

Reichenbach's Posits Revisited

David Atkinson · Jeanne Peijnenburg

Received: 10 October 2007 / Accepted: 29 October 2007 / Published online: 4 April 2008
© The Author(s) 2008

Abstract Reichenbach's use of 'posits' to defend his frequentistic theory of probability has been criticized on the grounds that it makes unfalsifiable predictions. The justice of this criticism has blinded many to Reichenbach's second use of a posit, one that can fruitfully be applied to current debates within epistemology. We show first that Reichenbach's alternative type of posit creates a difficulty for epistemic foundationalists, and then that its use is equivalent to a particular kind of Jeffrey conditionalization. We conclude that, under particular circumstances, Reichenbach's approach and that of the Bayesians amount to the same thing, thereby presenting us with a new instance in which chance and credence coincide.

1 Introduction

Many associate the frequency interpretation of probability with Richard von Mises, despite the fact that several years before he introduced his ideas in von Mises (1919), Hans Reichenbach had already developed his own frequentistic probability theory in his inaugural dissertation of 1915 (cf. Reichenbach 1978; Galavotti 2003). The reason why Reichenbach's theory was no match for that of von Mises is not difficult to discern. For while von Mises' pivotal notion of the *Kollektiv* has its difficulties, the objections to Reichenbach's idea of *posits* are even more telling.

Suppose that, in n repeated trials, there are m occurrences of a specified kind, so that the relative frequency of the occurrence in question is m/n . If we let n go to infinity, then either m/n has a limit or it has not. Reichenbach's opening gambit

D. Atkinson (✉) · J. Peijnenburg
Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, Groningen 9712 GL,
The Netherlands
e-mail: D.Atkinson@rug.nl

J. Peijnenburg
e-mail: Jeanne.Peijnenburg@rug.nl

(Reichenbach 1949) is the claim that, if it has a limit, then this limit, called p , is expected to be somewhere in the region of m/n . Specifically, for large n , p satisfies the inequality

$$\frac{m}{n} - \delta < p < \frac{m}{n} + \delta, \quad (1)$$

where δ is some margin of error. The problem is of course to find a value for δ . What exactly does ‘somewhere in the region of m/n ’ mean? If no prior knowledge of the system is available, Reichenbach argues, we simply guess what δ might be, and then the whole statement (1) becomes a so-called *blind* or *anticipative* posit. This statement can subsequently be made more precise, resulting in an *informed* or *appraised* posit.

One of Reichenbach’s reasons for introducing posits of this kind was no less than a pragmatic vindication of induction: a practical answer to Hume’s claim that the uniformity of nature is indemonstrable. We cannot prove that nature will be uniform with respect to a given sequence of events, but if she is, in the sense that the ratio m/n has a limit, then this method of the posit will eventually obtain that limit to any required degree of accuracy (Reichenbach 1949, p. 446, 475; cf. Salmon 1966, p. 86). Although Reichenbach had invoked the principle that nature is uniform as a synthetic a priori in his 1915 dissertation, he later repudiated this view (Reichenbach 1951, pp. 246–247). His way of vindicating its use was to throw everything into the conditional mood, as it were: *if* a sequence has a limit, *then* I have a way of nailing it down. Technically, he was forced to introduce a hierarchy of levels of probabilities: the second-order probability that the first-order probability, p , lies in the above interval tends to unity as n tends to infinity (Reichenbach 1949, p. 442).

Some have called Reichenbach’s theory ‘empirical frequentism’, thus suggesting that it is falsifiable. The theory has however been justly criticized on the grounds that, although m/n tends to p in the limit that n goes to infinity (on condition that the limit exists), the speed at which this limit is attained is unknown. Indeed wild fluctuations on the way to that limit cannot be ruled out. Moreover, *any* initial segment of a sequence of trials is consistent with *any* limit p : everything depends on the infinite tail of the sequence, not on a finite initial part of it! In practice then, since we can only work with a finite number of trials, and no estimate can be given of how many trials would be needed to come within a given δ of p , Reichenbach’s use of these posits is not falsifiable, a fatal shortcoming for any theory with empirical pretensions. Even Wesley Salmon, after a working lifetime largely devoted to the defence of his teacher, admitted defeat: “Reichenbach’s attempt to vindicate his rule of induction cannot be considered successful. ... My attempt to vindicate Reichenbach’s rule of induction cannot be considered successful.” (Salmon 1991, p. 105, 107).

These familiar failings have tended to obscure the fact that Reichenbach also used the concept of a posit in another, and altogether more defensible context, namely in his debate with the foundationalists of his day such as Bertrand Russell and Clarence I. Lewis. In Sects. 2 and 3 we will briefly sketch this debate, and then we will investigate this new role for posits by considering three concrete cases

(Sects. 4–6). Contrary to Reichenbach's intuitions, it turns out that in two of these three cases, posits are not needed at all (Sects. 4, 5), whereas in the third case they are indispensable (Sect. 6). In each of these cases a simplifying assumption has been made, namely that the conditional probabilities in question remain constant. In Sect. 7, and later again in the Appendix, we explain how to generalize our argument to cases where this assumption is dropped. Finally, in Sect. 8, we show that Reichenbach's method is identical to Jeffrey conditionalization under a certain restriction.

2 Reichenbach versus Lewis

From 1930 until his death in 1953, Reichenbach was strenuously engaged in a debate with Lewis, staunch defender of a 'strong foundationalist' program in epistemology. The debate apparently started with a letter that Reichenbach wrote to Lewis on July 29, 1930. Although this letter is now lost, we roughly know its content from Lewis's answer to it, written one month later, and today kept at the University of Pittsburgh. Lewis's letter makes clear that Reichenbach had objected to the idea, defended in Lewis's *Mind and the World Order* of 1929, that an event can only be probable if we assume other events to be certain; in Lewis's view, these other events consist of sense data.¹ Reichenbach's disagreement with this position is profound. He denies that sense data must be certain, and he disputes that an event can only be probable if it is ultimately grounded in events that are certain, sense data or otherwise. In Reichenbach's opinion there is nothing incoherent in the concept of an infinite sequence of events, where each event is made probable by its predecessor, never reaching certain ground.

