

# *Subjective Probabilities as Basis for Scientific Reasoning?*

Franz Huber

Penultimate version, please cite the *The British Journal for the Philosophy of Science* paper

## ABSTRACT

Bayesianism is the position that scientific reasoning is probabilistic and that probabilities are adequately interpreted as an agent's actual subjective degrees of belief, measured by her betting behaviour.

Confirmation is one important aspect of scientific reasoning. The thesis of this paper is the following: If scientific reasoning is at all probabilistic, the subjective interpretation has to be given up in order to get right confirmation—and thus scientific reasoning in general.

- 1 *The Bayesian Approach to Scientific Reasoning*
- 2 *Bayesian Confirmation Theory*
- 3 *The Example*
- 4 *The Less Reliable the Source of Information, the Higher the Degree of Bayesian Confirmation*
- 5 *Measure Sensitivity*
- 6 *A More General Version of the Problem of Old Evidence*
- 7 *Conditioning on the Entailment Relation*
- 8 *The Counterfactual Strategy*
- 9 *Generalizing the Counterfactual Strategy*
- 10 *The Desired Result—and a Necessary and Sufficient Condition for it*
- 11 *Actual Degrees of Belief*
- 12 *The Common Knock-Down Feature or Anything Goes*

## 1 The Bayesian Approach to Scientific Reasoning

Bayesianism is the position that scientific reasoning is probabilistic and that probabilities are adequately interpreted as an agent's actual subjective degrees of belief, measured by her betting behaviour.

Confirmation is one important aspect of scientific reasoning. The thesis of this paper is the following: Given that scientific reasoning—and thus confirmation—is at all probabilistic, the subjective interpretation of probability has to be given up in order to get right confirmation—and thus scientific reasoning in general.

This will be argued for as follows. First, an example will be considered that is an instance of a more general version of the problem of old evidence (POE). This suggests that we look whether the two existing solutions to POE—conditioning on the entailment relation (Garber [1983]) and the counterfactual strategy (Howson and Urbach [1993])—provide a solution to the more general problem (called C, for 'counterintuitive').

As a first result, we get that these two solutions to POE are not genuine solutions, because they do not provide a solution to the more general C.

More importantly, the solutions to C considered here all have in common that they depend on the agent's very first guess, her first degree-of-belief function  $Pr_0$ .

C thus leads to the problem of prior probabilities, POPP. However, the standard solution to POPP—the 'washing out of priors' relying on convergence to certainty and merger of opinion (cf. Earman [1992], especially pp. 57-9 and ch. 6)—is not applicable here, because the solutions to C never get rid of the agent's first degree-of-belief function  $Pr_0$ .

On the subjective interpretation of probability,  $Pr_0$  is any arbitrary assignment of values in  $[0, 1]$  to the atomic propositions of the underlying language. By choosing an appropriate  $Pr_0$ , one can obtain more or less any degree of confirmation.

The only way out is some kind of objective or logical probability function the agent could adopt as her first degree-of-belief function  $Pr_0$ . However, the difficulty of determining such a logical probability function was precisely the reason for turning to the subjective interpretation of probability.

## 2 Bayesian Confirmation Theory

According to Bayesian confirmation theory, an agent's degree of confirmation of hypothesis  $H$  by evidence  $E$  relative to background knowledge  $B$  is measured by some function  $C_{Pr}$  such that

$$C_{Pr}(H, E, B) > 0 \quad \Leftrightarrow \quad Pr(H | E, B) > Pr(H | B)$$

$$C_{Pr}(H, E, B) = 0 \quad \Leftrightarrow \quad Pr(H | E, B) = Pr(H | B)$$

$$C_{Pr}(H, E, B) < 0 \quad \Leftrightarrow \quad Pr(H | E, B) < Pr(H | B),$$

where  $Pr$  is the agent's degree-of-belief function. Any such function  $C_{Pr}$  is called a *relevance measure* (based on  $Pr$ ).

One example is the *distance measure*  $d_{Pr}$ ,

$$d_{Pr}(H, E, B) = Pr(H | E, B) - Pr(H | B).$$

## 3 The Example

An agent with degree-of-belief function  $Pr$  considers the hypothesis

$H =$  All Scots wear kilts.

At time  $t_1$  she has the impression of seeing her friend Stephen wearing a kilt. As the agent is not wearing her glasses, her degree of belief in

$E =$  Stephen wears a kilt

is not very high, say

$$Pr_1(E | B_1) = .6,$$

where  $Pr_1$  is her degree-of-belief function at  $t_1$ .  $B_1$  is her background knowledge at that time, which contains the information that Stephen is Scottish.

Because she knows that  $H$  and  $B_1$  logically imply  $E$ , the agent gets interested in whether Stephen is indeed wearing a kilt. So she puts on her glasses and has a careful second look at Stephen, who still seems to be wearing a kilt; this all happens at time  $t_2$ .

In passing from  $t_1$  to  $t_2$ , the only change in the agent's degrees of belief is in  $E$ . Moreover, for some reason the agent cannot express her observation in terms of a proposition. So her degree of belief in  $E$  increases *exogenously*, say to

$$Pr_2(E | B_2) = .9,$$

where  $Pr_2$  is the agent's degree-of-belief function at  $t_2$ . Her background knowledge  $B_2$  at  $t_2$  is the same as at  $t_1$ , because the only change is in  $E$  and that change is exogenous, *i.e.* not due to any proposition that is fully believed and could thus be conditioned on. So  $B_1$  is logically equivalent to  $B_2$ ,  $B_1 \equiv B_2$ .

