Introduction

This Element introduces and critically explores the Bayesian approach to the logic and epistemology of scientific reasoning. This approach is distinguished by its conceptual *medium* and formal *method*. Bayesians paint in the medium of what has variously been called "degree of belief," "credence," "confidence," and so on – which is to say that Bayesian accounts essentially involve a gradational notion of the doxastic attitude that agents have toward propositions. Some common terminology notwithstanding (e.g., "partial belief" and "degree of belief"), this concept is not a straightforward, gradational version of *belief*, which would be referring to something like proportion of outright belief. Such a notion would "max out" in the case of belief simpliciter, but Bayesians tend to be very quick in dismissing the identification of maximal degree of belief, credence, and so on with qualitative belief (Maher, 1993; Leitgeb, 2013; Buchak, 2013). Much more commonly, the Bayesian's gradational medium is described as maxing out in the special case of *certainty* (e.g., Ramsey 1926, §3; Jaynes 2003; Jeffrey 2004; Leitgeb 2013; Sprenger and Hartmann 2019, p. 26). Certainty is a doxastic, propositional attitude of agents generalizing naturally to less extreme attitudes of *uncertainty* or *confidence*. Accordingly, in an attempt to avoid confusion in this Element, I use these latter terms when referring to the concept at the core of Bayesian accounts.[1]

Regarding method, Bayesians conduct their investigations using the mathematical tools of probability. This formal approach enables Bayesians to pursue their philosophical work with rigor and precision. The use of probability theory goes hand in hand with Bayesianism's focus on confidence; indeed, Bayesians view the probability calculus as an apt formal tool because of their emphasis on confidence and uncertainty. The bridge between medium and method here is the Bayesian's epistemic interpretation of probabilities as degrees of (more or less operationalized and/or idealized conceptions of) confidence.

Terms of art in probability theory – like "probable" and "likely" – are often used in everyday language to communicate epistemic judgments of uncertainty. However, the Bayesian interpretation is far from uncontroversial. Probability theory, mathematically speaking, is a branch of measure theory, complete with an axiomatization and consequent structure (Kolmogorov, 1933). While the epistemic interpretation of probability as a guide to rational confidence was there from the beginning of this math's development, much of this development was driven by more objective, physical, and aleatory concepts, applications,

[1] The use of "partial belief" and "degree of belief" to refer to Bayesianism's central notion can indeed lead to confusion and criticisms of Bayesianism that are rather too easy. For example, see Horgan's (2017, p. 236) "conceptual confusion" objection to Bayesianism.

and interpretations (see Hacking 2006, ch. 2). The mathematics of probability was at best developed only in part as an explication of *confidence*. Thus, a case needs to be made for the validity and usefulness of the Bayesian's interpretation.

This Element explores the Bayesian approach to logic and epistemology in three parts. Section 1 provides a primer on the elementary mathematics of probability, motivated and presented through a Bayesian lens. We develop probability theory as a conservative generalization of classical logic, which is more readily applicable to reasoning under uncertainty. Probability theory provides a *logic of consistency* for attitudes of confidence. That is, probability theory describes how an agent's confidences (at a particular time) ought to relate in order for them to be internally consistent. Additionally, this section discusses the Bayesian interpretation of probability as a measure of confidence, and it critically evaluates some of the common arguments presented for and against this interpretation.

Section 2 explores a number of contentious principles put forward and debated by Bayesians. Unlike the rules discussed in Section 1, these principles are more characteristically and recognizably epistemological. This is because they go beyond the logic of consistency in theorizing about how our confidences ought to be, not just related to each other, but sensitive to and constrained by our experience of the world. For direction in our exploration, we turn to the patron saint of Bayesian epistemology, the good Reverend Thomas Bayes, and his seminal "Essay Towards Solving a Problem in the Doctrine of Chances" (1763). Bayes suggests three epistemological principles in this work, each of which is still much discussed, developed, and debated by contemporary Bayesian epistemologists. We discuss these principles in some detail along with some corresponding criticisms and complications that lie in wait for the Bayesian epistemologist.

