

Two Dogmas of Strong Objective Bayesianism

Prasanta S. Bandyopadhyay and Gordon Brittan, Jr

We introduce a distinction, unnoticed in the literature, between four varieties of objective Bayesianism. What we call ‘strong objective Bayesianism’ is characterized by two claims, that all scientific inference is ‘logical’ and that, given the same background information two agents will ascribe a unique probability to their priors. We think that neither of these claims can be sustained; in this sense, they are ‘dogmatic’. The first fails to recognize that some scientific inference, in particular that concerning evidential relations, is not (in the appropriate sense) logical, the second fails to provide a non-question-begging account of ‘same background information’. We urge that a suitably objective Bayesian account of scientific inference does not require either of the claims. Finally, we argue that Bayesianism needs to be fine-grained in the same way that Bayesians fine-grain their beliefs.

1. Overview

All Bayesians assign a central role to Bayes’s theorem in the development of an adequate account of scientific inference and all Bayesians understand the probability operators in that theorem in terms of degrees of belief. Probability theory measures something that is ‘in the head’, not ‘in the world’. In this sense, it is misleading to identify some Bayesians as ‘objective’, others as ‘subjective’. The traditional term ‘personalist’ is better suited to identify the latter.

Although all Bayesians are, in the sense indicated, subjectivists, not all subjectivists are Bayesians. There is a long tradition beginning with Laplace which equally takes probabilities as something ‘in the head’, namely, as measuring degrees of our ignorance. At the same time, all subjectivists, and *a fortiori* all Bayesians, insist that there is something ‘objective’ about the attribution and interpretation of probabilities. Laplaceans locate objectivity with respect to the principles of indifference and determinism.

Prasanta S. Bandyopadhyay and Gordon Brittan, Jr are at the Department of History and Philosophy and the Astrobiology Biogeocatalysis Research Center, Montana State University. Correspondence to: Department of History and Philosophy, Montana State University, Bozeman, MT 59717, USA. E-mail: psb@montana.edu, uhigb@montana.edu

Bayesians of different stripes characterize it in terms of varying consistency requirements placed on the agent's beliefs.

According to personalists and all objective Bayesians, an agent's belief must satisfy the rules of the probability calculus. Otherwise, on a familiar Dutch Book argument, the agent's degree of belief is incoherent. Personalist Bayesians take this (probabilistic) coherence to be both a necessary and a sufficient condition for the rationality of an agent's beliefs, and then (typically) argue that the beliefs of rational agents will converge over time. The point of scientific inference, and the source of its 'objectivity', is to guarantee coherence and ensure convergence. Objective Bayesians, on the other hand, *typically* insist that while the coherence condition is necessary, it is not also sufficient for the kind of objectivity which scientific methodologies are intended to make possible. At this point, they diverge rather markedly, different camps proposing different additional coherence conditions.

In this paper, we distinguish four varieties of objective Bayesianism (Section 1): strong objective Bayesianism (hereafter strong objectivism),¹ moderate objective Bayesianism (hereafter moderate objectivism), quasi-objective Bayesianism (hereafter quasi-objectivism), and necessitarian objective Bayesianism (hereafter necessitarian objectivism).² We then focus on strong objectivism. The correct characterization of scientific inference is the aim of any form of Bayesianism. They are intended to be descriptive of at least many scientific practices and normative with respect to all of it. On its account, scientific inference has or should have two features: all scientific inference is or should be deductive or probabilistic in character, and only one probability function is or should be considered rational for an agent based on her background knowledge. These two features come together in the standard strong objectivist claim that, in Roger Rosenkrantz's words, 'two reasonable persons in possession of the same relevant data should assign the same probabilities to the elements of a considered partition of hypotheses' (Rosenkrantz 1981, 4), or as Edwin Jaynes has it, 'if in two problems the robot's state of knowledge is the same ..., then it must assign the same plausibilities in both' (Jaynes 2003, 19; emphasis altered).³ The same idea has been expressed forcibly by Jon Williamson: '[t]he aim of objective Bayesianism is to constrain degrees of belief in such a way that only one value of $p(v)$ will be deemed rational on the basis of an agent's background knowledge. Thus, objective Bayesian probability varies as background knowledge varies but *two agents* with the same background knowledge *must adopt the same* probabilities as their rational degree of belief'. (Williamson 2005, 11–12; emphasis added). This claim allows us to understand just how a scientific inference is an *inference*, 'logical', and in what way it is *scientific*, i.e., 'agent- or observer-invariant'.

Our argument is twofold. We argue first that the strong objectivist account of scientific inference is *not* sufficient; certain basic dimensions of inferences commonly accepted as 'scientific' are not 'logical' (they do not conform to the rules of deductive or probability logic).⁴ Neither is it *necessary*; since two equally 'rational' agents⁵ having the same background knowledge need not have only one probability function. Put otherwise, our argument is that strong objectivism rests on two dogmas which are, in turn, incompatible with at least some scientific reasoning and not required by any of it.

Only a sensible objective form of Bayesianism what we call ‘quasi-objectivism’ frees us from them.

The argument proceeds in this way. We first point out that the strong objectivist fails to distinguish between belief and evidence relations, that the evidence relation is a prominent feature of much scientific reasoning, and that the evidence relation is not (again in the intended sense) ‘logical’ (Section 2). Indeed, failure to distinguish between belief and evidence leads directly to the ‘old evidence’ problem, often thought to be an insurmountable problem for Bayesians of every stripe (Section 3). We then argue that the strong objectivist’s demand that, given the same background information, two agents will assign the same hypothesis the same probability simply cannot be satisfied in all cases (Section 4). We consider a possible rejoinder by a strong objectivist (Section 5). Finally, we outline a version of Bayesianism without the dogmas (Section 6).

2. Varieties of Objective Bayesianism

Jaynes, Rosenkrantz, and Williamson are notable representatives of the strong objectivist camp.⁶ Jaynes, for example, while admitting that a probability assignment could be construed as ‘subjective in the sense that it describes only a state of knowledge, and not anything that could be measured in a physical experiment’, writes that ‘this assignment is completely objective in the sense that [it is] independent of the probability of the user. [...] It is “objectivity” in this sense that it is needed for a scientifically respectable theory of inference’ (Jaynes 2003, 44–45). He writes further that ‘our theme is simply: *probability theory as extended logic*’ (Jaynes 2003, xxiii); ‘[i]t is clear that not only is the quantitative use of the rules of probability theory as extended logic the only sound way to conduct inference, it is the failure to follow those rules strictly that for many years has been leading to unnecessary error, paradoxes, and controversies’ (Jaynes 2003, 143).

Influenced by Jaynes’s work, Williamson construes strong objectivism in terms of two norms; (a) a logical norm, something like the classical principle of indifference, that requires the agent’s assignment of priors to be maximally non-committal (in those cases in which the priors do not simply correspond to relative frequencies), and (b) a background information norm that further constrains her non-committal degree of belief (Williamson 2005). These two norms are intended to insure the ‘logical’ character of scientific inference and the eventual generation of unique probabilities with respect to its conclusion. Williamson contends that when there is a conflict between the logical norm and the background information norm, the latter should prevail. He explains this point about a possible conflict and its subsequent resolution by invoking Jakob Bernoulli’s well-known example.