$E$  is positively relevant for  $H$  given  $B$  (in the sense of  $Pr_1$ ). Furthermore, the agent's degree of belief in  $E$  increases from  $Pr_1(E | B_1) = .6$  at  $t_1$  to  $Pr_2(E | B_2) = .9$  at  $t_2$ . Therefore, by Jeffrey conditionalisation, her degree of belief in  $H$  must also increase, namely from  $Pr_1(H | B_1)$  to

$$\begin{aligned} Pr_2(H | B_2) &= Pr_1(H | E, B_1) \cdot Pr_2(E | B_2) + Pr_1(H | \neg E, B_1) \cdot Pr_2(\neg E | B_2) \\ &= Pr_1(H | E, B_1) \cdot Pr_2(E | B_2) \\ &= Pr_1(H | B_2) \cdot Pr_2(E | B_2) / Pr_1(E | B_1). \end{aligned}$$

Obviously, the increase in the agent's degree of belief in  $H$  is greater, the larger the increase is in her degree of belief in  $E$  (which is logically implied by and thus positively relevant for  $H$  given  $B_1$ —or its equivalent  $B_2$ —in the sense of  $Pr_1$ ).

#### 4 The Less Reliable the Source of Information, the Higher the Degree of Bayesian Confirmation

Let us compare the agent's degrees of confirmation at time  $t_1$  and at time  $t_2$ .<sup>1</sup>

The agent knows that the conjunction of  $H$  and  $B_1$  logically implies  $E$ , and, as time passes, she does not forget this nor that Stephen is Scottish. Thus

$$Pr_j(E | H, B_j) = 1, \text{ for all points of time } t_j, j \geq 0,$$

even if it is not assumed that she is logically omniscient in the sense that all logical truths are transparent to her (cf. Earman [1992], p. 122).

Given Jeffrey conditionalisation, *i.e.* assuming

$$Pr_1(H | \pm E, B_1) = Pr_2(H | \pm E, B_2),$$

( $B_1$  and  $B_2$  are logically equivalent), it follows that

$H$  is more confirmed by  $E$  relative to  $B_1$  at  $t_1$  than at  $t_2$  relative to  $B_2$  if and only if the agent's degree of belief in  $E$  given  $B_1$  at  $t_1$  is smaller than her degree of belief in  $E$  given  $B_2$  at  $t_2$ , *i.e.*

$$d_{Pr_1}(H, E, B_1) > d_{Pr_2}(H, E, B_2) \quad \Leftrightarrow \quad Pr_2(E | B_2) > Pr_1(E | B_1).$$

More generally,

$$C \quad d_{Pr_1}(H, E, B_1) > d_{Pr_2}(H, E, B_2) \Leftrightarrow Pr_1(E | H, B_1) > Pr_1(E | B_1) \ \& \ Pr_2(E | B_2) > Pr_1(E | B_1)$$

or

$$Pr_1(E | H, B_1) < Pr_1(E | B_1) \ \& \ Pr_2(E | B_2) < Pr_1(E | B_1),$$

where the only change in the agent's degrees of belief in passing from  $t_1$  to  $t_2$  is exogenous and in  $E$ , whence  $B_1$  is logically equivalent to  $B_2$ , and Jeffrey conditionalisation (JC) is used.

Here and in the following, the probabilities of all contingent propositions involved are assumed to be positive.

C is counterintuitive, because  $E$ —which is positive evidence for hypothesis  $H$ —should not provide less and less confirmation for  $H$  when it becomes more and more established, much less cease to provide any confirmation in the limiting case when it becomes a certainty (which is the problem of old evidence).

On the contrary, the more certain it becomes that such positive evidence  $E$  is true, the more this should support  $H$ .

If some  $E$  speaks in favour of some  $H$ —say, because it is a logical consequence of the latter—then getting to know that  $E$  is probably false should not provide confirmation for  $H$ ;

rather,  $H$  should be *disconfirmed* by that. On the other hand, getting to know that  $E$  is probably true should provide confirmation for  $H$ —and the more probable it is that  $E$  is true, the more it should do so.

Finally, instead of considering  $t_1$  and  $t_2$  as two successive points of time, one may alternatively view them as two possible situations or worlds differing from each other just in the respect that the agent's degree of belief in  $E$  is lower in  $t_1$  than in  $t_2$ .

If  $H$  and  $B$  ( $\equiv B_1 \equiv B_2$ ) logically imply  $E$ , or more generally, if

$$Pr_1(E | H, B) = Pr_2(E | H, B),$$

and  $Pr_1(H | B)$  and  $Pr_2(H | B)$  are assumed to equal each other<sup>2</sup>, the following holds,

independently of whether  $E$  is positively or negatively relevant for  $H$  given  $B$ :

$H$  is more confirmed by  $E$  relative to  $B$  in  $t_1$  than in  $t_2$  just in case the agent's degree of belief in  $E$  in  $t_1$  is lower than in  $t_2$ .

## 5. Measure Sensitivity

As shown by Fitelson ([2001]), many arguments in the literature on Bayesian confirmation theory are measure sensitive in the sense that their validity depends on which relevance measure one takes as measure of confirmation.

