Probability and arguments: Keynes's legacy

William Peden[*]

John Maynard Keynes's *A Treatise on Probability* is the seminal text for the logical interpretation of probability. According to his analysis, probabilities are evidential relations between a hypothesis and some evidence, just like the relations of deductive logic. While some philosophers had suggested similar ideas prior to Keynes, it was not until his Treatise that the logical interpretation of probability was advocated in a clear, systematic and rigorous way. I trace Keynes's influence in the philosophy of probability through a heterogeneous sample of thinkers who adopted his interpretation. This sample consists of Frederick C. Benenson, Roy Harrod, Donald C. Williams, Henry E. Kyburg and David Stove. The ideas of Keynes prove to be adaptable to their diverse theories of probability. My discussion indicates both the robustness of Keynes's probability theory and the importance of its influence on the philosophers whom I describe. I also discuss the Problem of the Priors. I argue that none of those I discuss have obviously improved on Keynes's theory with respect to this issue.

Key words: John Maynard Keynes, History of probability, Philosophy of probability induction, Logical probability
JEL classification: B10

1. Introduction

In *A Treatise on Probability*, John Maynard Keynes (1921) provided the first systematic, subtle and self-conscious theory of what philosophers now call 'logical probability'. I shall focus on a single thesis by Keynes and trace its influence. This thesis is that *probability is an evidential relation holding between an ordered pair of sets of statements*. Modern logicians generally define an 'argument' to be an ordered pair of sets of statements.[1] Thus, another way of understanding Keynes's thesis is that *probability is a feature of arguments*. Hypotheses have probabilities only derivatively, by being the conclusion of an argument with that probability relation. This is the key distinctive feature of the logical

*Erasmus University Rotterdam.

[1] Or propositions etc. I shall switch between these depending on the usage of the philosopher I am discussing. Differences between propositions and statements are important, but not here.

interpretation of probability.[2] For brevity, I shall use 'relationism' to denote this view. I shall argue for the following:

(1) Keynes's relationism influenced an intellectually diverse group of philosophers of probability. This variety indicates the robustness of relationism.
(2) The Problem of the Priors tends to be why these philosophers depart from Keynes, but their theories do not obviously improve on his theory in this respect.

I shall begin by explaining relationism, before discussing its influence among a variety of philosophers. My sample was chosen using criteria of (a) influence from Keynes, (b) the differences among the philosophers and (c) the extent to which they have not been discussed by historians. Consideration (a) excludes a number of famous relationists, including Gottfried Leibniz, W. E. Johnson and Harold Jeffreys, because they developed their ideas independently (or mostly independently) of Keynes. Consideration (c) excludes Rudolf Carnap, who has been discussed at length by historians of probability. My sample is not exhaustive, but it will establish my theses.

2. Keynes's interpretation

Throughout, by 'a theory of probability', I shall mean the combination of (1) an account of what statements about probabilities mean (or should mean) and (2) a theory of which probability statements are true. I shall call (1) an 'interpretation' of probability. For lack of a better term, I shall call (2) a 'semantics' of probability. There are many theories of probability. Some view probabilities as subjective degrees of belief, with just a few constraints on what constitutes rational beliefs (de Finetti, 1964). Others identify probabilities with *uniquely* rational degrees of belief (Williamson, 2010). And still others view probability as an objective feature of physical reality (Von Mises, 1951; Popper, 1957). Moreover, there are many pluralists, who adopt multiple interpretations (Carnap, 1962). An interpretation will often indicate, but not determine, a semantics for probability. For some interpretations, such as frequentism, which statements are true depends on logico-mathematically contingent facts, i.e. facts outwith logical facts about statements/arguments and mathematical facts about numbers, models, and other mathematical entities. (I shall use 'contingent' to mean logico-mathematically contingent.) However, in Keynes's theory, the semantics is a matter of abstract logical relations. Keynes (1921) argues that probability is relational: it holds between two propositions or sets of propositions. A proposition H is only probable or improbable in the derivative sense that H is probable or improbable in *relation* to a set of propositions. It follows that there can be no unconditional probabilities of the form $P(H) = r$ for some real number r.[3] Keynes also thought that probability is logical, in a broad sense of 'logic' that includes deductive and non-deductive argumentation. Deductive logicians study arguments with respect to whether they are deductively valid, i.e. whether their premises (or 'evidence') provide maximally strong evidence to their conclusions (or 'hypotheses'). For brevity, I shall use 'valid' for 'deductively valid'. Keynes's theory

---

[2] 'Logical probability' is sometimes used for the more general idea that there are objective epistemic probabilities. However, objectivity does not entail that probability is relational, since there could be objective unconditional probabilities, as in the classical theory of probability.

[3] Note that there are relational theories of probability in which this relation is not *logical* (Williamson, 1998).

generalises logic to all arguments, including evidential relations within invalid arguments. Keynes agrees with 'frequentist' probability theorists that the truth or falsity of a probability statement like (i) 'It will probably rain tomorrow' is objective, in the sense that its truth or falsity is independent of our opinions. He thus denied subjectivism. However, Keynes denies that the truth or falsity of (i) is determined by relative frequencies. Instead, if it is true, then it is true in the same way as the deductive logic statement (ii) 'The premises that it will be cloudy tomorrow and that it will be rainy tomorrow deductively imply that it will be rainy tomorrow'. According to a relationist interpretation, the truth-value of both (i) and (ii) depends on logical facts about arguments. Superficially, (i) seems to be very different from (ii). On Keynes's interpretation, this appearance is just because (ii) makes its premises explicit. Implicit premises are common in probabilistic reasoning, just like deductive reasoning. (Argumentation would be *very* tedious otherwise). Typically, a statement like (i) tacitly refers to a relation between our relevant evidence (weather reports, historical trends, that the air pressure is dropping etc.) and the statement said to be probable. In this example, this statement said to be probable is 'It will rain tomorrow'. However, the information is not always implicit. For instance:

(iii) 'Since the card was randomly selected from a normal deck, the probability that it is an Ace of Spaces is 1/52'.

—and:

(iv) 'This coin has landed Heads in all of the 50 times I've tossed it, so probably it's biased towards Heads'.