Consider three ships that set sail from a port. After some time, suppose that we have received the information that one of them suffered a shipwreck. Based on the logical norm, we should be justified in assigning probability $1/3$ to the one that really suffered a shipwreck because there were only three ships to begin with. The case of assigning an equal probability to each ship that could sustain the shipwreck represents the agent’s

maximal uncertainty about this situation. However, if we recall that in fact one of them was in a bad shape due to its old age, then we could safely assume that the latter was the one which was really shipwrecked. This is how our background information could be brought to bear on constraining logical features if and when the former is available. If it is not, and in the absence of frequency data, we fall back on sophisticated variants of the principle of indifference.

In contrast, statisticians like José M. Bernardo have approached objective Bayesianism in a slightly different way. Bernardo writes that ‘[i]t has become standard practice ... to describe as “objective” any statistical analysis which only depends on the [statistical] model assumed. In this precise sense (and only in this sense) reference analysis is a method to produce “objective” Bayesian inference’ (Bernardo 2005).

For Bernardo, the reference analysis which he has advocated to promote his brand of objective Bayesianism should be understood in terms of some parametric model of the form $M \equiv \{\Pr(x | w), x \in X, w \in \Omega\}$, which describes the conditions under which data have been generated. Here, the data x are assumed to consist of one observation of the random process $x \in X$ with probability distribution $\Pr(x | w)$ for some $w \in \Omega$. A parametric model is an instance of a statistical model. Bernardo defines $\theta = \theta(w) \in \Theta$ to be some vector of interest. All legitimate Bayesian inferences about the value θ are captured in its posterior distribution

$$\Pr(\theta | x) \propto \int_{\Lambda} \Pr(x | \theta, \lambda) \Pr(\theta, \lambda) d\lambda$$

provided these inferences are made under an assumed model. Here, λ is some vector of nuisance parameters and is often referred to as ‘model’ $\Pr(x | \lambda)$.

The posterior distribution combines the information provided by the data x along with some other information about θ contained in the prior distribution $\Pr(\theta, \lambda)$. The *reference prior function* for θ and given model M , and a class of candidate priors, P , according to Bernardo, then become the joint prior $\pi^{\theta}(\theta, \lambda | M, P)$. For him, the reference prior $\pi^{\circ}(\theta, \lambda | M, P)$ is objective in the sense that it is a well-defined mathematical expression of the vector of interest θ , the assumed model M , and the class of candidate priors P , with no additional subjective elements. Since reference priors are defined as a function of the entire model, $M \equiv \{\Pr(x | w), x \in X, \theta \in \Theta\}$, it is clearly not a function of the *observed* likelihood. The purpose of invoking reference priors is to represent lack of information on the part of an agent about the parameter of interest. To be faithful to this goal, a moderate Bayesian extracts information about that parameter by *repeated sampling* from a specific experimental design M .

Insofar as the foundation of Bayesian statistics is concerned, the theme of repeated sampling that underlies regeneration of reference priors is alleged to defy the conditionality principle (CP), a backbone of Bayesianism. When the CP is informally stated, it implies that what matters in making inferences about θ is actually performed experiments, and what the agent would have obtained if she had carried out a non-randomly chosen experiment in repeated samplings becomes irrelevant for the purpose of inference about that parameter. However, Bernardo thinks that the use of reference distributions is fully in agreement with the CP. The point is that the

reference posteriors provide a conditional answer and all probabilities are conditional namely what could be said about the quantity of interest if prior information was minimal compared to the information that particular model could possibly provide. Obviously in this situation one can make as many of these conditional statements as one wishes to do.

Coming back to the reference priors, often the reference priors turn out to be both *non-informative* and *improper*. They are non-informative because they convey ignorance about the parameter of interest. They are improper because they don't integrate to one. One such prior is Jeffreys's prior, which is taken to be proportional to the square root of the determinant of the information matrix.⁷ One needs to be careful that appropriate use of improper prior functions as reference priors fits perfectly within standard probability.⁸

The attraction of this kind of objectivism is its emphasis on 'reference analysis', which with the help of statistical tools has made further headway in turning its theme of objectivity into a respectable statistical school within Bayesianism. Bernardo writes, '[r]eference analysis may be described as a method to derive model-based, non-subjective posteriors, based on the information-theoretical ideas, and intended to describe the inferential content of the data for scientific communication' (Bernardo 1997). Here by the 'inferential content of the data', Bernardo means that the former provides 'the basis for a method to derive non-subjective posteriors'. Bernardo's objective Bayesianism consists of the following two claims.

(i) Bernardo thinks that the agent's background information should help the investigator build a statistical model, hence ultimately influence which prior the latter should assign to the model. Therefore, although Bernardo might endorse arriving at a unique probability value as a goal, he does not require that we need to have the unique probability assignment in all issues at one's disposal. He writes, '[t]he analyst is supposed to have a *unique* [emphasis added] (often subjective) prior $p(w)$, independently of the design of the experiment, but the scientific community will presumably be interested in comparing the corresponding analyst's personal posterior with the *reference* (consensus) posterior associated to the published experimental design' (Bernardo 2005, 29).

(ii) For Bernardo, statistical inference is nothing but a case of deciding among various models or theories, where decision includes, among other things, the utility of acting on the assumption of the model or theory being empirically adequate. Here the utility of acting on the empirical adequacy of the model or theory in question might involve some loss function (Bernardo and Smith 1994, 69).

We call Bernardo's objectivism 'moderate' because the additional conditions he proposes fall short of requiring that the investigator arrive at unique probability assignments. This feature distinguishes his version of objectivism from strong objectivism, although like both Jaynes and Williamson, he breaks with the personalists in constraining both priors and posteriors.

We now consider a third variety of objective Bayesianism, which we call quasi-objectivism. This version differs from the others in that it imposes a simplicity constraint on the agent's prior beliefs over and above their coherence with the axioms

of probability theory, but does not impose the further constraints suggested by Jaynes, Williamson, and Bernardo. In particular, it does not require, as do Jaynes and Williamson, that ‘objectivity’ be construed in terms of agent-invariance. The simplicity constraint guides us as to how priors should be assigned to different competing theories or models, but does not determine them completely. Consider the curve-fitting problem to understand how it works. The problem arises when an investigator is confronted with a trade-off between simplicity and goodness of fit. The investigator could consider several possible competing hypotheses, namely, $H_1, H_2,$ and H_3 for solving the optimization problem, and would like to choose one among three mutually exclusive hypotheses as the best trade-off. (The rest of the hypotheses are lumped together and called the *catch-all* hypothesis; it is left out of discussion because it generates well-known problems.) These possible hypotheses are:

$$H_1: E(Y|x) = \alpha_0 + \alpha_1 x$$

$$H_2: E(Y|x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2$$

$$H_3: E(Y|x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$$

Here, $E(Y|x)$ is the conditional expectation of Y given x . To say that these hypotheses are mutually exclusive is to say that the coefficients of x^k under H_k are non-zero. The simpler a hypothesis, the greater prior probability it will receive. On this view, the simplicity of a theory helps to constrain its prior probability. Two sorts of factors, formal and non-formal, determine the simplicity of a theory. The formal factor in question measures the paucity of adjustable parameters; functions of lower degree get higher prior probability than functions of higher degree.

Though the formal factor is the only factor that dictates one’s choice of one theory over others, it does not often suffice to determine the simplicity of a theory. A choice of one theory over others sometimes depends on epistemic and pragmatic considerations as well. For example, we disregard the negative value of a quadratic equation when confronted with a distance problem. For another example, whether an equation is mathematically tractable plays a vital role in theory choice, hence in the assignment of prior probabilities to theories.