Lewis in turn fiercely disagreed and re-explained his view with gusto, first in letters and conversations, later also in journals and at conferences. The dispute reached its climax at the forty-eighth meeting of the Eastern Division of the American Philosophical Association in December 1951, where both Lewis and Reichenbach read papers that were subsequently published in *The Philosophical Review* of April 1952.

The essence of the dispute concerns the very existence of a foundation in epistemology, rather than the specific nature thereof. The central question is not whether sense data are certain, even though Lewis would give an affirmative answer, and Reichenbach a negative one. The central question rather is: 'Can events be probable without being ultimately connected to a foundation that is certain, whether this foundation be sensory or not?'

In the next section we will present this disagreement in a formal and more precise way. We also explain how Reichenbach invokes a different type of posit to attack Lewis's stance.

¹ Rather than talking about probable or certain *events*, one might talk instead about the probability or certainty of *propositions*. In Atkinson and Peijnenburg (2006) it has been shown that, in the relevant modal systems, the two ways of talking are equivalent.

3 Alternative Posits

Suppose that the occurrence of an event A_0 is made probable by that of another event A_1 . The probability of A_0 is given by the rule of total probability:

$$P(A_0) = P(A_0|A_1)P(A_1) + P(A_0|\neg A_1)P(\neg A_1). \quad (2)$$

If A_1 is in turn made probable by A_2 , the rule must be applied a second time:

$$P(A_1) = P(A_1|A_2)P(A_2) + P(A_1|\neg A_2)P(\neg A_2). \quad (3)$$

Does it make sense to continue this procedure, allowing for events made probable by other events, made probable by still other events, and so on, *ad infinitum*? Of course the question is not whether we can go on applying the rule in practice, but whether we can do so in principle. Lewis's answer to this question is that we cannot. He claims that the iteration must stop eventually. For some finite n , if A_n is made probable by A_{n+1} , the latter must be certain: $P(A_{n+1}) = 1$. Once we have arrived at this certain point, the total probability rule has the simple form $P(A_n) = P(A_n|A_{n+1})$, so there is nothing more to iterate. Denying the necessity of such a termination, Lewis argues, would amount to abnegating the very concept of probability; in the above case, such a denial would imply that the probability of the event with which we began, $P(A_0)$, is equal to zero (Lewis 1952, p.172).

Reichenbach demurs, claiming that the above sequence not only can, but must go on indefinitely. This raises the question as to how we might calculate the probability of A_0 . If $P(A_0)$ is the outcome of an infinite regression, how can we compute its value? Again, the question is not just a matter of practice, but also of principle. Is not the calculation of such an infinite regression too complex for us to bring to completion? After all, insertion of Eq. 3, together with

$$P(\neg A_1) = P(\neg A_1|A_2)P(A_2) + P(\neg A_1|\neg A_2)P(\neg A_2) \quad (4)$$

into the right-hand side of Eq. 2 leads to an expression with four terms, namely

$$\begin{aligned} P(A_0) = & P(A_0|A_1)P(A_1|A_2)P(A_2) + P(A_0|\neg A_1)P(\neg A_1|A_2)P(A_2) \\ & + P(A_0|A_1)P(A_1|\neg A_2)P(\neg A_2) + P(A_0|\neg A_1)P(\neg A_1|\neg A_2)P(\neg A_2). \end{aligned} \quad (5)$$

A repetition of this manoeuvre to express $P(A_2)$ and $P(\neg A_2)$ in terms of $P(A_3)$ and $P(\neg A_3)$ produces eight terms, and after $n + 1$ steps the number of terms is 2^{n+1} . This yields a lengthy expression that seems at first sight hard to compute in simple closed form.

Perhaps Reichenbach saw this difficulty. At any rate he does not even attempt to engage in the task: nowhere does he try to find a usable expression in the limit as n goes to infinity. Instead, he chooses to make a guess as to what $P(A_{n+1})$ might be for a given, fixed n . If nothing is known about A_{n+1} , this guess is a blind posit. But if, on the other hand, some empirical information is available that serves to delimit the possible values of $P(A_{n+1})$, the guess is not a wild one, and the posit becomes *appraised*. The idea is that blind posits can become appraised by testing them in suitable empirical situations. Of course, even appraised posits will never be more than conjectural: their very nature as posits prevents them from ever becoming

certain or categorically true. A posit will never be more, so to speak, than the antecedent in a conditional statement.

It should be noted that the above use of blind and appraised posits is quite different from the one that we described in the first section. True, here as well as there, posits are used to determine the probability of an event, and here as well as there, posits are conjectures and as such subject to further adjustment. Nevertheless their functions are quite different. In Sect. 1 posits were used to determine the probability of an event on the basis of relative frequencies, and this use has been rightly criticized. But the posits that Reichenbach deploys to counter Lewis's position determine the probability of an event, $P(A_n)$, on the basis of the probability of another event, $P(A_m)$, and such a use is perfectly defensible. To keep the distinction clear, we will call the former 'posits of the first kind' and the latter 'posits of the second kind' or 'alternative posits'.

Reichenbach himself does not explicitly distinguish between posits of the first and second kind. However, that there are indeed two kinds of posit can be clearly distilled from his multifarious writings. In some of these, he refers merely to posits of the first kind (e.g. Reichenbach 1951); in others he is solely talking about posits of the second kind (Reichenbach 1952); and in Reichenbach (1949), which contains his considered opinions on probability, he uses first the one and then the other without mentioning the shift. But as we have seen, the two types clearly differ in character. Moreover, Reichenbach's motivation for using each type also appears to be quite different.