—both state some of the relevant evidence. Thus, a hypothesis's logical probability is always a relation between that hypothesis and some premises. These premises might *refer* to relative frequencies, but the probability is not *identical* with them.

Probabilities are also not identical to rational degrees of belief; instead, they are relations between propositions (Keynes, 1921, p. 5). Nonetheless, probabilities can guide our beliefs (Keynes, 1921, p. 2, p. 351.). In this respect, probabilities are again akin to deductive relations like entailment or contradiction. We should not believe contradictions if we want to know the truth, but it does not follow that 'contradiction' can be *defined* this way. Similarly, if we are rational and we want to have true beliefs, then we should be confident in hypotheses that are probable in relation to our evidence, but Keynes's interpretation of probability is not definable in terms of rational confidence.

Another notable feature of Keynes's theory is probabilities are often imprecise, in that they cannot be identified with a real number (Keynes, 1921, pp. 181–3). There is sometimes precision according to his theory. Indeed, when the premises feature a contradiction, there is no apparent probability relation at all (Keynes, 1921, pp. 127–8). A broad class of precise probabilities occur under the conditions of Keynes's carefully revised version of the Principle of Indifference (POI). If:

(1) The set of statements $A = \{H_1, H_2 \dots H_n\}$ are each possible given the evidence E.
(2) If E is true, only one hypothesis in A can be true.
(3) Each hypothesis in A can be expressed in the same logical form as the others. The logical 'form' or 'argument form' is what results from substituting variables for terms other than logical constants into an argument. Logical constants are terms like 'all', 'some', 'and', 'or' and 'if…then'.
(4) They cannot be further analysed into a different set of statements satisfying (1–3).

(5) The terms in A's members and E are all meaningful.

(6) E does not favour any hypothesis in A over any other member of A.

—then the probability in relation to E of each hypothesis H_i among the n-fold hypotheses in A is $P(H_i \mid E) = 1/n$ (Keynes, 1921, pp. 62–7). In contemporary notation, if there is a precise probability relation between H and E, Keynes advocates that $P(H \mid E) = r$ means that the logical strength of an argument from E to H is r.

Keynesian probabilities are independent of contingent facts. They share this feature with deductive relations. For instance, the following argument is valid:

(1a) All cats are kind to birds.

(2a) Keith is a cat. Therefore, (3a) Keith is kind to birds.

—even though (1a) is false. Its validity is independent of contingent facts about how the world happens to be

Similarly, in Keynes's theory, the inductive argument:

(1b) Under a great variety of circumstances, Keith and all the many other known cats are kind to birds.

Therefore, (2b) All cats are kind to birds

—has a probability relation between (1b) and (2b) that is independent of contingent facts. For instance, the probability relation is the same regardless of whether the sample described in (1b) is unrepresentative of cats in general.

Keynes's theory is pregnant with a whole litter of fascinating philosophical ideas. There is logical probability relations' objectivity (McCann, 1994; Gillies, 2006). There is his idea that probability relations can be inexpressible by real numbers (Runde, 1994; Brady and Arthmar, 2012). There are Keynes's views on the connections of probability with uncertainty and ontology (Lawson, 1985, 2003). There is probability's connection with his general epistemology (Carabelli, 1988; O'Donnell, 1989). Keynes's influence with respect to these ideas is an even wider topic. However, his legacy in the philosophy of probability is still largely unexplored. Additionally, philosophical and historical discussions of relationism tend to focus on Carnap (e.g. Salmon, 1967; for an exception that focuses on Keynes, see Rowbottom 2015, Chapter 3). Given Carnap's historical importance in philosophy, the focus on him is understandable. However, it creates a misleadingly narrow impression of Keynes's influence, and a misleadingly tight association between relationism and Carnap. My inquiry will help remedy these misconceptions. For each thinker I discuss, I shall briefly provide some biographical information, then explain their *interpretations*, then their *semantics*, then Keynes's *influence* on them, and finally the plausibility of their views in relation to the Problem of the Priors.

3. Frederick C. Benenson

Benenson advocated a relationist theory of probability during a period when Keynes's ideas were unfashionable among philosophers (Benenson, 1984, p. 1). Benenson studied at Harvard and Oxford (New York Times, 1978). He was a lecturer at the University of Birmingham until 2004 (Benenson Capital, 2014, p. 6). He wrote several publications in the late 1970s and early 1980s; his *Probability, Objectivity and Evidence* (Benenson, 1984) was widely reviewed. However, he has not published since that book.

3.1 Interpretation

Benenson's version of relationism is idiosyncratic, but he credits with Keynes for his interpretation of probability (Benenson, 1984, pp. 13–4), describing Keynes as the 'ancestor' of his theory (Benenson, 1984, p. 45). Like Keynes, Benenson thought that relationism can accommodate *all* the rational uses of probabilistic concepts. In contrast, most contemporary philosophers of probability are pluralists: they adopt different interpretations for different probability statements. For example, they might adopt frequentist interpretations of the use of 'probability' in scientific theories (sometimes called 'physical probability') but another interpretation for appraising theories' evidential support. Benenson argues that, for both the uses of 'probability' in scientific theories and these theories' appraisals, we should typically interpret the term as referring to an evidential relation between a hypothesis and what Benenson called 'total statistical evidence'. For a probability statement *S* being assessed by a particular scientist in a time period *t*, the total statistical evidence is what the entire human race could, in principle, ascertain during *t* (pp. 208–9). Total statistical evidence is thus distinct from the 'total evidence' that relationists like have typically thought appropriate for probabilistic reasoning: total evidence is just the evidence available to an *individual* at the moment of their reasoning.

