How objective is objective Bayesianism – and how Bayesian?

Olle Häggström

The dust jacket of Jon Williamson's *In Defence of Objective Bayesianism* is dominated by an ingenious drawing by early 20th century artist William Heath Robinson that beautifully illustrates the second word of the book's title. From there, the book goes quickly downhill, never to recover.

Objective Bayesianism, in Williamson's view, is an epistemology which prescribes that the degrees to which we believe various propositions should be

(a) probabilities,

(b) calibrated by evidence, and

(c) otherwise as equally distributed as possible among basic outcomes.

The task Williamson sets himself is, as the title suggests, to defend the idea that this is the right epistemology to guide how we acquire and accumulate knowledge, especially in science. This makes the book primarily a contribution to the philosophy of science rather than to mathematics, even though mathematical formalism – especially propositional and predicate logic, entropy calculations and probability – pervades it. The author masters such formalism fairly well, apart from the occasional lapse (such as when, on p 34, he implies that a dense subset of the unit interval must be uncountable).

Among the proposed requirements (a), (b) and (c) on the extent to which we should believe various propositions, (b) strikes me as the least troublesome (it would probably take a theologist to dissent from the idea that evidence should constrain and guide our beliefs), while (a) seems more open to controversy but not obviously wrong. I have more trouble with the final requirement (c) about equivocation between different outcomes.

Given that we accept the premise (a) about expressing degrees of belief in terms of probabilities, surely an unbiased thinker should follow (c) in spreading his belief uniformly over the possible outcomes, unless constrained otherwise by evidence? This may seem compelling, until we examine some examples. Williamson is aware of the mathematical obstacles to defining uniform distribution on various infinite sets, but seems unaware of how poorly assumptions of uniform distribution may perform even in finite situations.

Consider the following image analysis situation. Suppose we have a very fine-grained image with $10^6 \times 10^6$ pixels, each of which can take value black or white. The set of possible images then has $2^{10^{12}}$ elements. Suppose that we assign the same probability $1/2^{10^{12}}$ to each element. This is tantamount to assuming that each pixel, independently of all others, is black or white with probability 1/2 each. Standard probability estimates show that with overwhelming probability, the image will, as far as the naked eye can tell, be uniformly grey. In fact, the conviction of uniform greyness is so strong that even if, say, we split the image in four equally sized quadrants and condition on the event that the first three quadrants are pure black, we are still overwhelmingly convinced that the fourth quadrant will turn out grey. In practice, this can hardly be called unbiased or objective.

Intelligent design proponent William Dembski (2002) makes a very similar mistake in his attempt to establish the unfeasibility of Darwinian evolution by appealing to the so-called no free lunch theorems. In doing so, he implicitly assumes that the fitness landscape (a function which describes how fit for reproduction an organism with a given genome is) is randomly chosen from a large but finite set of possible such landscapes with a similar product structure as in the image example. Just as the uniform prior in the image example assigns probability very close to 1 to the event that the image is just grey, the uniform prior in the biology example assigns probability very close to 1 to the (biologically completely unrealistic) event that the fitness landscape is entirely unstructured. See Häggström (2007) for a more detailed discussion.

These examples show that the term "objective" for the habit of prefering uniform distributions whenever possible is about as suitable as the term "objectivist" for someone who favors the night watchman state and who has read and memorized *Atlas Shrugged*.

At this point, a defender of uniform distributions might suggest that the reason why requirement (c) can lead so badly wrong in these examples is the extremely large state spaces on which uniform distribution is applied. So let's look at an example with a smaller state space, with just 2 elements. In his first chapter, Williamson describes a situation where a physician needs to judge the probability that a given patient has a given disease S. All the physician knows is that there is scientific evidence that the probability that a patient with the given symptoms actually has disease S is somewhere in the interval $[0.1, 0.4]$. Williamson's suggestion is that the physician should settle for $P(\text{ill}) = 0.4$, because this is as close as he can get to uniform distribution $(0.5, 0.5)$ on the space $\{\text{ill}, \text{healthy}\}$ under the constraint given by the scientific evidence.