Reichenbach's main motivation for using his posits of the first kind springs from his adherence to a frequentistic theory of probability. As he sees it, there are two major philosophical objections to frequentism (Reichenbach 1951, pp. 236–237). The first is that it assumes inductive inference, and hence presupposes the unjustifiable principle of the uniformity of nature. The second is that a frequentistic probability theory cannot handle single cases: how can I ever come to know what *my* chances of surviving *my* cancer are, if these chances are stated in terms of relative frequencies? Reichenbach was convinced that both objections could be overcome by bringing his posits of the first kind into play (Reichenbach 1951, p. 241; cf. Reichenbach 1949, pp. vii–viii). We have expressed our reservations regarding these posits in Sect. 1, and we do not believe that they can resolve either of the two objections.

Posits of the second kind are mainly used in the context of the debate with Lewis. Here Reichenbach's motivation is not to defend a frequentistic theory of probability, but to attack foundationalism in epistemology, at least in the form in which it occurs in the writings of Lewis (and also Russell). In the next section we will show that Reichenbach was correct in criticizing Lewis's position. For Lewis was indeed mistaken: it is not true that an infinite sequence of probabilities, supported by probabilities indefinitely, necessarily converges to zero. However, we will also show that this does not imply that Reichenbach was right in claiming that we always need posits (of the second kind) when dealing with infinite sequences. As will become clear in Sect. 4, we can dispense with such posits when n goes to infinity, and as we show in Sect. 5, the same holds when n is large but finite. Reichenbach failed to notice this, perhaps because he did not realise that he had used his posits in

two essentially different ways. Or perhaps he did realise it, but was daunted by the exponential explosion of terms that occurs in calculating the outcome of the infinite regression. However this may be, and whatever philosophical or mathematical reasons Reichenbach might have had, we will introduce below a minimal but useful change of notation that enables us to complete the calculations without too much effort.

4 Dispensing with Posits, Part 1

The complication of the exponentially increasing number of terms, of which Eq. 5 was the first illustration, can be drastically reduced by replacing $P(\neg A_1)$ in Eq. 2 by $1 - P(A_1)$, and then write this equation as

$$P(A_0) = P(A_0|\neg A_1) + [P(A_0|A_1) - P(A_0|\neg A_1)]P(A_1). \quad (6)$$

A similar treatment can be applied to Eq. 3, which then becomes

$$P(A_1) = P(A_1|\neg A_2) + [P(A_1|A_2) - P(A_1|\neg A_2)]P(A_2), \quad (7)$$

and so on. These changes, small as they may be, turn out to have significant consequences. For they enable us to obtain a closed and usable expression for $P(A_0)$ in all situations, no matter whether the number of steps is finite or infinite. To see how this works out in detail, let us consider a concrete example.

Imagine colonies of a bacterium growing in a chemical environment known to be favourable to a particular mutation of practical interest. The bacteria reproduce asexually, so that only one parent, the ‘mother’, is sufficient to produce a child, the ‘daughter’. The probability that a mutated daughter descends from a normal, not mutated mother is known to be very small (say 0.02); but the probability that a mutated daughter descends from a mutated mother is on the other hand high (say 0.99). We are told that each colony, or batch, develops from a different, single ancestor; but it is not known, for a given batch, whether the ancestor was normal or mutated. Now we select a bacterium from a random batch. What is the probability that the selected bacterium is a mutant?

To answer this question, interpret A_0 in Eq. 6 as the event or proposition that the selected bacterium, a_0 , is a mutant, and A_1 as the event that its immediate ancestor, i.e. its mother, a_1 , was a mutant. Thus a_0 is the selected bacterium, a_1 is its immediate ancestor, and $P(A_0)$ is the probability that the selected bacterium is a mutant.

We know that $P(A_0|A_1) = 0.99$ and $P(A_0|\neg A_1) = 0.02$. In words: the probability that a_0 is a mutant is 0.99 if its mother, a_1 , was mutated, and it is 0.02 if a_1 was normal, i.e. not a mutant. If we insert

$$\alpha \equiv P(A_0|A_1) = 0.99 \quad \text{and} \quad \beta \equiv P(A_0|\neg A_1) = 0.02$$

into Eq. 6, we get as the probability that a_0 is a mutant

$$P(A_0) = \beta + (\alpha - \beta)P(A_1). \quad (8)$$

However, we can only use Eq. 8 to compute $P(A_0)$ if we know the value of $P(A_1)$, i.e. the probability that the mother is a mutant. How do we compute that value? The answer is of course that we must apply the same procedure to Eq. 7. That is, we must insert α and β into that equation, thus obtaining as the probability that a_1 is a mutant

$$P(A_1) = \beta + (\alpha - \beta)P(A_2). \tag{9}$$

But the story does not end here. In order to compute $P(A_2)$, we must know $P(A_3)$, and so on, *ad infinitum* (assuming for a moment that each bacterium has infinitely many ancestors). We have here a case of which the framework has already been sketched in the previous section, and it is not difficult to imagine how Lewis and Reichenbach would react to it. Lewis denies that an infinite sequence makes sense, for either it is incomputable or it will converge to zero. Reichenbach, on the other hand, believes that such a sequence is useful, although he does agree with Lewis that its infinite number of terms hinders its computability. As we have seen, he proposes to truncate the sequence by making a blind posit, and then to compute the probability on the basis of that.

Our position deviates from both. Against Lewis we claim that infinite sequences like the one above do make sense. For not only can they be computed, the resulting outcomes need not be zero either. Against Reichenbach we hold that the computation can be executed without truncations and without using any posit at all. The latter possibility arises from the small but significant change that we made at the beginning of the present section.