3.2 Semantics

What probability statements are true according to Benenson? We can understand his semantics by starting with an idea from Carnap (1952). For a simple and unambiguous formal language consisting of logically independent predicates *F*, *G* etc. and names of distinct individuals *a*, *b*, *c* etc., we can fully characterise a probability distribution via its values for the following schema:

Key

E_R: A sample report describing the joint frequency of *F* and *G* according to our total statistical evidence.
n: The joint occurrence of *F* and *G* described in E_R.
s: The size of the sample.
λ and k: Other parameters affecting the probability value. These are important in Carnap's probability theory, but not here.

$$P(Fa \mid Ga \ \& \ E_R) \ = \ \frac{n + \lambda/k}{s + \lambda}$$

Benenson argues for setting λ and k to zero, so that the probability of *Fa* in relation to *Ga* and E_R is the value that E_R reports for the sample frequency of *F*'s among *G*'s (Benenson, 1984, pp. 103–13). For example, if E_R just asserts that 7 of 10 *G* things are *F*, then $P(Fa \mid Ga \ \& \ E_R) = 0.7$.

Things are messier in the natural languages that we actually speak, but the spirit of Benenson's semantics is that probabilities are equal to sample frequencies reported in the total statistical evidence. Hence, there cannot be probabilities in relation to premises without sample reports. Benenson's semantics thus leaves many arguments (those lacking reported sample frequencies) without probability values.

In practice, we must estimate the frequency in the total statistical evidence using the merely partial evidence that is available to us as individuals. For these estimates,

Benenson endorsed confidence interval methods for (imprecise) 'second-order' probabilities that the relative frequency in the total statistical evidence is within a particular interval ([Benenson, 1984](), pp. 210–4).

3.3 Influence

A curious feature of Benenson's theory is that it is sample *frequencies* that matter for probabilities; sample *sizes* are irrelevant. Therefore, if two evidence reports E_{R1} and E_{R2} report the same frequencies in two different samples, then $P(Fa \mid E_{R1}) = P(Fa \mid E_{R2})$ even if E_{R1} describes a tiny sample and E_{R2} describes a sample of thousands. Benenson argues that $P(Fa \mid Ga$ & $E_R)$ is only relevant for our beliefs if E_R describes a large sample. To take an extreme example, if E_R just asserts that 'Fb and Gb', obviously we should not be maximally certain that a will be an F. Benenson never explained what constitutes a sufficiently large sample. Nonetheless, this features indicates the influence of Keynes's separation of the concepts of probability and belief. Keynes's separation makes it possible to talk about probabilities without corresponding rational degrees of belief, just as we can talk about deductive arguments that do not justify inferring their conclusions. For example, in classical logic, a contradiction implies every statement, but believing a contradiction does not entitle us to believe every statement. Thus, Benenson's views on the connection of probability with belief make crucial use of Keynes's relationism.

3.4 Priors

By the 'Problem of the Priors', I shall mean this issue: many relationist theories, such as Carnap's, imply that we can (should!) have strong beliefs about empirical hypotheses due to mere facts about logical possibilities. In a relationist theory, the *a priori* probabilities (the 'initial priors') are the probabilities of statements in relation to tautologies, i.e. logically true statements, since tautologies do not assert anything that we would need to know *a posteriori*. For example, if $P(H_i \mid t) = 0.5$ for each i of an n-fold exchangeable sequence of hypotheses, a tautology t, and n is large, then the hypothesis 'H_1 or H_2 or…or H_n' will be highly probable in relation to t. In other words, an equivocal (neutral) prior for each H_i implies a strong (non-neutral) prior for other hypotheses. For example, your priors might be equivocal towards each of a set of coin tosses landing Heads, but very unequivocal towards the conjecture that one of them will land Heads. I shall discuss possible Keynesian responses to this problem in my conclusion.[4]

Benenson avoids the Problem of the Priors, because logical probabilities only exist with evidence that contains relative frequency data. Yet there is a cost: the priors do not soften the impact of new evidence. For instance, according to Benenson's theory $P(Fa \mid Ga$ & Fb & $Gb) = 1$. If probability were rational belief, this would imply maximal certainty for Fa given the evidence. Benenson's sample size restrictions avoid this consequence. However, this solution seems to throw out the baby with the bathwater: small samples might be weak evidence, but they do indicate *something*. For example, if b is an F and a G, then F and G must be physically compatible. Benenson

---

[4] In a very different way, the Problem of the Priors also afflicts the more popular subjective Bayesian theories of probability, but that is beyond my scope.

might argue that his theory accommodates weak evidence via the imprecision of estimates, but in some cases no estimates will be necessary e.g. if we know that a and b are the only things that are G. In contrast, Keynes can represent weak evidence using imprecise probabilities, such as $P(Fa \mid Ga \ \& \ Fb \ \& \ Gb) > 0$, which acknowledges the weak but genuine relevance of such evidence. Benenson's theory thus seems to have no ultimately advantages for the Problem of the Priors.

4. Roy Harrod

Harrod (1900–78) was greatly influenced by Keynes in many respects. This influence began in 1922, when Harrod (then an Oxford Lecturer) visited for a term at King's College Cambridge and was guided in his research by Keynes (Caldentey, 2019, pp. 7–8). Harrod is best known today as an economist, but his philosophical research was widely reviewed and discussed in the 1950s. His overarching philosophical aim was to answer David Hume's sceptical challenge to induction. Harrod tried to establish that the premises of inductive arguments can have a strong probability relation in favour of their conclusions, even though the conclusions' truth is never guaranteed by the premises. I shall focus on his ideas about probability, rather than induction.

4.1 Interpretation

Harrod was an explicit relationist, defining probability 'as a relation between premisses and conclusion…' (Harrod, 1961, p. 45). According to Harrod, the probability relation is positive if the frequency of truth among statements with the conclusion's logical form would be greater than 50%, were the premises true. For instance, suppose we believe that we are conducting Bernoulli trials with a coin that lands Heads with a frequency of about r%. Predictions of the form 'The coin will land Heads about r% of the time in n trials', where n is a large number, for all of the possible trials, would be true more often than not, so such predictions are favourable relative to our premises. This analysis in terms of truth-frequencies differs from Keynes's relationism, because Harrod's analysis of the probability relation only makes sense in cases where the premises contain information about relative frequencies. By contrast, in Keynes's theory, arguments such as analogical arguments can probabilify a hypothesis using premises about similarities, with no mention of relative frequencies. However, Harrod thought that all good argumentation—including analogical arguments—could be interpreted in terms of evidence about relative frequencies (Harrod, 1956, p. 248; pp. 255–6), so the divergence is not important.