I must admit first thinking that the author was joking in suggesting such an inference, but no, further reading reveals that he is dead serious about it. Rather than giving the whole list of objections that come to my mind, let me restrict to one of them: what Williamson himself calls *language dependence*. Let us suppose that we refine the crude language which only admits the two possible states "ill" and "healthy" to account for the fact that a healthy person can be either susceptible or immune, so that the state space becomes $\{\text{ill}, \text{susceptible}, \text{immune}\}$, and Williamson's favored estimate goes down from $P(\text{ill}) = 0.4$ to $P(\text{ill}) = 1/3$. By further linguistic refinement (such as distinguishing between "moderately ill", "somewhat more ill", "very ill" and "terminally ill"), we can make $P(\text{ill})$ land anywhere we wish in $[0.1, 0.4]$. How's that for objectivity?

Williamson is aware of the language dependence problem and devotes Section 9.2 of his book to it. His answer is that one's language has evolved for usefulness in describing the world, and may therefore itself constitute evidence for what the world is like. "For example, having dozens of words for snow in one's language says something about the environment in which one lives; if one is going to equivocate about the weather tomorrow, it is better to equivocate between the basic states definable in one's own language than in some arbitrary other language" (Williamson, p 156–157). This argument is feeble, akin to noting that all sorts of dreams and prejudices we may have are affected by what the world is like, and suggesting that we can therefore happily and unproblematically plug them into the inference machinery.

So much for requirement (c) about equivocation; let me move on. Concerning requirement (a) that our degrees of beliefs should be probabilities, let me just mention that Williamson attaches much significance to so-called Dutch book arguments. These go as follows. For a proposition θ, define my belief $p(\theta)$ as the number p with the property that I am willing to enter a bet where I receive $\$a(1-p)$ if θ but pay $\$ap$ if $\neg\theta$ – regardless of whether a is positive or negative. Leaving aside the issues of existence and uniqueness of such a p, it turns out that I am invulnerable to the possibility of a Dutch book – defined as a collection of bets whose total effect is that I lose money no matter what – if and only if my beliefs satisfy the axioms of probability.

Let me finally discuss requirement (b) that beliefs should be calibrated by evidence. This, as mentinoned above, is in itself pretty much uncontroversial; the real issue is *how* this calibration should go about. Here, when reading the book, I was in for a big surprise. Having spent the last couple of decades in the statistics community, I am used to considering the essence of Bayesianism to be what Williamson calls *Bayesian conditionalization*: given my prior distribution (collection of beliefs), my reaction to evidence is to form my posterior distribution by conditioning the prior distribution on the evidence.

Not so in Williamson's "objective Bayesianism"! His favored procedure for obtaining the posterior distribution is instead to find the maximum entropy distribution among all those that are consistent with the evidence.

This is especially surprising given the significance that Williamson attaches to Dutch book arguments, because it is known that if the way I update my beliefs in the light of evidence deviates from what is consistent with Bayesian conditionalization, then I am susceptible a Dutch book in which some of the bets are made before the evidence is revealed, and some after (Teller, 1973). Even more surprisingly, it turns out that Williamson knows this. How, then, does he handle this blatant inconsistency in his arguments?

At this point he opts for an attempt to cast doubt on the use of sequential Dutch book arguments. On p 85 he claims that

> in certain situations one can Dutch book *anyone who changes their degrees of belief at all*, regardless of whether or not they change them by conditionalization. Thus, avoidance of Dutch book is a lousy criterion for deciding on an update rule.

Here emphasis is from the original, but I would have preferred if Willimanson, for clarity, had instead chosen to emphasize the words *"in certain situations"*. The force of his argument obviously hinges on what these situations are. The answer: "Suppose it is generally known that you will be presented with evidence that does not count against θ, so that your degree of belief in θ will not decrease" (Williamson, p 85). Here it must be assumed that by "generally known" he means "generally known by everyone but the agent", because as a Bayesian conditionalizer I would never find myself in a situation where I know beforehand in which direction my update will go, because then I would already have adjusted my belief in that direction. So what he's actually referring to is a situation where the Dutch bookmaker has access to evidence that I lack. A typical scenario would be the following. I have certain beliefs about how the football game Arsenal vs Real Madrid will end, and set my probabilities accordingly. Now, unbeknownst to me (who was confused about the game's starting time), the first half of the game has already been played, and Arsenal is down 0-3. The Dutch bookmaker approaches me for a bet, then reveals what happened in the first half, and offers a second bet. Well, of course he can screw me over in such a situation! But if we allow the Dutch bookmaker to peek at evidence that is currently unavailable to me, then we might just as well let him see the whole match in advance, in which case he could easily empty my wallet without even the need for a sequential betting procedure.

