One does not expect to obtain such high probabilities concerning such generalizations on no evidence at all; something seems to have gone wrong.

The subjectivist's answer is a shrug of the shoulders: "All this shows is that your original dispositions (degrees of belief), consistently carried out, lead to some surprising conclusions. But there is nothing unusual in that state of affairs in mathematics or in philosophy. If you find the conclusions intolerable, change the premises." The logical theorist will have much the same answer; he may regard the conclusion simply as a surprising consequence of initial statements that he takes to be incontrovertible; or else he will take the surprising conclusion to throw doubt on the initial measure function which led to that conclusion.

The third objection, which applies only to the subjectivist theory, is already suggested by the subjectivist's answer to the second objection: if you find the conclusion intolerable, change the premises. Since on the subjectivistic interpretation of probability I can adjust my probabilities in any way that I want to, provided only that they satisfy the rules of the calculus, there is nothing in principle that precludes my arranging my degrees of belief in such a way that I attribute a high degree of belief to the things that fit in with my preconceptions, and a low degree of belief to the things that don't fit in. That is: The relation of evidence to conclusion itself depends on the probability assignments that I make; I can in principle adjust these assignments so that the evidence supports (or does not badly undermine) the hypotheses that I want it to support, rather than those I don't want it to support. As Savage has observed, the probability calculus may show that we should modify some of our beliefs, but cannot show *which* of our beliefs should be modified.

Now it seems absurd that we should be able, so to speak, to change our beliefs retroactively after seeing to what further beliefs they lead. It is certainly poor science to decide which hypothesis to accept first and to evaluate the evidence concerning that hypothesis in the light of the hypothesis itself. The defense is that people just don't do that with their beliefs; people don't decide what to believe first and then calculate backwards, using the probability calculus, to find out what their prior beliefs must (should) have been. On the contrary, the way the probability calculus does generally function is to lead us to modify our beliefs in accordance with the evidence, rather than to modify the impact of the evidence in accordance with our beliefs. From an abstract point of view this may seem like no more than a fortunate accident. From the point of view of the subjectivist it is simply a fact, a datum, and for the subjectivist that is sufficient. For the epistemologist or the scientist, however, this is the way probability calculus *ought* to function, and indeed it is the whole point of having a probability calculus.

EXERCISES

1. The text contains an informal demonstration that the holder of a set of degrees of belief that violated the theorem $P(H) = 1 - P(\bar{H})$ could have a book made against him. Construct a similar demonstration for the necessity of the addition axiom.
2. Construct a similar demonstration for the multiplication axiom.
3. Construct a similar demonstration for the axiom of total probability.
4. Give a proof of the assertion on page 73 that the probability is 0.99 that in an arbitrarily long sequence of draws, there will be less than 10% purple balls. (Use the fact that

$$P\left[|f - P_1| > K\left(\frac{P_1}{H_1} + \frac{h-1}{h} P_1P_2 - P^2\right)^{1/2}\right] < \frac{1}{K^2},$$

where P_1 is the prior probability of a purple ball, P_2 is the conditional probability of getting a purple ball on the second draw, given that a purple ball resulted on the first draw, and f is the relative frequency of purple balls among h draws; K is arbitrary).
5. There are two urns, urn 1 and urn 2. One has ten white balls and five black balls; the other has ten black balls and five white balls. Suppose that I believe to the degree 0.75 that I have urn 1. What is the conditional probability that I should attach to the hypothesis that I have urn 1, given that I draw four balls (with replacement) and that they are all black? I now draw four balls, with careful mixing and replacement; all of them turn out to be black. What degree of belief should I now attach to the hypothesis that I have urn 1, according to Bayes' Theorem? Suppose in spite of this recommendation, my belief remains of degree 0.75 that I have urn 1; how can this be explained?
6. Suppose that the subjective probability for you that the Yankees will win the first game of the coming World Series is 0.4, and that the probability for you that they will win the Series is 0.6. Suppose the conditional probability that if they will have won the first game is 0.5. What is the conditional probability that they will win the Series, given that they win the first game? Suppose they win the first game and your belief that they will go on to win the Series is 0.3. How can this be explained?
7. Let H be the statement that the football team will win their game this afternoon, and E the statement that the field will be wet. My current credence function has the values

$$\mathrm{Cr}(H) = 0.4, \qquad \mathrm{Cr}(H|E) = 0.2, \qquad \mathrm{Cr}(H|\bar{E}) = 0.5.$$

