Philosophy Compass 9/12 (2014): 876–883, 10.1111/phc3.12182

Probability and Logic

Kenny Easwaran*
Texas A&M University

Abstract

Probability and logic are two branches of mathematics that have important philosophical applications. This article discusses several areas of intersection between them. Several involve the role for probability in giving semantics for logic or the role of logic in governing assignments of probability. Some involve probability over non-classical logic or self-referential sentences.

Logic and probability are two branches of mathematics that have very important philosophical applications. In logic, we study the relations of entailment, contradiction, and consistency among sentences. In probability, we assign more fine-grained distinctions allowing various consistent sentences to be more or less probable. Both areas have major applications in epistemology, though they are also relevant for considerations in metaphysics, philosophy of language, philosophy of science, and ethics, among other areas.

In this article, I only have space to discuss a few areas of research at the intersection of these two fields. Each section briefly outlines one area of research that involves themes from these two areas. Various of these research programs have connections to one another, though they are far from forming any sort of unified whole.

In this article, I assume familiarity with classical logic and its associated notions of validity, entailment, consistency, and logical necessity. I also assume some familiarity with the notion of an interpretation (or model) of a formal language and what it takes for such an interpretation to make a sentence (or proposition) true or false. At some points, I also mention non-classical logics, though the details of how they work is not essential for anything I say, except to note that advocates of non-classical logics argue that all of these concepts must be modified in some way. But I will give a brief discussion of the formal theory of probability in Section 2.

1. Inference and Implication

One popular view of logic is that it, in some way, provides rules for thought. It is assumed that we believe some set of sentences or propositions and that logic has some important normative power for these beliefs. Perhaps consistency tells us about what sets of beliefs are permissible to hold, or entailment can tell us about the ways in which some beliefs are justified by others.

These ideas dovetail nicely with one of the prominent applications of probability theory, on which probability measures the strength of some agent's beliefs, and the theory of probability tells us what strengths of belief are reasonable for an agent to hold. (There are other interpretations of probability on which the numerical values represent objective chances, or frequencies of some type of event, but those interpretations have fewer interactions with logic, so I largely omit the discussion about them in this article (Hájek, 2007)).

However, Gilbert Harman has cautioned us to be careful about the role of logic in epistemology. It is tempting to think that if p implies q, and one believes p, then one ought to believe q. However, as he points out (Harman, 1986, Ch. 1), it is just as reasonable to reject p on this basis as it is to accept q. Logic can be relevant to metaphysics, in that it says which things can be true or false together. And it can be relevant to semantics, in that it gives an important aspect of the meanings of sentences or propositions. But its role in epistemology, on Harman's view, is limited to giving us a 'wide scope' norm – if p entails q, then one shouldn't believe p while disbelieving q. (For caution on even this limited role for logic in epistemology, see MacFarlane (2004) and Easwaran and Fitelson (2014)).

This relates to an earlier point made by Frank Ramsey about the relation between logic and probability. In the final sections of the work of Ramsey (1926), he says that although both subjects give a 'logic of consistency', logic also gives a 'logic of truth' that probability does not. Logic tells us what the world could be like, or about the meanings of expressions in our language, and in virtue of that, puts constraints on what we ought to believe – we must not have beliefs that could not be true together. But probability just directly puts constraints on our beliefs, without saying anything about the world or our language (though see Sections 3 and 4 for applications of probability to meaning). Harman's point is that even this logic of consistency doesn't tell us what changes we must make to our beliefs if we discover that they violate its dictates. Although logic and probability both have an important role to play in telling us the constraints on our beliefs, logic tells us more about truth and meaning, while probability tells us more directly how our beliefs should be.

At this point, it is useful to give the formal theory of probability more explicitly in order to say in more detail what these constraints look like.