Hence, what Williamson's intended reductio shows is not that sequential Dutch book arguments should be avoided, but rather that we must insist on

Dutch bookmakers not having access to evidence that the agent lacks. If we do so, it follows from a straightforward martingale argument that an agent who sticks to Bayesian conditionalization is immune to sequential Dutch books with a bounded number of stages.

Dutch books aside, there is practically no end to the silliness of the author's further arguments for why his maximum entropy method is superior to Bayesian conditionalization. On p 80, he offers the following example.

> Suppose A is 'Peterson is a Swede', B is 'Peterson is a Norwegian', C is 'Peterson is a Scandinavian', and ε is '80% of all Scandinavians are Swedes'. Initially, the agent sets $P_{\mathcal{E}}(A) = 0.2$, $P_{\mathcal{E}}(B) = 0.8$, $P_{\mathcal{E}}(C) = 1$, $P_{\mathcal{E}}(\varepsilon) = 0,2$ and $P_{\mathcal{E}}(A \wedge \varepsilon) = P_{\mathcal{E}}(B \wedge \varepsilon) = 0.1$. All these degrees of belief satisfy the norms of subjectivism. Updating by [maximum entropy] on learning ε, the agent believes that Peterson is a Swede to degree 0.8, which seems quite right. On the other hand, updating by conditionalization on ε leads to a degree of belief of 0.5 that Peterson is a Swede, which is quite wrong.

Here Williamson obviously thinks the evidence ε constrains the probability of A to be precisely 0.8. This is plain false – unless we redefine C to say something like "Peterson was sent to us via some mechanism that picks a Scandinavian at random according to uniform distribution, and we have absolutely no other information about how he speaks, how he dresses, or anything else that may give a clue regarding his nationality". But this is not how the problem was posed.

Suppose however for the sake of the argument that ε does have the consequence that Williamson claims. Then in fact the choice of prior is incoherent, because $P_{\mathcal{E}}(A \wedge \varepsilon) = P_{\mathcal{E}}(B \wedge \varepsilon) = \frac{1}{2} P_{\mathcal{E}}(\varepsilon)$ *means* that given ε, the odds for Peterson being Swedish or Norwegian are fifty-fifty. Hence, this argument of Williamson against Bayesian conditionalization carries about as much force as if I would make the following argument against his objective Bayesianism: "Suppose that, in the course of working out his maximum entropy updating, Williamson assumes that $x < 3$ and that $x = 5$. This obviously leads to a contradiction, so there must be something fishy about objective Bayesianism."

I could go on and on about the weaknesses of Williamson's case for his pet epistemology, but this review has already grown too long, so I'll just finish by pointing to one more crucial issue. Namely, exactly *how* does evidence lead to constraints on what is reasonable to believe – constraints that serve as boundary conditions in the entropy maximization procedure that follows next. Williamson tends to treat this step as a black box, which seems to me very much like begging the issue. For instance, on p 83 he discusses what to expect of the 101th raven if we've already seen 100 black ravens – will it be black or non-black? Unconstrained entropy maximization yields the distribution $(0.5, 0.5)$ on {black, non-black}, but Williamson rejects this, claiming that the evidence constrains P(black) to be close to 1. And then this: "Exactly how this last constraint is to be made precise is a question of statistical inference – the details need not worry us here" (Williamson, p 83). An author who wishes to promote some particular philosophy of science but has no more than this to say about the central problem of induction has a long way to go. In his final chapter, Williamson does admit that "there is plenty on the agenda for those wishing to contribute to the objective Bayesian research programme" (p 163). To this, I would add that they face an uphill struggle.

References

Dembski, W. (2002) *No Free Lunch: Why Specified Complexity Cannot Be Purchased without Intelligence*, Roman & Littlefield, Lanham, MA.

Häggström, O. (2007) Intelligent design and the NFL theorems, *Biology and Philosophy* **22**, 217–230.

Teller, P. (1973) Conditionalization and observation, *Synthese* **26**, 218–258.

Williamson, J. (2010) *In Defense of Objective Bayesianism*, Oxford University Press.