To make this clear, let us first generalize Eqs. 6 and 7 to

$$P(A_m) = P(A_m|A_{m+1})P(A_{m+1}) + P(A_m|\neg A_{m+1})[1 - P(A_{m+1})] \tag{10}$$

which gives the probability that a_m is a mutant. With α and β in place Eq. 10 reads

$$P(A_m) = \beta + (\alpha - \beta)P(A_{m+1}). \tag{11}$$

It has been assumed here that the conditional probabilities, α and β , are the same from generation to generation. This assumption is reasonable (although by no means necessary) in the case of bacteria being grown in controlled laboratory conditions. In the wild, variations of temperature and nutrient availability could be accommodated by allowing α and β to change from one generation to another. In Sect. 7 and in the Appendix we shall discuss the generalized case where α and β may vary; in the present and in the next two sections, however, we suppose α and β to be constant.

Let us now apply the rule expressed in this equation to $m = 0,1,2,3,\dots,n$. The result is a finite series that can be summed, yielding

$$\begin{aligned} P(A_0) &= \beta + (\alpha - \beta)[\beta + (\alpha - \beta)[\beta + (\alpha - \beta)[\dots]]] \\ &= \beta \left[1 + (\alpha - \beta) + (\alpha - \beta)^2 + \dots + (\alpha - \beta)^n \right] + (\alpha - \beta)^{n+1} P(A_{n+1}) \\ &= \frac{\beta}{1 - \alpha + \beta} + (\alpha - \beta)^{n+1} \left[P(A_{n+1}) - \frac{\beta}{1 - \alpha + \beta} \right]. \end{aligned} \tag{12}$$

Here the value of $P(A_0)$ is ultimately derived from one single term, the remainder term $(\alpha - \beta)^{n+1}P(A_{n+1})$, containing the probability that the primal mother a_{n+1} of a certain batch of bacteria is a mutant (see Eq. 8). However, the value of this remainder term cannot be computed unless we know the value of $P(A_{n+1})$, the probability that the primal mother of a_0 is a mutant. Does this mean that Lewis was right in claiming that Eq. 12 can only be solved if we assume that $P(A_{n+1}) = 1$? Or that Reichenbach was right when he argued that we have to make a blind posit concerning $P(A_{n+1})$, in order to be able to calculate $P(A_0)$?

The answers are ‘no’ and ‘no’. To see this, let us consider the infinite case. The standard way to investigate the convergence of an infinite series is first to look at a finite series of, say, $n + 1$ terms only, with a remainder term, and then to investigate what happens as n tends to infinity. Applying this procedure to Eq. 12, we observe that, since $0 < \alpha - \beta < 1$, the factor $(\alpha - \beta)^{n+1}$ becomes smaller and smaller as n becomes larger and larger. In the formal limit that n tends to infinity, we find that the series has an infinite number of terms, and that the terms in the second and third lines of Eq. 12 that contain the unknown $P(A_{n+1})$ tend to zero, and hence disappear completely.

In the limit of an infinite number of terms in the series, corresponding to an indefinite iteration of Eq. 10, we find

$$P(A_0) = \frac{\beta}{1 - \alpha + \beta} = \frac{0.02}{1 - 0.99 + 0.02} = \frac{2}{3},$$

with the values given above for the conditional probabilities α and β . Thus we conclude that, after an infinite number of generations, the batch of bacteria is two thirds mutated. The series is, although infinite, perfectly computable, yielding a number that is not zero. Moreover, this conclusion required no truncation or blind posit at all (cf. Peijnenburg 2007).

5 Dispensing with Posits, Part 2

We have seen how Reichenbach would proceed when confronted with an infinite series. First he would truncate the series by making a blind posit (in our example: for the probability that the $(n + 1)$ st ancestor of a_0 was a mutant), and then he would use that posit to compute what he wanted to know (in our example: the probability that a_0 was a mutant). Apparently his philosophical devotion to the method of posits was so strong that he did not consider the possibility of calculating the outcome of an infinite number of iterations.

In addition, there is something else that Reichenbach did not mention. Not only can we generally dispense with posits when n is infinite, sometimes we do not even need posits when n is finite but large. The difference between the two situations is slight but subtle. In the infinite case, posits are not needed because the terms containing the unknown probability dwindle away to nothing. In the finite case, however, posits can be ignored if we have a satisfactory approximation for the probability of interest, i.e. as long as we have a value that, although imprecise, is

acceptable for our purposes, which is so when we can estimate an error or uncertainty that is sufficiently small.

Take again our batch of bacteria and imagine it to be sampled after, shall we say, 150 generations. Assuming that no record of the whole history exists, we need a posit here if we want a precise value for $P(A_0)$. However, this would be to forget that in most scientific contexts an imprecise value is all that is required. And such a value, with small error estimates, can often easily be supplied without any posit at all: it certainly can be readily produced in our example. This becomes clear when we substitute $n = 149$, $\alpha = 0.99$ and $\beta = 0.02$ in Eq. 12, thereby obtaining

$$P(A_0) = 0.667 + (0.97)^{150}[P(A_{150}) - 0.667]. \quad (13)$$

While it is true that we would now need $P(A_{150})$ in order to find a precise value for $P(A_0)$, in practice we can make an acceptable estimate without a posit. For the maximum value that $P(A_0)$ can have is obtained by replacing $P(A_{150})$ by 1, and the minimum by replacing it by 0. Thus $P(A_0)$ is certainly not greater than $0.667 + (0.97)^{150} \times 0.333 = 0.670$, and it is not less than $0.667 - (0.97)^{150} \times 0.667 = 0.660$. After 150 generations of growth, any batch will be within a percent or so of being two thirds mutated, whether the original mother bacterium was a mutant or not. This information would presumably be all that an experimenter, or a supplier of mutant bacteria, would need.

6 Using Posits

In the previous sections we saw how, in a favourable situation, we can get along without the use of blind posits. Now we will address a problem where, on the contrary, it is essential to make a blind posit and to replace it subsequently by an appraised one. As before, we will present our case by giving a concrete example.