2. *Axioms of Probability*

The most classic axiomatization of probability is due to Kolmogorov (1950). It starts with a set Ω and a field $\mathcal{F}$ of subsets of Ω. To say that it is a 'field' of subsets of Ω means that if x is a subset of Ω that is in $\mathcal{F}$, then its complement $\Omega \setminus x$ is also in $\mathcal{F}$, and if x and y are subsets of Ω that are in $\mathcal{F}$, then their union $x \cup y$ is also is in $\mathcal{F}$. Kolmogorov then states that a probability function is a function P assigning a real number to each element of $\mathcal{F}$, such that $P(\Omega) = 1$, $P(x) \geq 0$, and $P(x \cup y) = P(x) + P(y)$ whenever x and y are disjoint sets. (It is common to supplement both the definition of a field of sets, and the definition of a probability function, so that they respect unions of countably infinite collections of sets, and not just unions of finite collections of sets.)

One standard way to apply this formalism is to let Ω be the set of possible worlds and to let $\mathcal{F}$ be the set of propositions. Then each proposition has a probability. However, if propositions are individuated more finely than sets of possible worlds (for instance, the proposition that Superman can fly may be distinct from the proposition that Clark Kent can fly, even though they may be true in the same set of possible worlds), then this can lead to some difficulty. This arises especially in cases where probability is taken to represent the degrees of belief of some agent, who may find one claim more plausible than another if she doesn't realize that they are necessarily equivalent. These problems can be mitigated if the elements of Ω are taken to represent some sort of possibility other than a metaphysically possible world. But it is not always clear what such possibilities might be.

An alternative, logic-based axiomatization of probability avoids these interpretive issues by taking the objects of the probability function to be not sets, but sentences of some formal language. One such formulation would let $\mathcal{L}$ be a language that is closed under syntactic operations of negation, conjunction, and disjunction and then state that a probability function P assigns a real number to each element of $\mathcal{F}$, such that $P(T) = 1$ whenever T is a logical necessity, $P(x) \geq 0$, and $P(x \vee y) = P(x) + P(y)$ whenever x and y are logically incompatible. It is a straightforward consequence of these axioms that if x and y are logically equivalent, then they have the same probability. This is because $x \vee \neg x$ and $y \vee \neg x$ are both logical truths, which thus have probability 1, and because $\neg x$ is logically incompatible both with x and with y, so both x and y must have probability $1 - P(\neg x)$.

With this observation, there is a natural way to relate these two separate axiomatizations of probability. If $\mathcal{L}$ is a formal language, we can let Ω be the set of interpretations of the language, and let the members of $\mathcal{F}$ be sets of interpretations that make a given sentence true. Logically equivalent sentences will correspond to the same set of interpretations. Because logical necessities are true on all interpretations and because logically incompatible sentences are true in disjoint sets of interpretations and because their disjunction is true in that union, the axioms of one formulation correspond to axioms of the other formulation. Thus, the logic-based axiomatization of probability can always be interpreted in the set-theoretic framework, which has greater generality, because it can also apply in cases where there is not an obvious language to use. But, as will be discussed in Section 6, the logic-based formulation can be generalized more easily to account for non-classical logics.

One distinction is important to note here. On both axiomatizations, the theory of probability is the set of logical consequences of some axioms. Thus, logic can play a role in the metatheory for probability. However, on the second axiomatization, there is also a logical language that probability applies to. So probability can also play a role in the metatheory for some logic.

3. Adams' Probability Logic for Conditionals

One particular application of a probabilistic metatheory for logic has been in the interpretation of conditionals. Conditionals have posed problems for both logic and probability. In probability, the notion of conditional probability has been standardly defined by saying that the conditional probability of A given B, notated as $P(A \mid B)$, is equal to $P(A \wedge B)/P(B)$ (or in the Kolmogorov axioms, $P(A \cap B)/P(B)$). There are some serious worries about this definition when $P(B) = 0$ (see Hájek (2003) and Easwaran (2008)), though most of these worries have led to supplementation of this definition, rather than any claim that it is extensionally inadequate in the cases where it gives an answer.