The final Section 3 displays the potential fruitfulness of the Bayesian approach for the study of scientific reasoning. We introduce just a handful of the many topics Bayesians have discussed in the epistemology of science: specifically, we focus on the epistemology of confirmation, explanatory reasoning, evidential diversity and robustness analysis, hypothesis competition, and Ockham's Razor. Via our discussions of these topics, we aim to show that our understanding of some important concepts, methods, and strategies commonly used in scientific practice can be improved by taking a Bayesian approach.

For the sake of keeping this work Element-length, I've often had to refrain from delving into interesting issues and obvious questions left standing by my presentations. At one level, this necessity has bothered me, since it opens me up to criticisms that I might have tried to preempt and limits me to simplistic overviews of certain topics. However, at another level, I think and hope that the result is more conducive to promoting further discussion (be it in classrooms or research settings), and indeed to inspiring a wider variety of future research on the relevant issues.

I first learned about Bayesianism through a series of courses I took from Timothy McGrew as an MA student at Western Michigan University (from 2004 to 2006). To this day, I owe Tim an enormous debt of gratitude, not only for introducing me to an enjoyable and fascinating field of study, but also for being such a patient, caring, and careful instructor. The "dartboard representations" I employ in Section 1 trace their roots to similar diagrams Tim developed and used when introducing me to the field. My deepest thanks additionally go out to several colleagues and students who willingly and graciously gave their time to read drafts of this work and provide me with feedback. Jonathan Livengood and Joshua Barthuly in particular each carefully read and provided substantial feedback on complete drafts of this Element; this work is much better because of their gracious help. Other readers who gave me invaluable feedback on portions of the text include Jean Berroa, Liam Egan, Samuel Fletcher, Konstantin Genin, David H. Glass, Qining (Tim) Guo, Robert Hartzell, Daniel Malinsky, Conor Mayo-Wilson, Lydia McGrew, Jacob Stegenga, Michael Titelbaum, Jon Williamson, and two anonymous reviewers for Cambridge University Press. The material for §3.3.2 and a large part of §3.4 was developed in collaboration with David H. Glass. My heartfelt gratitude also goes out to Abbot Silouan and all of the monks at the Monastery of the Holy Archangel Michael (Cañones, New Mexico) for providing me with gracious hospitality, good conversation, and the unimaginably peaceful environment of their guesthouse, where I wrote the bulk of this Element.

While working on this Element, I was supported by the Charles H. Monson Mid-Career Award administered by the University of Utah's Philosophy Department and by grant #61115 from the John Templeton Foundation (which I co-directed with David H. Glass). The opinions expressed in this Element are those of the author and do not necessarily reflect the views of either of these funding sources.

Finally, "thank you" is not enough when I try to imagine how I should respond to the support, love, and joy that I receive everyday from my wife and children. I don't deserve such blessings. Truth be told, my confidence is not very high that any one of them will ever read the entirety of this Element. Nonetheless, it wouldn't have happened were it not for them, and I dedicate it to them.

1 Probability Theory, a Logic of Consistency

Bayesianism is characterized by its focus on the notion of confidence ("degree of belief," "credence," "partial belief," etc.) and its corresponding epistemic interpretation of probability theory.[2] Bayesians apply the probability calculus, so interpreted, to a wide variety of epistemological principles, concepts, and puzzles. This section introduces the mathematics of probability theory through a Bayesian lens. By first touching on some features of classical, deductive logic, we highlight the need to generalize this formal logic in order to deal directly with the inevitable uncertainties of scientific (and everyday) inferences. The probability calculus, when given a Bayesian interpretation, provides a particularly appealing generalization, resulting in a compelling logic of consistency for the confidences of uncertain agents. We finish this section by considering some arguments for and against this way of construing probability theory.