Consider the male inhabitants of Northern Ireland, who may or may not be Anglicans. Let α be the probability that a man is an Anglican, given that his father is one, and β the probability that he is an Anglican, given that his father is not an Anglican. We assume again that α and β are the same from generation to generation. Although this assumption is not as reasonable as it was in the case of the bacteria, we will make it for reasons of simplicity. A more realistic situation, in which these conditional probabilities change from generation to generation, could be accommodated, but we will not do that in this example.

Let $P(A_0)$ be the probability that a man, selected at random in Northern Ireland today, is an Anglican. Let $P(A_1)$ be the probability that his father, and $P(A_2)$ that his grandfather were Anglicans. Finally, going back to the time of the foundation of Eire and the beginning of what became known as the Irish Troubles, let $P(A_3)$ be the probability that the great-grandfather of our man was baptized in the Anglican tradition. This situation can be represented as a short finite series, namely as an equation like (12), which we here rewrite for $n = 2$:

$$\begin{aligned}
 P(A_0) &= \beta + (\alpha - \beta)[\beta + (\alpha - \beta)[\beta + (\alpha - \beta)P(A_3)]] \\
 &= \beta \left[1 + (\alpha - \beta) + (\alpha - \beta)^2 \right] + (\alpha - \beta)^3 P(A_3) \\
 &= \frac{\beta}{1 - \alpha + \beta} + (\alpha - \beta)^3 \left[P(A_3) - \frac{\beta}{1 - \alpha + \beta} \right].
 \end{aligned}
 \tag{14}$$

Suppose that $\alpha = 0.8$ and $\beta = 0.1$. In this case we certainly need to make a guess for $P(A_3)$, the probability that the great-grandfather was an Anglican, in order to find out the value of $P(A_0)$. Moreover, it matters very much which posit is used for $P(A_3)$, because $P(A_3) = 0$ results in $P(A_0) = 0.219$, while $P(A_3) = 1$ yields $P(A_0) = 0.562$. To obtain an accurate value for $P(A_0)$, we evidently need an accurate posit for $P(A_3)$. How are we to obtain it?

One way of doing that is the following. Imagine that we already know, for example on the basis of an Irish ecclesiastical census, the percentage of males who at the present time are Anglicans, i.e. we already have an accurate value of $P(A_0)$ at hand. Then of course we do not need $P(A_3)$ to compute $P(A_0)$, but it might happen that we wish to know the value of $P(A_3)$ for the calculation of other quantities of interest. If so, we can use our knowledge of $P(A_0)$ to compute $P(A_3)$, and to do that we invert Eq. 14:

$$P(A_3) = \frac{\beta}{1 - \alpha + \beta} + \left[P(A_0) - \frac{\beta}{1 - \alpha + \beta} \right] / (\alpha - \beta)^3.
 \tag{15}$$

Here $P(A_3)$ plays the role of an appraised posit that is moreover perfectly successful: it is the value of great-grandfather’s probability that would have yielded precisely the correct value for $P(A_0)$. For example, if one in two males are currently Anglicans in Northern Ireland, i.e. $P(A_0) = 0.5$, we find from Eq. 15 that $P(A_3) = 0.82$; thus 82% of the male population in great-grandfather’s day were Anglicans.

This appraised posit, $P(A_3) = 0.82$, cannot of course be employed to determine $P(A_0)$ without circularity; but we could use it to evaluate certain other things we might wish to know, for example the probable income of the Anglican church in Northern Ireland shortly after partition, or the degree of emigration to England in the following generation, and so on.

7 Dispensing with Uniformity

Up to this point, all our arguments have been made under the uniformity assumption that α and β remain constant throughout the entire chain of reasoning. As we have seen, this assumption comes naturally in the case of bacteria cultivated in the laboratory, but it is somewhat artificial when dealing with Anglican Irishmen. More often than not, the conditional probabilities α and β will change from generation to generation.

Let us therefore now drop the uniformity assumption, and suppose that α and β vary with n . We can express this by adding an index:

$$\alpha_n = P(A_n|A_{n+1}) \quad \text{and} \quad \beta_n = P(A_n|\neg A_{n+1}),
 \tag{16}$$

where α_n and β_n may now depend nontrivially on n . In the limit as n tends to infinity, the second line of Eq. 12 implies

$$P(A_0) = \beta \left[1 + (\alpha - \beta) + (\alpha - \beta)^2 + (\alpha - \beta)^3 + \dots \right]. \tag{17}$$

In (17) α and β are constant, and the generalization to non-uniform conditional probabilities is

$$P(A_0) = \beta_0 + \sum_{n=1}^{\infty} \gamma_0 \gamma_1 \dots \gamma_{n-1} \beta_n, \tag{18}$$

where we have put $\gamma_n = \alpha_n - \beta_n$ for convenience. While it is clear that Eq. 17 is a special case of Eq. 18, the correctness of the latter equation still needs to be established. In fact, the proof of (18) requires some computational efforts, and these are given in the Appendix, notably in Eqs. 26–28.

Equation 18 is only correct under the condition that the series converges. This condition puts some relatively mild restrictions on the allowable expressions that we use for β_n and γ_n . However, many examples can be given in which the entire series is not only convergent, but also explicitly summable. An example is $\beta_n = bz^n$ and $\gamma_n = a(c + n)/(1 + n)$, where a, b, c and z are constants, all lying in the interval (0,1). For then we find

$$P(A_0) = b(1 - az)^{-c},$$

with further restrictions to guarantee that $P(A_0) < 1$. With these formulas for β_n and γ_n , it is the case that both $\alpha_n = bz^n + a(c + n)/(1 + n)$ and β_n depend on n . As n tends to infinity, α_n tends to a and β_n to 0, on condition that $z < 1$. More complicated, but still explicitly summable forms for β_n and γ_n can be given that generate a hypergeometric series. However, we will not give the details, since they are irrelevant for the main point that we are making, namely that our argument in no way requires α_n and β_n to be uniform, i.e. to be independent of n .