However, in logic, the notion of conditional has standardly been defined as the material conditional, where 'if B then A', notated as $(B \supset A)$, is true iff either A is true or B is false. This leads to various well-known 'paradoxes of material implication', such as the claim that $(B \supset A)$ is logically entailed by A or by $\neg B$. Many philosophers have noted that these entailments, although valid for the material conditional, do not appear to be valid for the conditional in ordinary language. Thus, many different theories of the meaning of the ordinary language conditional have been given to avoid these paradoxes. One such theory is of particular to note here, because it comes out of probability theory.

Adams (1975) proposed that we treat the conditional as a new connective, $(B \rightarrow A)$. He proposed giving semantics to this connective by using a probabilistic metatheory. In particular, he proposed that $P(B \rightarrow A) = P(A \mid B)$. To give a theory of logical implication, he said that a set of premises entails a conclusion iff for every $\varepsilon > 0$, there is a $\delta > 0$, such that, whenever all the premises have probability greater than $1 - \delta$, the conclusion must have probability greater than $1 - \varepsilon$. As long as no conditionals are involved in the argument, this is equivalent to the claim that whenever all the premises have probability 1, the conclusion must also have probability 1. But when conditionals are involved, the paradoxes of material implication come out valid on the latter definition, but not on Adams' definition. Thus, he suggests that this new logic can better account for the ordinary notion of conditional.

Soon after Adams made this proposal, David Lewis pointed out that assuming that $P(B \rightarrow A) = P(A \mid B)$ leads to triviality, given some other basic assumptions about the logic of conditionals (Lewis, 1976; Hájek & Hall, 1994). For instance, one result shows that if $(A \rightarrow (B \rightarrow C))$ is always equivalent to $((A \wedge B) \rightarrow C)$, then $P(A \mid B) = P(A)$ for all A and B. A more general result shows that if there are only finitely many worlds, then there are more distinct conditional probabilities than probabilities, so conditionals can't all be propositions that correspond to sets of worlds. In any case, some severe modification of the logic or semantics of conditionals must be made in order to preserve Adams' account. Some have suggested that conditionals cannot be nested (so that the first argument is blocked) or that conditionals don't express propositions that can be true or false (so that the second is blocked). For details of these views, see Bennett (2003) and Edgington (1995). At any rate, the move that is made initially to give a more plausible probabilistic semantics for conditionals requires further modifications in order to work.

4. Probability Semantics for Logic

Another use for probability in giving part of the semantics for logic is Hartry Field's conceptual role semantics. The idea of conceptual role semantics has existed in many forms. The general idea is that rather than basing a theory of meaning on conditions under which various sentences are true, we should base a theory of meaning on conditions under which one sentence can be inferred from another. This may be because some sentences are said to lack truth conditions, as some have suggested for conditionals and others have suggested for ethical or esthetic vocabulary. Or it may be for reasons like those of Michael Dummett, who argues that the metaphysical notion of truth going beyond epistemic verification is inherently inaccessible, and that once we retreat to verification, we will reject classical logic and accept intuitionist logic (Dummett, 1974; Dummett, 1973).

Field's response to this idea uses probability to get around Dummett's arguments. In his (1977), Field suggests that the right way to think of inferential role relies not on full belief, as Dummett suggests, but on strength of partial belief. Logical relations among sentences can just be taken to be those relations that competent speakers of a language are required to have among their partial beliefs expressed in that language. Because these relations are the foundation for logic, they must not be axiomatized by means of the axioms for probability I mentioned above, which take logical relations as granted. But they also can't be based on truth conditions, if they are to respond to Dummett's worries, so the Kolmogorov axioms most likely can't be used either.

Instead, Field appeals to a set of axioms for probability given by Popper (1955). These axioms take the bearers of probability to be uninterpreted sentences of a formal language. These axioms are 'autonomous' in the sense that they don't assume anything about the truth conditions corresponding to formal symbols in the language. Instead, they just give direct constraints on how one's conditional degrees of belief in one set of sentences should relate to one's conditional degrees of belief in another set of sentences, where the sentences are treated purely syntactically.