For illustrative purposes, we consider a range of putatively plausible priors for the hypotheses in the curve-fitting problem (Table 1). Each prior in Table 1 satisfies the

Prior	Equation
$\text{Pr}_1(H_k)$	$2^{-k}; k = 1, 2, \dots$
$\text{Pr}_2(H_k)$	$(e - 1) e^{-k}; k = 1, 2, \dots$
$\text{Pr}_3(H_k)$	$(\sqrt{n} - 1) n^{-k/2}; k = 1, 2, \dots$

Table 1 A range of putatively plausible priors for the hypotheses in the curve-fitting problem

rules of the probability calculus. One could use any of the priors from Table 1 and still be consistent with the quasi-objective Bayesian's account of simplicity.

Since these priors are not unique, given the data, the investigator is not expected to get a unique posterior probability for her models. At best, she can rank-order them. Yet the simplicity condition does add a further element of objectivity, which in this context comes to restrain the agent's initial degrees of belief. Different possible priors along with the likelihoods under different competing hypotheses yield a unique choice, but not a unique probability (for more on the mathematical aspects of how a unique choice could be achieved in the curve-fitting problem, see Bandyopadhyay, Boik, and Basu 1996; Bandyopadhyay and Boik 1999; Bandyopadhyay and Brittan 2001).

The fourth and final variety of objective Bayesianism we call necessitarian objectivism. The latter takes probabilities to represent an objective logical relation between a set of sentences. The degree of confirmation for the hypothesis, 'the die will show face six on the next throw' will take exactly one value in between 0 and 1 relative to a set of sentences that consists of an agent's total body of evidence. Probability is a syntactical, meta-linguistic notion, which is an analogue of the notion of provability. Even though premises and a conclusion of an argument could be empirical, the relation between premises and the conclusion must be analytic or logical. The originators of this view are Rudolf Carnap (1950) and John Maynard Keynes (1921). Carnap's early work (1950) consists in finding a measure m such that the degree of confirmation of the hypothesis given the data is a function of that measure of the ratio of the possible worlds in which both the joint distribution of the hypothesis and the data hold to those in which the data hold. Two measures he considered were called m^+ and m^* in which the former does not allow learning from experience, whereas the latter does. What is significant is that since at this stage he thought that a unique probability could be achieved for a proposition given the total body of evidence, any reference to subjectivity turns out to be redundant.

Carnap's monumental work on inductive logic provides a rationale for making inductive inference analogous to deductive reasoning. In his early works, we find that the probability the agent would have for a proposition independent of any evidence is a matter of logic. On a total body of evidence, the agent's degree of belief would be determined by the rule of conditionalization. However, the later Carnap (1971) takes constraints on the agent's degree of belief *not* to determine *one*, but rather a plethora of permissible belief functions *not yielding a unique probability* for a specific conclusion conditional on given premises of an argument.

For the sake of discussion, it is helpful to refer to Table 2, summarizing how varieties of objective Bayesians differ with regard to issues pertaining to scientific inference. The first column represents names for varieties of objective Bayesians:⁹ strong objective Bayesians (SOB), moderate objective Bayesians (MOB), quasi-objective Bayesians (QOB), and necessitarian objective Bayesians (NOB). The second column stands for how they differ with regard to the coherence constraint on an agent's degree of belief. The third represents the tools different objective Bayesians use for their inference and decision machines. The fourth is the name of the machine each uses for achieving her divergent goals. The fifth and final column captures different stances objectivists adopt

Varieties	Coherence constraint	Tools for machine	Machine	Unique probability
SOB	Necessary, but not sufficient	Probability theory and deductive logic	Inference	Yes
MOB	Necessary, but not sufficient	Probability theory, deductive logic and loss function	Decision	Need not be
QOB	Necessary, plus simplicity consideration	Probability theory, deductive logic and likelihood function	Inference	Need not be, but unique choice
NOB	Necessary	Probability theory and deductive logic	Inference	Yes (early Carnap) Need not be (later Carnap)

Note: MOB, moderate objective Bayesians; NOB, necessitarian objective Bayesians; QOB, quasi-objective Bayesians; SOB, strong objective Bayesians.

Table 2 Varieties of Bayesians on issues pertaining to scientific inference

toward how or whether one is able to arrive at a unique probability for the conclusion of an argument.

Most Bayesians agree that it is necessary that an agent's degree of belief must satisfy the probability calculus, but differ regarding whether it is sufficient. As already noted, both strong and moderate objective Bayesians take the satisfaction of the calculus to be necessary. Quasi-objectivists think that when one makes an inference, one draws a conclusion from a set of premises that contains an agent's prior knowledge of the situation in question and her data information about it; the utility of accepting or rejecting a theory raises very different sorts of questions. Indeed, quasi-objectivists do not think that scientific inference can be understood entirely in probabilistic terms. Both moderate and quasi-objectivism break with strong objectivism on the issue of requiring unique probability distributions. Quasi-objectivism differs from moderate objectivism with regard to the latter's view that scientific inference is primarily a form of a decision among contending theories.¹⁰

3. Belief and Evidence: Epistemological Significance

Consider two hypotheses: H , representing that a patient suffers from tuberculosis, and $\sim H$, its denial. Assume that an X-ray, which is administered as a routine test, comes out positive for the patient. Based on this simple scenario, following the work of Richard M. Royall (1997), one could pose three questions that underline the epistemological issue at stake:

- (i) Given the datum, what should we *believe* and to what degree?
- (ii) What does the datum say regarding *evidence* for H against its alternative?
- (iii) Given the datum what should we *do*?

The first question we call the belief question, the second the evidence question, and the third the decision question. In this paper, we address the first two questions. Bayesians

have developed distinct accounts to answer them; the first is an account of confirmation and the second of evidence. For Bayesians, an account of confirmation explicates a relation, $C(D, H, B)$ among data, D , hypothesis, H , and the agent's background knowledge, B . Because the confirmation relation so explicated is a belief relation, it must satisfy the probability calculus, including the rule of conditional probability, as well as some reasonable constraints on one's *a priori* degree of belief in an empirical proposition.

For Bayesians, belief is fine-grained. In the case of empirical propositions, belief admits of any degree of strength between 0 and 1, except insofar as the proposition is conditionalized (in which case it has probability equal to one). A satisfactory Bayesian account of confirmation captures this notion of degree of belief. In formal terms, it may be stated as:

Confirmation condition (CC): D confirms H if and only if $\Pr(H|D) > \Pr(H)$

The posterior/prior probability of H could vary between 0 and 1 exclusive. Confirmation becomes strong or weak depending on how high or low the difference is between the posterior probability, $\Pr(H|D)$, and the prior probability of the hypothesis, $\Pr(H)$:

$$\Pr(H|D) = \Pr(H)\Pr(D|H) / \Pr(D) \quad (1)$$

While this account of confirmation is concerned with belief in a single hypothesis embodied in equation (1), an account of evidence in our framework must compare the merits of two simple statistical hypotheses, H_1 and H_2 (or $\sim H_1$) relative to the data D , auxiliaries A , and background information, B . The justification for the account to be comparative is that the evidence question is comparative and the account of evidence has been developed to handle the evidence question. However, because the evidence relation is not a belief relation it need not satisfy the probability calculus. Bayesians use the Bayes factor (BF) to make this comparison, while others use the likelihood ratio (LR) or other functions designed to measure evidence. For simple statistical hypotheses, the Bayes factor and the likelihood ratio are identical, and capture the bare essentials of an account of evidence without any appeal to prior probability (Berger 1985, 146).¹¹ The BF/LR can be represented in this case by equation (2).

$$\text{BF/LR} = \Pr(D|H_1, A_1 \& B) / \Pr(D|H_2, A_2 \& B) \quad (2)$$

is called the Bayes factor (likelihood ratio) in favour of H_1 , A_1 , and B .¹²

D becomes E (evidence) for $H_1 \& A_1 \& B$ against $H_2 \& A_2 \& B$ if and only if their ratio is greater than one. An immediate corollary of equation (2) is that there is equal evidential support for both hypotheses only when $\text{BF/LR} = 1$. Note that in equation (2) if $1 < \text{BF/LR} \leq 8$, then D is often said to provide weak evidence for H_1 against H_2 , while when $\text{BF/LR} > 8$, D provides strong evidence. This cut-off point is otherwise determined contextually and may vary depending on the nature of the problem that confronts the investigator. The range of values for BF/LR varies from 0 to infinity inclusive.