The example of the preceding section (C) is no exception. C holds for the distance measure  $d_{Pr}$ , the *log-likelihood ratio*  $l_{Pr}$ , and the *ratio measure*  $r_{Pr}$ ,

$$\begin{aligned} l_{Pr}(H, E, B) &= \log[Pr(E | H, B)/Pr(E | \neg H, B)] \\ &= \log[(Pr(H | E, B) \cdot Pr(\neg H | B))/(Pr(\neg H | E, B) \cdot Pr(H | B))], \end{aligned}$$

$$r_{Pr}(H, E, B) = \log[Pr(H | E, B)/Pr(H | B)].$$

C does not hold for  $s_{Pr}$  (Christensen [1999])<sup>3</sup>,

$$s_{Pr}(H, E, B) = Pr(H | E, B) - Pr(H | \neg E, B) = d_{Pr}(H, E, B) \cdot [1/Pr(\neg E | B)],$$

because the latter is invariant with regard to exogenous belief changes in  $E^4$  (which yield  $B_1$  logically equivalent to  $B_2$ ), *i.e.*

$$s_{Pr1}(H, E, B_1) = Pr_1(H | E, B_1) - Pr_1(H | \neg E, B_1)$$

$$= Pr_2(H | E, B_2) - Pr_2(H | \neg E, B_2) = s_{Pr_2}(H, E, B_2).$$

Indeed, the same holds true of *every* function of  $Pr(H | E, B)$  and  $Pr(H | \neg E, B)$ .

In case of  $c_{Pr}$  (Carnap [1962]),

$$c_{Pr}(H, E, B) = Pr(H, E, B) \cdot Pr(B) - Pr(H, B) \cdot Pr(E, B) = d_{Pr}(H, E, B) \cdot Pr(E, B) \cdot Pr(B),$$

something different—but not much better—holds:

$$C' \quad c_{Pr_1}(H, E, B_1) > c_{Pr_2}(H, E, B_2) \Leftrightarrow$$

$$Pr_1(E | H, B_1) > Pr_1(E | B_1) \ \& \ Pr_1(E, B_1)/Pr_2(E, B_2) > Pr_2(\neg E, B_2)/Pr_1(\neg E, B_1)$$

or

$$Pr_1(E | H, B_1) < Pr_1(E | B_1) \ \& \ Pr_1(E, B_1)/Pr_2(E, B_2) < Pr_2(\neg E, B_2)/Pr_1(\neg E, B_1).$$

## 6 A More General Version of the Problem of Old Evidence

C is a more general version of the problem of old evidence (POE).

C says that evidence  $E$ —which is positively relevant for hypothesis  $H$  given background knowledge  $B^5$ —provides more confirmation for  $H$  relative to  $B$ , the less the agent believes in  $E$ . In the limiting case of POE where  $E$  is known,  $E$  ceases to provide any confirmation at all.

Conversely, if  $E$  is negatively relevant for  $H$  given  $B$ ,  $E$  provides the less *disconfirmation* for  $H$  relative to  $B$ , the more the agent believes in  $E$ . In the limiting case of POE where  $E$  is known,  $E$  ceases to provide any disconfirmation at all.

POE is that evidence  $E$  that is old in the sense of being assigned a degree of belief of 1 cannot provide any confirmation, since for any  $Pr$ ,  $H$ ,  $E$  and  $B$ :

$$Pr(H | E, B) = Pr(H | B), \text{ if } Pr(E | B) = 1.$$

POE is a problem, because there are historical cases in which old evidence did provide confirmation (for hypotheses, both old and new—cf. chapter 5 of Earman [1992] for an excellent discussion).

If POE is a problem, so is C.

This is important, because a Bayesian could simply refuse to consider  $C$  as counterintuitive. Is it not rational, she might say, that I take positively relevant  $E$  to provide the less confirmation for  $H$ , the more I already believe in  $E$  and have built this belief into my degree of belief in  $H$ ?<sup>6</sup>

This reply is perfectly reasonable, but it applies equally well to POE. However, a brief look at the literature shows that Bayesians *do* take POE to be a problem.

So let us look whether the existing solutions to POE give rise to a solution to  $C$ .

Generally, there are two ways of approaching POE:

- 1) Conditioning on the entailment relation: Garber ([1983]), Jeffrey ([1983]), Niiniluoto ([1983])<sup>7</sup>
- 2) Counterfactual strategy: Howson and Urbach ([1993])

Each of them will be considered in turn.

### 7 Conditioning on the Entailment Relation

The idea here is to distinguish between a historical and an ahistorical POE and to solve the former by noting that

what *increases* [the agent]'s confidence in [ $H$ ] is not  $E$  itself, but the *discovery* of some generally logical or mathematical relationship between [ $H$ ] and  $E$  (Garber [1983], p. 104).

Then one shows that even if  $Pr(E | B) = 1$ ,

the discovery that [ $H$  entails  $E$ ] *can* raise [the agent]'s confidence in [ $H$ ] (Garber [1983], p. 123).

Conditioning on the entailment relation does not provide a solution to  $C$ , because in our example the agent is only interested in  $E$  because she knows that the conjunction of  $H$  and  $B_1$  logically implies  $E$  (and does not forget this and that Stephen is Scottish), whence

$$Pr_j(H \text{ entails } E | B_j) = 1, \text{ for every point of time } t_j, j \geq 0.$$

Moreover, by substituting ‘ $H$  entails  $E$ ’ for  $E$ , one gets another instance of C: Given that ‘ $H$  entails  $E$ ’ is positively relevant for  $H$  given  $B$ , it provides more confirmation for  $H$ , the less the agent believes in it.