1.1 Logic and Uncertainty

Logic is the science of inference. This discipline systematically studies the relation by which conclusions follow from premises. As such, while logic is fundamental to epistemology, the two disciplines are distinct. Epistemology would surely be lacking if it gave no role to inference in the theory of knowledge, but there are other factors besides inference that play crucial epistemological roles.[3] Logic is also distinct from another science of inference, the psychology of inference, which observes, predicts, and models the inferential behavior of humans actually drawing conclusions from premises.[4] Logic, by contrast, is more of a theoretical science. It theorizes about the existence and nature of general principles relating conclusions and premises in such a way as to legitimize, guide, and regulate such behavior. These principles, which the renowned logician George Boole (1854) called the "laws of thought," concern the fundamental nature of consequence, or what follows from what.

[2] Bayesianism is often additionally characterized as being committed to **Bayes's Rule**, a principle legislating how confidences (explicated as probabilities) should evolve over a diachronic process of learning new information. We postpone our discussion of this principle until Section 2.

[3] There are other conceptions of logic, some of which encompass epistemology. For example, Ramsey (1926, p. 87) views logic as consisting of two parts, "formal logic" or the explicative logic of consistency, and "inductive logic" or the ampliative logic of discovery and truth. I accept Ramsey's distinction but am trying to keep terminology tidy by reserving "logic" for Ramsey's notion of formal logic. What Ramsey calls "inductive logic" is part of epistemology as I characterize it – cf. Jeffreys's (1939, p. 1) remark: "The theory of learning in general is the branch of logic known as epistemology."

[4] The wisdom of distinguishing the logical from the psychological may be questioned, however. Kimhi (2018), for example, argues that the sundering of the two is a recent, Fregean development that has had detrimental philosophical effects.

Logic is also sometimes thought of as the study of *valid* inference, in which the truth of a target conclusion follows *inescapably* from the collective truth of a set of corresponding premises (i.e., such that the conclusion cannot possibly be false if the premises are all true).[5] But while this seems a fair characterization of classical, deductive logic, the science-of-inference and study-of-valid-inference conceptions of logic may come apart for at least two related reasons. First, some inferences are logical in the sense that their conclusions follow from corresponding premises, even though they don't follow as a matter of necessity (i.e., even though the inferences aren't valid). Second, it may be useful to examine a notion of inference that relates propositions with respect to something other than their truth values. Both of these considerations underlie the approach to *inductive* logic that we take in this Element.

At the heart of both considerations is the attempt to accommodate the ubiquitous presence of uncertainty in scientific reasoning (and human reasoning more generally). Apart from some very special contexts in which truth values may legitimately be directly known with certainty, or stipulated, or constructed, and so on, reasoners are plagued by uncertainty. They simply do not have omniscient access to the truth values of propositions. What they have instead are inferior surrogate attitudes toward propositions, including attitudes of confidence and uncertainty.

Knowing how propositions relate to each other in terms of their truth values can still be very useful information for an uncertain agent. However, an alternative, inductive approach to logic seeks to shed light more directly on how propositions relate to each other in terms of the uncertain attitudes agents actually have toward them.[6] Instead of studying what truths follow from other truths, this approach studies what confidences, expectations, uncertainties, and the like follow from other such attitudes. And on this approach, it may be that some *invalid* inferences are nonetheless good insofar as (and possibly to the extent that) their conclusions are made more nearly certain by their corresponding premises. We seek a logic better suited for mortals,

[5] While some writers use the term "valid" to apply to good inductive arguments (e.g., Priest 2006; Sprenger and Hartmann 2019), in this Element, I reserve the term strictly for the notion of deductive validity.

[6] Alternatively, one might think of inductive logic as relating propositions in terms of their truth values, while generalizing the notion of *consequence* itself. From this perspective, inductive logic is still about what truths follow from other truths; however, the salient notion of consequence is generalized to allow for degrees of "partial entailment" and/or non-monotonic inference. Bolzano's (1837) notion of "relative satisfiability" is an early example of this perspective (see Howson 2011, §2). Other examples include Wittgenstein's interpretation of probability in the *Tractatus* (1922, §5.15) (see Williamson 2017, §1.1), "logical interpretations" of probability in terms of "partial entailment" (Keynes, 1921; Jeffreys, 1939; Carnap, 1962), and Priest's (2006, pp. 189–190) account of "inductive validity" using a non-monotonic logic.

a logic of uncertain inference for agents like us who don't have omniscient access to truth values.