Let us continue examining the tuberculosis (TB) example described earlier to understand why belief and evidence are distinct concepts.¹³ We will suppose from a large

number of data that we are nearly certain that the propensity for having a positive X-ray for members without TB is 0.7333, and the propensity for a positive X-ray for population members without TB 0.0285. Call this background theory of the propensities for X-ray outcomes B . Again, let H represent the hypothesis that an individual is suffering from TB and $\sim H$ the hypothesis that she is not. These two hypotheses are mutually exclusive and jointly exhaustive. In addition, assume D represents a positive X-ray test result. We would like to find $\Pr(H|D)$, the posterior probability that an individual who tests positive for TB actually has the disease. Bayes's theorem helps to obtain that probability. However, to use the theorem, we need to know first $\Pr(H)$, $\Pr(\sim H)$, $\Pr(D|H)$, and $\Pr(D|\sim H)$.

$\Pr(H)$ is the prior probability that an individual in the general population has TB. Because the individuals in different studies need not be chosen from the population at random, the correct frequency based prior probability of the hypothesis couldn't be obtained from them. Yet in a 1987 survey, there were 9.3 cases of TB per 100,000 population (Pagano et al. 2000). Consequently, $\Pr(H) = 0.000093$. Hence, $\Pr(\sim H) = 0.999907$. Based on a large data kept as medical records, we are certain about these following probabilities. $\Pr(D|H)$ is the probability of a positive X-ray given that an individual has TB. $\Pr(D|H) = 0.7333$. $\Pr(D|\sim H)$, the probability of a positive X-ray given that a person does not have TB, is $1 - \Pr(\sim D|\sim H) = 1 - 0.9715 = 0.0285$.

Using all this information, we compute $\Pr(H|D) = 0.00239$. For every 100,000 positive X-rays, only 239 signal true cases of TB. The posterior probability is very low, although it is slightly higher than the prior probability. Although CC is satisfied, the hypothesis is not very well confirmed. Yet at the same time, the BF, i.e., $0.7333/0.0285$ (i.e., $\Pr(D|H)/\Pr(D|\sim H)$) = 25.7, is very high. Therefore, the test for TB has a great deal of evidential significance.

There is little point in denying that the meanings of 'evidence' and 'belief' (or its equivalents) often overlap in ordinary English as well as among epistemologists. In probabilistic epistemology, the motivation for some epistemologists to use belief and evidence interchangeably could be justified by the theorem, $\text{BF} > 1$ if and only if $\Pr(H|D) > \Pr(H)$. However, belief and evidence are conceptually different, because strong belief does not imply strong evidence and conversely, as illustrated by the TB example. Our case for distinguishing them rests not on usage, but on the clarification in our thinking that is achieved and is supported by inferences frequently made in diagnostic studies (for a discussion on the belief and evidence distinction, see Bandyopadhyay 2007; Bandyopadhyay and Brittan 2006).

Our argument at this point against strong objectivism is simple. Many scientific inferences depend on weighing the evidence for and against particular hypotheses and their alternatives. But such weighing of the evidence does not involve reasoning, whether probabilistic or deductive, from data premises to a hypothetical conclusion. It is therefore not, as the strong objectivist requires and the modest objectivist takes as worthy, 'logical'. On the one hand, it is (in the sense indicated) contextual in character. The piece of data provides evidence for a hypothesis over its rival relative to background information. Because of the contextual nature of this evidential relation

between data and hypotheses, strong evidential strength for one hypothesis over its rival does not imply that we have more reasons to believe that hypothesis for which we have strong evidence. The TB example convincingly shows that evidence is not 'logical' in the sense strong objectivists use the term. On the other hand, the ratio between hypotheses in terms of which evidence is characterized is not a probability measure. Therefore, the strong objectivist position must be rejected. From a slightly different point of view, the logical norm of the strong objectivist position precludes a distinction between confirmation and evidence. But failure to make this distinction leads directly to the old evidence problem. And unless the old evidence problem is resolved, Bayesianism generally fails as an adequate account of scientific inference.

4. The Old Evidence Problem and Its Diagnosis

Perhaps the most celebrated case in the history of science in which old data have been used to construct and vindicate a new theory concerns Einstein. He used Mercury's perihelion shift (M) to verify the general theory of relativity (GTR). The derivation of M is considered the strongest classical test for GTR. However, according to Clark Glymour's old evidence problem, Bayesianism fails to explain why M is regarded as evidence for GTR. For Einstein, $\Pr(M) = 1$ because M was known to be an anomaly for Newton's theory long before GTR came into being. But Einstein derived M from GTR; therefore, $\Pr(M|GTR) = 1$. Glymour contends that given equation (1), the conditional probability of GTR given M is therefore the same as the prior probability of GTR; hence, M cannot constitute evidence for GTR. Since Glymour agrees that the old evidence problem arises even when $\Pr(M) \approx 1$, we will take $\Pr(M) \approx 1$, but not equal to 1.

We will cite the passage where Glymour discusses the problem, and for the sake of clarity divide it into two parts, A and B.

- (A) 'The conditional probability of T [i.e., theory] on e [i.e., datum] is therefore the same as the prior probability of T '.
- (B) ' e cannot constitute evidence for T '. (Glymour 1980, 86)

Concerning (A), Glymour assumes that conditional and prior probabilities measure an agent's degrees of belief and that conditional probability prescribes a rule for belief revision. On the other hand, (B) states that e can't be regarded as evidence for T . We would argue that (A) pertains to the belief question, while (B) pertains to the evidence question and that confusion (in this case, paradox) only results when one conflates them.

Although some Bayesians would disagree, we hold that the standard Bayesian account of confirmation cannot solve the old evidence problem (Bandyopadhyay 2002).¹⁴ It suffices here to show how our account of evidence in terms of Bayes factor is able to solve the problem. Consider GTR and Newton's theory relative to M with different auxiliary assumptions for respective theories. Two reasonable background assumptions for GTR are (i) the mass of the earth is small in comparison with that of the sun so that the earth can be treated as a test body in the sun's gravitational field and,

(ii) the effects of the other planets on the earth's orbit are negligible. Let A_1 represent those assumptions.