4.2 Semantics

Harrod admired Keynes's work on probability, but disagreed on many questions of semantics. Firstly, Harrod rejected initial priors (Harrod, 1961, pp. 44–6).[5] He even rejected initial comparative priors, such as the principle that simpler laws have a greater initial probability than more complex laws (Harrod, 1956, Chapter 6). In

---

[5] It is not clear if Harrod would accept initial priors for arguments with contradictory or tautologous conclusions.

logical terms, Harrod requires logically contingent information in the argument's premises, if there is to be a probability relation. Therefore, tautologies, which are not logically contingent, cannot be the premises. Harrod thus rejects the POI, in favour of his 'the Principle of Experience'. He formulates this principle in mathematical detail, but I shall just describe his informal intuition. According to Harrod, given some continuous process (such as a sequence of events, traversing a region of space in a straight line, and so on) we are more likely *ceteris paribus* to be far from the processes' end of that process than near or at it (Harrod, 1956, p. 78). Harrod's main example is someone travelling along a large expanse, such as an unfamiliar forest or desert. If the traveller supposes that they are not at the edge, then they would be right more often than not. Therefore, without additional information, there is a probability in favour of the hypothesis that they are not at the edge of the expanse, relative to their knowledge. The Principle of Experience was Harrod's fundamental tool for constructing induction-friendly prior probability distributions to answer Hume's scepticism. Finally, another divergence of Harrod from Keynes is that the former's probabilities are always precise (Harrod, 1956, pp. 34–5). However, this difference is not very significant because Harrod argues that precise probabilities are typically unknowable for humans, due to the immense complexity of the evidential relations in more interesting arguments in science and ordinary life (Harrod, 1956, p. 36). Whether precise probabilities in such cases are unknowable or non-existent makes little difference for our practical reasoning, nor even philosophical questions of epistemology.

4.3 Influence

Surprisingly, Harrod does not directly note Keynes's influence in his intellectual autobiography in (Harrod, 1956, pp. ix–xix). However, since Harrod does not attribute relationism to anyone other than Keynes and given that he was very familiar with the *Treatise on Probability*, it is highly plausible that Keynes was the source of Harrod's interpretation. Another sign is that, in Harrod's first foray from economics into the philosophy of probability (Harrod, 1942), Keynes is the only philosopher of probability cited (p. 56)— and the only citation! Keynes's influence on Harrod is clearest insofar as Harrod sees him as a point of departure. Harrod typically uses contrasts with Keynes to explain the distinctive elements of his own theory, which suggests that Harrod saw Keynes as the first word on probability. It is true that Harrod sometimes discounts Keynes's work as 'just a sketch' (Harrod, 1956, p. 19) or 'sketchy' (p. 27). He even criticises Keynes's combination of relationism with the POI as 'a confusion of thought' (p. 27). However, these remarks show that Harrod sees Keynes's theory as a prototype: an early version needing improvements, but still the proper starting point for Harrod's own inquiries. Furthermore, Harrod saw the rejection of initial priors as a consequence of taking Keynes's relationism seriously. Just as many economists have regarded themselves as more Keynesian than Keynes, so Harrod regarded himself as more faithful to Keynes's philosophy of probability than John Maynard. Despite his general admiration for Keynes's ideas in both economics and philosophy, he was willing to disagree with Keynes on some matters in order to preserve their shared relationist interpretation of probability. He did not try to muddy the waters to protect Keynes from (perceived) problems; Harrod's intellectual honesty was even stronger than his admiration for Keynes.

4.4 Priors

Harrod regarded it as obvious there can be no logical deductive or probabilistic relation between a contingent conclusion and *a priori* continent premises, like a tautology *t*. Given the absence of such a relation and assuming relationism, it follows that there can be no probability for such arguments, and hence no initial priors. However, his 'Principle of Experience' redistributes the same basic worry: what justification do we have, without evidence, to believe this strong claim about the universe? Additionally, Harrod's argument against initial priors are not convincing. They are based on the alleged absence of a probabilistic logical relation between a tautologous premise *t* and a contingent conclusion *H*. Yet, even in classical deductive logic, *H* and *t* do have a logical relation: a relation of asymmetric implication, whereby *H* implies *t* but *t* neither implies nor contradicts *H*. If such statements can have a deductive logical relationship, then why not a probabilistic relationship? It is when we face the problem of trying to characterise this relationship that the Problem of the Priors re-emerges.

5. Donald C. Williams

Williams (1899–83) is now principally known for his work on metaphysics and the justification of induction (Campbell et al., 2019). However, his views on probability are unusual and wide-ranging. From his institutionally influential position at Harvard, Williams shaped much of American philosophy in the later 20th century, with massively influential doctoral students, such as Roderick Chisholm and Donald Davidson. He wrote in the mid-20th century, when many philosophers thought that philosophical problems were due to linguistic confusions and needed to be 'dissolved' rather than solved. In contrast, Williams—like Keynes—took a traditional view of philosophical problems: questions like 'How should we interpret probability statements?' were meaningful questions that could be solved by careful theorising, plus rigorous attention to empirical and logico-mathematical facts. See (Campbell et al., 2019) for more biographical details.

5.1 Interpretation

Williams agrees with Keynes that probability is a generalisation of deductive logic to invalid but evidentially relevant arguments (Williams, 1947, Chapter 2). However, like Harrod, Williams analyses logical probabilities in terms of truth-frequencies. To understand 'truth-frequencies', consider this example. If we substituted a unique name for each swan *x* into the argument form below and (1c) were true, then the *deductive* argument:

(1c) 100% of swans are white.
(2c) *x* is a swan.

Therefore, (3c) *x* is white.

—would have a truth-frequency of 100%—we would always be right using the argument form for each swan. Williams analyses the argument's validity as this truth-frequency. Intuitively, (3c) has a probability of 100% in relation to (1c) and (2c).

   Similarly, if (1d) were true, then the non-deductive argument form:

(1d) 90% of swans are white.
(2d) *x* is a swan.

Therefore, (3d) *x* is white.

—would have a truth-frequency of 90%, because if we used that argument form for every swan (once each) we would be right 90% of the time. Thus, according to Williams, (3d) has a probability of 90% in relation to (1d) and (2d).

