International Journal of Approximate Reasoning 54 (2013) 1351–1360

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

A consistent set of infinite-order probabilities

David Atkinson *, Jeanne Peijnenburg

University of Groningen, Faculty of Philosophy, Oude Boteringestraat 52, 9712 GL Groningen, Netherlands

ARTICLE INFO

Article history:
Received 18 January 2013
Received in revised form 24 April 2013
Accepted 26 April 2013
Available online 6 May 2013

Keywords:
Model
Higher-order probability
Infinite regress

ABSTRACT

Some philosophers have claimed that it is meaningless or paradoxical to consider the probability of a probability. Others have however argued that second-order probabilities do not pose any particular problem. We side with the latter group. On condition that the relevant distinctions are taken into account, second-order probabilities can be shown to be perfectly consistent.

May the same be said of an infinite hierarchy of higher-order probabilities? Is it consistent to speak of a probability of a probability, and of a probability of a probability of a probability, and so on, *ad infinitum*? We argue that it is, for it can be shown that there exists an infinite system of probabilities that has a model. In particular, we define a regress of higher-order probabilities that leads to a convergent series which determines an infinite-order probability value. We demonstrate the consistency of the regress by constructing a model based on coin-making machines.

1. Introduction

Let q_0 be a proposition with a probability v_0:

$$P^1(q_0) = v_0, \tag{1}$$

where P^1 stands for an ordinary unconditional probability of the first order, and where v_0 is some number between 0 and 1. We now might try to assert a second-order probability, P^2, to the effect that the probability of q_0, given (1), is v_0:

$$P^2\big(q_0|P^1(q_0) = v_0\big) = v_0. \tag{2}$$

Is this coherent? Some have denied that it is. Bruno de Finetti famously claimed that second-order probabilities are devoid of meaning, whereas David Miller has argued that they lead to an absurdity [1,11]. According to Miller, if we substitute $P^1(\neg q_0)$ for v_0 in Eq. (2), we obtain

$$P^2\big(q_0|P^1(q_0) = P^1(\neg q_0)\big) = P^1(\neg q_0),$$

which is the same thing as

$$P^2\big(q_0|P^1(q_0) = \tfrac{1}{2}\big) = P^1(\neg q_0).$$

However, with $v_0 = \frac{1}{2}$, we see that (2) yields $P^2(q_0|P^1(q_0) = \frac{1}{2}) = \frac{1}{2}$. Therefore $P^1(\neg q_0) = \frac{1}{2}$, and thus $P^1(q_0) = \frac{1}{2}$. So if Eq. (2) were unrestrictedly valid, we could prove that the probability of an arbitrary proposition q_0 is equal to one-half, which is absurd. This absurdity is known as the Miller paradox.

* Corresponding author.
E-mail address: d.atkinson@rug.nl (D. Atkinson).

0888-613X/$ – see front matter
http://dx.doi.org/10.1016/j.ijar.2013.04.015

Brian Skyrms pointed out that the Miller paradox "rests on a simple *de re–de dicto* confusion" [18, p. 111]. One and the same expression is used both referentially and attributively, so that a scalar or number (here v_0) is wrongly put on a par with a random variable (here $P^1(\neg q_0)$) that takes on a range of possible values [5, pp. 399–400]. So long as we recognize this confusion and keep the two levels apart, the notion of a second-order probability is harmless.

Another objection to higher-order probabilities can be discerned in de Finetti's work. As is well known, de Finetti holds that probability judgements are expressions of attitudes that lack truth values. However, as Skyrms has pointed out, de Finetti's work is less hostile to a theory of higher-order probabilities than might at first appear [18, p. 113]:

> "For a given person and time there must be, after all, a proposition to the effect that that person then has the degree of belief that he might evince by uttering a certain probability attribution.
>
> De Finetti grants as much:
>
> > The situation is different of course, if we are concerned not with the assertion itself but with whether 'someone holds or expresses such an opinion or acts according to it,' for this is a real event or proposition. [1, p. 189]
>
> With this, de Finetti grants the existence of propositions on which a theory of higher order personal probabilities can be built, but never follows up this possibility."[1]

It seems, then, that de Finetti was not opposed to higher-order personal probabilities in such an uncompromising way as might at first sight seem to be the case.

Skyrms is not alone in having seen that second-order probabilities need not pose any particular problem. Several others recognized that, when the relevant distinctions are taken into account, second-order probabilities can be shown to be formally consistent [19,10,2,8,3]. This is not to say that such probabilities are mandatory. As Pearl has explained, second-order probabilities, although consistent, can be dispensed with, for one can always express them by using a richer first-order probability space [12].

These findings on second-order probabilities can be easily extended to probabilities of any finite order. But what about a hierarchy of higher-order probabilities that is infinite? Is the idea of a probability of a probability *ad infinitum* just as consistent as the idea of a probability of a finite order? As far as we know, nobody has ever claimed that it is, let alone has anyone given a proof of consistency. The reason is not difficult to find. A proof of consistency should involve a demonstration that the infinite series is convergent, and it is not immediately clear how that can be done.

As early as 1738, David Hume objected to an infinite hierarchy of probabilities of probabilities, on the grounds that it would tend to zero in the end:

> "Having thus found in every probability ... a new uncertainty ... and having adjusted these two together, we are oblig'd ... to add a new doubt .... This is a doubt ... of which ... we cannot avoid giving a decision. But this decision, ... being founded only on probability, must weaken still further our first evidence, and must itself be weaken'd by a fourth doubt of the same kind, and so on in infinitum: till at last there remain nothing of the original probability, however great we may suppose it to have been, and however small the diminution by every new uncertainty."[2]

Doubts about the consistency of an infinite regress of higher-order probabilities have also been expressed in more recent times. Savage wrote that such a hierarchy is beset by "insurmountable difficulties" and his worries were anticipated by Russell ([17, p. 58]; [16, p. 385 ff]). Modern objections against infinite-order probabilities are however not always rooted in Humean worries. For example, Nicholas Rescher claims that an infinite hierarchy of probabilities is inconsistent, not because it goes to zero, but because it will forever remain indeterminate [15, pp. 36–37].

Against these objections, we shall argue in the present paper that an infinite hierarchy of probabilities can indeed be consistent. The above-mentioned difficulty of finding a consistency proof will be overcome in two stages. First we construct a structure for which the proof *can* be given; we do this by making sure that this structure is subject to a Markov condition. Then we demonstrate that this structure is a model of an abstract system of infinite-order probabilities; in other words, it is a structure that makes all the sentences of the abstract system true. In this way we show that one can calculate a definite probability value from an endless hierarchy of probability statements. *Pace* Hume, this probability is in general not zero.

It should be noted that our purpose in this paper is not to give general conditions on probability distributions over probability distributions *ad infinitum*. Spelling out the details of those conditions would be very interesting; but we shall not attempt such an ambitious project in the present paper. Rather we content ourselves with a simple Bernoulli distribution, and with single events at each order.

This paper is set up as follows. We will start, in Section 2, by describing our model for a regress of higher-order probabilities; the model in question consists in an infinite set of machines that produce biased coins. In Section 3 we set up an abstract system of equations that produces the same probabilities as those in the model of Section 2, thereby showing that

[1] Skyrms [18, pp. 113–114].

[2] Hume [7, Book I, Part IV, Section I]. See also Lehrer [9].