Motivating Objective Bayesianism: From Empirical Constraints to Objective Probabilities

Jon Williamson

In William L. Harper and Gregory R. Wheeler (eds)
Probability and Inference: Essays in Honour of Henry E. Kyburg Jr
London: College Publications, 2006
Draft of August 8, 2006

Abstract

Kyburg goes half-way towards objective Bayesianism. He accepts that frequencies constrain rational belief to an interval but stops short of isolating an optimal degree of belief within this interval. I examine the case for going the whole hog.

Contents

§1
Partial Beliefs

Bayesians argue that an agent's degrees of belief ought to satisfy the axioms of the probability calculus. Thus for example if A is the outcome that it will snow in Stroud today, and $p(A)$ is the agent's degree of belief in A, then $p(A)+p(\bar{A}) = 1$, where $\bar{A}$ is the complement of A, i.e. the outcome that it will not snow in Stroud today.

But Bayesians differ as to whether degrees of belief should satisfy any further constraints.[1] Suppose our agents know only that the physical probability (frequency or propensity) of it snowing in Stroud on a day like today is between 0.2 and 0.3. Then the three main views can be formulated thus:

Subjective Bayesianism maintains that an agent can set her degree of belief to any value between 0 and 1. Thus one agent can choose degree of belief $p(A) = 0$, another can choose $q(A) = 0.25$ and a third can choose $r(A) = 1$—all are equally rational.[2]

Empirically-Based Subjective Bayesianism maintains that an agent's degrees of belief ought to be constrained by empirical knowledge such as knowledge of frequencies. In our example, agents should set their degrees of belief between 0.2 and 0.3, but degrees of belief $p(A) = 0.2, q(A) = 0.25, r(A) = 0.3$ are equally rational.[3]

Objective Bayesianism maintains not only that an agent's degrees of belief ought to be constrained by empirical knowledge, but also that degrees of belief should be as middling as possible—as far away as possible from the extremes of 0 and 1. In our example there are two outcomes A and $\bar{A}$, the middling assignment of belief gives $p(A) = p(\bar{A}) = 1/2$, and the value in the interval $[0.2, 0.3]$ that is closest to the middling value is 0.3. Thus an agent should assign $p(A) = 0.3$, and agents that assign other degrees of belief are irrational. The agent's degrees of belief are objectively determined by her background knowledge and there is no room for subjective choice.[4]

There is also an important non-Bayesian position that is related to empirically-based subjective Bayesianism. Under this view (advocated for instance by Henry Kyburg)[5] empirical knowledge should constrain an agent's partial beliefs, but these partial beliefs are not in general real numbers—they are intervals instead. Thus in our example an agent should adopt the interval $[0.2, 0.3]$ as her partial

[1] I am concerned here with constraints on prior degrees of belief. Diachronic constraints, e.g. Bayesian conditionalisation, will be discussed in §4.

[2] Bruno de Finetti was an influential subjective Bayesian—see de Finetti (1937).

[3] This view was adopted by Frank Ramsey: 'it will in general be best for his degree of belief that a yellow toadstool is unwholesome to be equal to the proportion of yellow toadstools which are in fact unwholesome.' (Ramsey, 1926, p. 50). Colin Howson is a recent proponent of this type of position—see Howson and Urbach (1989, §13.e) for instance. Salmon (1990) also advocates a version of this view.

[4] Edwin Jaynes is an influential objective Bayesian—see Jaynes (2003). The essential feature of objective Bayesianism is not that degrees of belief be uniquely determined by background knowledge—this is too much to ask in infinite domains (Williamson, 2006, §19)—but that constraints on degrees of belief go beyond purely empirical constraints.

[5] (Kyburg Jr and Teng, 2001)

belief in A. Because of their common ground, I shall classify empirically-based subjective probability and Kyburg's probability-interval position as *empirical-constraint theories*.

Subjective Bayesianism is attractive because it is easy to justify: one only needs an argument that degrees of belief ought to be probabilities, and the Dutch book argument does this job quite well.[6] However, many applications of probability demand objectivity; the other positions, though philosophically more taxing, are more appealing in that respect.[7] Here I shall take it for granted that empirical constraints on partial belief are desirable.

The aim of this chapter is to examine the motivation for moving beyond empirical constraints to the stronger constraints advocated by objective Bayesianism. Is it enough to restrict partial beliefs to probability intervals, as recommended by Kyburg and empirically-based subjective Bayesians? Or should one strive for the extra objectivity afforded by the most middling degrees of belief within such intervals?[8]

The plan is to introduce the objective Bayesian framework and its standard justifications in §2 and §3. While these standard justifications are lacking for our purposes, one can appeal to considerations of objectivity (§4), efficiency (§6) or caution (§8) to try to decide between objective Bayesianism and empirical constraint theories. As we shall see in §8, objective Bayesianism clearly surpasses empirically-based subjective Bayesianism in terms of caution. On the other hand, Kyburg's approach appears to be superior to objective Bayesianism in this respect (§11). However I shall argue that, taking several considerations into account, a case can be made for objective Bayesianism.

§2
The Maximum Entropy Principle

Objective Bayesianism appeals to the *maximum entropy principle* to determine the degrees of belief that an agent ought to adopt on the basis of background knowledge β.[9] In this section I shall introduce the principle and, in §3, its key justifications.

Given a finite outcome space $\Omega = \{A_1, \ldots, A_n\}$, i.e. a set of mutually exclusive and exhaustive outcomes, the most middling probability function, the *central* function, assigns each outcome the same probability, $c(A_i) = 1/n$ for $i = 1, \ldots, n$. Let $\mathbb{P}_\beta$ be the set of all probability functions that satisfy constraints imposed by the agent's background knowledge β. Objective Bayesians

[6](Ramsey, 1926; de Finetti, 1937)

[7]De Finetti argued that subjective Bayesians can account for objectivity since under certain conditions (e.g. degrees of belief must satisfy an *exchangeability* assumption and must be updated by *Bayesian conditionalisation*) different agents' degrees of belief will converge to a single objective value in the limit. However these conditions are controversial and are by no means guaranteed to hold. Moreover, this line of argument does nothing to allay worries about a lack of objectivity in the short run: those whose degrees of belief fail to reflect their empirical knowledge may simply do worse in the short run.

[8]The focus of this chapter is purely the motivation behind objective Bayesianism. There are number of other interesting challenges facing objective Bayesianism—e.g. does it suffer from representation dependence? does it apply to infinite as well as finite domains? Interesting as they are, these questions are beyond the scope of this chapter. See Williamson (2006) for an overview of these other challenges.

[9](Jaynes, 1957)

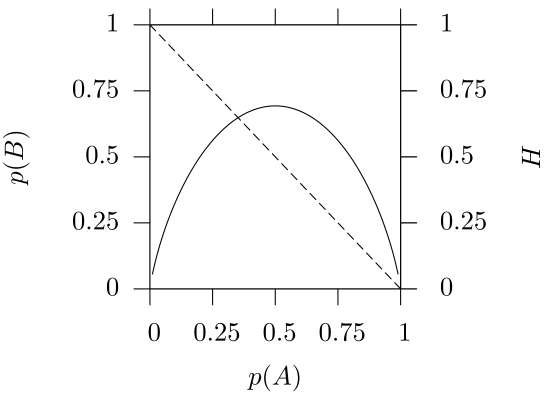

Figure 1: Probability functions (dotted line) and their entropy H (solid curve) in two dimensions.

suggest that the agent should adopt as a representation of her degrees of belief a probability function in $\mathbb{P}_\beta$ that is closest to the central function. Now distance between probability functions is normally measured by *cross-entropy*,

$$d(p,q) = \sum_{i=1}^{n} p(A_i) \log \frac{p(A_i)}{q(A_i)}.$$

Thus the function in $\mathbb{P}_\beta$ that is closest to c is the function that minimises

$$d(p,c) = \sum_{i=1}^{n} p(A_i) \log p(A_i) + \sum_{i=1}^{n} p(A_i) \log n = \sum_{i=1}^{n} p(A_i) \log p(A_i) + \log n.$$

This distance is minimised by the function p that has maximum *entropy*

$$H(p) = -\sum_{i=1}^{n} p(A_i) \log p(A_i).$$

Thus

MAXIMUM ENTROPY PRINCIPLE Suppose $\mathbb{P}_\beta$ is the set of probability functions that satisfies constraints imposed by the agent's background knowledge β. The agent should select the member of $\mathbb{P}_\beta$ that maximises entropy as her belief function.