The truth of the premises may be hypothetical, but Williams denies that there is anything peculiar about truth-frequencies:

The [truth-frequency] in the class of possible inferences whose conclusions are propositions concerning biting dogs… our whole treatment… can be reformulated without mention of anything more unearthly and ethereal than the behavior of the flesh-and-blood animals themselves… (Williams, 1947, p. 40).

Williams noted the similarities between his theory and Keynes's, but regards the 'classical' tradition of theorists like Pierre-Simon Laplace as his main influence (Williams, 1947, pp. 50–1). Nonetheless, Williams's *relationism* apparently comes from Keynes, because his works only refer to Johnson in passing, nor at all to Leibniz on probability, and he does not take credit for relationism in particular.[6] Williams *was* aware of Charles S. Peirce's parallels between probability and deductive logic (Williams, 1947, pp. 196–200) and therefore his relationism might be from Peirce. Even if so, it was Keynes who first presented Williams with a thorough, systematic and explicitly relationist theory.

5.2 Semantics

We have already seen examples of Williams's probabilities. In general, he contends that the probability relation in an argument possessing this form:

(1) r% of F's are G's.
(2) x is an F.

Therefore, (3) x is a G

—has a value of r% (Williams, 1947, p. 35). While r can be a real number, it can also be an interval or qualitative expression like 'generally'. Hence, Williams agreed with Keynes that there are imprecise probabilities.

There many names for this type of argument, but I shall use Carnap's (1962, pp. 492–8) expression, 'direct inference'. Many philosophers, like Keynes and Carnap, think that there is *something* right about direct inference. Williams's atypical claim is that he thought that *all* probabilities were reducible (including via rules like the finite additivity axiom) to direct inferences (Williams, 1947, pp. 48–9 and Chapters 4–5). Williams regarded this claim as his principal disagreement with Keynes (Williams, 1947, pp. 50–1).

A version of 'the Problem of the Reference Class' applies to Williams's theory. Consider the swans example, but suppose we add two statements to the premises:

(1e) 90% of swans are white.
(2e) x is a swan.
(3e) 10% of animals in the zoo are white.
(4e) x is an animal in the zoo.

Therefore, (5e) x is white.

---

[6] Williams could not inherit relationism from classical probabilists, who believed in non-relational probabilities: for a discussion and examples, see Carnap (1962, pp. 47–51).

Ninety percent is no longer an intuitive probability for this argument; there is not any intuitively obvious objective value. Yet such ambiguous statistical evidence is common in science and everyday life. In the context of Williams's theory, the Problem of the Reference Class is the challenge of selecting and combining relevant statistical data in the premises about multiple reference classes. Until recently, it seemed that Williams was unaware of this problem, but at least some cases of this problem are addressed by Williams in a recently published posthumous paper (Williams, 2018) though not the particular type above.

5.3 Influence

Both Williams and Keynes saw that relationism implies that induction's rationality is independent of any contingent fact (Keynes, 1921, p. 254; Williams, 1947, Chapter 6). The universe might be hospitable to induction by being very regular or inhospitable by being very irregular. Yet according to Keynes and Williams, in either case it will be rational to reason inductively. They would say that, even when we know that induction is unreliable and irrational (e.g. we know that our sampling is misleading), we still know this unreliability via induction. Therefore, the probabilities of inductive arguments are independent of contingent facts, just like deductive validity/invalidity. This idea is crucial for Williams's views on Hume's Problem of Induction. Hume argued that we cannot justify the hypothesis that the universe is induction-friendly by *a priori* intuition (which does not provide knowledge of contingent facts) or by appeal to induction's past successes (which would be circular). Williams agreed, but argued that we can justify induction using probabilities from direct inference. Like deductive relations, these probabilistic relations do not depend on nature's actual regularity. Consequently, relationism is vital to Williams's theory of induction.

5.4 Priors

Williams (1947, p. 192) claims that his theory avoids the problem of assigning priors, at least in the case of populations' compositions, which seems to be why he rejects Keynes's version of the POI. Yet Williams's analysis of probability entails that such priors exist. For example, suppose that there are n mathematically possible distributions of a binomial random variable in a population. The relative frequency of a particular distribution i is $1/n$. Williams's usage of direct inference contains no clauses to rule out such applications, and thus the prior according to his theory is $1/n$, which raises the issues discussed in Section 3.4. Therefore, he has not escaped the problem.

6. Henry E. Kyburg

Kyburg (1928–2007) was a key figure in the development of what is now called 'formal epistemology'—the application of logic and mathematics to clarifying and helping to answer problems in the theory of knowledge. For instance, his 'Lottery Paradox' (Kyburg, 1961) proved that a contradiction follows from apparently common sense claims about when high probability enables us to accept a hypothesis. Keynes and his Cambridge contemporaries were the subjects of Kyburg's PhD thesis (Kyburg, 1955) and were Kyburg's principal influences in the philosophy of probability, along with his frequentist PhD supervisor, Ernest Nagel. Kyburg saw his theory as similar to

Keynes's (Kyburg, 1995). For an intellectual and personal autobiography, see Bogdan (1982, Part One).

6.1 Interpretation

Kyburg wanted to combine what he saw as insights in both Keynes's theory and frequentist theories: in his view, all true probability statements must *mention* statements about relative frequencies, but probability statements are not *assertions* about relative frequencies. Instead, probability is about logical relations, as Keynes said. The relative frequencies can be empirical (e.g. that a vaccine protects about 95% of people against a disease) or mathematical (e.g. the Central Limit Theorem, Bernoulli's theorem, and other combinatoric principles of statistics). For the estimation of relative frequencies, Kyburg mostly follows frequentist ('classical') statistics (Kyburg and Teng, 2001, pp. 261–7). However, unlike frequentists, he believes in single-case probabilities: for instance, the probability that *this particular* coin toss will land Heads, given the premise that about half of the coin's tosses land Heads, is about 50%.[7] In general, Kyburgian probabilities are relations between (1) premises describing frequencies in populations and (2) a conclusion describing a single member of those populations. The member could be a single event, but also a set of events.