8 Bayesian Updating and Appraised Posits

It has been known since at least the 19th century, and perhaps even the 17th century, that the word ‘probable’ is ambiguous: it can be either objective (‘ontological’) or epistemological (‘subjective’). It is moreover common knowledge that both are interpretations of Kolmogorov’s axiomatic scheme. However, when it comes to the question of the relation between the two interpretations, there is much dissent. Are they basically the same, in the sense that the one can be reduced to the other? Or are they completely disjunct, and is it merely a coincidence that we use the word ‘probability’ for both? Each view has had its advocates. Whereas Carnap for example embraced a ‘disparity conception’ of probability, others claimed an ‘identity view’, espousing a reduction in one of the two possible directions.²

² Frank Ramsey is often ranked with Carnap in this matter. Galavotti however argues that, according to Ramsey, probability in physics can be accounted for in terms of belief of a special sort. Therefore it would be a mistake to call Ramsey a dualist (Galavotti 2005, p. 204).

Today many people realise that neither a disparity nor an identity view is particularly fruitful, since the former makes it difficult to understand why both interpretations obey the same ‘syntax’ (Kolmogorov’s axioms), while the latter neglects their far-reaching differences in the field of semantics. In recent years several attempts have been made to pave a third way, one that incorporates the strong and eliminates the weak aspects of both interpretations. The most promising of these attempts provide bridges between the objective notion of *chance* and the subjective concept of *credence* or degree of belief. Thus David Lewis’s Principal Principle states that, if the chance that an event e occurs were known to be r , and no other relevant evidence were available, then your credence that e occurs had better be r as well (Lewis 1980). And Howson and Urbach, to mention another example, have deftly clothed von Mises’ objective theory of chance in subjectivist, Bayesian raiment (Howson and Urbach 1989, pp. 344–347).

In this section we propose to contribute to the general project of connecting chance and credence by showing that Reichenbach’s objectivistic approach is intimately linked to subjectivistic Jeffrey conditionalization. More specifically, we will prove that, notwithstanding substantial epistemological and methodological differences, the two approaches can yield the same result: what we shall call *ideal* posits à la Reichenbach are equal to what we shall call *invariant* Jeffrey updates.

For simplicity we show this equivalence here for a single step, rather than a sequence of them. The general proof, for any number of steps, and also without the assumption of uniformity, is to be found in the Appendix.

If it is not sure whether A_1 has occurred or not, we can assign a certain probability to that occurrence, which we designate $P_{\text{old}}(A_1)$. Supposing the conditional probabilities $P(A_0|A_1) = \alpha$ and $P(A_0|\neg A_1) = \beta$ to be known, the Bayesian updating of $P_{\text{old}}(A_1)$ to $P_{\text{new}}(A_1)$ is made by identifying the latter with $P(A_1|A_0)$, i.e. $P_{\text{new}}(A_1) \equiv P(A_1|A_0)$, where

$$\begin{aligned}
 P(A_1|A_0) &= \frac{P(A_0|A_1)P_{\text{old}}(A_1)}{P(A_0|A_1)P_{\text{old}}(A_1) + P(A_0|\neg A_1)P_{\text{old}}(\neg A_1)} \\
 &= \frac{\alpha P_{\text{old}}(A_1)}{\beta + (\alpha - \beta)P_{\text{old}}(A_1)}.
 \end{aligned}
 \tag{19}$$

The above classic or Bayesian updating is based on the assumption that there is no doubt that A_0 has indeed occurred. A_0 is, as it were, incoming indubitable evidence that is used to improve the estimate of A_1 ’s probability of occurrence.

Jeffrey’s generalization of this updating starts from the idea that incoming evidence, A_0 , will always carry its own modicum of uncertainty. Instead of $P_{\text{new}}(A_1) \equiv P(A_1|A_0)$, we write, following Jeffrey’s lead,

$$P_{\text{new}}(A_1) = P_{\text{old}}(A_1|A_0)P_{\text{new}}(A_0) + P_{\text{old}}(A_1|\neg A_0)P_{\text{new}}(\neg A_0),
 \tag{20}$$

where $P_{\text{old}}(A_1|A_0)$ is modelled on the Bayesian update (19), namely

$$\begin{aligned}
 P_{\text{old}}(A_1|A_0) &= \frac{P(A_0|A_1)P_{\text{old}}(A_1)}{P(A_0|A_1)P_{\text{old}}(A_1) + P(A_0|\neg A_1)P_{\text{old}}(\neg A_1)} \\
 &= \frac{\alpha P_{\text{old}}(A_1)}{\beta + (\alpha - \beta)P_{\text{old}}(A_1)}.
 \end{aligned}
 \tag{21}$$

Similarly, $P_{\text{old}}(A_1|\neg A_0)$ is defined by the Bayesian update

$$\begin{aligned}
 P_{\text{old}}(A_1|\neg A_0) &= \frac{P(\neg A_0|A_1)P_{\text{old}}(A_1)}{P(\neg A_0|A_1)P_{\text{old}}(A_1) + P(\neg A_0|\neg A_1)P_{\text{old}}(\neg A_1)} \\
 &= \frac{(1 - \alpha)P_{\text{old}}(A_1)}{1 - \beta - (\alpha - \beta)P_{\text{old}}(A_1)}.
 \end{aligned}
 \tag{22}$$

The suffix ‘old’ on $P_{\text{old}}(A_1|A_0)$ and $P_{\text{old}}(A_1|\neg A_0)$ is intended to stress the fact that they are functions of $P_{\text{old}}(A_1)$, the old or pre-update value of the probability of A_1 ’s occurrence. It is supposed that the known conditional probabilities $\alpha = P(A_0|A_1)$ and $\beta = P(A_0|\neg A_1)$ are unequal, $\alpha \neq \beta$, for otherwise the events A_0 and A_1 would be probabilistically independent of one another, and it would be senseless to try to use knowledge about the one to update knowledge about the other.