## 8 The Counterfactual Strategy

Concerning POE, Howson and Urbach ([1993]) write:

the support of [ $H$ ] by  $E$  is gauged according to the effect which one believes a knowledge of  $E$  would now have on one’s degree of belief in [ $H$ ], on the (counterfactual) supposition that one does not yet know  $E$  (Howson and Urbach [1993], pp. 404-5).

Suppose  $B - E$  is the logically weakest proposition such that

$(B - E) \wedge E$  is logically equivalent to  $B$ ,

so that  $Pr(X | B - E)$  is the agent’s degree of belief in  $X$  ‘on the (counterfactual) supposition that [she] does not yet know  $E$ ’.

Then, if  $Pr(E | B) = 1$ , the agent’s degree of confirmation is given by

$$d'_{Pr}(H, E, B) = Pr(H | B) - Pr(H | B - E) \quad \Rightarrow \quad g_5$$

‘actual’ – ‘counterfactual’,

which is positive if and only if

O      $Pr(H | B) > Pr(H | B - E)$ ,

(‘O’ for ‘obvious’) and also if and only if (‘P’ for ‘positive’)

P      $Pr(E | H, B - E) > Pr(E | B - E)$ .

However, if  $E$  is not known, it cannot be dropped from  $B$ . Therefore one has to generalize from the case of POE where  $Pr(E | B) = 1$  to the case of C where  $Pr(E | B)$  need not be 1.

The question is, of course, how the counterfactual strategy can be adequately generalized. Apart from the above, there are the following (and uncountably many more) formulations of  $d'_{Pr}(H, E, B)$ :

$$\begin{aligned}
d'_{Pr}(H, E, B) &= Pr(H | B - E, E) \cdot Pr(E | B) + Pr(H | B - E, \neg E) \cdot Pr(\neg E | B) - Pr(H | B - E) \\
&\Rightarrow g_1 \\
&= Pr(H | (B - E) \wedge E) \cdot Pr(E | B) - Pr(H | B - E) \quad \Rightarrow g_2 \\
&= Pr(H | B - E, E) - Pr(H | B - E) \quad \Rightarrow g_3 \\
&= Pr(H | B, E) - Pr(H | B - E) \quad \Rightarrow g_4
\end{aligned}$$

(The  $g_i$  refer to the generalisations considered in the next section.)

## 9 Generalizing the Counterfactual Strategy

Instead of considering

the (counterfactual) supposition that one does not yet know  $E$  (Howson and Urbach [1993], p. 405)

the quote suggests considering

the (counterfactual) supposition that one does not yet *believe in  $E$  to degree  $Pr(E | B)$* .

However, in our example the background knowledge at  $t_1$  and at  $t_2$  is the same, because the change in the agent's degree of belief in  $E$  is exogenous. Therefore one cannot just drop something (say, all information bearing on  $E$ ) from  $B_2$  to get a counterfactual supposition that could play a role analogous to that of  $B_2 - E$  in the special case where  $Pr_2(E | B_2) = 1$ .

Instead, one really has to adopt a new probability function  $Pr^E$ . Suppose, therefore, that  $Pr^E(X | B)$  is the agent's degree of belief in  $X$  on the counterfactual supposition that she does not yet believe in  $E$  to degree  $Pr(E | B)$ .

Then there are the following (and uncountably many more) ways of generalizing  $d'$ :

$$g_{1Pr}(H, E, B) = Pr^E(H | B, E) \cdot Pr(E | B) + Pr^E(H | B, \neg E) \cdot Pr(\neg E | B) - Pr^E(H | B)$$

$$g_{2Pr}(H, E, B) = Pr^E(H | B, E) \cdot Pr(E | B) - Pr^E(H | B)$$

$$g_{3Pr}(H, E, B) = Pr^E(H | B, E) - Pr^E(H | B)$$

$$g_{4Pr}(H, E, B) = Pr(H | B, E) - Pr^E(H | B)$$

$$g_{5Pr}(H, E, B) = Pr(H | B) - Pr^E(H | B)$$

### 10 The Desired Result—and a Necessary and Sufficient Condition for it

Instead of arguing for or against any of these generalisations, let us first have a look at where we want to arrive. According to Bayesian intuitions, the desired result is that

*H* is more confirmed by *E* relative to *B*<sub>2</sub> at *t*<sub>2</sub> than relative to *B*<sub>1</sub> at *t*<sub>1</sub> if and only if the agent's degree of belief in *E* given *B*<sub>2</sub> at *t*<sub>2</sub> is greater than her degree of belief in *E* given *B*<sub>1</sub> at *t*<sub>1</sub>, *i.e.*

$$C_{Pr_2}(H, E, B_2) > C_{Pr_1}(H, E, B_1) \quad \Leftrightarrow \quad Pr_2(E | B_2) > Pr_1(E | B_1),$$

provided *E* is positively relevant for *H* given *B*<sub>1</sub> ( $\equiv B_2$ ).

