



Rational inference: deductive, inductive and probabilistic thinking
Author(s): RJ Ormerod
Source: *The Journal of the Operational Research Society*, Vol. 61, No. 8 (August 2010), pp. 1207-1223
Published by: Palgrave Macmillan Journals on behalf of the Operational Research Society
Stable URL: https://www.jstor.org/stable/40802205
Accessed: 05-11-2019 22:10 UTC

REFERENCES
Linked references are available on JSTOR for this article:
https://www.jstor.org/stable/40802205?seq=1&cid=pdf-reference#references_tab_contents
You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms



Operational Research Society, *Palgrave Macmillan Journals* are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of the Operational Research Society*

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

Journal of the Operational Research Society (2010) 61, 1207–1223 © 2010 Operational Research Society Ltd. All rights reserved. 0160-5682/10

www.palgrave-journals.com/jors/

Rational inference: deductive, inductive and probabilistic thinking

RJ Ormerod*

University of Warwick, Coventry, UK

The paper reviews three modes of rational inference: deductive, inductive and probabilistic. Many examples of each can be found in scientific endeavour, professional practice and public discourse. However, while the strengths and weaknesses of deductive and inductive inference are well established, the implications of the emerging probabilistic orientation are still being worked through. The paper discusses some of the recent findings in psychology and philosophy, and speculates about the implications for scientific and professional practice in general and OR in particular. It is suggested that the probabilistic orientation and Bayesian approach can provide an epistemological lens through which to view the claims of different approaches to inference. Some suggestions for further research are made.

Journal of the Operational Research Society (2010) **61,** 1207–1223. doi:10.1057/jors.2009.96
Published online 26 August 2009

Keywords: philosophy of science; philosophy of mathematics; philosophy of statistics; method of OR; Bayes's theorem

Introduction

Operational research (OR) consultants help clients draw inferences relevant to the situation they face and the decisions they have to make. In doing so, they lay claim to rationality, drawing on the ethos of science and the language of mathematics. Deductive logic and mathematics have, since the Greeks, been taken to be the corner stone of rational thinking. Descartes, for example, built his entire philosophy on the deductive method. However, the great success of science in the 18th and 19th centuries was built on observation, empirical testing and inductive reasoning; the scientist themselves often described their activities in terms of inductive thinking following Francis Bacon and John Locke. However, David Hume, the 18th century philosopher, argued convincingly that induction could not be justified. In the 20th century, Popper, in his seminal work, *The Logic of Scientific Discovery* (Popper, 1934; 1959), sought to solve the 'problem of induction' by avoiding it altogether. In doing so, he challenged the view that science progressed by accumulating knowledge through inductive reasoning from empirical observations. His insight that progress in science could be understood in terms of empirically falsifying erroneous ideas rather than confirming correct ones was based on the application of deductive logic rather than the philosophically flawed inductive logic. His ideas both gained support and met opposition in scientific and philosophical communities across the world (Ormerod, 2009). It gave rise to a wide ranging debate drawing in contributors from many other disciplines including historians and social scientists.

More recently, a number of academic disciplines have challenged the primacy of logical deductive thinking. In psychology and cognitive science, researchers have concluded that deductive logic provides an inadequate explanation of the way that people think; in computer science and artificial intelligence, researchers have concluded that deductive logic cannot reproduce the ability of humans to respond to their environment; in the philosophy of science, researchers have concluded that theories of scientific method based on deductive logic fail to provide an adequate explanation or normative guidelines. A number of researchers in these areas (and others including, of course, statistics) have concluded that the way forward is to replace deductive logic with probabilistic thinking; they have taken a 'probabilistic turn'. To support this probabilistic approach researchers have increasingly employed Bayes's theory. Although the Bayesian approach can be used with objective probabilities, it is when it is combined with a subjective epistemological approach that researchers find that Bayesian analysis sheds light on a wide range of previously intractable problems. According to Howson and Urbach (1989, p 10) Bayesianism has experienced a strong revival among statisticians and philosophers in recent years, due, in part, to an intrinsic plausibility and also to the weaknesses that have gradually been exposed in the standard methodologies.

Analysis (quantitative and qualitative) is one of the core competences of OR practice (Ormerod, 2002). The ability to draw inferences from statements, observations and data is one of the key elements of that competence (see Figure 1).