With the identifications (21) and (22), the Jeffrey update (20) becomes

$$P_{\text{new}}(A_1) = \frac{\alpha P_{\text{old}}(A_1)P_{\text{new}}(A_0)}{\beta + (\alpha - \beta)P_{\text{old}}(A_1)} + \frac{(1 - \alpha)P_{\text{old}}(A_1)[1 - P_{\text{new}}(A_0)]}{1 - \beta - (\alpha - \beta)P_{\text{old}}(A_1)}.
 \tag{23}$$

We say that the Jeffrey updating is *invariant* with respect to A_0 if the new value of the probability of A_1 ’s occurrence is equal to the old value, i.e.

$$P_{\text{new}}(A_1) = P_{\text{old}}(A_1).
 \tag{24}$$

Evidently this is the best possible value of $P_{\text{old}}(A_1)$, in the sense that updating it by means of $P_{\text{new}}(A_0)$ has no effect at all. When this condition of invariance is satisfied, one can cancel $P_{\text{old}}(A_1)$ out from both sides of Eq. 23, on condition of course that $P_{\text{old}}(A_1) \neq 0$. After some algebra one finds

$$(\alpha - \beta)[1 - P_{\text{old}}(A_1)][P_{\text{new}}(A_0) - \beta - (\alpha - \beta)P_{\text{old}}(A_1)] = 0.$$

Under our assumption that $\alpha \neq \beta$, and on condition that $P_{\text{old}}(A_1) \neq 1$, the first two factors above do not vanish, and so the third factor must be zero, i.e.

$$P_{\text{new}}(A_0) = \beta + (\alpha - \beta)P_{\text{old}}(A_1).
 \tag{25}$$

For a single step, Eq. 25 is precisely Reichenbach’s formula for the perfectly successful appraised posit, i.e. the appraised posit $P_{\text{old}}(A_1)$ that leads to the exact probability $P_{\text{new}}(A_0)$ (cf. Eq. 8). Such an appraised posit that leads to the exact probability we will call an *ideal posit*. The example of the Irish great-grandfather’s probability of being an Anglican, which was tailored to the known probability associated with contemporary Irish males, was an ideal posit in the case of three steps rather than one.

As to the inverse proposition, i.e. that Eq. 25 implies Eq. 24, this is also readily proved. If Eq. 25 holds, the denominators in Eq. 23 are respectively $P_{\text{new}}(A_0)$ and $1 - P_{\text{new}}(A_0)$, so the latter equation degenerates into

$$P_{\text{new}}(A_1) = \alpha P_{\text{old}}(A_1) + (1 - \alpha)P_{\text{old}}(A_1) = P_{\text{old}}(A_1).$$

Thus if the Reichenbach posit is ideal, then the corresponding Jeffrey update is invariant, and the converse.

Note that the equivalence between ideal posits and invariant updates also applies if $\alpha = \beta$. In fact, the proof of the equivalence is then completely trivial. For if $\alpha = \beta$, then A_0 and A_1 are independent, as we remarked. And once we have independence, the Jeffrey update is trivially invariant and the Reichenbach posit is trivially ideal: $P_{\text{new}}(A_0)$ is always equal to β . For this reason, we are always interested in the case where $\alpha \neq \beta$, so that A_0 and A_1 are not independent of one another.

Earlier we remarked that the equivalence between ideal posits and invariant updates, together with such rules as the Principal Principle, might contribute to a rapprochement between talk about credence and talk about chance. The equivalence to which we lay claim is an implication of Reichenbach’s and Jeffrey’s assumptions (together with an implicit use of the Principal Principle), inasmuch as the former apply to chance and the latter to credence, receiving its justification from the proof that has been sketched above and that in its generality is given in the Appendix below.

Acknowledgements We thank the members of the Groningen research group PCCP (Promotion Club Cognitive Patterns) for helpful remarks, in particular Theo Kuipers and Erik Krabbe. We also gratefully acknowledge comments from Igor Douven, that led to considerable improvements in our notation. Useful criticism from three reviewers has been incorporated into the final version of the paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix

It will be shown that if Reichenbach’s ideal posit is imposed, then the corresponding Jeffrey update for the probability of the occurrence of the event A_{n+1} is invariant; and, conversely, if the Jeffrey update is invariant, then $P(A_{n+1})$ is given by Reichenbach’s ideal posit.

Jeffrey’s update of the probability of A_{n+1} , from an old value, $P_{\text{old}}(A_{n+1})$, to a new value, $P_{\text{new}}(A_{n+1})$, can be written

$$P_{\text{new}}(A_{n+1}) = \left\{ \frac{P(A_0|A_{n+1})P_{\text{new}}(A_0)}{P_{\text{old}}(A_0)} + \frac{P(\neg A_0|A_{n+1})P_{\text{new}}(\neg A_0)}{P_{\text{old}}(\neg A_0)} \right\} P_{\text{old}}(A_{n+1}). \tag{26}$$

Here $P_{\text{old}}(A_0)$ is not in general equal to $P_{\text{new}}(A_0)$, rather it is the following function of the conditional probabilities, and of $P_{\text{old}}(A_{n+1})$:

$$P_{\text{old}}(A_0) = \sum_{m=0}^n Q_m P(A_m|\neg A_{m+1}) + Q_{n+1} P_{\text{old}}(A_{n+1}), \tag{27}$$

where $Q_0 = 1$ and

$$Q_n = \prod_{m=0}^{n-1} P(A_m|A_{m+1}),$$

for $n \geq 1$. Here the Jeffrey relevance of A_{m+1} to A_m is defined by

$$P(A_m|A_{m+1}) = P(A_m|A_{m+1}) - P(A_m|\neg A_{m+1}). \tag{28}$$

If $P_{\text{old}}(A_{n+1})$ were the correct value for the probability of the occurrence of the event A_{n+1} , then $P_{\text{old}}(A_0)$, as defined by Eqs.(27)–(28), would be the correct value for the probability of the occurrence of the event A_0 (see Atkinson and Peijnenburg (2006) Appendix).