Scientific practice provides any number of instances of compelling arguments in contexts of uncertainty. For example, in *On the Heavens* Aristotle builds a case from multiple lines of evidence for accepting the sphericity of the earth. His arguments provide a compelling interplay of observational evidence and inference.[7] He cites the evidence of the earth's shadow's circular shape as observed during a lunar eclipse. He then infers from this evidence and from his (accurate) understanding of lunar eclipses that the earth is a sphere, since that shape would account for this observation, given that theory. Similarly, Aristotle records observations of an "alteration of the horizon" relative to the fixed stars effected by northward or southward travel. He then infers from this evidence that the earth is a sphere, since again that shape would account for differences in the visible stars depending on an observer's latitude.[8] Distilled into a more organized form, the arguments are as follows:

Argument 1:

A1. If the earth is spherical, it would cast a circular shadow.
A2. The earth casts a circular shadow.
C. Thus, the earth is spherical.

Argument 2:

B1. If the earth is spherical, northward and southward travel would alter the range of visible stars.
B2. Northward and southward travel alters the range of visible stars.
C. Thus, the earth is spherical.

Both of these arguments are still used today and considered to provide powerful reasons for accepting the earth's sphericity. And they both exemplify an

[7] It's a longstanding and stubbornly persisting myth (still taught in schools, despite repeated corrections) that Christopher Columbus courageously embarked on his 1492 journey in the face of fears that his ship would fall off the edge of a flat earth. The truth is that virtually *no* educated Western European at the time of Columbus shared in such fears. Not only did the ancient Greeks discover the earth's shape, but they also made impressively accurate calculations of its size. This enduring myth was invented in the 1820s by the American writer, Washington Irving, and propagated in his *History of the Life and Voyages of Christopher Columbus* (Lindberg, 2007, p. 161). For more on the invention of this flat earth myth, see Russell (1991) and Garwood (2008).

[8] Aristotle's own presentation of these arguments runs as follows: "How else would eclipses of the moon show segments shaped as we see them? […I]n eclipses the outline is always curved; and, since it is the interposition of the earth that makes the eclipse, the form of this line will be caused by the form of the earth's surface, which is therefore spherical. Again, our observations of the stars make it evident, not only that the earth is circular, but also that it is a circle of no great size. For quite a small change of position to south or north causes a manifest alteration of the horizon. There is much change, I mean, in the stars which are overhead, and the stars seen are different, as one moves northward or southward." – *De Caelo* 297^b24-34, as translated in Barnes (1984).

extremely common style of scientific reasoning: we confirm a hypothesis by observing evidence that we expect to find if the hypothesis is true. But uncertainty manifests itself in these arguments in both of the aforementioned ways. First, the conclusion doesn't follow inescapably in either case; that is, the inference in both cases is invalid and thus uncertain. In fact, as any astute intro to logic student will quickly recognize, both arguments – if construed as attempts at deductive argumentation – commit the ostensibly pernicious, *elementary* fallacy of affirming the consequent! The earth could have some nonspherical shape (e.g., a flat disc) that still casts a circular shadow on the moon because of its orientation with respect to the Sun; and north/southward travel would result in a change of horizon with respect to the fixed stars if the earth were, for example, a properly oriented cylinder instead of a sphere. So neither argument's premises force it to be true that the earth is a sphere. A classically deductive approach, by focusing exclusively on valid inference, will thereby neglect any sense in which these invalid arguments are nonetheless good.

Second, reasoners aren't handed the truth values of these premises. For example, in **Argument 1**, A2 is particularly questionable since the shadow observed during lunar eclipses is, at any one time, partial with fuzzy boundaries. A1 is also dubious; even if the earth were a perfect sphere (which, of course, it's not), we wouldn't expect it to cast a *perfectly* circular shadow onto the moon insofar as the surface of the moon is itself curved and includes significant elevation changes (mountains and craters). At best, we are highly confident in the approximate truth of both premises, but classical, deductive logic tells us nothing about what to do with such attitudes and what they may or may not imply about the attitude we should take toward conclusion C.