6.2 Semantics

Like Williams, Kyburg aimed to base his theory on direct inference, but this aim raised the Problem of the Reference Class: what if there is ambiguous information in an argument's premises? (See Section 5.) Initially, Kyburg pursued rules for identifying a single appropriate reference class among those described in the premises. The reported frequency or interval-valued estimate for this reference class would be the conclusion's probability. Kyburg ultimately decided that this was impossible. Instead, he thought that such rules could be winnowed down to a *set* of statements (potentially a unit set) plus a rule for combining these statements to generate an interval-valued probability (Kyburg and Teng, 2001, Chapter 9). In some cases, these intervals can be degenerate intervals, e.g. (0.5, 0.5) for a coin landing Heads, given that it is fair. The technical details of Kyburg's rules for reference class selection are too complex to describe here, but these are the core ideas (Kyburg and Teng, 2001):

Conditioning: *Statistical data with more relevant conditions takes precedence over data with fewer relevant conditions, where their intervals conflict.* Suppose if we know both (1) the proportion of red balls in an urn and (2) the joint distribution of our non-random selection process and the red balls' frequency, then (1) should be ignored in favour of (2) for the logical probability of selecting a red ball.

Specificity: *Statistical data about reference classes that we know to be subsets of broader reference takes precedence over the data the broader reference classes, if their intervals conflict.* For example, we know that white bears are a small proportion of the total population of bears, but if we observe a bear-like white object within the Arctic Circle, then it is the approximate frequency of bears among bear-like white objects within the Arctic Circle is relevant for the conjecture that a bear is approaching.

---

[7] Kyburg often talks of 'rational corpora' instead of premises or evidence (Kyburg 1983, p. 209). For simplicity, I shall stick to more standard terminology.

Precision: *Precise statistical data takes precedence over imprecise statistical data, if their intervals do **not** conflict.* By 'conflict', I mean that neither interval is a subinterval—possibly improper—of the other. Imagine that we do not know the relative frequency with which a particular coin lands Heads. Yet we know that, generally, coins land Heads with a frequency close to 50%. We should use the latter data, since it is more informative, so that the probability interval is roughly [0.5, 0.5].

The Kyburgian probability is the narrowest interval covering the relative frequency data that survives the application of Kyburg's formal versions of these principles. Note that **Precision** is consistent with **Specificity**, even though they can favour broader and narrower reference classes respectively, because **Precision** only applies in the *absence* of conflict, whereas **Specificity** only applies in its *presence*. If we had information about the particular coin that conflicted with the general data for coins, the former would have priority.

Kyburg rejected all versions of the POI. In relation to premises containing all logico-mathematical truths, his theory has only these priors:

(i) [1, 1] if the conclusion is a logico-mathematical truth, e.g. a mathematical fact or a tautology.
(ii) [0, 0] if the conclusion is a logico-mathematical falsehood.
(iii) [0, 1] otherwise.

—because these priors follow from direct inference using just information from logic and mathematics. For instance, (iii) follows from the mathematical fact that a unit-valued relative frequency must be between 0 and 1, inclusive.

Keynes and Kyburg both advocated imprecise probabilities, but they differed about when imprecision occurs. In Keynes's theory, precision usually occurs when our evidence is *sparse* and unambiguous, so that his version of the POI applies. In Kyburg's theory, precision usually occurs when our evidence is *rich* and unambiguous, so that we have detailed and non-conflicting statistical evidence.

6.3 Influence

Relationism is crucial for Kyburg's theory, because direct inferences require premises. Keynes also influenced Kyburg's theory of physical probability statements. These statements can be *empirically* wrong. According to Kyburg, physical probabilities are the logical probabilities that would be relevant for a demon-like entity that knew all the *general* facts about frequencies in long-run trials, e.g. the frequency of Tails in tossing a coin to its limiting relative frequency (Kyburg, 1990, p. 50). Such demons are psychologically impossible because they would have infinitely many beliefs. However, Kyburg can consistently adopt this view, due to Keynes's legacy: that the requisite probability relations are abstract logical relations, not psychological facts.

6.4 Priors

Apart from (i–iii) in Section 6.2, Kyburg's theory implies no priors. This austerity avoids the Problem of the Priors, since these intervals do not suggest any unequivocal *a priori* beliefs. However, a supporter of Keynes could reply that there are plausible initial priors, e.g. the probability of a coin landing Heads is 1/2. Here, my own intuitions are with Kyburg. Although 1/2 seems appropriate in ordinary life, we have a lot

of evidence about coins. In contrast, if the subject matter is totally unfamiliar ('What is the probability that a ẟ is ꓕ?'), then something imprecise is perhaps more appropriate, but the resolution of this disagreement is beyond this article's scope.

7. David Stove

Stove (1927–94) is perhaps best known for his idiosyncratic polemics against socialism, feminism, social Darwinism, Karl Popper, Christianity etc. (Stove, 2002). However, much of his academic writing is in the philosophy of probability and induction. Stove is sometimes identified as a 'Carnapian' (Miller 1988, p. 287) and he cited Williams as a major influence (Stove, 1986, pp. 62–3). However, an unpublished paper[8] reveals that he held Keynes in highest regard. Stove describes his theory as 'Keynesian', by which he means the 'conception of logical probability as degree of *conclusiveness*, a property ascribable to, and only to, arguments' (Stove, 1960, p. 3). Like Harrod and Williams, Stove's main interest in probability was due to his interest in Hume's Problem of Induction. I shall put that mostly aside and focus on Stove's philosophy of probability. For biographical details, references and intellectual context, see Franklin's (2003) history of philosophy in Australia, in which Stove is an important and recurring character.

7.1 Interpretation

Stove closely follows Keynes on issues of interpretation, but with some idiosyncratic emphases. One is Stove's insistence that the probabilities of arguments are not *just* relative to their premises, but also to their conclusions (Stove, 1960, pp. 17–8). He cites several reasons for this emphasis: one was that the idea that 'The probability of *H* is relative to *E*' could encourage the idea that *H* had a probability that is dependent on the truth of a contingent *E*. Relationists deny that probabilities depend on *any* contingent facts. One departure from Keynes is that he adopts Carnap's pluralism about probability: there are 'factual' probabilities, i.e. physical probabilities, which he believed require an additional interpretation of probability[9] (Stove, 1960, p. 2). In writing, Stove never committed to a particular theory of factual probability, though he briefly criticised frequentism (Stove, 1986, p. 57). However, this difference is not philosophically very important for Stove's philosophy, since he was mainly interested in factual probability statements as premises in arguments, rather than analysing factual probability's meaning.