For Newton, on the other hand, the auxiliary assumption is that there are no masses other than the known planets that could account for the perihelion shift. Let A_2 stand for Newton's assumption. When the Bayes factor is applied to this situation, there results a factor of infinity of evidential support for GTR against its rival. For, $\Pr(M|GTR \& A_1 \& B) \approx 1$, where $\Pr(M|Newton \& A_2 \& B) \approx 0$. Hence, the ratio of these two expressions goes to infinity as $\Pr(M|Newton \& A_2 \& B)$ goes to zero. The likelihood function does not satisfy the rules of the probability calculus.

Two comments are in order. The first concerns the criterion of evidence on which the account of the old evidence problem implicitly rests. The criterion of evidence Glymour exploits to generate the old evidence problem is that D is evidence for H if and only if $\Pr(H|B \& D) > \Pr(H|B)$, where the relevant probability is just the agent's current probability distribution as an ideal rational agent. According to the old evidence problem, the agent's belief in H should increase when the criterion holds, as shown by the theorem $BF > 1$ iff $\Pr(H|D) > \Pr(H)$. However, there is an ambiguity in the criterion. The criterion could mean two things. (i) There is a linear relationship between belief and evidence, that is, they increase or decrease in the same rate and it (the criterion) measures the same concept. Or (ii) there is a relationship between the two concepts, but if one increases or decreases, the other does not increase or decrease at a same rate and they are different concepts. The belief/evidence distinction rests on disambiguating these two notions (i from ii). In fact, the TB example elucidates this aspect of the distinction clearly by contending that it is the second sense (ii) that is implied by the belief/evidence distinction. This makes clear how the issue of what the agent should believe is correctly embedded in the criterion for evidence, although the criterion hides an ambiguity correctly diagnosed by the belief/evidence distinction.

The second comment concerns the implication of Glymour's evidence criterion regarding the relationship between evidence, knowledge, and belief. According to Glymour's old evidence problem, whenever the criterion of evidence is satisfied, D becomes evidence for H relative to an agent. D being evidence for H does not, however, turn out to be independent of whether the agent knows or believes D to be the case. By contrast, in our account whether D is evidence for H is independent of whether the agent believes or knows D to be true. In the TB example, D , the datum about the positive X-ray screening, provides evidence for H , the hypothesis about the presence of the disease, without the probability of D being one. So it is possible that D could be evidence for H without an agent knowing or believing D to be true.

5. The Unique Probability Problem

One purpose of objective Bayesianism is to develop an inductive logic which is observer-invariant. To achieve this aim, objective Bayesians impose constraints on the agent's beliefs. Agents are bound to agree with regard to their determination of probabilities when they hold the same background information, provided the agents satisfy the other constraints already imposed. Jaynes's recommendation is to choose a

probability distribution that maximizes entropy subject to the constraint of two agents having the same background information (Jaynes 1963; see also Seidenfeld 1987). Consider a fair six-sided die whose ‘bias’ to constrain our expectation for its next roll is taken to be:

$$E[\text{number of spots on next roll}] = 3.5 \quad (3)$$

The problem is to find a probability distribution for the set $X = \{1, \dots, 6\}$ representing possible outcomes of its roll. Shannon’s definition for the entropy, which is a measure of uncertainty in a discrete probability distribution, is:

$$U_s = -\sum_{i=1}^n Pr_i \cdot \log(Pr_i) \quad (4)$$

Jaynes’ maximizing entropy principle forces us to select the distribution over X ($Pr_i > 0$ when the probability sum of Pr_i will equal 1) which maximizes the quantity (4), subject to the constraint (3). If we satisfy the constraint, the distribution that we will arrive at is the uniform distribution, $Pr_i = 1/6$ when $i = 1, 2, 3 \dots 6$. For a different constraint on the expectation for the next roll of a die, however, his principle of maximizing entropy would choose another unique uniform probability distribution. This simple example¹⁵ has motivated the strong objectivist’s slogan that two agents holding the same background information will arrive at the same probability distribution.

One could hardly deny the force of the logic behind arriving at a unique probability distribution given the same background information for two agents with regard to this canonical example. However, there is plenty of qualitative information available in our day-to-day life which is too vague to be quantified. Keynes emphasizes this point when he writes:

[t]he statistical result is so attractive in its definiteness that it leads us to forget the more vague though more important considerations which may be, in a given particular case, within our knowledge. *To a stranger the probability that I shall send a letter to the post unstamped may be derived from the statistics of the Post Office; for me those figures would have but the slightest bearing on the question.* (Keynes 1921, 322; emphasis added)

For a strong objectivist the question is, how could she sharpen this vague knowledge to arrive at a knowledge which is quantitative and to which probability theory can be applied? Williamson suggests that some interval-based approach with bounds should be able to dispel Keynes’s worry. He writes, ‘I know that I am scattier than the average member of the populace, but I do not know quantitatively how much scattier, so this knowledge only constrains my degree of belief that I have posted a letter unstamped to lie between the Post Office average and 1’ (Williamson 2005). He is aware that he might have a different mental constitution than being a scatter-brained philosopher so that he has realized that in over 90% of the cases, the addressee received the letter. Being careful of this statistic, his degree of belief would fall anywhere between the Post Office average and 1. This is a suggestion as to how an objective Bayesian could respond to Keynes.

However, the bounds of these intervals could vary depending on the nature of an agent, leading to some subjectivity in the probability assignments. Williamson is aware of this problem. He thinks that the strong objectivist demand that the agent's degree of belief be determined on the basis of her background knowledge overcomes it. But without further elaboration, he does not show how an interval-based probability will give us the sort of numbers needed to reach a unique final probability assignment.

There are further problems for strong objectivism. Consider the slogan of objective Bayesianism, two agents must agree with regard to their determination of probabilities when they share the same background information. The problem is that it is hard to pin down whether or when two agents must agree with regard to a matter. What does it mean to say, in a non-trivial way, that two agents share the same background information? An answer to this question is crucial for strong objective Bayesianism; otherwise the slogan begs the question.

Consider two sets of agents. The first set of agents holds the same background information including the belief that simpler theories should be assigned high prior probabilities. Here the former takes simplicity to consist in paucity of adjustable parameters. In contrast, the second set of agents holds the same background information including the fine grained belief that the simplest theory which is to be counted as having the least number of adjustable parameters should be assigned prior probability $1/2$. Then, the second simplest theory should receive $1/4$, followed by the more complex theory $1/8$ and so on. Does the first set of agents share the same background information? Does the second? One needs to have a non-arbitrary way of determining which set of agents share the same background information. If an objective Bayesian remains vague as to what counts as two agents having the same background information, then the first set of agents might differ markedly with regard to the values to be assigned the 'simplest' hypothesis. And so on.

The problem is that having the same background information seems to be neither a necessary nor a sufficient condition for what counts as two agents having the same background information. On the one hand, it is logically possible for two agents to have the same probability for each and every hypothesis although they might not share the same background information. Of course, this is very unlikely. But we know that in individual cases people assigned the same (low) prior, say the probability of Barack Obama winning the US presidency two years ago, on the basis of very different information.

On the other hand, let's assume that there is a way to characterize what counts as two agents having the same background information without further ado. Yet one could argue that having the same background information for two agents does not necessarily ensure the uniqueness of probability assignments. Suppose two agents could share some sort of Jeffreys' simplicity postulate along with the same kind of background information, yet still assign different priors to the three competing hypotheses in the curve-fitting problem. The first one chooses the first set of priors from Table 1. For example, $1/2$, $1/4$, $1/8$ to H_1 , H_2 , and H_3 respectively and so on. The second one could choose the second set of priors from the same table for those models and so on. It could be that their mental make-up is different; that is why they come up with different

priors. It is, therefore, possible for these two agents to come up with two distinct consistent sets of priors, even though they might have the same background information.