If there are two outcomes $\Omega = \{A, B\}$ then $p(B) = 1 - p(A)$, as depicted by the dotted line in Fig. 1. Entropy H is depicted by the solid line. Clearly the closer to the centre of the dotted line, the higher the entropy.

Fig. 2 is the corresponding visualisation of the three outcome case, $\Omega = \{A, B, C\}$. The probability functions are depicted by the plane and their entropy by the curved surface.

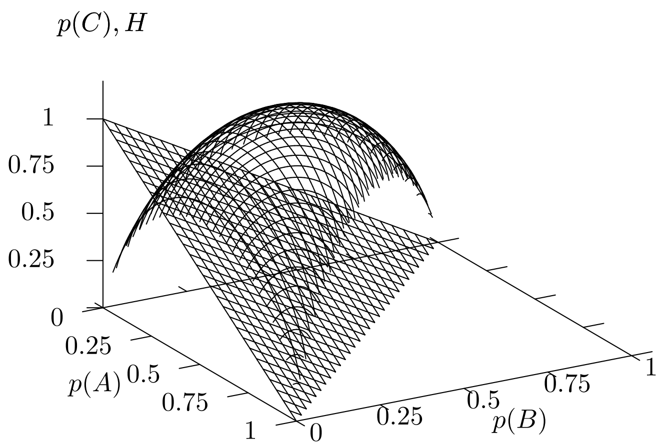

Figure 2: Probability functions (plane) and their entropy H (curved surface) in three dimensions.

§3
Standard Justifications

There are two major arguments in favour of the maximum entropy principle, but neither of these conclusively decide between objective Bayesianism and the empirically-based approach, as we shall now see.

The original justification of the maximum entropy principle in Jaynes (1957) is perhaps best known. This justification appeals to Claude Shannon's use of entropy as a measure of the uncertainty embodied in a probability function.[10] Jaynes maintains that an agent's belief function should be informed by background knowledge but should otherwise be maximally uncertain or noncommittal—thus it should have maximum entropy according to Shannon's measure:

> in making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased assignment we can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have.
>
> ...
>
> The maximum entropy distribution may be asserted for the positive reason that it is uniquely determined as the one that is maximally noncommittal with regard to missing information.[11]

[10](Shannon, 1948, §6)
[11](Jaynes, 1957, p. 623)

The gap in the argument is clear. Jaynes *assumes* that maximally non-committal, unbiased degrees of belief are most desirable and argues that they should then be found by maximising entropy. Even if we grant Jaynes that entropy measures lack of commitment, we still need some reason to accept his premiss. Why should maximally non-committal degrees of belief be any better than, say, maximally committal degrees of belief? Why is bias bad?

The second key line of argument in favour of the maximum entropy principle takes the form of an axiomatic derivation. There are various versions, but the derivation of Jeff Paris and Alena Vencovská is perhaps most compelling.[12] Their argument takes the following form:

- ◦ An *inference process* is a function which maps a domain and background knowledge involving that domain to a probability function on the domain that satisfies the background knowledge.
- ◦ If the selected probability function is to be construed as a representation of the degrees of belief that one ought to adopt on the basis of the background knowledge, then the inference process ought to satisfy some common-sense desiderata. For example, given two logically equivalent knowledge bases the inference process should select the same probability function.[13]
- ◦ The only inference process that satisfies these desiderata is the maximum entropy principle.
- ▷ The only commonsensical inference process is the maximum entropy principle.

Again, even if we grant that the only inference process satisfying the desiderata is the maximum entropy principle, this argument does not do enough for our purposes. It assumes from the outset that we need an inference process, i.e. that we need to select a single probability function that satisfies the background knowledge (Kyburg would disagree with this) and that this probability function must be uniquely determined by domain and background knowledge (empirically-based subjectivists would disagree with this, arguing that different individuals are free to choose different belief functions on the basis of the same background knowledge).

Paris and Vencovská (2001, §3) do relax the uniqueness requirement when they consider the case in which background knowledge imposes non-linear constraints on degrees of belief. Their modified argument is unlikely to satisfy the empirical constraint theorist, however, because some of their desiderata go significantly beyond the empirical constraints imposed by background knowledge. The Renaming Principle, for instance, dictates that (background knowledge permitting) degrees of belief should be invariant under permutations of the domain; the Independence Principle holds that in certain cases degrees of belief should be independent if there is no evidence of dependence. While these desiderata may be commonsensical, they are not merely empirical, and more justification is required to convince the proponent of an empirical constraint theory.[14]

[12](Paris and Vencovská, 1990, 2001; Paris, 1994)

[13]See e.g. Paris (1994) for a full list of the desiderata.

[14]Hosni and Paris (2005) explore a line of justification of the desiderata, claiming that they are commonsensical because they force us to assign similar probabilities, and that conformity is some kind of rational norm. Again, many would take issue with these claims. Assigning

So we see that current justifications of the maximum entropy principle do not fully motivate the move from empirical constraints to objective Bayesianism—they presume that an agent's degrees of belief should be maximally non-committal or that they should be fully determined by background knowledge and domain. Further argument is needed.

§4
The Argument from Objectivity

One might try to construct an argument around the suggestion that empirically-based subjective Bayesianism is *not objective enough* for many applications of probability.[15] If applications of probability require full objectivity—i.e. a single probability function that fits available evidence—then the axiomatic justification kicks in and one can argue that the maximum entropy function is the only rational choice.

The typical Bayesian methodology involves the following process, called *Bayesian conditionalisation*. First a *prior* probability function p must be produced. Then empirical evidence E is collected. Finally predictions are drawn using the *posterior* probability function $p'(A) = p(A|E)$. Now the prior function is determined before empirical evidence is available; this is entirely a matter of subjective choice for empirically-based subjectivists. However, the ensuing conclusions and predictions may be sensitive to this initial choice of prior, rendering them subjective too. If, for example, $p(A) = 0$ then $p'(A) = p(A|E) = p(E \wedge A)/p(E) = 0$ (assuming $p(E) > 0$). On the other hand, if $q(A) = 1$ then $q'(A) = q(E|A)q(A)/(q(E|A)q(A) + q(E \wedge \bar{A})) = q(E|A)/q(E|A) = 1$.

It is plain to see that two empirically-based subjectivists can radically disagree as to the conclusions they draw from evidence. If they adopt the opposite prior beliefs they will draw the opposite conclusions. Under the empirically-based subjectivist approach, if conditionalisation is adopted then all conclusions are relative to initial opinion.

If the empirically-based subjectivist is to avoid such strong relativity, she must reject Bayesian conditionalisation as a universal rule of updating degrees of belief. The standard alternative is *cross entropy updating*.[16] Here the agent adopts prior belief function p and her posterior p' is taken to be the probability function, out of all those that are compatible with the new evidence, that is closest to p in terms of cross-entropy distance. (Note that, unlike the Bayesian conditionalisation case, the evidence does not have to be representable as a domain event E.)

Suppose $\Omega = \{A, B\}$ for instance. Two empirically-based subjectivists may set $p(A) = 0$ and $q(A) = 1$ respectively, while an objective Bayesian is forced to set $r(A) = 1/2$ in the absence of any empirical evidence. Now suppose evidence is collected that constrains the probability of A to lie in the interval $[0.6, 0.9]$.

similar probabilities may be of pragmatic advantage, but hardly seems to be requirement of rationality except in special cases (see §4 and Gillies (1991)). Nor need this consequence of the desiderata render the desiderata commonsensical in themselves—the end doesn't justify the means.

[15]Note that Salmon (1990, §4) argues for a version of empirically-based subjective Bayesianism on the grounds that subjective Bayesianism is not objective enough for science.

[16](Williams, 1980; Paris, 1994, pp. 118–126)

The objective Bayesian must adopt $r'(A) = 0.6$ while the empirically-based subjectivists now adopt posteriors $p'(A) = 0.6$ and $q'(A) = 0.9$ respectively. Thus with cross entropy updating evidence can shift degrees of belief away from 0 and 1.