7.2 Semantics

Along with his pluralism, Stove's most important departure from Keynes was his semantics for logical probability. He used 'statement of logical probability' (SLP) to refer to an assertion about a probability relation. SLPs can be equalities, such as $P(H \mid E) = 0.5$, but also inequalities like $P(H \mid E) > 0.5$ or categorical statements like '*H* is very probable in relation to *E*'. According to Stove, there are no SLPs that

---

[8] I am grateful to James Franklin, Stove's literary executor, for providing me access to this source.
[9] Carnap was the first *relationist* to be a pluralist, but not the first pluralist, e.g. (Cournot, 1851, Chapters III and IV).

are both (a) non-trivial and (b) of a high level of generality (Stove, 1986, Chapter 9). Here, 'generality' means that the SLPs have no restrictions on any variable terms. For instance:

(i) 'The probability is 1/3 that Abe is a raven in relation to the premise that just one of Abe, Barry, and Charlie are ravens'.

—has a very low degree of generality, while:

(ii) 'The probability is 1/3 that x is an F in relation to the premise that just one of x, y, and z ($x \neq y \neq z$) is an F'.

—has a high degree of generality, since F, x, y and z are so unspecified.

Stove grants that, while there are some very general SLPs, these are largely uninformative claims, e.g. that *some* inductive arguments have a high probability relation. Nor does Stove deny that there are some very informative SLPs, such as (i). However, such informative SLPs are apply only to suitable predicates, such as 'raven', and do not apply for others, e.g. predicates such as 'rog', which I am arbitrarily defining as 'A raven if observed prior to 2020 AD or a frog if observed from 2020 AD onwards'.[10] Consequently, sweeping generalisations such as Williams's principle of direct inference and the POI can only be, at best, useful rules-of-thumb.

7.3 Influence

Keynes's relationism permeates Stove's entire work on probability and reasoning in general. Following Williams, Stove believed Hume's Problem of Induction to be a challenge to establish an SLP: that the premises of *some* inductive arguments confirm their conclusions (Stove, 1986, Chapter III). Stove does not want to justify *all* inductions, because induction can be unreasonable e.g. if we know that our sampling methods are misleading. Another influence from Keynes is Stove's idea of 'misconditionalisation' (Stove, 1960, Sections (ii)–(iv); Stove, 1972). Where p is a deductive logical claim or SLP, and c is a contingent statement, then misconditionalisations have the form 'If c, then p' or 'If p, then c'. The danger of misconditionalisations is that they can lead us to assert false statements when we mean true statements. For example:

(i) 'That all men are mortal and Socrates is a man implies that Socrates is mortal'

—is true, but:

(ii) 'That Socrates is a man implies that Socrates is mortal'

—is false. Yet it is easy to misconditionalise:

(iii) 'If all men are mortal, then that Socrates is a man implies that Socrates is mortal'

—but (iii) is also false. (It has a true antecedent and a false consequent.) Stove explored how misconditionalisation can lead us from truth to falsity.

---

[10] See Goodman (1954, pp. 72–73ff).

Stove argued that misconditionalisations occur in an number of places in philosophy (Stove, 1972). In particular, he argued that the common claim that 'Induction's rationality presupposes that nature is regular' is a misconditionalisation. According to Stove, induction's rationality consists in non-contingent facts about logical probabilities, whereas nature's regularity is contingent (Stove, 1986, Chapter 1). Knowing that nature is regular would be favourable to induction, but that is not the same claim. Stove's reasoning uses part of Keynes's legacy: SLPs are not contingent.

7.4 Priors

Stove never explicitly addressed the Problem of the Priors. However, in all his work, he never asserted statements like $P(H \mid t) = 0.5$. He did assert numerical inequalities such as $P(H \mid t) < 1$ (Stove, 1986, p. 43) but these do not imply strong *a priori* beliefs. Since Stove's overall theory and judgements are very akin to Keynes's, his implicit answer seems quite coherent as a modest modification of Keynes's position in order to address the Problem of the Priors. However, Stove's answer leaves an important question for future researchers: what function should such imprecise initial priors have in non-deductive learning? Standard Bayesian learning methods assume precise priors, so an alternative would be necessary.

8. Conclusion

Kyburg and Stove offer one strategy for Keynesian probabilists to answer the Problem of the Priors: use only imprecise priors for empirical hypotheses. Another strategy, which Keynes develops in the *Treatise* (Chapters VI and XXVI), is to exploit the distinction between logical probability and degrees of belief. Thus, unequivocal initial priors for a hypothesis only warrant cautious beliefs, due to low 'weight of argument' (see Runde, 1990; Brady, 1993; Franklin, 2001, pp. 281–4). The interpretation of 'weight' is controversial, but the basic idea is that strong confidence requires both a strong relation of evidence to hypothesis *and* ample evidence. This epistemologically sophisticated approach to probability and belief is also an attractive tool for understanding uncertainty, as Post-Keynesian economists have long recognised. As we have seen, none of the philosophers that I have discussed has an obviously superior approach to the problem.

I have been able to indicate Keynes's influence without even discussing Carnap, the most prominent and influential example in the philosophy of probability. As this table shows, my sample has been variegated:

	Initial priors?	Principle of indifference?	Informative general SLPs?	Truth-frequency analysis?	Imprecise probabilities?
Keynes	Yes	Yes	Yes	No	Yes
Benenson	No	No	Yes	No	No[a]
Harrod	No	No	Yes	Yes	No
Williams	Yes	Yes	Yes	Yes	Yes
Kyburg	Yes	No	Yes	No	Yes
Stove	Yes	No	No	No	Yes

[a]Except second-order probabilities.

This variety indicates the robustness of Keynes's relationism. Thus, criticisms of a particular theory might not generalise across all theories influenced by him. Keynes's relationist interpretation of probability is currently unpopular among philosophers, but its robustness should make them wary of dismissing it.