*Correspondence: RJ Ormerod, 33 Normandy Hill, Plymouth PL5 1LF, UK.
E-mail: richard@rormerod.freeserve.co.uk

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

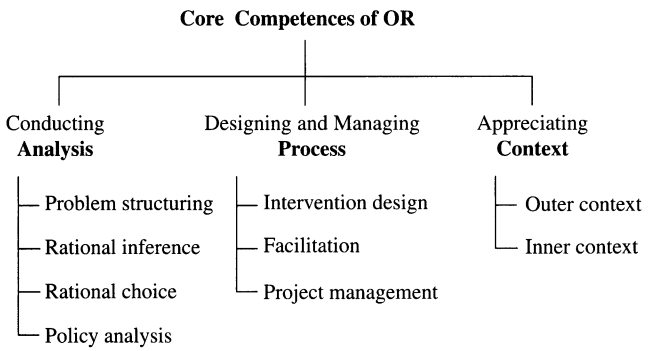


Figure 1 OR competences.

OR practitioners seek to draw inferences rationally in the light of uncertainty. This article reviews the arguments and findings generated by those researchers who have taken a probabilistic turn, and examines their relevance for OR practice and research. The first section of this paper characterises science and professional practice in terms of values, rules and aims. The next three sections explain some of the characteristics, strengths and weaknesses of the three modes of inferential reasoning: deductive, inductive and probabilistic. The following section looks at some intractable problems of standard inductive reasoning from a Bayesian perspective. Finally, some of the implications for OR are discussed and conclusions are drawn. Various versions of Bayes's theorem are given in the Appendix.

Scientific method and professional practice

Scientific method refers to the values that scientists pursue, the rules of method they adopt and the logic they deploy; it is the means that scientists adopt to achieve their aims. Individual scientists might have all sorts of aims that are extrinsic to science, such as forwarding their career, satisfying curiosity, increasing man's power over nature, contributing to employment and wealth and combating medical ailments, or the pursuit of humanitarian goals, such as flood or famine control. However, here, it is the intrinsic aims, values, rules and logic of the conduct of science that concern us; as Nola and Sankey (2007, p 34) put it, 'our main concern is what values, if any, are involved in the epistemological task of critically evaluating the truth or evidential support of hypotheses'.

The values and rules of science

Philosophers use various terms to refer to values. Discussion of values is part of the general theory of *axiology*, which includes accounts of the *aesthetic* and *moral* values as well as the *theoretical* values (values pertaining to theories) that our sciences and belief systems ought to exhibit. Values can be the intellectual and cognitive virtues that people ought to show in their actions or the theoretical virtues that our sciences and belief systems should embody. Theoretical values can be considered as coming in two broad varieties: epistemic and pragmatic (alternatively some authors, such as Hilary Putman (2003) take pragmatic values to be a subset of epistemic values). *Epistemic* values guide the choice made between theories according to their truth or justification. There are two species of epistemic values. In the first species, truth is the prime *alethic* value. In the second, the evidential justification concerns their *confirmatory* virtues. Some reject the involvement of alethic values, or more generally epistemic values, advocating instead *pragmatic* virtues, such as a theory's usefulness, beauty, simplicity, scope, unification, explanation and so on. van Frassen observes, 'In so far as they [pragmatic virtues] go beyond consistency, empirical adequacy, and empirical strength, they do not concern the relation between the theory and the world, but rather the use and usefulness of the theory; they provide reasons to prefer the theory independently of questions of truth.' (Nola and Sankey, 2007, pp 35–56).

Philosophers of science have offered many possible values (virtues, desiderata, aims and so on) of science, their lists often being overlapping. Carl Hempel suggests that a theory should yield precise, preferably quantifiable predictions, should be accurate in terms of experimental tests, should be consistent both internally and with currently accepted theories in neighbouring fields, should have broad scope, should predict phenomena that are novel, should be simple and should be fruitful (Fetzer, 2001, pp 384–385). Nola and Sankey (2007, pp 37–45), drawing on Willard Quine and Joseph Ullian (1978), give five theoretical virtues that they claim any scientific hypothesis ought to exemplify: conservativeness, modesty, simplicity, generality and refutability (explanatory and predictive power are presupposed). To Quine and Ullian's list, Nola and Sankey add a sixth virtue suggested by others, namely consistency. Kuhn (1962, pp 184–186) lists some of these along with others including: accuracy, explanatory and predictive power, ability to form puzzles and provide solutions, a preference for quantitative over qualitative, scope and fruitfulness. Further, Kuhn claims that although his list of values can be shared by all scientists this does not ensure that all scientists will make the same choice; scientists can understand the values differently and can also attribute a different relative weight to each value. In other words, scientists can opt for both different scores and weights, both across paradigms and within a paradigm.