While cross entropy updating may be an improvement over Bayesian conditionalisation, one might argue that the remaining relativity in empirically-based subjective Bayesianism is too still much for applications of probability if applications demand full objectivity.

But such an argument would be hard to execute. First, one would expect that the amount of tolerable subjectivity would vary from application to application—it is unlikely to be the case that *all* applications of probability demand full objectivity. While objectivity of conclusions seems desirable in a computer system for controlling nuclear retaliation, it is clearly less desirable in an automated oenophile.

Second, it is difficult to judge the need for objectivity. Scientists often emphasise the objectivity of their disciplines, but it can be difficult to say whether their claims reflect their sciences' needs for objectivity or their own. (Of course, a perceived objectivity of science is rhetorically very useful to scientists—their conclusions appear more forceful.) Moreover, even if scientific methodology does assume an inherent objectivity, such an assumption may simply be erroneous. There may be less objectivity to be found than commonly supposed.[17] These difficulties have led to widespread disagreement between sociologists of science on the one hand, many of whom view scientific conclusions as highly relative, and philosophers of science and scientists on the other hand, many of whom view science as an objective activity by-and-large.[18] Until some (objective) common ground can be found in the study of science, we are a long way from determining whether the extreme objectivity of objective Bayesianism is required in science, or whether the more limited objectivity of empirically-based subjective Bayesianism is adequate.

Finally, even in cases where objectivity is required, that objectivity need not necessarily incline one towards objective Bayesianism. In order to run the axiomatic derivation of the maximum entropy principle (§3), one must first accept that the common-sense desiderata are indeed desirable. If not, one may be led to alternative implementations of the objectivity requirement. Consider the *minimum entropy principle*: if $\Omega = \{A_1, \ldots, A_n\}$ choose as belief function the probability function, out of all those that satisfy constraints imposed by background knowledge β, that is as close as possible to the function p that sets $p(A_1) = 1, p(A_2) = \cdots = p(A_n) = 0$. This is objective in the sense that once the domain and its ordering has been fixed, then so too is the prior probability function. Moreover this principle is not as daft as it may first look—there may be semantic reasons for adopting such an unbalanced prior. Consider a criminal trial setting: $\Omega = \{A_1, A_2\}$ where A_1 represents innocence and A_2 represents guilt, and where there is no background knowledge—then the minimum entropy principle represents a prior assumption of innocence; this is in fact the recommended prior.[19]

Note too that while the issue of objectivity might help decide between

[17] Press and Tanur (2001) present a case for subjectivity in science.

[18] (Gottfried and Wilson, 1997)

[19] I am very grateful to Amit Pundik for suggesting this example. See §8 for more on the criminal trial setting.

empirically-based subjective Bayesianism and objective Bayesianism, it can not decide between Kyburg's probability interval approach and objective Bayesianism, since neither of these approaches permit subjective choice of partial beliefs: background knowledge fully determines the partial beliefs that an agent ought to adopt.

In sum, while it may be tempting to argue for objective Bayesianism on the grounds that applications of probability demand objective conclusions, there are several hurdles to be overcome before a credible case can be developed.

§5
Rationality and Evidence

One thing should be clear by now: it can not be empirical warrant that motivates the selection of a particular belief function from all those compatible with background knowledge, since all such belief functions are equally warranted by available empirical evidence. If objective Bayesianism is to be preferred over empirical-constraint approaches, it must be for non-evidential reasons. (Equally, one can't cite evidence as a reason for abandoning objective Bayesianism in favour of an empirical-constraint approach.)

Thus the problem of deciding between objective Bayesianism and empirical-constraint approaches hinges on the question of whether evidence exhausts rationality. Objective Bayesianism supposes that there is more to rationality than evidence: a rational agent's degrees of belief should not only respect empirical evidence, they should also be as middling as possible. For empirical-constraint approaches, on the other hand, empirical warrant is sufficient for rationality. (This puts the empirical constraint theorist at a disadvantage, because from the empirical perspective there simply are no considerations that can be put forward to support an empirical constraint theory over objective Bayesianism since both are empirically optimal; in contrast, the objective Bayesian can proffer non-empirical reasons to prefer objective Bayesianism over empirical constraint theories.)

If rationality goes beyond evidence, what else does it involve? We have already discussed one form of non-evidential reason that might decide between the two types of approach—a demand for *objectivity*. But there are other more overtly pragmatic reasons that can be invoked to the same end. In §8 we shall see whether *caution* can be used to motivate objective Bayesianism. But first we turn to *efficiency*.

§6
The Argument from Efficiency

One might be tempted to appeal to efficiency considerations to distinguish between objective Bayesianism and empirically-based subjective Bayesianism. If objective Bayesian methods are more efficient than empirically-based subjective Bayesian methods then that would provide a reason to prefer the former over the latter.

One possible line of argument proceeds as follows. If probabilities are in any way subjective then their measurement requires finding out which particular degrees of belief have been chosen by some agent. This can only be done by elicitation: asking the agent what her degrees of belief are, or perhaps inducing them from her behaviour (e.g. her betting behaviour). But, as developers of expert systems in AI have found, elicitation and the associated consistency-checking are prohibitively time-consuming tasks (the inability of elicitation to keep pace with the demand for expert systems is known as *Feigenbaum's bottleneck*). If subjective Bayesianism or empirically-based subjective Bayesianism is to be routinely applied throughout the sciences it is likely that a similar bottleneck will be reached. For example, determining the most probable statistical model given evidence would first require eliciting model assumptions (are the agent's degrees of belief normally distributed, for instance? are certain degrees of belief probabilistically independent?) and also the agent's prior degree of belief in each model—a daunting task. On the other hand, if objective Bayesianism is adopted, degrees of belief are objectively determined by background knowledge and elicitation is not required—degrees of belief are calculated by maximising entropy. Therefore objective Bayesianism is to be preferred for reasons of efficiency.

Of course this argument fails if the objective Bayesian method is itself more computationally intractable than elicitation. Indeed Judea Pearl rejected the maximum entropy principle on the grounds that computational techniques for maximising entropy were usually intractable (Pearl, 1988, p. 463). However, while that was indeed the case in 1988, it is not the case now. Pearl's own favoured computational tool, *Bayesian nets*, can be employed to vastly reduce the complexity of entropy maximisation, rendering the process tractable in a wide variety of natural settings—see Williamson (2005a, §§5.5–5.8) and Williamson (2005b). Thus despite Pearl's doubts, efficiency considerations do lend support to objective Bayesianism after all.

But efficiency considerations on their own fail to distinguish between objective Bayesianism and other procedures for selecting a unique probability function. The maximum entropy principle is no more efficient than the minimum entropy principle. Worse, consider the *blind minimum entropy principle*, where one ignores background knowledge and minimises entropy straight off: if $\Omega = \{A_1, \ldots, A_n\}$ choose as belief function the probability function p such that $p(A_1) = 1$ and $p(A_2) = \cdots = p(A_n) = 0$. This modified principle avoids elicitation and is far easier to implement than the maximum entropy principle. Should one really blindly minimise entropy?

If not, efficiency can't be the whole story. At best one can say something like this: efficiency considerations tell against elicitation and motivate some procedure for mechanically determining an agent's degrees of belief; other desiderata need to be invoked to determine exactly which procedure; the axiomatic derivation can then be used to show that the maximum entropy principle is the only viable procedure.

While this is an improvement over the argument from objectivity, it is still rather inconclusive as it stands. Further work needs to be done to explain why efficiency isn't the whole story, i.e. to explain why the other desiderata override efficiency considerations. The other desiderata thus play an important role in this new argument, and stand in need of some form of justification themselves. Given that empirically-based subjectivists would find several of these desiderata dubious, their justification may turn out to be more of an ordeal than elicitation.

§7
Derivation versus Interpretation

We have seen that the arguments from objectivity and efficiency require an appeal to the axiomatic derivation of the maximum entropy principle, and hence to the desiderata presupposed by that derivation. Some of these desiderata are hard to justify—empirically-based subjectivists would simply deny their desirability. Hence the arguments from objectivity and efficiency flounder.