More generally, this means either D<sub>C</sub> or D<sub>A</sub> ('C' for 'counterfactual', 'A' for 'actual'), depending on how one construes 'positively relevant': D<sub>C</sub>

$$C_{Pr_2}(H, E, B_2) > C_{Pr_1}(H, E, B_1) \Leftrightarrow Pr_1^E(E | H, B_1) > Pr_1^E(E | B_1) \ \& \ Pr_2(E | B_2) > Pr_1(E | B_1)$$

or

$$Pr_1^E(E | H, B_1) < Pr_1^E(E | B_1) \ \& \ Pr_2(E | B_2) < Pr_1(E | B_1)$$

$$D_A \quad C_{Pr_2}(H, E, B_2) > C_{Pr_1}(H, E, B_1) \Leftrightarrow Pr_1(E | H, B_1) > Pr_1(E | B_1) \ \& \ Pr_2(E | B_2) > Pr_1(E | B_1)$$

or

$$Pr_1(E | H, B_1) < Pr_1(E | B_1) \ \& \ Pr_2(E | B_2) < Pr_1(E | B_1).$$

Before continuing, note that it is plausible to assume that counterfactual degrees of belief are stable over time, *i.e.*

$$E \quad Pr_1^E(H | B_1) = Pr_2^E(H | B_2).$$

The reason is that in going from *t*<sub>1</sub> to *t*<sub>2</sub> the only change in the agent's degrees of belief is exogenous and in *E*, and  $Pr_i^E(H | B_i)$  just is the agent's degree of belief in *H* on the counterfactual supposition that she does not yet believe in *E* to degree  $Pr_i(E | B_i)$ .

Interestingly, E sheds positive light on *g*<sub>1</sub> and *g*<sub>5</sub>, in which *B*<sub>1</sub> and *B*<sub>2</sub> are assumed to be logically equivalent:

- 1) E is necessary and sufficient for  $g_1$  to satisfy  $D_C$ , assuming ‘counterfactual Jeffrey conditionalisation’, *i.e.*  $Pr_1^E(H | \pm E, B_1) = Pr_2^E(H | \pm E, B_2)$ , and
- 2) E is necessary and sufficient for  $g_5$  to satisfy  $D_A$ , assuming Jeffrey conditionalisation.

Moreover, E reflects badly on  $g_i$ ,  $i = 2, 3, 4$ . Given counterfactual JC,

- 3) E is necessary and sufficient for  $g_2$  to satisfy F, and
- 4) E is necessary and sufficient for  $g_3$  to satisfy  $G_C$ .

Given JC,

- 5) E is necessary and sufficient for  $g_4$  to satisfy  $G_A$ .

Here

$$F \quad C_{Pr_2}(H, E, B_2) > C_{Pr_1}(H, E, B_1) \quad \Leftrightarrow \quad Pr_2(E | B_2) > Pr_1(E | B_1),$$

$$G_C \quad C_{Pr_2}(H, E, B_2) = C_{Pr_1}(H, E, B_1) = Pr_i^E(H | B_i, E) - Pr_i^E(H | B_i) = g_{3Pr_i}(H, E, B_i),$$

$$G_A \quad C_{Pr_2}(H, E, B_2) = C_{Pr_1}(H, E, B_1) = Pr_i(H | B_i, E) - Pr_i^E(H | B_i) = g_{4Pr_i}(H, E, B_i).$$

F is odd because it says that it does not matter whether  $E$  is positively relevant for  $H$  given  $B_1$  ( $\equiv B_2$ ) in the sense of  $Pr_1$  or  $Pr_2$ .

$G_C$  and  $G_A$  are odd for a Bayesian, because they have confirmation being invariant with regard to exogenous belief changes in  $E$ . They yield that the differences in the agent’s degree of belief  $Pr_i(E | B_i)$  in  $E$  at different times  $t_i$  are irrelevant for the comparison of her degrees of confirmation of  $H$  by  $E$  relative to  $B_i$  at the times  $t_i$ . For this reason the knock-down feature that confirmation is dependent on the agent’s first degree of belief function  $Pr_0$  is also true for any measure satisfying  $G_A$  or  $G_C$ .

All things considered, it seems fair to say that the proper generalisation of  $d'$  is  $g_1$  or  $g_5$ . In order to get confirmation right they both require counterfactual degrees of belief to be stable over time.

$g_1$  and  $g_5$  reduce to

$$g_{1Pr_i}(H, E, B_i) = Pr_0^E(H | B_0, E) \cdot Pr_i(E | B_i) + Pr_0^E(H | B_0, \neg E) \cdot Pr_i(\neg E | B_i) - Pr_0^E(H | B_0),$$

$$g_{5Pr_i}(H, E, B_i) = Pr_i(H | B_i) - Pr_0^E(H | B_0),$$

where the only changes in the agent's degrees of belief in going from  $t_0$  to  $t_i$  are exogenous and in  $E$ , making  $B_0$  logically equivalent to  $B_j$  for any  $j$ ,  $0 \leq j \leq i$ .

Obviously,  $g_5(H, E, B)$  is positive if and only if

$$O' \quad Pr_i(H | B_i) > Pr_0^E(H | B_0),$$

which generalizes O.

$g_1(H, E, B)$  is positive if and only if

$$P_C \quad Pr_0^E(E | H, B_0) > Pr_0^E(E | B_0) \ \& \ Pr_i(E | B_i) > Pr_0^E(E | B_0)$$

or

$$Pr_0^E(E | H, B_0) < Pr_0^E(E | B_0) \ \& \ Pr_i(E | B_i) < Pr_0^E(E | B_0),$$

which seems to be the appropriate generalisation of P in terms of counterfactual degrees of belief.

## 11 Actual Degrees of Belief

Whether or not the preceding generalisations are appropriate, they are not satisfying, because it remains questionable how the agent's counterfactual degree of belief function  $Pr^E(\cdot | B)$  is determined and related to her actual degree of belief function  $Pr(\cdot | B)$ . This question being unanswered, the counterfactual strategy does not provide a genuine solution to C.

Let us therefore consider an account couched solely in terms of actual degrees of belief (and providing a possible answer to the aforementioned question).