The basic point remains the same: either ‘sameness of background information’ is characterized in a question-begging way or it is so vague as to admit of a number of different possibilities, at least some of which allow for the assigning of very different prior probabilities.

There are thus four interrelated problems. The first problem arises from an inability to quantify many of our day-to-day intuitions. The second is associated with the arbitrariness involved in characterizing the notion of two agents having the same background information. The third stems from difficulties in providing a sufficient condition for when two agents share the same background information. The fourth and final problem is the possibility of failure in arriving at a unique probability distribution in model selection, even though two agents are assumed to share the same background information. Together, they undermine the second dogma of strong objective Bayesianism, that sameness of background information will insure uniqueness of probability distributions.

6. A Strong Objective Bayesian Rejoinder

We considered two objections to strong objectivists. The first one is that since they take all statistical inferences to be logical, they fail to account for the difference between the belief and evidence questions; the evidence measure is not ‘logical’ in the strong objectivist’s sense. The second one is that strong objectivists ensure the unique probability of a conclusion given a set of premises only by appealing in a problematic way to the notion of ‘same background information’.

One response to the first objection is that to say that it is at best misleading. Evidence is equivalently to be measured by the posterior-prior ratio $\Pr(H|D)/\Pr(H)$ (call it PPR) or the posterior-prior difference $\Pr(H|D) - \Pr(H)$, and therefore, it is incorrect to contend that strong objectivists could not handle the evidence question while they could handle the belief question. In reply, we would like to say that there is some truth to this rejoinder, because in some cases, the posterior-prior ratio could yield the same quantitative evidential support as provided by the BF/LR measure. In the TB case, there is an agreement between the number yielded by PPR and BF, that is, 25.7 times. There is a theorem, $\text{BF/LR} = \Pr(D|H)/\Pr(D|\sim H) = \Pr(H|D)/\Pr(H) \times \Pr(D)/\Pr(D|\sim H)$, which shows this connection. When $\Pr(D)/\Pr(D|\sim H) = 1$, both the BF and the PPR yield the same numerical value. In the TB case this means that $\Pr(D)/\Pr(D|\sim H) = 0.628/0.628 = 1$. Although it might be difficult to get an intuitive grasp of $\Pr(D)/\Pr(D|\sim H)$, another theorem, $\Pr(D)/\Pr(D|\sim H) = 1 = \Pr(\sim H)$ or $\text{BF} = 1$, in fact helps to provide an intuitive understanding of this condition. If $\Pr(D)/\Pr(D|\sim H) = 1$, then $\Pr(D|H) = \Pr(D|\sim H)$ and each of these is equal to $\Pr(D)$. If the probability of D does not depend on H , then D and H are statistically independent. This merely says that observing D tells us nothing about how likely H or $\sim H$ is. That is, when D and H are statistically independent, $\text{BF/LR} = 1$.¹⁶

But the PPR and the BF/LR measure don't express the same concept. The former measure provides an answer to the question, 'how likely is the hypothesis when D is learned as compared to how likely the hypothesis is by itself'. In contrast, the latter measure provides an answer to the question, 'how likely are the data when the hypothesis is learned as compared to how likely are the data when the competing hypothesis is learned?' The objector could further show the connection between evidence and belief by referring to a straightforward theorem, $\Pr(D|H) > \Pr(D)$ iff $\Pr(H|D) > \Pr(H)$.

Our answer will be in the same spirit as before and point out that they express different concepts; the former (the left-hand side of the equivalence) measures 'evidence' (which is, uncontroversially, agent-independent) and the latter (the right hand side of the equivalence) measures 'belief' (independent of however objectively priors are determined), still agent-dependent in the sense that it depends on an agent's degree of belief). Finally, the basic point remains that both the posterior-prior ratio and the BF/LR are ratios, and do not very well fit the strong objectivist demand that all scientific inference is 'provable' or 'probable'.

One possible response to our second objection, that sameness of background information does not force different agents to arrive at one unique probability for a conclusion is that it not a genuine problem for strong objectivism. The responder argues that it is the *goal* of strong objectivism to arrive at such probability under suitable conditions (Williamson 2007), rather than a description of what agents do in a real world situation. The rejoinder adds that if background knowledge is the same, then all degrees of belief are the same in the sense that the latter converges if the background knowledge converges. Some objectivists in the Carnapian tradition similarly ensure convergence by imposing something like exchangeability on an agent's beliefs. A sequence of exchangeable events is a sequence in which the probability of any particular distribution of a given property in a finite set of events of the sequence depends on the number of events in that distribution that have that property. If strong objectivists assume something like exchangeability to arrive at a unique probability, then the further question arises, 'how does a strong objectivist differ from a personalist Bayesian who subscribes to de Finetti's representational theorem?' (de Finetti 1980 [1937]; Kyburg 1983). The idea behind this theorem is that if two people agree that events of a certain sequence are exchangeable, then with whatever opinions they start, given a long enough segment of the sequence as data, they will come arbitrarily close to agreement regarding the probability of future events, provided that neither of the agents assign zero or one probability to an empirical proposition.

However, there are numerous discussions about the tenability of the exchangeability assumption. Some think that it is a controversial assumption, because many sequences are not exchangeable where the order in which the agent receives information plays a crucial role. If a Bayesian personalist defends the inter-subjectivity of scientific inference by invoking de Finetti's theorem, then the former can be criticized for holding a strong *a priori* assumption about the world. And there is no reason to suppose that the world is like that. It seems that if a strong objectivist imposes such an assumption as this to arrive at convergence of agent's degree of belief, then her assumption is as strong as a subjectivist's assumption of a sequence of events being exchangeable. Hence, there

is no difference between a strong objective Bayesianism and subjective Bayesian insofar as both fall back on convergence and the same sort of constraints on an agent's degrees of belief.

7. Bayesianism without the Dogmas¹⁷

What constraint(s) need one impose on a non-deductive theory of inference so that an agent is able to reach a correct conclusion most of the time?¹⁸ Like most people, we think that this question, posed in just these terms, is irresolvable. A more tractable, and certainly useful, question is whether there are weaker constraints than those imposed by strong objectivists such that resulting scientific inferences, although not invariably or even for the most part 'correct', are in some plausible sense of the term 'objective'.

Personalist Bayesians like Bruno de Finetti (1980 [1937]) and Leonard J. Savage (1972) claim that there is only one such condition, coherence. This condition is clearly necessary; coherence precludes sure loss, and sure loss is incompatible with rationality and in a plausible sense with 'objectivity'. But this requirement is too easy to satisfy; there are many inferences which satisfy it whose conclusions at least initially conflict and which converge over the long run only if some frequently untenable assumptions are made. Strong objectivists like Jaynes, Rosenkrantz, and Williamson swing to the other end of the spectrum and impose conditions that insure agreement on probability distributions from the outset. But their conditions can't plausibly be necessary; they disallow, as in the case of weighing the evidence for and against particular hypotheses, otherwise accepted forms of scientific reasoning.