Funding

Bibliography

Benenson Capital. 2014. A Century of Experience, http://www.benensoncapital.com/pdf/BenensonBrochure.pdf [date last accessed 21 January 2021]

Benenson, F. C. 1984. *Probability, Objectivity, and Evidence*, London, Routledge & Kegan Paul

Bogdan, R. 1982. *Henry E. Kyburg and Isaac Levi*, Dordrecht, D. Reidel

Brady, M. E. 1993. J. M. Keynes's theoretical approach to decision-making under conditions of risk and uncertainty, *The British Journal for the Philosophy of Science*, vol. 44, no. 2, 357–76

Brady, M. E. and Arthmar, R. 2012. Keynes, Boole, and the Interval Approach to Probability, History of Economic Ideas, vol. 20, no. 3, 65–84

Caldentey, E. P. 2019. *Roy Harrod*, London, Palgrave Macmillan UK

Campbell, K, Franklin, James and Ehring, Douglas. 2019. Donald Cary Williams in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2019 ed., Stanford, CA, Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/williams-dc/ [date last accessed 27 February 2020]

Carabelli, A. 1988. *On Keynes's Method*, London, Macmillan

Carnap, R. 1952. *The Continuum of Inductive Methods*, Chicago, IL, University of Chicago Press

Carnap, R. 1962. *The Logical Foundations of Probability*, Chicago, IL, University of Chicago Press

Cournot. 1851. *Essai Sur Les Fondements de Nos Connaissances Philosophique*, Paris, Librairie de L. Hachette

de Finetti, B. 1964. Foresight: its logical laws, its subjective sources, pp. 93–158 in Kyburg, H. E. and Smokler, H. E. (eds.), *Studies in Subjective Probability*, New York, John Wiley & Sons

Franklin, J. 2001. Resurrecting logical probability, *Erkenntnis*, vol. 55, no. 2, 277–305.

Franklin, J. 2003. *Corrupting the Youth: A History of Philosophy in Australia*, Sydney, Australia, Macleay Press

Gillies, D. 2006. Keynes and probability, pp. 199–216 in Backhouse, R. E. and Bateman, B. W. (eds.), *Cambridge Companion to Keynes*, Cambridge, Cambridge Companion to Keynes

Goodman, N. 1954. *Fact, Fiction, and Forecast*, London, The Athlone Press

Harrod, R. 1942. Memory, *Mind*, vol. 51, no. 201, 47–68

Harrod, R. 1956. *Foundations of Inductive Logic*, London, Macmillan

Harrod, R. 1961. The general structure of inductive argument, *Proceedings of the Aristotelian Society*, vol. 61, 41–56

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

Kyburg, H. E. 1955. *Probability and Induction in the Cambridge School*, Ann Arbor, MI, University Microfilms

Kyburg, H. E. 1983. *Epistemology and Inference*, Minneapolis, MN, Minnesota University Press

Kyburg, H. E. 1990. *Science and Reason*, New York, Oxford University Press

Kyburg, H. E. 1995. Keynes as a philosopher, pp. 7–32 in Cottrell, A. F. and Lawlor, M. S. (eds.), *New Perspectives on Keynes*, Durham and London, Duke University Press

Kyburg, H. E. and Teng, C. M. 2001. *Uncertain Inference*, Cambridge, Cambridge University Press

Lawson, T. 1985. Uncertainty and economic analysis, *Economic Journal*, vol. 95, no. 380, 909–27

Lawson, T. 2003. Keynes's realist orientation, pp. 159–69 in Runde, J. and Mizuhara, S. (eds.), *Perspectives on the Philosophical Underpinnings of Keynes's Economics: Probability, Uncertainty and Convention*, London and New York, Routledge

McCann, C. R., Jr. 1994. *Probability Foundations of Economic Theory*, London, Routledge

Miller, D. 1988. The rationality of induction by D. C. Stove, *Philosophy*, vol. 63, no. 244, 286–88

New York Times. 1978. *Donna Mileti Plans To Be Wed Aug. 5*. New York, New York Times, July 2, 36

O'Donnell, R. M. 1989. *Keynes: Philosophy, Economics and Politics, The Philosophical Foundations of Keynes's Thought and Their Influence on His Economics and Politics*, London, Macmillan

Popper, K. R. 1957. The propensity interpretation of the calculus of probability and of the quantum theory, pp. 65–70 in Körner, S. (ed.), *Observation and Interpretation*, New York, Buttersworth Scientific Publications

Rowbottom, D. 2015. *Probability*, Cambridge, Polity Press

Runde, J. 1990. Keynesian uncertainty and the weight of arguments, *Economics and Philosophy*, vol. 6, no. 2, 275–92

Runde, J. 1994. Keynes after Ramsey: in defence of *A Treatise on Probability*, *Studies in History and Philosophy of Science*, vol. 25, no. 1, 97–121

Salmon, W. C. 1967. *The Foundations of Scientific Inference*, Pittsburgh, PA, University of Pittsburgh Press

Stove, D. C. 1960. *Bernoulli's Theorems and Logical Probability*, unpublished.

Stove, D. C. 1972. Misconditionalisation, *The Australasian Journal of Philosophy*, vol. 50, no. 2, 173–83

Stove, D. C. 1986. *The Rationality of Induction*, Oxford, Oxford University Press

Stove, D. C. 2002. *On Enlightenment* Irvine, A. (ed.), New Brunswick and London, Transaction

von Mises, R. 1951. *Positivism: A Study in Human Understanding*, Harvard, Harvard University Press

Williams, D. C. 1947. *The Ground of Induction*, Cambridge, MA, Harvard University Press

Williams, D. C. 2018. How reality is reasonable, pp. 80–103, in Fisher, A. R. J. (ed.), *The Elements and Patterns of Being: Essays in Metaphysics*, Oxford, Oxford University Press

Williamson, J. 2010. *In Defence of Objective Bayesianism*, Oxford, Oxford University Press

Williamson, T. 1998. Conditionalizing on knowledge, *The British Journal for the Philosophy of Science*, vol. 49, no. 1, 89–121