Generally, the example in section 3 is one in which evidence  $E$  is positively relevant for hypothesis  $H$  given the agent's current background knowledge  $B$  according to her current degree of belief function  $Pr$ ; and her degree of belief in  $E$  changes exogenously as time goes on. If there is an increase (decrease) in the agent's degree of belief in  $E$  given  $B$ , her degree of belief in  $H$  given  $B$  increases (decreases), too—and conversely, if  $E$  is negatively relevant for  $H$  given  $B$  according to  $Pr$ .

All Bayesian accounts of (incremental) confirmation measure in some way the difference between

$$Pr(H | E, B) \text{ and } Pr(H | B).$$

Given Bayes or strict conditionalisation, this is just the difference between the agent's prior and posterior degree of belief in  $H$  given  $B$  when she learns  $E$  and nothing else.

The counterfactual strategy measures the difference between the agent's actual or posterior degree of belief in  $H$  given  $B$  and her counterfactual one—the latter replacing her prior. The reason is that the prior and posterior degrees of belief in  $H$  given  $B$  coincide if  $E$  was already known.

Solving C requires something more general, because in C the agent does not learn or know  $E$ ; there is only a change in the agent's degree of belief in  $E$  given  $B$ .

This suggests considering the agent's prior and posterior degree of belief in  $H$  given  $B$  when the only change in her degrees of belief is exogenous and in  $E$ . In other words, one replaces strict conditionalisation by Jeffrey conditionalisation.

However, one cannot simply take the difference between

$$Pr_i(H | B_i) \text{ and } Pr_{i-1}(H | B_{i-1}).$$

For suppose the agent's degree of belief in  $E$  increases enormously between  $t_{i-2}$  and  $t_{i-1}$ , say from

$$Pr_{i-2}(E | B_{i-2}) = .01 \text{ to } Pr_{i-1}(E | B_{i-1}) = .9;$$

and then it increases again in going to  $t_i$ , but only slightly, say to

$$Pr_i(E | B_i) = .91.$$

Then the difference between

$$Pr_{i-2}(H | B_{i-2}) \text{ and } Pr_{i-1}(H | B_{i-1})$$

is much greater than the difference between

$$Pr_{i-1}(H | B_{i-1}) \text{ and } Pr_i(H | B_i).$$

Consequently, the difference between the prior and posterior degree of belief in  $H$  at  $t_{i-1}$  is much greater than that at  $t_i$ , although the agent's degree of belief in  $E$  at  $t_{i-1}$  is smaller than at  $t_2$ , *i.e.*

$$\left| Pr_i(H | B) - Pr_{i-1}(H | B) \right| < \left| Pr_{i-1}(H | B) - Pr_{i-2}(H | B) \right| \text{ and } Pr_i(E | B) > Pr_{i-1}(E | B).$$

The absolute value is needed for the case in which  $E$  is not positively but rather negatively relevant for  $H$  given  $B$  in the sense of  $Pr_{i-2}$ .

What one must consider instead is the difference between the agent's current degree of belief in  $H$ ,  $Pr_i(H | B_i)$  and her *first* degree of belief in  $H$ ,  $Pr_0(H | B_0)$ , where the only changes in her degrees of belief in going from  $t_0$  to  $t_i$  are exogenous and in  $E$ .

The proposal, therefore, is:

The agent's degree of (incremental) confirmation of  $H$  by  $E$  relative to  $B_i$  at time  $t_i$  is given by a *generalized* relevance measure, *i.e.* some function  $g(0,i) =: g$  such that

$$g(H, E, B) > 0 \quad \Leftrightarrow \quad Pr_i(H | E, B_i) > Pr_0(H | B_0)$$

$$g(H, E, B) = 0 \quad \Leftrightarrow \quad Pr_i(H | E, B_i) = Pr_0(H | B_0)$$

$$g(H, E, B) < 0 \quad \Leftrightarrow \quad Pr_i(H | E, B_i) < Pr_0(H | B_0),$$

where the only changes in the agent's degrees of belief in going from  $t_0$  to  $t_i$  are exogenous and in  $E$  (in which case  $B_0 \equiv B_j$ , for every  $j$ ,  $0 \leq j \leq i$ ).

An example is the *generalized* distance measure  $g_6$ ,

$$\begin{aligned} g_6(H, E, B_i) &= Pr_i(H | B_i) - Pr_0(H | B_0) \\ &= Pr_0(H | E, B_0) \cdot Pr_i(E | B_i) + Pr_0(H | \neg E, B_0) \cdot Pr_i(\neg E | B_i) - Pr_0(H | B_0) \end{aligned}$$

JC  $i$  times, and  $B_0 \equiv B_j$ , for every  $j$ ,  $0 \leq j \leq i$ .

$g_6$  satisfies  $D_A$ , and it is positive if and only if

$$P_A \quad Pr_0(E | H, B_0) > Pr_0(E | B_0) \ \& \ Pr_i(E | B_i) > Pr_0(E | B_0)$$

or

$$Pr_0(E | H, B_0) < Pr_0(E | B_0) \ \& \ Pr_i(E | B_i) < Pr_0(E | B_0),$$

which seems to be the appropriate generalisation of P in terms of actual degrees of belief.