The option, we suggest, is to find a 'middle way' between personalism and strong objectivism. We label it 'quasi-objective Bayesianism'. Quasi-objective Bayesianism, recall, is a theory of non-deductive inference which, in addition to coherence imposes simplicity considerations on the assignment of prior probabilities (typically within a context provided by a range of alternative hypotheses). These simplicity considerations are, at least in part, 'objective' in character. But they do not, in most cases, lead to the assignment of unique probabilities to the hypotheses being tested. Most often, we can do no more than rank order the hypotheses with respect to their priors. But this more sensible (if not also moderate) goal is, we believe, closest to the actual practice of science than its Bayesian rivals. Despite the vast influence exercised by Euclid's *Elements* on subsequent scientific methodology, only rarely are the inferences made in science 'logical' in the sense that data premises support a hypothetical conclusion in ways conforming to deductive or probability logic. The evidence for a hypothesis is never in this same sense 'conclusive' hypotheses are always tested in a context involving other hypotheses (whose unique partition can never be guaranteed from the outset) and a variety of epistemic and pragmatic assumptions, and the result most frequently is acceptance of one on the basis of the judgement that it is 'better' than the others (which, again, are determined contextually). Unique probability assignments are the exception, not the rule, no matter how plentiful and persuasive the data that we might at a given time have in hand. Only a mistaken assimilation of 'objectivity' with 'truth',

or even more implausibly with ‘certainty’, would lead it to be understood as a methodological requirement.¹⁹

Inside this methodological debate over the correct form of Bayesianism there is also hidden an epistemological message. An entrenched belief among both philosophers and statisticians is that a Bayesian has to be either a subjective or objective, and there is nothing in between.²⁰ The epistemological message aims to challenge this conventional wisdom. Bayesianism is not an all or nothing affair. As Bayesians take belief to be fine-grained, we argue, Bayesianism itself needs to be fine-grained too. As such, the subjective/objective distinction does not have a sharp boundary in Bayesianism. Bayesians may come in different shades and shapes in between these two extremes. We distinguish among four varieties of objective Bayesians.²¹ They are (i) strong objectivists, (ii) moderate objectivists, (iii) quasi-objectivists and (iv) and finally, necessitarian objectivists. While the strong objectivist lies at the extreme end of the spectrum, the quasi-objectivist leans more toward the personalist or subjectivist side. In a similar vein, necessitarian objectivists come closer to strong objectivists or moderate objectivists, depending on which Carnap (early or late) one has in mind. However, whether the lack of a sharp distinction between subjectivism and objectivism transcends Bayesianism and affects epistemology in general remains a desideratum for further investigation.

Acknowledgements

We would like to acknowledge the help given by us by John Bennett, José Bernardo, Robert Boik, Dan Flory, Jack Gilchrist, Megan Higgs, Tasneem Sattar, Teddy Seidenfeld, and C. Andy Tsao for their comments related to several parts of the paper. An earlier version of the paper was presented at the ‘Objective Bayesianism’ conference held in the London School of Economics and Political Science in 2005. Both authors very much appreciate the comments they received there. They are especially thankful both to two anonymous referees for their insightful comments regarding the contents of the paper and to the editor of this journal for his helpful suggestions. We also thank Jan-Willem Romeijn, John Welch, and Gregory Wheeler for their encouragement regarding the paper. Special thanks owe to John Bennett for his several e-mails containing suggestions about the paper. The research for the paper has been financially supported both by our university’s NASA’s Astrobiology Center (Grant number 4w1781) with which both of us are associated and the Research and Creativity grant from our university.

Notes

- [1] Williamson (2005) is presently the most vocal proponent of this position. See also Williamson (2007, 2009).
- [2] Here we borrow the term ‘necessitarian’ from Levi (1980) who credits Savage with this coinage. For Levi’s position on objective necessitarianism see Levi (1980), especially chapters 15–17.
- [3] Jaynes also expressed the same idea in his earlier writing: ‘the most elementary requirement of consistency demands that two persons with the same relevant prior information should assign

the same prior probabilities' (Jaynes 1968). Interestingly, Howson (Howson and Urbach 1993), who is a subjective Bayesian, accepts the first thesis of the strong objective Bayesian, but rejects her second thesis.

- [4] In this respect we agree with Fisher's judgement that '[a]lthough some uncertain inferences can be rigorously expressed in terms of mathematical probability, it does not follow that mathematical probability is an adequate concept for the rigorous expression of *uncertain inferences of every kind*' (Fisher 1956, 40; emphasis added).
- [5] Or robots, perforce 'rational' on Jaynes's account.
- [6] Our discussion is not intended to imply that there is no difference among these three objectivists. Although both Rosenkrantz (1977, ix and 254) and Williamson (2005, ch. 11) discuss why the probability theory is one significant tool for understanding objective Bayesianism, they are not as emphatic as Jaynes is to take the probability logic as an extension of deductive logic and then contend that this is the only way to approach scientific inference. However, what matters for our purpose is a general agreement among strong objective Bayesians about these two theses earlier attributed to them. We owe this point to one of the referees of this journal.
- [7] The information matrix is the negative expected value of the second derivative of the log-likelihood function with respect to the parameter.
- [8] Some of these ideas go back to Hartigan (1983). For details of a recent, very general mathematical proof of this fact, see Berger, Bernardo, and Sun (2009).
- [9] Jeffreys (1961) has sometimes been considered an objective Bayesian. See Bandyopadhyay (2007) for a discussion of Jeffreys's views on scientific inference.
- [10] For a nice historical discussion on the subjective and objective distinction in the probability theory dating back to 18 and 19th centuries please consult Zabell, 2010.
- [11] A Bayesian of our kind should very well be able to construe an evidence relation to be independent of what the agent believes about those two competing hypothesis. The present-day subjective-Bayesian-dominated philosophy of science is, however, slow to appreciate this point. For a subjective Bayesian, since the evidence relation is defined in terms of a ratio of two probabilities, those probabilities have to be understood in terms of an agent's degree of belief. Therefore, according to her, the evidence relation must be subjective. However the situation need not be like that. A subjective Bayesian-flavoured textbook on econometrics explains this point well and discusses how an objective Bayesian could approach the likelihood function which is at the core of our evidential account: '[t]he likelihood function was simply the *data generating process* (DGP) viewed as a function of the unknown parameters, and it was thought of having an "objective existence" in the sense of a meaningful repetitive experiment. The parameters themselves, according to this view, also had an objective existence independent of the DGP. ... It is possible to adopt the so-called *objective Bayesian* viewpoint in which the parameters and the likelihood maintain their objective existence ...' (Poirier 1995, 288–289).
- [12] For some recent review of the literature on Bayes factor see Ghosh, Delampady, and Samanta (2006, chapter 6). This book also contains some interesting discussion and evaluation of objective Bayesianism.
- [13] This way of reconstructing the example owes a great deal to the help we received from James Hawthorne, Teddy Seidenfeld, Peter Gerdes, and Robert Boik. These names are arranged in the order we received their help.
- [14] We owe this point to one of the referees. We thank the editor for suggesting how to formulate the point of disagreement some Bayesians might have with us at this point.
- [15] Our comments on entropy cannot properly be generalized to continuous variables. However, the purpose of this example addressing entropy needs to be taken as an illustrative case for explaining the intuition behind strong objectivism, and not in the spirit whether this example is generalizable. As our discussion of the example explaining entropy holds in discrete cases, this example serves its intuitive function.