Such difficulties are bound to beset any derivation of the maximum entropy principle. If the principle *MEP* is a logical consequence of assumptions A, i.e. $\models A \rightarrow MEP$, then the assumptions A must be at least as strong as the maximum entropy principle and are unlikely to be trivially true. Empirically based subjectivists, who reject the maximum entropy principle, can just use the contrapositive of the derivation, $\models \overline{MEP} \rightarrow \bar{A}$ as an argument for the falsity of the assumptions. Thus what is an argument in favour of the maximum entropy principle for the objective Bayesian is nothing of the sort for the empirically-based subjectivist. Consequently a derivation of the maximum entropy principle is unlikely to be of help to us in our quest to motivate objective Bayesianism.

But Jaynes' original justification, which proceeds by *interpreting* lack of commitment as entropy rather than *deriving* the maximum entropy principle, does not suffer from these difficulties. It requires an assumption, namely that one ought to be maximally non-committal, but that assumption is relatively weak—much of the work is being done by the act of interpretation.

Consider an analogy: the principle, adopted by all Bayesians, that degrees of belief should satisfy the axioms of probability. If we try to *derive* the axioms of probability from some assumptions then those assumptions will have to be at least as logically strong as the axioms, and hence at least as controversial.[20] Instead it is more usual to *interpret* degrees of belief as betting quotients[21] and then to show that the axioms of probability must hold on pain of incoherence.[22] This argument is not simply a drawing-out of logical consequences; the act of interpretation is doing significant work. Yet this interpretation is rather natural and can itself be justified. As Ramsey notes,

> this will not seem unreasonable when it is seen that all our lives we are in sense betting. Whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay at home.[23]

Similarly, the interpretation of the uncertainty of a probability function as its entropy is, if not obvious, fully justifiable. Indeed Shannon provided two justifications, a derivation from desiderata, and—more importantly for Shannon—a pragmatic justification:

[20] Cox's derivation of the axioms of probability has been the topic of substantial controversy. Paris (1994, pp. 24–32) shows that this type of derivation requires strong assumptions; see also Halpern (1999a,b).

[21] A *betting quotient* for event E is a number q that the agent selects on the presumption that she will make a bet, paying qS with return S if E occurs, where the stake S is to be chosen after she selects q and may be positive or negative.

[22] (Ramsey, 1926; de Finetti, 1937)

[23] (Ramsey, 1926, p. 183)

> This theorem [the derivation from desiderata], and the assumptions required for its proof, are in no way necessary for the present theory. It is given chiefly to lend a certain plausibility to some of our later definitions [including the interpretation of uncertainty as entropy]. The real justification of these definitions, however, will reside in their implications.[24]

Indeed Shannon's definitions have been very fruitful in communication and information theory and are now well entrenched in several branches of science.

The interpretation of uncertainty as entropy plays a substantial role in Jaynes' argument for the maximum entropy principle. The key assumption (that a lack of commitment is desirable) is relatively meagre, in the sense that it does not on its own presuppose entropy maximisation. We shall see now that this assumption can be justified by an appeal to caution.

§8
The Argument from Caution

As Ramsey notes in the above quote, our degrees of belief guide our actions and our actions are tantamount to bets. To embark on a course of action (such as going to the station) a degree of belief (in the train running, in this case) must be sufficiently high. Now, every course of action has its associated risks. Going to the station only to find that the train is not running wastes time and effort, and one may miss an important appointment. Such a course of action is not to be embarked upon lightly, and prudence is required. The trigger for action will vary according to risk—if a lot hangs on it, one may only go to the station if one has degree of belief at least 0.95 in the train running, but if the consequences are less dire, a lower trigger-level, say 0.85, may be appropriate. Suppose one knows only that the local train operating company has passed the minimum threshold of eighty percent of trains running. According to empirically-based subjective Bayesianism, one's degree of belief in the train running can be chosen from anywhere in the interval $[0.8, 1]$. According to objective Bayesianism the least extreme value in this interval must be chosen: i.e. 0.8. So the empirically-based subjectivist may decide to go to the station while the objectivist decides not to. Thus the objective Bayesian decision is more cautious and is to be preferred since there is no empirical evidence to support a less cautious decision.

In sum, extreme degrees of belief trigger actions and open one up to their associated risks. In the train example the objective Bayesian strategy of adopting the least extreme degree of belief seems to be the most prudent. Can one abstract from this case to argue that objective Bayesianism is to be preferred in general? There are some potential difficulties with such a move to generality, as we shall now see.

The first potential problem stems from the fact that it is not only extreme degrees of belief that trigger actions—middling degrees of belief can also trigger actions. Consider a case where a patient has one of two possible diseases A and B. A high degree of belief in A will trigger a course of medication to treat A. Similarly, a high degree of belief in B will trigger treatment of B.

[24](Shannon, 1948, §6)

However, a more middling degree of belief—a degree of belief that triggers neither treatment of A nor treatment of B—will trigger another action, namely the gathering of more empirical evidence in order reach a conclusive diagnosis. Collecting further symptoms from the patient also has its associated costs and risks: it requires time, effort and money to perform more tests, and some tests might harm the patient. Now objective Bayesianism advocates setting middling degrees of belief, thereby exposing the diagnoser to the risks associated with collecting more symptoms. It seems that objective Bayesianism is not such a risk-averse position after all.

But this apparent problem does not in fact scupper the prospect of a general argument in favour of objective Bayesianism. While there are indeed cases in which middling degrees of belief trigger actions, these actions are always *less risky* than those triggered by extreme degrees of belief. Suppose in the above example that A and B are two types of meningitis, requiring very different treatment. The risks associated with either outcome are so high that the risks associated with collecting more symptoms pale into insignificance in comparison. Suppose on the other hand that A and B are just two strains of cold; A responds best to a nightly dose of rum toddy while B requires a whisky toddy taken three times a day; in either case a full recovery will be made even if no treatment is taken. In this case the risks associated with either diagnosis are so low that in the absence of a diagnosis it simply is not worth doing the blood test that would provide conclusive evidence: a more middling degree of belief does not trigger any action. The point is that collecting further evidence will only be an option if the resulting knowledge is worth it, i.e. if the costs and risks associated with the primary outcomes A and B outweigh those associated with collecting new evidence. Hence it will still be most cautious to have more middling degrees of belief.

It might be thought that there is a more serious variant of this problem. For a politician the risks associated with appearing non-committal outweigh those of committing to an unjustified course of action or making a promise that can't be kept. People like their politicians bold and it would seem that a non-committal objective Bayesian politician would not get elected. But this objection hinges on a mistaken conflation of appearance and reality. The fact that a politician should not *appear* non-committal does not mean that her beliefs should not actually *be* non-committal—politicians simply need to mask their beliefs. Their beliefs need to be as cautious as anyone else's though: they need to be shrewd judges of which promises the electorate will swallow, and should not commit to one lie over another unless they have a justified belief that they can get away with it.

A second type of problem confronts any attempt to generalise the argument from caution. This concerns cases in which risks are known to be imbalanced. In the diagnosis example, take A to be meningococcal meningitis and B to be 'flu. In this case, the risks associated with failing to diagnose A when it is present are so great that it may be prudent to *assume* A and prescribe antibiotics, unless there is conclusive evidence that decides in favour of B. We have already come across another example of an imbalance of risks: the risk of judging the innocent guilty is considered to outweigh that of judging the guilty innocent, and this motivates a presumption of innocence in criminal cases. Such presumptions seem far from non-committal, yet rational.

Perhaps the best way of resolving this difficulty is to distinguish appearance and reality again. It is important in these cases to *act as if* one believed one of the alternatives—to prescribe antibiotics or to release the suspect—not to actually believe that alternative. In these cases, the imbalanced risks motivate imbalanced trigger levels rather than imbalanced degrees of belief. If degree of belief in 'flu is higher than say 0.95 prescribe paracetamol, otherwise, if degree of belief in meningococcal meningitis is at least 0.05, prescribe antibiotics. If guilt is not proved beyond reasonable doubt (degree of belief of guilt is not higher than 0.99 say, in which case degree of belief in innocence is at least 0.01) then trigger action that corresponds to innocence, i.e. release the suspect. In both these cases the trigger level for one alternative is very high while the trigger level for the other alternative is very low.