Given these axioms, we can define an inferential relation $A \succcurlyeq B$ meaning that for all X, $P(A|X) \geq P(B|X)$. That is, one can infer A from B if one is certain that regardless of what evidence one might come across in future, one will always be at least as confident in A as in B. Importantly, as Popper had already shown, whenever B classically entails A by mere syntactic manipulation, his syntactic axioms entail that $A \succcurlyeq B$. Field shows that this holds for an extended language including quantifiers as well. Thus, Field argues that full classical logic, and not merely intuitionist logic, must be among the inferential relations for any competent speaker of a language.

5. Logical Probabilities

Some philosophers have suggested a reversal of this interaction of probability and logic. Rather than using probability to give a semantics for logic, they have tried to use logic to give further requirements for probability. Rather than allowing any probability function satisfying the axioms above (which give some constraints from logic, but not many), they suggest that a rational agent's degrees of belief must be fully determined by her evidence. On this interpretation, probability doesn't just give a notion of partial belief generalizing the notion of full belief, but conditional probability is a sort of 'inductive logic' giving a notion of 'partial entailment' generalizing the notion of full entailment that we get from deductive logic. Something like this has often seemed to be needed in order for there to be objective facts about evidential support. This idea is most prominent in the work of Keynes (1921) and Carnap (1950, 1952).

Although Keynes generally suggested that many of these inductive probability relations would be imprecise and non-numerical, Carnap tried throughout his career to give various precise numerical criteria for what the relevant probability function should look like. Versions of this idea generally rely on some form of the 'principle of indifference', stating that equal possibilities should have equal probability. However, the application of this idea changed greatly over the course of his career. In each case, he considered a language with finitely many predicates and finitely many names for objects. He then defined the notion of a 'state description', which is a sentence stating, for each object, whether or not each of the predicates applies to it. If there are m names and n predicates in the language, then there are $m \cdot n$ sentences applying a predicate to a name, and there are thus $2^{m \cdot n}$ state descriptions.

Carnap's first proposal stated that each state description should get equal probability. This is perhaps the most natural application of the principle of indifference. However, this logical probability function gives no justification for inductive learning. Every attribution of a predicate to an object has probability 1/2 conditional on any other. Evidence saying that there are a million black ravens, and that this particular object is a raven, gives no evidence that this particular object is black.

Carnap's second proposal involved a notion of a 'structure description', which is a disjunction of all state descriptions that are identical up to a permutation of the names of objects. If there is only one predicate, then structure descriptions just amount to statements of how many objects bear that predicate, while if there are more, then they state how many objects bear each particular combination of predicates.

On his second proposal, each structure description has equal probability, and then each state description within a given structure description has the same fraction of that probability. If the language contains one predicate and two names, then there are four state descriptions and three structure descriptions. The state descriptions saying that neither object bears the predicate, or that both do, each get probability 1/3, while the state descriptions saying that one of the objects does and the other doesn't, each get probability 1/6. Thus, conditional on one object bearing the predicate, the probability that the other does is 2/3, so there is quite a large confirmation from induction. However, it turns out that the same is true regardless of how many names or predicates there are in the language – a single instance confirms all others to this same large degree.

In his later work, Carnap came up with successive modifications to these proposals that involved additional parameters and complications. His 'continuum of inductive methods' is formally equivalent to a proposal by Johnson (1932). Skyrms (1991, 1996) argues in his work that had Carnap continued his research, he eventually would have developed more and more patterns of Bayesian statistical inference, but they would have been less and less suited to count as truly logical.

6. Non-classical Logic and Probability

Many philosophers of logic have argued that classical logic has the wrong concepts of validity, consistency, entailment, interpretation, and so on. Instead, they propose that some non-classical logic is correct, either in general or for some applications. If any of these arguments is correct, then a modification of probability theory is needed whenever probability applies to one of these contexts. On the Kolmogorov axiomatization, some interpretation of the set Ω must be taken that properly accounts for the non-classical aspects of the propositions that the probability applies to. But the other axiomatization allows for more direct insight into the modifications that the theory of probability must have.