Interestingly,  $g_1$ ,  $g_5$ , and  $g_6$  coincide, if

$$Pr_0^E(H | \pm E, B_0) = Pr_0(H | \pm E, B_0) \text{ and } Pr_0^E(H | B_0) = Pr_0(H | B_0).$$

As counterfactual degrees of belief  $Pr^E(X | B)$  in  $X$  are required to be invariant with regard to exogenous belief changes in  $E$ , this is also the promised possible answer to the question of how an agent's counterfactual degree of belief function at any time  $t_i$ ,  $Pr_i^E(\cdot | B)$ , should be related to her actual degree of belief function at that time,  $Pr_i(\cdot | B)$ : It should equal her first degree of belief function  $Pr_0(\cdot | B)$  when the only changes in her degrees of belief in going from  $t_0$  to  $t_i$  are exogenous and in  $E$  (in which case  $B \equiv B_j$ , for every  $j$ ,  $0 \leq j \leq i$ ).

## 12 The Common Knock-Down Feature *or* Anything Goes

All three measures  $g_1$ ,  $g_5$ , and  $g_6$  (and also  $g_3$ ,  $g_4$ ,  $s$ , and every function of  $Pr(H | E, B)$  and  $Pr(H | \neg E, B)$ ) have in common that their values essentially depend on the agent's first degree of belief function  $Pr_0$ .

In case  $E$  is known and logically implied by  $H$  and  $B$ , the agent's degree of confirmation of  $H$  by  $E$  relative to  $B$  at time  $t_i$  (measured by  $g_6$ ) is even uniquely determined by her initial guesses in  $E$  and  $H$ ,  $Pr_0(E | B)$  and  $Pr_0(H | B)$ !

Why the exclamation mark?

First, because this shows that the idea behind any Bayesian theory of confirmation—namely to determine an agent's degree of confirmation by her actual subjective degrees of belief—is shown to fail.

Second, because—by the subjective interpretation— $Pr_0$  is any arbitrary assignment of values in  $[0, 1]$  to the atomic propositions of the underlying language, and thus by choosing an appropriate  $Pr_0$ , one can obtain more or less any degree of confirmation.

For let  $r$  be any value in the interval

$$[Pr_i(H | B) - Pr_i(H | E, B), Pr_i(H | B) - Pr_i(H | \neg E, B)],$$

if  $Pr_i(H | E, B) > Pr_i(H | \neg E, B)$ ; and let  $r$  be any value in the interval

$$[Pr_i(H | B) - Pr_i(H | \neg E, B), Pr_i(H | B) - Pr_i(H | E, B)],$$

if  $Pr_i(H | E, B) < Pr_i(H | \neg E, B)$ , where the index to  $B$  is dropped, because all changes in the agent's degrees of belief are exogenous.

This means that  $r$  can always be chosen to be positive or negative or 0!

Then the function  $Pr_0$ ,

$$Pr_0(E | B) := [Pr_i(H | B) - Pr_i(H | \neg E, B) - r] / [Pr_i(H | E, B) - Pr_i(H | \neg E, B)],$$

$$Pr_0(\cdot | \pm E, B) := Pr_i(\cdot | \pm E, B),$$

$$Pr_0(\cdot | B) := Pr_i(\cdot | E, B) \cdot Pr_0(E | B) + Pr_i(\cdot | \neg E, B) \cdot Pr_0(\neg E | B),$$

is a conditional probability function (defined on the same  $(\sigma)$ -field as  $Pr_i$  and conditional on the same background knowledge  $B$ ) that yields that

$$g_{6Pr_i}(H, E, B) = r,$$

where  $Pr_i$  results from  $Pr_0$  by Jeffrey conditioning  $i$  times on  $E$  and where the agent's degrees of belief changed exogenously and only in  $E$  in going from  $t_0$  to  $t_i$ .

Indeed, under this assumption that  $E$  is not independent of  $H$  given  $B$  (in the sense of  $Pr_i$ ) one can have, for every generalized relevance measure, whatever one pleases:

confirmation, disconfirmation, or irrelevance!<sup>8</sup> Simply choose  $r$  from the above interval  $> 0$  for confirmation,  $< 0$  for disconfirmation, and  $= 0$  for irrelevance. Then  $Pr_0$  as defined above yields the desired result, for any generalized relevance measure, since

$$Pr_i(H | B_i) = Pr_0(H | B_0) + r.$$

### 13 The Problem of Prior Probabilities

Thus we are back at the problem of prior probabilities, POPP. According to Earman ([1992]), there are three answers to this problem:

The first is that the assignment of priors is not a critical matter, because as the evidence accumulates, the differences in priors "wash out." [...] it is fair to say that the

formal results apply only to the long run and leave unanswered the challenge as it applies to the short and medium runs. [...] The second response is to provide rules to fix the supposedly reasonable initial degrees of belief. [...] We saw that, although ingenious, Bayes's attempt is problematic. Other rules for fixing priors suffer from similar difficulties. And generally, none of the rules cooked up so far are capable of coping with the wealth of information that typically bears on the assignment of priors. [...] The third response is that while it may be hopeless to state and justify precise rules for assigning numerically exact priors, still there are plausibility considerations that can be used to guide the assignments. [...] This response [...] opens the Bayesians to a new challenge[.] [...] That is, Bayesians must hold that the appeal to plausibility arguments does not commit them to the existence of a logically prior sort of reasoning: plausibility assessment. Plausibility arguments serve to marshal the relevant considerations in a perspicuous form, yet the assessment of these considerations comes with the assignment of priors. But, of course, this escape succeeds only by reactivating the original challenge. The upshot seems to be that some form of the washout solution had better work not just for the long run but also for the short and medium runs as well ([Earman 1992], pp. 57-9).

I take the standard Bayesian answer to be that differences in the priors do not matter, because they are 'washed out' in the long run.