One might respond to this move by accepting this proposed resolution for the diagnosis example, but rejecting it for the legal example. One might claim that in the legal case there should not only be a high standard of reasonable doubt for guilt and a corresponding low standard of doubt for innocence, but there should also be prior degree of belief 1 in innocence, in order to make it as hard as possible for the prosecution to sway degree of belief to beyond reasonable doubt. This response seems natural enough—it seems right to make it as hard as possible to trigger guilt. However, this response should not be acceptable to any Bayesian—whether subjective, empirically-based or objective—because it does not sit well with Bayesian methods of updating. Recall that in §4 we saw that there are two standard options for updating an agent's degree of belief in new evidence, Bayesian conditionalisation and cross-entropy updating. If Bayesian conditionalisation is adopted then a prior degree of belief 0 of guilt can never be raised above 0 by evidence, and it will be impossible to convict anybody. On the other hand, if cross-entropy updating is adopted then a presumption of innocence will make no difference in the legal example. A presumption of innocence corresponds to prior degree of belief 0 in guilt, while a maximally non-committal probability function will yield a prior degree of belief of $\frac{1}{2}$. In either case degree of belief can only be raised above 0.99 if empirical evidence constrains degree of belief to lie in some subset of the interval $[0.99, 1]$ (because the cross-entropy update is the degree of belief, from all those that are compatible with evidence, that is closest to $\frac{1}{2}$). As long as the prior degree of belief is lower than the trigger level for guilt, triggering is dependent only on the evidence, not on prior degree of belief. In sum, whichever method of updating one adopts there is no good Bayesian reason for setting prior degree of belief in guilt to be 0.

There is a third, more substantial, problem that besets an attempt to generalise the argument from caution: the maximum entropy principle is not *always* the most cautious policy for setting degrees of belief. Consider a case in which there are three elementary outcomes, $\Omega = \{A, B, C\}$, a risky action is triggered if $p(\{A, B\}) \geq 7/8$ and another risky action is triggered if $p(\{B, C\}) \geq 7/8$. Suppose there is background knowledge $\beta = \{p(B) \geq 3/4\}$. Then the set $\mathbb{P}_\beta$ of probability functions that are consistent with this knowledge is represented by the shaded area in Fig. 3. (The triangular region represents the set of all probability functions—these must satisfy $p(A) + p(B) + p(C) = 1$.) Now the maximum entropy principle advocates adopting the probability function p from $\mathbb{P}_\beta$ that is closest to the central probability function c which sets $c(A) = c(B) = c(C) = 1/3$. Thus the maximum entropy function sets $p(A) = 1/8, p(B) = 3/4$ and $p(C) = 1/8$. This triggers both $\{A, B\}$ and $\{B, C\}$ since $p(\{A, B\}) = 7/8 = p(\{B, C\})$. On the other hand, the minimum entropy prin-

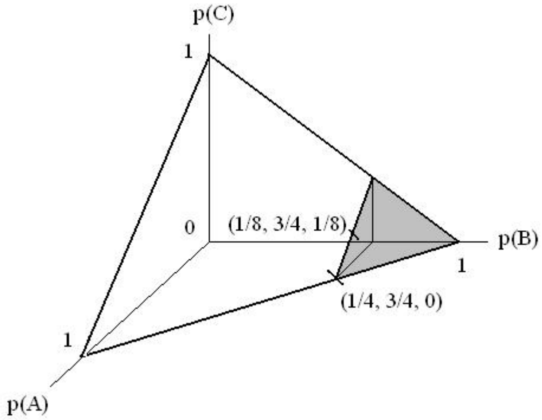

Figure 3: Maximum and minimum entropy principle probability functions.

ciple advocates adopting the probability function q from $\mathbb{P}_\beta$ that is closest to the probability function a which sets $a(A) = 1$ and $a(B) = a(C) = 0$. Thus the minimum entropy principle sets $q(A) = 1/4, q(B) = 3/4, q(C) = 0$.[25] This triggers $\{A, B\}$, since $q(\{A, B\}) = 1 \geq 7/8$, but does not trigger $\{B, C\}$, since $q(\{B, C\}) = 3/4 < 7/8$. Hence in this case the minimum entropy principle licences the most cautious course of action while the maximum entropy principle seems to throw caution to the wind.

While this example shows that the most non-committal probability function is not the most cautious *in every situation*, it may yet be the most cautious *on average*. If it is most cautious when averaging over background knowledge β and decisions (i.e. $D \subseteq \Omega$ and trigger levels $\tau_D \in [0, 1]$ such that a course of action is triggered if $p(D) \geq \tau_D$) then adopting the maximum entropy principle will be the best policy in the absence of any knowledge about β, D and τ_D.[26]

The average caution for a policy for setting belief function p_β can be measured by the proportion of β, D, τ_D that result in a course of action being triggered. Define the *trigger function* $T(\beta, D, \tau_D) = 1 \Leftrightarrow p_\beta(D) > \tau_D$, and 0 otherwise. Then the average caution of a policy is the measure of β, D, τ_D for which $T(\beta, D, \tau_D) = 1$, i.e.

$$C = \frac{1}{z} \sum_{D \subseteq \Omega} \int_\beta \int_{\tau_D} T(\beta, D, \tau_D) d\beta d\tau_D$$

where z is a normalising constant, $z = 2^{|\Omega|} \int_\beta \int_{\tau_D} 1 d\beta d\tau_D$. The smaller the value of C, the more cautious the policy is on average.

[25] N.b. the 'minimum entropy principle' is a bit of a misnomer—while the probability function q commits most to A, it does not actually have minimum entropy in this example. Here entropy is in fact minimised by the probability function r that commits most to B, defined by $r(A) = 0, r(B) = 1, r(C) = 0$.

[26] I leave it open here as to the mechanism that is used to select the trigger levels. These may be set by experts perhaps, by maximising expected utility, or by some other decision-theoretic procedure.

However it turns out that the maximum entropy principle is no more cautious than the minimum entropy principle, even if one considers average caution. By way of example consider the two-dimensional case, $\Omega = \{A, B\}$. Suppose β constrains $p(A)$ to lie in a closed interval $[x, y]$ where $0 \leq x \leq y \leq 1$, and that the trigger levels are all the same, $\tau_D = \tau \in [0, 1]$. We are interested in the proportion of values of x, y, τ that trigger a decision. i.e. the volume of the part of the cube $[0, 1]^3$ defined by these parameters that triggers some decision. We are only concerned with the half of the cube such that $y \geq x$, so $z = 4 \times 1/2 = 2$. Trivially, $\{A, B\}$ always gets triggered and $\emptyset$ only gets triggered if $\tau = 0$. Consider first the entropy maximisation policy. In this case $T(\beta, A, \tau) = 1 \Leftrightarrow x \geq \tau$ or $\tau \leq 1/2$ and $y \geq \tau$, while $T(\beta, B, \tau) = 1 \Leftrightarrow x \leq 1 - \tau$ or $\tau \geq 1/2$ and $y \leq 1 - \tau$. Then

$$C = z^{-1} \int_\beta \int_\tau [T(\beta, \{A, B\}, \tau) + T(\beta, A, \tau) + T(\beta, B, \tau) + T(\beta, \emptyset, \tau)] \, d\beta d\tau$$

$$= 2^{-1} [1/2 + 1/4 + 1/4 + 0] = 1/2.$$

Next consider the entropy minimisation policy. In this case $T(\beta, A, \tau) = 1 \Leftrightarrow y \geq \tau$, while $T(\beta, B, \tau) = 1 \Leftrightarrow y \leq 1 - \tau$. Now

$$C = 2^{-1} [1/2 + 1/3 + 1/6 + 0] = 1/2.$$

Thus we see that what entropy minimisation loses in caution by committing to A, it offsets by a lack of commitment to B. On average, then, entropy minimisation is just as cautious as entropy maximisation.

While it might appear that caution can not after all be used as an argument in favour of the maximum entropy principle, such a conclusion would be too hasty. In fact the maximum entropy principle *is* most cautious where it matters, namely in the face of *risky* decisions, as we shall now see.

It is important to be cautious when a course of action should not be taken lightly, and as we have seen, the importance attached to a decision tends to be reflected in its trigger level. Thus when deciding between meningococcal meningitis and 'flu, a high trigger level for 'flu indicates that the ensuing course of action (treatment of 'flu rather than meningitis) is risky. Similarly when deciding between innocence and guilt in a criminal case, there is a high trigger level for guilt because the consequences of a mistaken judgement of guilt are considered dire. Thus when we consider the average caution of a policy for setting degrees of belief, it makes sense to focus on the decisions where caution is important, namely those decisions D with a *high* trigger level τ_D.