The motivation behind the approach taken here is to develop a logic that can make sense of uncertain inferences like Aristotle's. If our logic is going to make room for such inferences, then it must allow for a sense in which deductively invalid arguments can be cogent. And if our logic is to guide agents like us in reasoning similarly, it should instruct us with respect to the uncertain attitudes we are working with in such cases.

1.2 From Deductive Logic to Probability Theory

Here, we introduce the probability calculus as a promising inductive logic. While we ultimately want to move beyond a logic that relates propositions in terms of truth values, we will presently find reasons to think that the formal semantics of deductive logic should be retained as a limiting case of our inductive logic. Accordingly, it will prove useful for us to begin with a quick review of the formal language of classical, *propositional* logic.

Once we've defined the vocabulary and grammar of propositional logic, we have all we need to distinguish well-formed (grammatical) from ill-formed (ungrammatical) parts of a formal-logical language. In this Element, we'll use italicized, capital letters like E, H, K, P, Q, R, and S as the basic atoms of our formal language. They count as grammatical by themselves. More complex statements of propositional logic can be formulated using standard "connectives," like $\neg$, $\wedge$, and $\vee$. These don't just connect to atomic formulae (capital letters) but may be used to connect any grammatical statements of the formal language, according to the following recursive definition (here and throughout, we use lowercase Greek letters as metavariables standing in for *any* grammatical formula of the language):

Grammatical formulae for propositional logic:

- Capital letters (possibly with subscripts) are grammatical formulae;
- If ϕ is a grammatical formula, then so is $\neg\phi$;
- If ϕ and ψ are grammatical formulae, then so are $(\phi \vee \psi)$ and $(\phi \wedge \psi)$;[9]
- Only formulae that can be shown to be grammatical by the above conditions are grammatical.

These capital letters and connectives are supposed to mean something to us; they're specifically meant to be more exact versions of familiar components of our natural language. For example, $\neg$ is supposed to be like the English word "not," $\wedge$ like "and," and $\vee$ like the English inclusive "or" ("inclusive" meaning that the "or" statement doesn't rule out the possibility of both statements it connects being true). But it's important to note that, at *this* point in the development of the logical language, a grammatical formula like "$(P \wedge (\neg Q \vee R))$" means no more than complete gibberish, like "$((\vee \wedge \divideontimes PD$". In order to make our grammatical statements mean something, we need to go one more crucial step and specify their semantics.

Classical logic's reliance on truth values becomes apparent (and incredibly helpful) at this point. The formal semantics of propositional logic is straightforwardly "truth-functional," meaning that its connectives' meanings are specified precisely by articulating the truth values they output as a function of all possible combinations of truth values they take in. The connective $\neg$ denotes *negation*, the truth-functional operation resulting in a true statement if and only if

[9] The connective for the material conditional $\rightarrow$ is conspicuously absent from this list. In the rare cases where this connective makes an appearance, we'll think of it as part of our non-primitive vocabulary, $\ulcorner(\phi \rightarrow \psi)\urcorner$ being short for $\ulcorner\neg(\phi \wedge \neg\psi)\urcorner$. In fact, $\rightarrow$ will not play a substantial role in our discussions.

Though all grammatical disjunctions and conjunctions officially are enclosed in parentheses, we will follow a standard practice and drop *outermost* parentheses – i.e., any parentheses that have the entire remainder of the formulae within their scope.

the statement negated is false. The $\vee$ denotes *disjunction*, the truth-functional operation resulting in a false statement if and only if both of the connected "disjuncts" are false. The $\wedge$ denotes *conjunction*, the truth-functional operation resulting in a true statement if and only if both of the connected "conjuncts" are true. Articulating this formal semantics in terms of the standard truth tables (and again using lowercase Greek letters as metavariables standing in for any grammatical formula of the language), we have:

ϕ	$\neg\phi$
T	F
F	T

ϕ	ψ	$\phi \vee \psi$
T	T	T
T	F	T
F	T	T
F	F	F

ϕ	ψ	$\phi \wedge \psi$
T	T	T
T	F	F
F	T	F
F	F	F

From the viewpoint of inductive logic, there is nothing amiss thus far. Defining these connectives by specifying their truth-functional roles makes sense even if you deny that reasoners have direct access to such truth values. Nonetheless, while classical logic may offer an appropriate truth-functional semantics, that formal semantics falls short of providing us with an inductive logic of confidences.