Some preliminary work on this front was done by Weatherson (2003), who argued that the axioms should be modified slightly to properly account for the ways in which non-classical logics behave. In particular, the axioms should state that if x entails every sentence in the language, then $P(x) = 0$; if x is entailed by every sentence in the language, then $P(x) = 1$; if x entails y then $P(x) \leq P(y)$; and that $P(x) + P(y) = P(x \vee y) + P(x \wedge y)$. In the classical case, these axioms are equivalent to the ones given above. But in non-classical logics, these are more generally applicable. He particularly considers the application involving intuitionist logic, which may be appropriate on certain views of evidence.

In some recent papers, Robbie Williams has given more detailed arguments for these axioms (Williams, 2012; Williams, 2014). However, his arguments primarily apply if the logic involved is given by means of an expanded set of truth values, each of which comes with some appropriate degree of belief that an agent should ideally bear to a proposition of that truth value. For instance, his argument works for the Strong Kleene three-valued logic, if the third truth value is interpreted as one where agents should have degree of belief 1/2 to a proposition with this truth value. It also works for Graham Priest's Logic of Paradox, which is formally identical to the Strong Kleene logic, but where the third truth-value is interpreted as the sentence being true as well as false, so that agents should have degree of belief 1 to a proposition with this truth value.

Interestingly, intuitionist logic does not fall into this class, so Williams' justification for these axioms is separate from Weatherson's work. And plenty of other logics (particularly substructural logics) are not addressed by either. At any rate, much more work is needed to apply probability to the very broad class of non-classical logics that exist.

7. *Probability and Self-Reference*

One important motivation for non-classical logic has been the paradoxes of self-reference, including the liar and its relatives. Until recently, this issue has not been directly brought up in probability theory. On either of the traditional axiomatizations of probability, there is no direct way for self-reference to arise. On the Kolmogorov axioms, the bearers of probability are just sets, and there is no particular interpretation of the elements of the set, so no statement is really 'about' anything, let alone itself. Even on the logic-based axiomatization, the axioms themselves don't put any constraints on the semantics of the terms in the language beyond the sentential connectives, and these alone are also not enough to generate paradoxes of self-reference.

However, in formal epistemology, many of the arguments in favor of using probability theory as an account of degree of belief are sensitive to self-reference. In particular, one important recent argument for using probability theory is James Joyce's accuracy-based argument. Joyce (1998) argues that the goal of degree of belief is accuracy, which is a certain sort of closeness to truth value. He shows that if accuracy is measured in a particular sort of way, then degrees of belief that violate the probability axioms are guaranteed to be farther from the truth than ones that satisfy them.

This argument works as long as the propositions over which the agent is assumed to have degrees of belief correspond to sets of possibilities, as in the Kolmogorov axioms. But as Caie (2013) has shown, if self-referential propositions are included, then it seems that certain apparently non-probabilistic degrees of belief can be guaranteed to do better than any that satisfy the sentential probability axioms. He also argues that self-reference is not essential and that similar issues arise in any cases in which an agent knows that the propositions she has attitudes toward bear some causal dependence to her actual degrees of belief. Greaves (2013) has further shown that this sort of self-reference can cause problems for updating probabilities in light of evidence.

8. Conclusion

This is only a brief selection from many topics in the intersection of probability and logic. For greater detail on each of these projects, readers should consult the cited works. Work on many of these topics is ongoing.

Short Biography

Kenny Easwaran is an associate professor of Philosophy at Texas A&M University. His research is primarily on issues related to epistemology and decision theory and, particularly, the role of infinity and zero in calculations involving these concepts, as well as the epistemology of mathematical proof. He received his PhD working under Branden Fitelson in the Group in Logic and Philosophy of Science at UC Berkeley in 2008 and spent 6 years as an assistant professor at the University of Southern California before moving to Texas A&M.