So let us assume in our two-dimensional example that $\tau \geq 1 - \varepsilon$ where $0 \leq \varepsilon \leq 1/2$ is small. This extra constraint means that now $z = 4 \times \varepsilon/2 = 2\varepsilon$. As before $T(\beta, \emptyset, \tau) = 1$ has measure 0, but now $T(\beta, \{A, B\}, \tau) = 1$ has measure $\varepsilon/2$. In the entropy maximisation case both $T(\beta, A, \tau) = 1$ and $T(\beta, B, \tau) = 1$ have measure $\varepsilon^3/6$, and

$$C = (2\varepsilon)^{-1} \left[\varepsilon/2 + \varepsilon^3/6 + \varepsilon^3/6 + 0\right] = (1 + 2\varepsilon^2/3)/4.$$

On the other hand in the entropy minimisation case $T(\beta, A, \tau) = 1$ has measure $\varepsilon/2 - 1/6[1 - (1 - \varepsilon)^3]$, while $T(\beta, B, \tau) = 1$ has measure $\varepsilon^3/6$, and

$$C = (2\varepsilon)^{-1} \left[\varepsilon/2 + \varepsilon/2 - 1/6[1 - (1 - \varepsilon)^3] + \varepsilon^3/6 + 0\right] = (1 + \varepsilon)/4.$$

Thus entropy maximisation offers the smaller average caution: if $\varepsilon = 1/2$ then entropy minimisation is about 30% more cautious, while if $\varepsilon = 1/10$ it is about 10% more cautious.

In sum, one can, after all, appeal to caution to make a case for objective Bayesianism. The maximum entropy principle is *on average the more cautious policy when it comes to risky decisions.* This caution is explained by the fact that the more middling one's degrees of belief, the smaller the number of triggered decisions on average, when trigger levels are high.

§9
Sensitivity Analysis

We have seen that an appeal to caution can be used to motivate the move from empirically-based subjective Bayesianism to objective Bayesianism: the latter is more cautious on average with respect to risky decisions. In this section we shall consider a possible response. Arguably there is an extension of empirically-based subjective Bayesianism that is more cautious still than objective Bayesianism.

Suppose background knowledge constrains a degree of belief to an interval. Suppose too that empirically-based subjective Bayesianism is adopted, so that an agent may choose any degree of belief within this interval. One may want to be extra cautious and avoid taking a course of action if the decision to do so depends on the degree of belief chosen. This leads to the following modification of the decision rule: instead of embarking on a course of action iff one's own degree of belief triggers the action, embark on it iff every possible agent would too, i.e. iff the whole interval of possible degrees of belief is greater than the trigger level. This decision rule might be called the *sensitivity analysis* rule—a decision is only taken if it is not sensitive to subjective choice of prior probability.[27]

Under this view, degrees of belief are partly a matter of subjective choice, but, once trigger levels are chosen,[28] it is an objective matter as to whether a decision will be triggered. In two dimensions this decision procedure exhibits the same average caution as objective Bayesianism over risky decisions.[29] However, in higher dimensions it will be more cautious in general. Recall the example of Fig. 3: here the maximum entropy principle triggers both decisions, the minimum entropy principle triggers one decision, while the sensitivity analysis decision procedure triggers neither.

Thus it seems that by appealing to caution the objective Bayesian is shooting herself in the foot. While such an appeal favours objective Bayesianism over empirically-based subjective Bayesianism as normally construed, it also favours the sensitivity analysis modification of empirically-based subjective Bayesianism over objective Bayesianism.

However, the sensitivity analysis approach is conceptually rather unattractive. This is because it divorces the connection between belief and action: under the sensitivity analysis approach one's degrees of belief have no bearing on whether one decides to take a course of action. It matters not a fig the extent to which one believes a patient has meningitis, because the decision as to what

[27] This type of rule is sometimes also called a *robust Bayesian* rule.

[28] The trigger levels themselves may depend on an agent's utilities and thus be subjective.

[29] Assuming, as before, that β constrains $p(A)$ to lie in a closed interval.

treatment to give is based on the *range* of permissible beliefs one might adopt, not on one's own actual beliefs. Arguably this is an unacceptable consequence of the sensitivity analysis approach which more than offsets its merit with respect to considerations of caution.

But perhaps there is some way of putting the sensitivity analysis approach on firmer footing. Perhaps one can retain its caution while re-establishing the link between belief and action. We shall investigate two possible strategies for salvaging this approach in the next two sections.

§10
Higher-Order Beliefs

The sensitivity analysis approach bases decisions on the range of beliefs that an agent might adopt, rather than on the agent's own beliefs. If one wants to retain this decision mechanism but also to insist that an agent's decisions be made on the basis of her own beliefs then one must re-interpret her beliefs as somehow encapsulating this whole range.

Suppose for example that empirical constraints force $p(A) \in [0.7, 0.8]$. Under the sensitivity analysis approach the agent is free to choose whichever degree of belief she likes from this interval, but a decision on A will only be triggered if the whole interval triggers, i.e. if 0.7 is greater than the trigger level. Consider an alternative viewpoint—a Bayesian who is uncertain as to which degree of belief x to adopt from within this interval. In the face of this uncertainty higher-order degrees of belief, such as $p(x \in [0.7, 0.75])$, become relevant. Indeed the agent may form a prior belief distribution over x, and base her decision for A on various characteristics of this distribution. One decision rule, for instance, involves triggering A if $p(x \geq \tau_A) = 1$.

This alternative viewpoint yields a type of empirically-based subjective Bayesianism: an agent's degrees of belief are constrained just by empirical knowledge. It is also compatible with cautious decision rules, such as that exemplified above. Moreover by admitting higher order degrees of belief it reinstates the link between belief and action: decisions are triggered by features of these higher order beliefs. Thus this approach appears to combine the best of all worlds—perhaps for that reason it is very popular in Bayesian statistics.

But all is not plain sailing for higher-order degrees of belief. A decision rule, such as that given above, is only cautious under some priors over x. If x is uniformly distributed over $[0.7, 0.8]$ then $p(x \geq \tau_A) = 1$ iff $0.7 \geq \tau_A$, the same cautious decision rule as the sensitivity analysis approach. On the other hand, if the probability of x is concentrated on the maximal point, $p(x = 0.8) = 1$, then the decision on A triggers just when $0.8 \geq \tau_A$—in this case the decision rule is considerably less cautious, and in particular less cautious than the maximum entropy principle. Now the empirically-based subjective Bayesian can not advocate setting one prior rather than another, since there is no extra empirical evidence to constrain choice of prior. Indeed the agent is free to choose any prior she wishes, and if she sets $p(x = 0.8) = 1$ she is far from cautious. Suggesting that the agent forms a prior over priors only defers the problem, leading to a vicious regress.

Thus we see that while the higher-order belief approach is compatible with cautious decision rules, it is also compatible with rash decision rules. It certainly can not be argued that this approach is any more cautious than the objective Bayesian methodology. Higher-order beliefs do not, after all, lead to the salvation of sensitivity analysis.

§11
Interval-Valued Beliefs

There is another way one might try to salvage the cautiousness of sensitivity analysis. Again, this involves re-interpreting an agent's beliefs as encapsulating the whole range of empirically constrained values. But this time, rather than invoking uncertainty as to which degree of belief to adopt, one instead rejects the Bayesian idea that an agent's partial beliefs are numerical point-valued degrees of belief, i.e. probabilities. Under this approach an agent's partial belief in A is identified with the whole interval yielded by empirical constraints, $bel(A) = [0.7, 0.8]$ in our example. Kyburg Jr (2003) provides a recent exposition of this strategy.[30]

Interval-valued beliefs offer a very appealing reconstruction of the sensitivity analysis approach. A natural decision rule proceeds thus: trigger a course of events on A iff the agent's partial belief in A is entirely above the trigger level for A, in our example, iff $[0.7, 0.8] \geq \tau_A$, i.e. iff $0.7 \geq \tau_A$. Not only does the resulting framework capture the cautiousness of the sensitivity analysis approach, it also ties the triggering of a decision to the agent's own partial belief, rather than the beliefs of other possible agents.