In working toward such a logic, the following interesting point is crucially important. Classical logic's truth-functional semantics doubles as a plausible *certainty*-functional semantics. That is, we can interpret "T" as "certainty of truth" as opposed to truth simpliciter and "F" as "certainty of falsehood" as opposed to falsehood simpliciter, and all of the tables still turn out right. For example, this reading of the negation sensibly associates certainty that $\neg\phi$ is true [false] with certainty that ϕ is false [true]. This shift in interpretation is substantial; instead of trading directly in truth values, it trades in attitudes that we might have toward propositions, albeit extreme attitudes that we may have only rarely. To mark the shift in interpretation, let's rewrite the above tables using different values; we'll use "1" for certainty of truth and "0" for certainty of falsity:

ϕ	$\neg\phi$
1	0
0	1

ϕ	ψ	$\phi \vee \psi$
1	1	1
1	0	1
0	1	1
0	0	0

ϕ	ψ	$\phi \wedge \psi$
1	1	1
1	0	0
0	1	0
0	0	0

Using numerical values marks our shift in interpretation, but this move also – regardless of interpretation – allows us to represent deductive logic's formal semantics algebraically. Let Val be the "valuation function" that assigns to classical logic's grammatical formulae a member of the set $\{0, 1\}$.

The first and third tables are then straightforwardly summarized in terms of Val as follows:

- $Val(\neg\phi) = 1 - Val(\phi)$.
- $Val(\phi \wedge \psi) = Val(\phi) \times Val(\psi)$.

The case of $\vee$ is somewhat less obvious, and it proves instructive to flounder a bit trying to represent its table algebraically. Were it not for the first line of $\vee$'s table, the following straightforward sum operation would do the trick: $Val(\phi \vee \psi) = Val(\phi) + Val(\psi)$. In other words, in the special case where the first line can be ignored, this simple sum operation would be correct. But we're right to ignore the first line as a genuine possibility exactly when ϕ and ψ cannot possibly be jointly true – that is, when they are "mutually exclusive." This is worth emphasizing:

- If ϕ and ψ are mutually exclusive, then $Val(\phi \vee \psi) = Val(\phi) + Val(\psi)$.

Of course, we still want a general algebraic representation for disjunction. The reason that the above rule doesn't work for the case when we are certain that both ϕ and ψ are true is because $Val(\phi \vee \psi)$ would equal two instead of the appropriate value of one according to the straightforward sum. To correct for this, we could just subtract out a function that takes value one when $Val(\phi) = Val(\psi) = 1$ and 0 otherwise. We already have such a function in $Val(\phi \wedge \psi)$! Thus, our general algebraic representation for $\vee$ is:

- $Val(\phi \vee \psi) = Val(\phi) + Val(\psi) - Val(\phi \wedge \psi)$.

To recap, we are seeking a logic that deals in confidences and uncertainties instead of dealing directly in truth-values. Nonetheless, we have found it useful to start with the truth-functional semantics for classical logic, since this semantics doubles as a plausible *certainty*-functional semantics. This makes intuitive sense since extreme cases in which we have certainty of the truth or falsity of some proposition correspond exactly with those cases where we at least take ourselves to be dealing directly with truth-values. The upshot is that our inductive logic should retain the above algebraic rules, at least as limiting case rules that hold in contexts of certainty (i.e., degenerate cases of uncertainty or extreme cases of confidence).

Probability theory can then be thought of as a *conservative* departure from classical logic in the sense that it preserves all of the above algebraic rules for the connectives – and indeed uses two of these rules in particular to ground all of its mathematics. Probability generalizes the bivalent, classical semantics, not