BRUNO DE FINETTI. ***Philosophical Lectures on Probability***. Collected, edited, and annotated by Alberto Mura. Translated by Hykel Hosni. Synthese Library; 340. Dordrecht: Springer, 2008. ISBN 978-1-4020-8201-6. Pp. xxii + 212.

Reviewed by JON WILLIAMSON*

This posthumous work was produced by transcribing audio recordings of lectures that Bruno de Finetti gave at the National Institute for Advanced Mathematics in Rome in 1979. Alberto Mura attended the course, recorded the lectures, took notes, and edited the resulting volume, which was first published in Italian in 1995. Hykel Hosni translated the lectures for this English edition, which appears in the Synthese Library series of volumes on epistemology, logic, methodology and philosophy of science. The book contains an introductory essay about de Finetti by Maria Carla Galavotti. Three of the twenty-two lectures and part of a fourth are lost, but the remaining lectures have many useful editorial comments. Moreover, interesting discussion between de Finetti and those attending the course is also included. So we have many people to thank for this important text.

De Finetti wrote the following notice to advertise the course at the Institute:

> The course, with a deliberately generic title ['On Probability'] will deal with the conceptual and controversial questions on the subject of probability: questions which it is necessary to resolve, one way or another, so that the development of reasoning is not reduced to a mere formalistic game of mathematical expressions or to vacuous and simplistic pseudophilosophical statements or allegedly practical claims.

Since de Finetti was a key figure in the development of the conceptual foundations of probability, these lectures will be of great interest to philosophers of probability in particular, and to epistemologists and philosophers of mathematics and science in general. This new English edition is very welcome indeed.

De Finetti is known as a champion of the strictly subjective interpretation of probability. According to this view, probabilities are to be construed as degrees of belief, and are thus defined in relation to an agent holding those beliefs. These degrees of belief are subject to a rather weak normative constraint—coherence, which merely demands that degrees of belief satisfy the axioms of probability—but otherwise it is left up to the agent as

* Department of Philosophy, SECL, University of Kent, Canterbury CT2 7NF, U.K. j.williamson@kent.ac.uk

to how to apportion her degrees of belief. According to de Finetti there is no viable non-epistemic notion of single-case probability: ‘it is senseless to speak of the probability of an event unless we do so in relation to the body of knowledge possessed by a given person’ (p. 3).

De Finetti’s subjective interpretation—often called subjective *Bayesianism*—is typically contrasted with *objective Bayesianism* (see, *e.g.*, [Rosenkrantz, 1977; Jaynes, 2003; Williamson, 2010]). Objective Bayesians are also concerned with degrees of belief and also hold that these degrees of belief should satisfy the axioms of probability and are relative to the knowledge or evidence of a given agent. But objective Bayesians go further by arguing that the agent’s degrees of belief should also satisfy constraints imposed by evidence—for example they should be set to known frequencies in certain situations—and that, where evidence fails to determine degrees of belief fully, an agent should equivocate as far as possible between the basic possibilities under consideration; hence she should not adopt extreme degrees of belief unless forced to by her evidence.

Here we shall briefly examine the volume with regard to the distinction between subjective and objective Bayesianism.

That de Finetti is firmly in the subjectivist camp is witnessed by the following colourful expression of his views:

> Even if an individual were self-coherent, I might still find his opinions absurd and I might say: “Is he crazy?” Likewise he might say that I am crazy because my evaluations differ from his own. Of course, a third person might have reasons, on the grounds of what happened subsequently, to judge either my evaluations or the other person’s ones more reasonable. But this is the sort of “hindsight” of which we are told “ditches are full.” I have never seen ditches full of hindsight, yet there is probably a rough kernel of truth in this sentence. (p. 18)

Objectivists hold that it is not just in hindsight that we can judge an agent’s degrees of belief to be unreasonable. For de Finetti, if one has no evidence that bears on an elementary proposition A then one is reasonable to believe it to any degree at all (as long as one remains coherent, setting $P(\neg A) = 1 - P(A)$ for example). But objectivists would—*in advance*—judge an agent to be unreasonable if she were strongly to believe A or its negation in this absence of evidence. According to objectivists, the agent should equivocate between A and $\neg A$, awarding each of these propositions degree of belief $1/2$ or thereabouts.