The crunch point is this. The partial belief approach appears to be superior to objective Bayesianism with respect to caution; does this gain outweigh any difficulties that accompany the rejection of point-valued degrees of belief in favour of interval-valued beliefs?

I would argue not. First of all, there are qualifications to be made about the cautiousness of interval-valued beliefs that diminish their supposed superiority over objective Bayesianism. Second, the problems that accompany interval-valued beliefs arguably outweigh any remaining benefit in terms of caution.

First to the qualifications. The typical way of generating interval-valued beliefs runs as follows.[31] Sample an attribute A from a population; say the sample frequency is 0.75; under certain assumptions about the sampling mechanism and the population, one might form a confidence interval, say $[0.7, 0.8]$ for the population frequency; set one's partial belief in A to this confidence interval. The key problem with this approach is that the confidence interval will depend upon the chosen confidence level as well as the sampling assumptions—thus the endpoints of the interval are somewhat arbitrary. But the decision procedure depends crucially on the endpoints: a 95% confidence interval $[0.7, 0.8]$ may trigger a course of action for A while a 99.9% confidence interval $[0.5, 1]$ may fail to trigger the action.

There is no non ad hoc way of determining a suitable confidence interval and so this approach must be abandoned if one wants an objective, cautious decision

[30] See Kyburg Jr and Teng (2001) for the formal details of the this approach. Kyburg Jr and Teng (1999) argue that the interval approach performs better than the subjective Bayesian approach in the short run in a betting set-up.

[31] Kyburg Jr and Teng (2001, §11.3) adopt this sort of approach.

procedure. Perhaps the best alternative strategy is just to set one's partial belief in A to the sample frequency 0.75—this is, after all, the most probable candidate for the population frequency. But then the belief is not interval-valued after all, it is point-valued. Thus the interval-valued approach loses its edge. (Arguably this is a qualification to the interval-valued belief approach rather than a reason to dismiss it altogether: one can still adopt an interval-valued belief in certain circumstances, for example if there are two samples with sample frequencies 0.75 and 0.77 respectively then it seems natural to view the whole interval $[0.75, 0.77]$ as a candidate for partial belief.)[32]

There is a further qualification to be made to the supposed superiority of the interval approach over the objective Bayesian approach. In a fully-blown decision-theoretic setting where potential losses are quantifiable, there is a sense in which interval-valued beliefs perform no better than objective Bayesian degrees of belief. The maximum entropy principle and the sensitivity analysis approach *both behave optimally* in the sense that they both succeed in minimising worst-case expected logarithmic loss. (Again arguably this is merely a qualification: one may not care very much about minimising worst-case expected logarithmic loss. On the other hand Grünwald and Dawid (2004) show that a generalised version of the maximum entropy principle is equivalent to the sensitivity analysis approach in that they both minimise loss for an arbitrary loss function, not just logarithmic loss.)

We have seen then that the advantages of interval-valued beliefs with respect to caution are somewhat diminished. Next we turn to the problems that accompany interval-valued beliefs. As noted in §5, it is not empirical evidence that adjudicates between the two approaches; the problems with interval-valued beliefs are pragmatic and conceptual.

From the pragmatic point of view, it is harder to obtain and work with interval-valued beliefs than point-valued beliefs. Roughly speaking it is twice as hard, since there are twice as many numbers to have to deal with: to each point-valued degree of belief there are the two endpoints of the corresponding interval-valued belief.[33] Intervals also make it hard to do things that are simple using numbers. For instance, suppose one wants to either trigger a course of action for A, or otherwise to trigger another course of action for $\bar{A}$: to give antibiotics if meningococcal meningitis is suspected, otherwise paracetamol. In the case of point-values degrees of belief one simply needs to ensure that $\tau_A + \tau_{\bar{A}} = 1$, for then if $p(A) \geq \tau_A$ one action is taken, otherwise $p(\bar{A}) = 1 - p(A) > \tau_{\bar{A}}$ and the other action is taken.[34] On the other hand if partial beliefs are intervals then for any non-extreme trigger levels there are partial beliefs ($[x, y]$ where $x < \tau_A$ and $y > 1 - \tau_{\bar{A}}$) such that *neither* action triggers.[35]

There are thus pragmatic reasons to favour point-valued degrees of belief over interval-valued beliefs. Accordingly one might reason something like this: the formalism of point-valued degrees of belief offers a first approximation to how one should reason; the formalism of interval-valued beliefs offers a second

[32] See Williamson (2005a, §5.3) for discussion of this type of constraint on rational belief.

[33] Cozman (2000) develops a computational framework for interval-valued beliefs.

[34] One would of course also need a policy to decide which action is triggered if $p(A) = \tau_A$ (in which case $p(\bar{A}) = \tau_{\bar{A}}$).

[35] One can get round this problem by insisting that trigger levels be functions of the partial beliefs themselves as well as the decision outcomes A or $\bar{A}$. The point is not that intervals make things impossible, just that intervals make things more complicated.

approximation; the first approximation tends to be easier to use in practice and there is little to be gained (in terms of caution) by using the second approximation; thus one should use the first approximation by default. But such a view assumes that there is essentially more to interval-valued beliefs than point-valued beliefs—that conceptually they add something. There are reasons for doubting this perspective, as we shall now see.

From the conceptual point of view, the interval-valued belief approach is caught in a dilemma: it either lacks the intuitively compelling connection between beliefs and betting quotients that underpins the Bayesian approach, or it fails to add anything conceptually to the Bayesian approach. One of the key points in favour of the Bayesian approach is that an agent's partial belief in E is interpretable as a betting quotient, 'the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum pS'.[36] One cannot simply identify an interval-valued partial belief with a betting quotient—a betting quotient is a single number but an interval is a set of numbers. One might try, as Borel (1943, §3.8) did, to interpret the interval as a *set* of acceptable betting quotients.[37] To do so one must adopt a different betting set-up to that of the Bayesian, one without the requirement that the agent buys and sells bets at the same rate, i.e. one in which S must be positive rather than either positive or negative. (Otherwise, if the agent has more than one betting quotient in the same outcome then she can simply be Dutch booked—a set of bets can be found that forces her to lose money whatever happens.) But Adams (1964, §7) showed formally that when this strategy is pursued one can identify a *single* probability function that can be thought of as representing betting quotients that the agent regards as fair.[38] Thus this new betting set-up is a dead end for the proponent of interval-valued beliefs: a *set* of betting quotients $[0.7, 0.8]$ for an outcome in the new set-up turns out to be equivalent to a *single* betting quotient 0.75 in the original set-up—it is just a more complicated way of saying the same thing. In sum, by trying to provide a betting interpretation for interval-valued partial beliefs one just ends up motivating point-valued degrees of belief; betting fails to provide a distinctive foundation for interval-valued partial beliefs after all, and interval-valued beliefs should not be thought of as a second approximation or refinement of point-valued beliefs. Thus the proponent of interval-valued beliefs must either accept that they are essentially just a complication of point-valued degrees of belief, or, if they are to differ conceptually, they lack any link with betting behaviour that accounts for that difference. In the absence of a viable betting interpretation, the question arises as to how interval-valued beliefs are to be interpreted: what does it mean to believe A with value $[0.7, 0.8]$?[39] In either case, it is hard to see how intervals could be better candidates for partial beliefs than numbers.

The proponent of intervals may respond here that the Bayesian link between partial beliefs and betting quotients is less attractive than one might

[36] (de Finetti, 1937, p. 62). Ramsey (1926, p. 172) proposed a similar betting set-up: 'The old-established way of measuring a person's belief is to propose a bet, and see what are the lowest odds which he will accept'.

[37] See also Smith (1961) and Walley (1991, §1.6.3).

[38] See also Koopman (1940a,b).