What does coherence require of my belief in E? Suppose I look out the window and my belief in E becomes 0.7, because it is drizzling lightly. How should my other beliefs change, according to Jeffrey? Suppose my eggs are overdone and my coffee is cold and this causes my belief in H to change to 0.2. How should my other beliefs change?

BIBLIOGRAPHICAL NOTES FOR CHAPTER 6

A general source book on subjective probability is Kyburg and Smokler (eds.) *Studies in Subjective Probability*, John Wiley and Sons, New York, 1964. It includes reprints and translations of some of the items mentioned below.

F. P. Ramsey's original penetrating discussion of subjectivistic (the word is not his) probability was written in 1926 but did not appear in print until 1931 in the posthumous collection of essays, *The Foundations of Mathematics*, Humanities Press, London, 1931, 1950. This essay is reprinted in Kyburg and Smokler.

The work of Bruno de Finetti is altogether independent of that of Ramsey, and appears in a number of publications from 1930 onward. An early philosophically oriented essay is "Probabilismo: Saggio critico sulla teoria delle probabilità e sul valore della scienze," *Biblioteca di Filosofia diretta da Antonio Aliota*, Perrella, Naples, 1931. The most influential of de Finetti's works is "La Prévision: ses lois logiques; ses source subjectives," *Annales de l'Institute Henri Poincaré*, Volume 7, 1937, pp. 1–68; a translation of the latter article, with new (1964) footnotes by de Finetti, appears in Kyburg and Smokler.

The big book of subjective probability theory, including both logical foundations and statistical applications, is L. J. Savage, *The Foundations of Statistics*, John Wiley and Sons, New York, 1954.

The best book for the philosophy student, presenting both the general theory of subjectivistic probability and a detailed and comprehensive treatment of utilities, probabilities, and coherence, is Richard Jeffrey's *The Logic of Decision*, McGraw-Hill, New York, 1965. This book also contains Jeffrey's substitute for Bayes' Theorem.

Carnap's discussion of the relationships between credence functions, rational credence functions, conditional rational credence functions, and credibility functions is to be found in "The Aim of Inductive Logic," in *Logic Methodology and Philosophy of Science* (Nagel, Suppes, and Tarski, eds.) Stanford University Press, Stanford, California, 1962, pp. 303–18. The same volume also contains a presentation of I. J. Good's ideas concerning upper and lower probabilities, in "Subjective Probability as the Measure of a Non-Measureable Set," pp. 319–29. A similar approach is to be found in C. A. B. Smith, "Personal Probability and Statistical Analysis," *Journal of the Royal Statistical Society*, Series A, **128**, 1965, pp. 469–99.

7

The Epistemological Interpretation of Probability

An epistemological interpretation of probability is an interpretation which takes as the fundamental clue to the nature of probability the role it plays, both formally and informally, in scientific inference. Any of the preceding theories of probability can be understood in this way: in the cases of several of the authors whose views we have considered, epistemological considerations were uppermost. This is most clearcut for Keynes and Carnap. Each of them is explicitly concerned with the relation that obtains between evidence and conclusion when the evidence does not entail the conclusion, but when the evidence does provide strong support for the conclusion. As the words 'evidence' and 'conclusion' suggest, the importance of the relation in question lies in its relevance to the grounds we have for believing that quinine is a specific for malaria or that the present forms of life on the earth evolved from