The main point of departure of this later work of de Finetti’s from his earlier work on probability is at the level of foundations. Formerly he based his view of probability on betting considerations: these betting considerations lead to the so-called *Dutch book* or *Ramsey–de Finetti* *theorem* which is used to argue that an agent's assignment of degrees of belief are coherent (*i.e.*, not exploitable to yield sure loss in a betting situation) if and only if they satisfy the axioms of probability. In this work, de Finetti rejects that line of argument on the grounds that it presumes that anyone out to exploit the agent's bets must have the same degrees of belief as the agent herself (p. 29). Instead, de Finetti bases his interpretation of probability on *scoring rules*. A scoring rule is a particular kind of loss function. The loss one makes when one chooses course of action a if state ω of the world is the case can be denoted by $L(\omega, a)$; here L is a *loss function* and we may identify ω with an atomic state $\pm A_1 \wedge \cdots \wedge \pm A_n$ of the elementary propositions $A_1, \ldots, A_n$ expressible in the agent's language. In our context the agent must choose a probability function Q on which to base her predictions, in which case the loss is written $L(\omega, Q)$ and L is called a *scoring rule*. De Finetti maintains that minimising expected loss is a decisive desideratum and, given this desideratum, one should choose a probability function that accurately represents one's degrees of belief (p. 18 and Chapter 3).

I will argue here that by basing his approach on scoring rules, de Finetti shoots himself in the foot: minimising expected loss with respect to a scoring rule leads more naturally to objective Bayesianism than to subjective Bayesianism.

Given a loss function L,

$$L(P, a) \stackrel{\text{df}}{=} E_P L(\Omega, a) = \sum_{\omega \in \Omega} P(\omega) L(\omega, a)$$

is the *expected loss* for P. Here interpret P as a probability function that is a reasonable choice of belief function for the agent, and suppose that the agent only knows that $P \in \mathbb{E}$, where $\mathbb{E}$, the set of belief functions compatible with the agent's evidence, is some (not necessarily proper) subset of the set $\mathbb{P}$ of probability functions. $\Omega = \{\pm A_1 \wedge \cdots \wedge \pm A_n\}$ is the set of all atomic states. Define

$$H(P) \stackrel{\text{df}}{=} \inf_{a \in \mathcal{A}} L(P, a),$$

the *Bayes loss* or *generalised entropy* of P. In our case L is a scoring rule and the agent must choose some $Q \in \mathbb{P}$ as a basis for prediction. We shall assume, as de Finetti did, that the scoring rule L is *proper*: for all P, the choice $Q = P$ minimises $L(P, Q)$. (Such a choice Q is sometimes called a *Bayes act*.) Then $H(P) = L(P, P)$ for $P \in \mathbb{E}$. Next let

$$d(P, Q) \stackrel{\text{df}}{=} L(P, Q) - H(P),$$

which we shall call the *divergence* of P from Q. A divergence function is interpreted as a measure of the distance of P from Q, but it need not be a distance function in the strict mathematical sense, since it may be asymmetric or may not satisfy the triangle inequality. The *equivocator*, $P_=$, is a probability function defined by

$$P_=(\omega) \stackrel{\text{df}}{=} \frac{1}{2^{|\Omega|}}$$

for all $\omega \in \Omega$. This function equivocates between the basic possibilities $\omega \in \Omega$ expressible in the agent's language. We shall assume that the scoring rule L is *equivocator-neutral*: *i.e.*, $L(P, P_=) = k$, a constant, for all P. This holds for many common scoring rules including Brier score $L(\omega, Q) = (1 - Q(\omega))^2 + \sum_{\omega' \neq \omega} P(\omega')^2$ (which was the scoring rule advocated by de Finetti), logarithmic loss $L(\omega, Q) = -\log Q(\omega)$, and zero-one loss $L(\omega, Q) = 1 - Q(\omega)$. (This assumption is not essential for the main point developed here; it merely allows us to talk in terms of divergence rather than entropy.) As Grünwald and Dawid [2004] show, under natural conditions

$$\arg \inf_{Q \in \mathbb{P}} \sup_{P \in \mathbb{E}} L(P, Q) = \arg \inf_{P \in \mathbb{E}} d(P, P_=).$$

In particular if $\mathbb{E}$ is convex and under certain closure conditions, the functions in $\mathbb{P}$ minimising worst-case expected loss are those in $\mathbb{E}$ closest to the equivocator. But minimising worst-case expected loss is rather natural given de Finetti's desideratum of minimising expected loss. And choosing a function in $\mathbb{E}$ that is closest to the equivocator is exactly the recipe of objective Bayesianism [Williamson, 2010]. So de Finetti's advocacy of the desideratum leads one to objective, rather than subjective, Bayesianism.