The *ideal* Reichenbach posit is defined by

$$\text{IRP} : P_{\text{new}}(A_0) = P_{\text{old}}(A_0), \tag{29}$$

where the right-hand side is to be understood through its definition (27)–(28). The *invariant* Jeffrey update is specified by

$$\text{IJU} : P_{\text{new}}(A_{n+1}) = P_{\text{old}}(A_{n+1}). \tag{30}$$

We shall show that IRP and IJU are equivalent.

Proof of the implication IRP \rightarrow IJU : If $P_{\text{new}}(A_0) = P_{\text{old}}(A_0)$, then these probabilities may be cancelled out from the numerators and denominators in Eq. 26, yielding

$$P_{\text{new}}(A_{n+1}) = \{P(A_0|A_{n+1}) + P_{\text{new}}(\neg A_0|A_{n+1})\}P_{\text{old}}(A_{n+1}) = P_{\text{old}}(A_{n+1}),$$

which is IJU. □

Proof of the implication IJU \rightarrow IRP : The Jeffrey update (26) can be rewritten

$$P_{\text{new}}(A_{n+1}) = \left\{ 1 + \frac{[P_{\text{new}}(A_0) - P_{\text{old}}(A_0)][P(A_0|A_{n+1}) - P_{\text{old}}(A_0)]}{P_{\text{old}}(A_0)P_{\text{old}}(\neg A_0)} \right\} P_{\text{old}}(A_{n+1}),$$

so if $P_{\text{new}}(A_{n+1}) = P_{\text{old}}(A_{n+1})$, then one or other of the factors in the numerator of the above fraction must vanish. Either $P_{\text{new}}(A_0) = P_{\text{old}}(A_0)$, which is IRP, or $P(A_0|A_{n+1}) = P_{\text{old}}(A_0)$. In the latter case, since

$$P_{\text{old}}(A_0) = P(A_0|A_{n+1})P_{\text{old}}(A_{n+1}) + P(A_0|\neg A_{n+1})P_{\text{old}}(\neg A_{n+1}),$$

which can be rewritten

$$P_{\text{old}}(A_0) = P(A_0|A_{n+1}) - \left[P(A_0|A_{n+1}) - P(A_0|\neg A_{n+1}) \right] P_{\text{old}}(\neg A_{n+1}),$$

it follows that $P_{\text{old}}(A_0) = P(A_0|A_{n+1})$ implies $P(A_0|A_{n+1}) = P(A_0|\neg A_{n+1})$, since $P_{\text{old}}(\neg A_{n+1}) \neq 0$. This means that A_0 and A_{n+1} are independent. In this case $P_{\text{new}}(A_0) = P(A_0|A_{n+1})$ independently of the value of $P_{\text{old}}(A_{n+1})$, so that $P_{\text{new}}(A_0) = P_{\text{old}}(A_0)$, which is again IRP. □

Summarizing, the Ideal Reichenbach Posit for chances is equivalent to the Invariant Jeffrey Update for credences, subject only to the constraints that neither $P_{\text{old}}(A_{n+1})$ nor $P_{\text{old}}(A_0)$ have either of the extreme values 0 or 1.

References

- Atkinson, D., & Peijnenburg, J. (2006). Probability without certainty. Foundationalism and the Lewis–Reichenbach debate. *Studies in History and Philosophy of Science Part A*, 37(3), 442–453.
- Galavotti, M. C. (2003). Kinds of probabilism. In: P. Parrini, W. C. Salmon, & M. H. Salmon (Eds.), *Logical empiricism* (pp. 281–303). Pittsburgh: University of Pittsburgh Press.
- Galavotti, M. C. (2005). *Philosophical introduction to probability*. Stanford: CSLI Publications.
- Howson, C., & Urbach, P. (1989). *Scientific reasoning, The Bayesian approach*. Second Printing of Second Edition. Chicago: Open Court, 1996.
- Lewis, C. I. (1929). *Mind and the world-order. Outline of a theory of knowledge*. New York: Charles Scribner's Sons, Reprinted by Dover 1956.
- Lewis, C. I. (1952). The given element in empirical knowledge. *The Philosophical Review*, 61(2), 168–172.
- Lewis, D. (1980) A Subjectivist's guide to objective chance. In: Jeffrey, R. C. (Ed.), *Studies in inductive logic and probability* (Vol. II, pp. 263–293). Berkeley: University of California Press.
- Peijnenburg, J. (2007). Infitism regained. *Mind*, 116, 597–602.
- Reichenbach, H. (1915). Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit. Inaugural Dissertation, University of Erlangen, Leipzig, Barth.
- Reichenbach, H. (1949). *The theory of probability* (2nd ed.). Berkeley: University of California Press, 1971.
- Reichenbach, H. (1951). *The rise of scientific philosophy*, Berkeley etc., University of California Press, Second Edition, 1971.
- Reichenbach, H. (1952). Are phenomenal reports absolutely certain?. *The Philosophical Review*, 61(2), 147–159.
- Reichenbach, H. (1978). A letter of March 28, 1949 from Hans Reichenbach to Bertrand Russell. In: M. Reichenbach & R. S. Cohen (Eds.), *Hans Reichenbach, selected writings: 1909–1953* (Vol. II, pp. 405–411). Dordrecht: Reidel.
- Salmon, W. C. (1966). *The foundations of scientific inference*. Pittsburgh: University of Pittsburgh Press, Third Printing, 1971.
- Salmon, W. C. (1991). Hans Reichenbach's vindication of induction. *Erkenntnis*, 35(1–3), 99–122.
- Von Mises, R. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 5(12), 52–99.