[39] As Walley (1991, p. 22) himself notes, 'unless the conclusions have a behavioural interpretation, it is not clear how they can be used for making decisions of for guiding future inquiry and experimentation.'

think. Of course, if there are grounds for abandoning Bayesian betting set-up then the absence of a viable betting interpretation is not such a disadvantage for interval-valued beliefs. One objection to the Bayesian betting set-up is that human agents can't always evaluate their betting quotients in terms of unique real numbers—human beliefs are simply not so precise.[40] Another is that it is rather impractical to elicit degrees of belief using a series of bets, since as noted in §6 this is a time-consuming operation, and in any case people are often either reluctant to bet at all or happy to lose money whatever happens. These points are well made, but by-the-by in our context because they only trouble the subjective and empirically-based versions of Bayesianism, not objective Bayesianism. Under objective Bayesianism, agents do not need to search their souls for real numbers to attach to beliefs, nor is the betting set-up required to measure those degrees of belief. Degrees of belief are determined by the maximum entropy principle and they are measured by maximising entropy. (In this age of mechanisation human agents and artificial agents alike can use computing power to work out the extent they ought to believe various outcomes.) Thus for objective Bayesianism the betting interpretation is only important for the *meaning* it gives to degrees of belief. The fact that the betting set-up is simplistic, or an idealisation, or impractical, is neither here nor there—the objective Bayesian does not go on to invoke the betting set-up, as Ramsey did, as an elicitation or measurement tool.

There is another conceptual problem that besets the interval-valued belief approach. One might argue, as de Cooman and Miranda do in this volume, that interval-valued beliefs are superior to point-valued beliefs because they can represent the amount of knowledge that grounds a belief. Under the objective Bayesian account, a degree of belief 0.5 that it will snow in Stroud today may be based on total ignorance, or it may be based on good evidence, e.g. the knowledge that the frequency of it snowing in Stroud on a day like today is $\frac{1}{2}$. In contrast, the interval-valued approach distinguishes between these two cases: while in the latter case the belief might have value 0.5, in the former case the belief would have value $[0, 1]$. In a sense, then, interval-valued beliefs tell us about the knowledge on which they are based; this is supposed to be an advantage of the interval-valued approach.

But this is evidence of a conceptual problem with the interval-valued approach, not an advantage. The problem is this. The question 'how much knowledge does the agent have that pertains to it snowing in Stroud today?' is a question about knowledge, not belief. Consequently, it should not be the model of belief that answers this question; it should be the knowledge component of the agent's epistemic state. But on the interval-valued approach, it is the belief itself that is used to answer the question (and typically there is no separate model of the agent's knowledge). Thus on this approach, the belief model conflates concerns to do with knowledge and belief. On the other hand, the objective Bayesian approach maintains a nice distinction between knowledge and belief: an agent has background knowledge which is then used to constrain her choice of degrees of belief; the former component contains information about the extent of the agent's knowledge, while the latter contains information about the strength of the agent's beliefs; neither can be used to answer questions about the other. The interval-valued approach, then, muddles issues to do with knowledge

[40]See e.g. Kyburg Jr (1968).

and belief, while the objective Bayesian approach keeps them apart. But the goal of both approaches is to model a rational agent's epistemic state, and this requires a sharp distinction between knowledge and belief. It is a fact of the matter that some of our full beliefs are of a higher grade than others and are more entrenched in the sense that we are less willing to revise them. I have full belief that a point chosen at random in a ball will not be its centre; I also have a full belief that I am alive; I'm quite willing to revise the former belief in the face of evidence, but not the latter—it is knowledge, not belief, that accounts for this distinction. Hence the objective Bayesian approach, by maintaining a proper distinction between knowledge and belief, offers a better model of an agent's epistemic state.

To conclude, the limited gains that interval-valued beliefs offer in terms of caution are arguably offset by the conceptual as well as pragmatic advantages of point-valued beliefs.

§12
Summary

We see then that there is a case for preferring objective Bayesianism over an empirical-constraint theory of rational belief. Objective Bayesianism is more cautious than empirically-based subjective Bayesianism; while it may be less cautious than the interval-valued belief approach, it offers pragmatic advantages and a simple interpretation in terms of betting quotients.

References

Adams, E. W. (1964). On rational betting systems. *Archiv für mathematische Logik und Grundlagenforschung*, 6:7–29 and 112–128.

Borel, E. (1943). *Probabilities and life*. Dover, New York, 1962 edition.

Cozman, F. G. (2000). Credal networks. *Artificial Intelligence*, 120:199–233.

de Finetti, B. (1937). Foresight. its logical laws, its subjective sources. In Kyburg, H. E. and Smokler, H. E., editors, *Studies in subjective probability*, pages 53–118. Robert E. Krieger Publishing Company, Huntington, New York, second (1980) edition.

Gillies, D. (1991). Intersubjective probability and confirmation theory. *British Journal for the Philosophy of Science*, 42:513–533.

Gottfried, K. and Wilson, K. G. (1997). Science as a cultural construct. *Nature*, 386:545–547. With discussion in vol. 387 p. 543.

Grünwald, P. and Dawid, A. P. (2004). Game theory, maximum entropy, minimum discrepancy, and robust Bayesian decision theory. *Annals of Statistics*, 32(4):1367–1433.

Halpern, J. Y. (1999a). A counterexample to theorems of Cox and Fine. *Journal of Artificial Intelligence Research*, 10:67–85.

Halpern, J. Y. (1999b). Cox's theorem revisited. *Journal of Artificial Intelligence Research*, 11:429–435.

Hosni, H. and Paris, J. (2005). Rationality as conformity. *Synthese*, 144(3):249–285.

Howson, C. and Urbach, P. (1989). *Scientific reasoning: the Bayesian approach.* Open Court, Chicago IL, second (1993) edition.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.

Jaynes, E. T. (2003). *Probability theory: the logic of science.* Cambridge University Press, Cambridge.

Koopman, B. O. (1940a). The axioms and algebra of intuitive probability. *Annals of Mathematics*, 41(2):269–292.

Koopman, B. O. (1940b). The bases of probability. *Bulletin of the American Mathematical Society*, 46:763–774.

Kyburg Jr, H. E. (1968). Bets and beliefs. *American Philosophical Quarterly*, 5(1):54–63.

Kyburg Jr, H. E. (2003). Are there degrees of belief? *Journal of Applied Logic*, 1:139–149.

Kyburg Jr, H. E. and Teng, C. M. (1999). Choosing among interpretations of probability. In *Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence (UAI-99)*, pages 359–365. Morgan Kaufmann Publishers, San Francisco, CA.

Kyburg Jr, H. E. and Teng, C. M. (2001). *Uncertain inference.* Cambridge University Press, Cambridge.

Paris, J. B. (1994). *The uncertain reasoner's companion.* Cambridge University Press, Cambridge.

Paris, J. B. and Vencovská, A. (1990). A note on the inevitability of maximum entropy. *International Journal of Approximate Reasoning*, 4:181–223.

Paris, J. B. and Vencovská, A. (2001). Common sense and stochastic independence. In Corfield, D. and Williamson, J., editors, *Foundations of Bayesianism*, pages 203–240. Kluwer, Dordrecht.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann, San Mateo CA.

Press, S. J. and Tanur, J. M. (2001). *The subjectivity of scientists and the Bayesian approach.* Chichester : Wiley.

Ramsey, F. P. (1926). Truth and probability. In Kyburg, H. E. and Smokler, H. E., editors, *Studies in subjective probability*, pages 23–52. Robert E. Krieger Publishing Company, Huntington, New York, second (1980) edition.

Salmon, W. C. (1990). Rationality and objectivity in science, or Tom Kuhn meets Tom Bayes. In Wade Savage, C., editor, *Scientific theories*, pages 175–204. University of Minnesota Press, Minneapolis. Minnesota Studies in the Philosophy of Science 14.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423 and 623–656.

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society Series B*, 23:1–37. With discussion.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* Chapman and Hall, London.

Williams, P. M. (1980). Bayesian conditionalisation and the principle of minimum information. *British Journal for the Philosophy of Science*, 31:131–144.

Williamson, J. (2005a). *Bayesian nets and causality: philosophical and computational foundations.* Oxford University Press, Oxford.

Williamson, J. (2005b). Objective Bayesian nets. In Artemov, S., Barringer, H., d'Avila Garcez, A. S., Lamb, L. C., and Woods, J., editors, *We Will Show Them! Essays in Honour of Dov Gabbay*, volume 2, pages 713–730. College Publications, London.

Williamson, J. (2006). Philosophies of probability: objective Bayesianism and its challenges. In Irvine, A., editor, *Handbook of the philosophy of mathematics.* Elsevier. Handbook of the Philosophy of Science volume 4.