Note here that the geometry of the space of probability functions depends on the loss function, in the sense that the notion of distance varies according to the loss function. As a default loss function, de Finetti considered Brier score. This was apparently on the grounds of a mechanical analogy: it permits one to think of probabilities $P(A)$ and $P(\neg A)$ as masses hung at the ends of a uniformly dense bar (p. 5). But it is more typical to take logarithmic loss as the default loss function. This policy can be justified on the basis of considerations to do with information theory, information geometry, gambling (in particular a gambling set-up called Kelly gambling), game theory where penalties are proportional to code lengths, or the following axiomatic approach [Williamson, 2010, § 3.4]: formalise the agent's language as a finite propositional language $\mathcal{L} = \{A_1, \ldots, A_n\}$ on propositional variables $A_1, \ldots, A_n$, and assume the following of the default loss function L:

L1: By default, $L(\omega, Q) = 0$ if $Q(\omega) = 1$.

L2: By default, loss strictly increases as $Q(\omega)$ decreases from 1 towards 0.

L3: By default, loss $L(\omega, Q)$ depends only on $Q(\omega)$, not on $Q(\omega')$ for $\omega' \neq \omega$.

L4: By default, losses are presumed additive when the language is composed of independent sublanguages: if $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$ and $\mathcal{L}_1 \perp\!\!\!\perp_Q \mathcal{L}_2$ then $L_{\mathcal{L}}(\omega_1 \wedge \omega_2, Q) = L_{\mathcal{L}_1}(\omega_1, Q\downharpoonright_{\mathcal{L}_1}) + L_{\mathcal{L}_2}(\omega_2, Q\downharpoonright_{\mathcal{L}_2})$. (Here ω_1 and ω_2 are atomic states of $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively.)

If these conditions are to hold for any finite predicate language $\mathcal{L}$ and any probability function Q defined on this language, then the default loss function L must be logarithmic loss, $L(\omega, Q) = -\log_b Q(\omega)$ to some base $b > 1$.

Taking loss as logarithmic by default, the divergence function is just the well-known *Kullback-Leibler divergence*

$$d(P, Q) = \sum_{\omega \in \Omega} P(\omega) \log \frac{P(\omega)}{Q(\omega)},$$

and minimising divergence from the equivocator is equivalent to maximising entropy in the information-theoretic sense, $H = -\sum_{\omega \in \Omega} P(\omega) \log P(\omega)$. Applying the desideratum of minimising worst-case expected loss, we are led to the *maximum entropy principle* of [Jaynes, 1957]: one's degrees of belief should be representable by a function in $\mathbb{E}$ that has maximum entropy. This is the classic formulation of objective Bayesianism.

In [Williamson, 1999] I pointed out that, somewhat ironically, de Finetti's rejection of the axiom of countable additivity makes subjectivism more elusive. Here we see that the irony extends to his adoption of scoring rules as a foundation for subjectivism. In my view de Finetti was right to take minimising expected loss as a desideratum and right to pay attention to scoring rules; but in hindsight—of which ditches may well be full—these moves do not sit well with subjectivism.

I have only dwelt on one aspect of de Finetti's lectures but the book touches on almost all the major concerns of his theory of probability. It is full of fascinating material and all involved have done a great job in bringing it to a wider audience.

References

Grünwald, P., and A.P. Dawid [2004]: 'Game theory, maximum entropy, minimum discrepancy, and robust Bayesian decision theory', *Annals of Statistics* **32**, 1367–1433.

Jaynes, E. T. [1957]: 'Information theory and statistical mechanics', *The Physical Review* **106**, 620–630.

JAYNES, E. T. [2003]: *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press.

ROSENKRANTZ, R.D. [1977]: *'Inference, Method and Decision: Towards a Bayesian Philosophy of Science*. Dordrecht: Reidel.

WILLIAMSON, J. [1999]: 'Countable additivity and subjective probability', *British Journal for the Philosophy of Science* **50**, 401–416.

________ [2010]: *In Defence of Objective Bayesianism*. Oxford: Oxford University Press.

doi:10.1093/philmat/nkp019
Advance Access publication November 5, 2009

Books of Essays

IVOR GRATTAN-GUINNESS. ***Routes of Learning: Highways, Pathways, and Byways in the History of Mathematics***. Baltimore: Johns Hopkins University Press, 2009. ISBN-13 978-08018-9247-9 (hbk), 978-0-8018-9248-6 (pbk). Pp. xii + 372.

TITLES (PREVIOUSLY PUBLISHED)