There is reasonable agreement about the high-level aims of science. For example, Popper (1972) says that it is to give satisfactory explanations that also display increasing depth of explanation. Kuhn suggests the overall aim of a scientist is to understand the world and extend the precision and scope with which it has been ordered. However, it is in the more detailed descriptions that the differences show up. Positivists saw empirical science as a system of statements that satisfy certain logical criteria, such as meaningfulness or verifiability. In contrast, Popper (1959, pp 49–50) aligned himself with 'those who tend to see ... the distinguishing characteristic of empirical statements [scientific theories] in their susceptibility to revision—in the fact that they can be criticized, and

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

superseded by better ones; and who regard it as their task to analyse the characteristic ability of science to advance the characteristic manner in which a choice is made, in crucial cases, between conflicting systems of theories'. Indeed, it is difficult to imagine an approach to science that did not involve critical evaluation. Pierre Duhem (1954) contrasts his theory of explanation with the aim of the realists; realists aim to develop theories to explain experimentally established laws in order to get closer to bare reality itself. Although we might feel a pull towards the realists' direction, Duhem argues that it cannot be realized and lacks a clear criterion to tell us when we have realized it. He, therefore, drops the requirement that the fundamental hypotheses be true of hidden reality. He argues that agreement with experiment is the sole criterion of truth for physical theory: the aim of science is empirical adequacy. This approach led Duhem to adopt the value 'economy of thought' supported by the values of unification and unity, which are close to the Quine–Ullian virtues. Duhem suggests that the pragmatic values of beauty and elegance can be realized through unity, generality and systematicity of classification (Nola and Sankey, 2007, pp 52–55).

The values and rules of professional practice

In moving from scientific to professional practice, individuals will again differ in the weight they place on particular values. In general, it is likely that more weight will be placed on the practical values in professional practice than in science. However, there is still an attachment to epistemic values, in many cases a strong one. The extent of the differences is related to the fact that the natural and social sciences seek to achieve an understanding of nature and society: getting close to the truth is highly desirable. On the other hand, professional practice is only concerned with the instrumental, practical purposes of prediction and control; theories that provide satisfactory predictions will be good enough.

Much professional practice has embedded within it science-like activity; for example, medical science within medicine, materials science within engineering and forensic science within criminal investigation. One of the differences between the different types of professional activity is the timescale and resources available for investigation. Another is the opportunity for critical scrutiny by other professionals. In the case of OR, the consultant will be working to deadlines with limited resources and there may be limited opportunity for critical scrutiny. In such situations, shortcuts may be tempting; experience and self-criticism become the main control mechanisms.

Deductive inference

Deductive logic refers to arguments that are certain to be true by definition. It includes philosophical logic and mathematics. No empirical evidence is required for a proof. No inductive reasoning (see below) is involved. Proofs of theorems within Euclidean geometry (the angles in a triangle add to 180 degrees) are examples of deductive thinking.

Philosophical logic

Historically, rationality has been associated with deductive or philosophical logic (notably by Descartes and French philosophers). Until relatively recently, this meant the logic system of Aristotle (who drew on the work of Socrates, Zeno, Plato and others). Aristotle's great innovation in logic was the syllogism, which he originally defined as 'discourse in which, certain things being posited, something else necessarily follows'. It came to have the narrower meaning typified by 'All men are mortal (the first premise); Greeks are men (the second premise or the middle term); therefore Greeks are mortal (the conclusion, logically entailed by the first two premises)' (Williamson, 1995). The study and development of logic continued sporadically down to the 19th century when George Boole, Gottlob Frege, and later, Bertram Russell, Alfred Whitehead and Ludwig Wittgenstein developed the formal relationships between logic, language and mathematics. The logical positivists (centred on the Vienna Circle, which included Otto Neurath, Moritz Schlick, Rudolf Carnap and Kurt Gödel) continued to develop the theory. After the shock of Gödel's incompleteness theorem, logicians turned their attention to the notions of computability and effectiveness. During the 1930s, a number of characterizations of computability were developed by, among others, Gödel (recursiveness), Turing (the Turing machine) and Markov (the Markov algorithm). Logic formed the backbone of the philosophy of many of the major figures in the second half of the 20th century, including Willard Quine, Donald Davidson and Michael Dummett (King and Shapiro, 1995).

Deductive inference is the gold standard of rational thinking, but it can only tell us about the truth of the deductive process; it has nothing to say about the truth of the premises. However, deduction can be used to examine the relationship between universal theories and empirical observation by employing the hypothetico-deductive (H-D) method.