Note

* Correspondence: Philosophy, Texas A&M University, 4237 TAMU, YMCA Building 301, College Station, TX 77843, USA. Email: easwaran@tamu.edu

Works Cited

Adams, E. *The Logic of Conditionals: an Application of Probability to Deductive Logic*. Dordrecht: Reidel, 1975.

Bennett, J. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press, 2003.

Caie, M. 'Rational Probabilistic Incoherence.' *The Philosophical Review* 122.4 (2013): 527–575.

Carnap, R. *Logical Foundations of Probability*. University of Chicago Press, 1950.

——. The Continuum of Inductive Methods. University of Chicago Press, 1952.

Dummett, M. 'The Philosophical Basis of Intuitionistic Logic.' *Philosophy of Mathematics*, second edition. Ed. Benacerraf P., Putnam H. chapter 14. Cambridge: Cambridge University Press, 1973.

——. 'The Justification of Deduction.' Truth and Other Enigmas. Oxford: Oxford University Press, 1974.

Easwaran, K. The Foundations of Conditional Probability. PhD thesis, University of California, Berkeley, 2008.

Easwaran, K., Fitelson, B. *Accuracy, Coherence, and Evidence*. Oxford Studies in Epistemology, 2014.

Edgington, D. 'On Conditionals.' *Mind* 104.414 (1995): 235–329.

Field, H. 'Logic, Meaning, and Conceptual Role.' *The Journal of Philosophy* 74.7 (1977):379–409.

Greaves, H. 'Epistemic Decision Theory.' *Mind* (2013).

Hájek, A. 'What Conditional Probability Could Not Be.' *Synthese* 137 (2003): 273–323.

——. 'Interpretations of Probability.' *Stanford Encyclopedia of Philosophy* (2007). http://plato.stanford.edu/entries/probability-interpret/

Hájek, A., Hall, N. 'The Hypothesis of the Conditional Construal of Conditional Probability.' *Probability and Conditionals: Belief Revision and Rational Decision*. Ed. Eells E., Skyrms B., Cambridge University Press, 1994. 75–112.

Harman, G. *Change in View: Principles of Reasoning*. Cambridge: MIT Press, 1986.

Johnson, W. E. 'Probability: the Deductive and Inductive Problems.' *Mind* 41 (1932): 409–423.

Joyce, J. 'A Nonpragmatic Vindication of Probabilism.' *Philosophy of Science* 65.4 (1998): 575–603.

Keynes, J. M. *A Treatise on Probability*. London: Macmillan and co., 1921.

Kolmogorov, A. N. *Foundations of the Theory of Probability*. New York: Chelsea, 1950.

Lewis, D. 'Probabilities of Conditionals and Conditional Probabilities.' *The Philosophical Review* 85.3 (1976): 297–315.

MacFarlane J. 'In What Sense (if any) is Logic Normative for Thought?' unpublished, 2004.

Popper, K. 'Two Autonomous Axiom Systems for the Calculus of Probabilities.' *The British Journal for the Philosophy of Science* 6.21 (1955): 51–57.

Ramsey, F. P. 'Truth and Probability.' *The Foundations of Mathematics and Other Logical Essays (1931)*. Ed. Braithwaite R. B. New York: Harcourt, Brace and Company, 1926.

Skyrms, B. 'Carnapian Inductive Logic for Markov Chains.' *Erkenntnis* 35.1 (1991): 439–460.

——. 'Carnapian Inductive Logic and Bayesian Statistics.' Statistics, Probability and Game Theory: Papers in Honor of David Blackwell. Hayward: Institute of Mathematical Statistics, 1996. 321–336.
Weatherson, B. 'From Classical to Intuitionistic Probability.' *Notre Dame Journal of Formal Logic* 44(2) (2003): 111–123.
Williams, J. R. G. 'Generalized Probabilism: Dutch Books and Accuracy Domination.' *Journal of Philosophical Logic* 41.5 (2012): 811–840.
——. 'Non-classical Logic and Probability.' Oxford Handbook to Probability and Philosophy. Ed. Hájek A., Hitchcock C. Oxford: Oxford University Press, 2014.