However, this solution is not applicable here—and would not even be, if the limit theorems of convergence to certainty and merger of opinion worked for the short and medium runs as well. For  $g_6$  and company never get rid of the agent's first degree of belief function  $Pr_0$ .

The example shows that differences in the priors do matter. Unless  $E$  is irrelevant for  $H$  given  $B$  according to the agent's actual degree of belief function  $Pr_i$ , the agent's first degree

of belief function  $Pr_0$  can be used to obtain a positive or a negative value (or 0) for any generalized relevance measure  $g(H, E, B)$ —provided  $E$  is among the atomic statements.

The only way out is some kind of objective or logical probability function the agent could adopt as her first degree of belief function  $Pr_0$ .

Yet the difficulty of determining such a logical probability function just was the reason for turning to the subjective interpretation!

*Philosophy, Probability, and Modeling Research Group*

*Centre for Junior Research Fellows*

*PO Box M682*

*University of Konstanz*

*D-78457 Konstanz, Germany*

*franz.huber@uni-konstanz.de*

#### Acknowledgements

I am grateful to Luc Bovens for comments on several earlier versions of this paper and to Josh Snyder. My research was supported by the Alexander von Humboldt Foundation, the Federal Ministry of Education and Research, and the Program for the Investment in the Future (ZIP) of the German Government through a Sofja Kovalevskaja Award to Luc Bovens.

#### References

Carnap, R. [1962] *Logical Foundations of Probability*, 2<sup>nd</sup> ed., Chicago: University of Chicago Press.

Christensen, D. [1999] 'Measuring Confirmation', *Journal of Philosophy*, **XCVI**, pp. 437-61.

Earman, J. (ed.) [1983] *Testing Scientific Theories*, Minnesota Studies in the Philosophy of Science, vol. 10, Minnesota: University of Minnesota Press.

Earman, J. [1992] *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, Mass.: The MIT Press.

Fitelson, B. [2001] *Studies in Bayesian Confirmation Theory*, University of Wisconsin-Madison: Dissertation.

Garber, D. [1983] 'Old Evidence and Logical Omniscience in Bayesian Confirmation Theory', in J. Earman (ed.), *Testing Scientific Theories*, Minnesota Studies in the Philosophy of Science, vol. 10, Minnesota: University of Minnesota Press, pp. 99-131.

Howson, C. and Urbach, P. [1993] *Scientific Reasoning: The Bayesian Approach*, 2<sup>nd</sup> ed., Chicago/La Salle: Open Court.

Jeffrey, R.C. [1983] 'Bayesianism with a Human Face', in J. Earman (ed.), *Testing Scientific Theories*, Minnesota Studies in the Philosophy of Science, vol. 10, Minnesota: University of Minnesota Press, pp. 133-56.

Niiniluoto, I. [1983] 'Novel Facts and Bayesianism', *The British Journal for the Philosophy of Science*, **34**, pp. 375-79.

---

<sup>1</sup> It is crucial to note that what is compared here are *not* the degrees of confirmation obtained by two distinct pieces of evidences  $E_1$  and  $E_2$  but the degrees of confirmation obtained by one and the same piece of evidence  $E$  at two successive points in time (which may also be viewed as two possible worlds, in which case the use of Jeffrey conditionalisation is not justified).

As mentioned above, these different degrees of belief in  $E$  at two successive points in time lead to different degrees of belief in  $H$  (unless  $E$  is irrelevant for  $H$ ).  $Pr_1(H | B_1)$  and  $Pr_2(H | B_2)$  are related as dictated by Jeffrey conditionalisation. This means, in particular, that they are not assumed to be the same (unless  $E$  is irrelevant for  $H$ ).

<sup>2</sup> If  $t_1$  and  $t_2$  are interpreted as two possible worlds (and not as two successive points in time), Jeffrey conditionalisation cannot be used to obtain  $Pr_2$  from  $Pr_1$ . *For this reason* one has to assume that  $Pr_1(H | B)$  and  $Pr_2(H | B)$  equal each other, otherwise one cannot compare the corresponding two degrees of confirmation.

This is *not* assumed in the example of section 3, for there  $t_1$  and  $t_2$  are two successive points of time, and  $Pr_2(H | B)$  is as dictated by Jeffrey conditionalisation.

<sup>3</sup> The references to Christensen ([1999]) and Carnap ([1962]) are taken from Fitelson ([2001]). It is important to note that the common knock-down feature (cf. section 12)—namely dependence on the agent's first degree of belief function  $Pr_0$ —is also true of  $sPr$ .

<sup>4</sup> Cf. section 10 below.

<sup>5</sup> By the Duhem-Quine thesis, confirmation is always relative to a set of auxiliaries. Instead of interpreting  $B$  as background knowledge,  $B$  may be viewed as such a set of auxiliaries. This is particularly attractive if one considers the background knowledge to be summarized by the degree-of-belief function  $Pr$  (and not as a proposition). I owe this view of the background knowledge to Christopher Hitchcock.

<sup>6</sup> This point was made by Luc Bovens in personal correspondence.

<sup>7</sup> The references to Jeffrey ([1983]) and Niiniluoto ([1983]) are taken from Earman ([1992]).

<sup>8</sup> However, one cannot have confirmation, disconfirmation, or irrelevance to *any* degree (within some interval). This depends on the generalized relevance measure under consideration. For instance, the generalized relevance measure could be such that it takes on the value 1 when there is confirmation, the value  $-1$  when there is disconfirmation, and the value 0 when there is independence.