- [16] We owe this point of clarification to Robert Boik.
- [17] Both the title of this section and the title of the paper have been motivated by Quine's celebrated article "Two Dogmas of Empiricism" (Quine, 1953).
- [18] One of the referees has raised worries about this question.
- [19] One possible objection to our version of Bayesianism is that the rejection of the both strong objectivist's theses can't satisfactorily be combined with quasi-Bayesianism. According to the objection, in the case of the confirmation/evidence accounts, we assumed simple statistical hypotheses, whereas for the possibility of allowing infinitely many possible posterior probabilities for competing hypotheses we assumed statistical hypotheses to be complex. As we know, values of simple statistical hypotheses are independent of an agent's degree of belief, but values of complex hypotheses are dependent on an agent's degree of belief. Thus, the objector concludes our version is inadequate to be unified under one Bayesian framework. Our response is that both points, (i) assuming simple statistical hypotheses for the confirmation/evidence accounts and (ii) taking complex statistical hypotheses to be dependent on prior beliefs, are consistently accommodated within a Bayesian framework while keeping two issues distinct. Quasi-Bayesianism acts like a Swiss army knife that has various functions. When we are called upon to perform a specific function we apply a specific aspect of our version to address it. Our version is far from the idea of 'one size fits all' philosophy. Whether other statistical schools have that flexibility is beyond the scope of the present paper.
- [20] This belief is current among both Bayesian and non-Bayesian philosophers: see Swinburne (2002) and Sober (2002). Although Swinburne is a Bayesian and Sober is a non-Bayesian, they share the view about a sharp division into subjective and objective camps. For a sample of this rampant oversimplification among statisticians, look at any of the statistical literature on objective Bayesianism listed in the references.
- [21] Interestingly, Good (1983) has some similarities with what we have been arguing here. Like Good, we are arguing that there are several varieties of Bayesianism. However, unlike him, we don't have any specific number regarding how many such varieties exist or are possible.

References

- Bandyopadhyay, P. S. 2002. The old evidence problem and beyond. Paper presented at the American Philosophical Association, Eastern Division Meeting, Philadelphia, Dec.
- . 2007. Why Bayesianism? A primer on a probabilistic philosophy of science. In *Bayesian statistics and its applications*, edited by S. K. Upadhyay, U. Sing, and D. K. Dey, 42–62. New Delhi: Anamaya Publishers.
- Bandyopadhyay, P. S., and R. J. Boik. 1999. The curve-fitting problem: A Bayesian rejoinder. *Philosophy of Science* 66 (suppl.): S391–S402.
- Bandyopadhyay, P. S., R. J. Boik, and P. Basu. 1996. The curve-fitting problem: A Bayesian approach. *Philosophy of Science* 63 (suppl.): S264–S272.
- Bandyopadhyay, P. S., and B. Brittan. 2001. Logical consequence and beyond: A look at model selection in statistics. In *Logical consequence*, edited by J. Wood and B. Hepburn, 1–17. Oxford: Hermes Science.
- . 2006. Acceptance, evidence, and severity. *Synthese* 148: 259–293.
- Berger, J. O. 1985. *Statistical decision theory and Bayesian analysis*. 2nd ed. New York: Springer.
- Berger, J. O., J. M. Bernardo, and D. Sun. 2009. The formal definition of reference priors. *Annals of Statistics* 37: 905–938.
- Bernardo, J. M. 1997. Non-informative priors do not exist: A dialogue with José Bernardo. *Journal of Statistical Planning and Inference* 65: 157–189.
- . 2005. Reference analysis. In *Handbook of statistics*, edited by D. K. Dey and C. R. Rao, 17–90. Amsterdam: Elsevier.
- Bernardo, J. M., and A. Smith. 1994. *Bayesian theory*. New York: Wiley.

- Carnap, R. 1950. *Logical foundations of probability*. London: Routledge & Kegan Paul.
- . 1971. A basic system of inductive logic, part 1. In *Studies in inductive logic and probability*, edited by R. Carnap and R. Jeffrey, vol. 1, 33–165. Berkeley: University of California Press.
- de Finetti, B. 1980 [1937]. Foresight, its logical laws, its subjective sources. In *Studies in subjective probability*, edited by H. Kyburg and H. Smokler, 53–118. Huntington, NY: Robert E. Krieger.
- Fisher, R. 1956. *Statistical methods and scientific inference*. New York: Hafner Press.
- Ghosh, J., M. Delampady, and T. Samanta. 2006. *An introduction to Bayesian analysis: Theory and methods*. New York: Springer.
- Glymour, C. 1980. *Theory and evidence*. Princeton, NJ: Princeton University Press.
- Good, I. J. 1983. 46656 varieties of Bayesians. Rpr. in I. J. Good, *Good thinking: The foundations of probability and its applications*. Minneapolis: University of Minnesota Press.
- Hartigan, J. A. 1983. *Bayes theory*. New York: Springer.
- Howson, C., and P. Urbach. 1993. *Scientific reasoning: The Bayesian approach*. 2nd ed. La Salle, IL: Open Court.
- Jaynes, E. T. 1963. Information theory and statistical mechanics. In *1962 Brandeis Summer Institute in Theoretical Physics*, edited by K. Ford, 181–218. New York: Benjamin.
- . 1968. Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics* 4: 227–241.
- . 2003. *Probability theory: The logic of science*. Cambridge: Cambridge University Press.
- Jeffreys, H. 1961. *Theory of probability*. 3rd ed. Oxford: Clarendon Press.
- Keynes, J. M. 1921. *A treatise on probability*. London: Macmillan.
- Kyburg, H. E. 1983. *Epistemology and inference*. Minneapolis: University of Minnesota Press.
- Levi, I. 1980. *The enterprise of knowledge*. Cambridge, MA: MIT Press.
- Pagano, M., and K. Gauvreau. 2000. *Principles of biostatistics*. Pacific Grove, CA: Duxbury Thompson Learning.
- Poirier, D. 1995. *Intermediate statistics and econometrics: A comparative approach*. Cambridge, MA: MIT Press.
- Quine, W. V. O. 1953. Two dogmas of empiricism. In *From a logical point of view*. New York: Harper and Row.
- Rosenkrantz, R. 1977. *Inference, method, and decision: Towards a Bayesian philosophy of science*. Dordrecht: D. Reidel.
- . 1981. *Foundations and applications of inductive probability*. Atascadero, CA: Ridgeview.
- Royall, R. M. 1997. *Statistical evidence: A likelihood paradigm*. New York: Chapman & Hall.
- Savage, L. J. 1972. *The foundations of statistics*. 2nd ed. New York: Dover.
- Seidenfeld, T. 1987. Entropy and uncertainty. In *Foundations of statistical inference*, edited by I. B. MacNeill and G. J. Umphrey, 259–287. Dordrecht: D. Reidel.
- Sober, E. 2002. Bayesianism—Its scope and limits. In *Bayes's theorem*, edited by R. Swinburne, 21–38. Oxford: Oxford University Press.
- Swinburne, R. 2002. Introduction. In *Bayes's theorem*, edited by R. Swinburne, 1–20. Oxford: Oxford University Press.
- Williamson, J. 2005. *Bayesian nets and causality*. Oxford: Oxford University Press.
- . 2007. Motivating objective Bayesianism: From empirical constraints to objective probabilities. In *Probability and inference: Essays in honour of Henry E. Kyburg Jr*, edited by W. L. Harper and G. R. Wheeler, 155–183. London: College Publications.
- . 2009. Philosophies of probability. In *Handbook of the philosophy of mathematics*, edited by A. D. Irvine, 493–533. Amsterdam: North Holland.
- Zabell, S. L. 2010. The subjective and the objective. In *Handbook of the Philosophy of Statistics*, edited by P. S. Bandyopadhyay and M. Forster. North Holland: Elsevier.