The H-D method and falsification

The H-D method has a long history going back to Plato who referred to it by various names including the 'method of hypotheses'. Karl Popper, although he was not the first to do so (eg, the method was central in the 19th century debate between William Whewell and John Stuart Mill), placed much emphasis on the use of the method to establish the validity of scientific theories in the light of experimental evidence. The method starts with a hypothesis (a theory), the consequences of which are logically deduced. These consequences are then compared with the evidence. For example, a unifying theory (such as Newton's laws of motion) can be used to predict a variety of consequences, which are then compared with empirical laws (which summarize the observations). It might be thought that if the evidence agrees with the consequences,

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

the hypothesis would be confirmed. But, this is not the case as it is an example of what philosophers refer to as 'the fallacy of confirming the consequent'; other hypotheses could also explain the same evidence. However, if the consequences are inconsistent with the evidence, the hypothesis is falsified. Popper made much of this asymmetry between (inductive) confirmation and (deductive) falsification. He argued strongly that by concentrating on falsification, one did not need to rely on induction. However, there are some difficulties with this approach.

In practice (and indeed in logic), there are many possible explanations for any inconsistencies that arise between consequences and evidence; it may not be that the hypothesis is erroneous. First, there may be problems with the method used to collect the evidence, faulty apparatus, for example. However, such problems apply equally to any other proposed method. Second, there could be something wrong with the *auxiliaries* that are employed in the application of the hypothesis. Auxiliaries are any factors that are assumed by the hypothesis under scrutiny, for example, initial conditions and background theories. The H-D method can only indicate that there is a problem with either the hypothesis or with the auxiliaries, but it cannot distinguish where the problem lies. The Quine–Duhem thesis states that there are always some implicit auxiliaries, and therefore, no falsification can be decisive. Some believe that the Quine–Duhem thesis puts paid to Popper's falsification approach; others argued that the objections raised by Quine–Duhem can very often be met. Popper recognized that experiments actually test a system of interrelated theories rather than an isolated hypothesis. Thus, a convincing falsification ought to bring the whole system into question. However, if the system in question is well established, attention turns back to the hypothesis itself and the reliability of the experimental evidence.

One way to respond to a negative result is to introduce an *ad hoc* qualification to the hypothesis. Thus, if the hypothesis 'all swans are white' is falsified by observations of black swans in Australia, the hypothesis could be amended to, for example, 'all European swans are white'. Popper argues that *ad hoc* hypotheses should be avoided as they weaken the universality of scientific statements and immunize hypotheses from empirical testing; he proposed the principle of parsimony in the use of *ad hoc* hypotheses. In addition, Popper argues that scientists should adopt a number of methodological rules and stratagems, including making tests as severe as possible, aiming for the highest level of universality and precision and avoiding 'psychologism' (statements that involve belief, expectation, subjective views and the like).

Despite an impressive defence of his logical position, Popper had to address the fact that in the end it is the community of scientists who collectively, by a process of critical debate, determine whether a theory or system of theories (or empirical data used to test a theory) can be accepted into the scientific canon to be considered (for the time being) as knowledge. Scientists bring to the task a variety of beliefs and thus psychologism enters the picture. However, the critical process undertaken by the community of scientists is taken by Popper and his followers to iron out the biases and preferences of individuals over time to arrive at an objective view. Thus, psychologism is driven out again.

Despite its merits, the H-D method is considered lacking in certain respects. First, it lacks a confirmatory mechanism, a theory of confirmation; one has to be added to it. Second, it has a problem in distinguishing between alternative hypotheses both of which may be consistent with the experimental evidence. Third, there is what is referred to as the 'tacking problem' in which an irrelevant hypothesis is added to the theory (eg, a hypothesis in Bible scholarship about who authored, what chapter could be tacked on to a theory in physics); the tacked on theory picks up confirmation (if the H-D method is deemed to provide confirmation) from evidence relating to the main theory. Clearly, this is an unsatisfactory outcome. Finally, an objection often raised is that it has no way of dealing with statistical hypotheses or any non-deductive inference (Nola and Sankey, 2007, pp 172–179).

Deductive inference and professional practice

All professional practitioners make as much use of deductive logic as they can. After all, to be deductively illogical is simply to be wrong. Lawyers construct and deconstruct legal documents; accountants apply double-entry bookkeeping; engineers analyse structures. Of course, deductively logical thinking is likely to be involved in many other day-to-day professional activities, such as organizing activities, gathering evidence or data, writing reports and arguing a case.

The same can be said for operational researchers, but their specialist deductive expertise lies in their modelling activities, in Linear Programming (LP), simulation, critical path analysis, business models, conceptual models and so on. Given the premises, the models deduce the logical conclusions. However, that is all they do; such models may well serve all sorts of useful purposes, but the conclusions they reach are entirely dependent on the premises they are based on: the parameters, the constraints, the objectives, the costs, the times to complete a task and so on. It may well be that the operational researcher's activities include establishing the premises required, but this cannot be based on deductive logic. Of course, operational researchers would like it to be otherwise, so that their results could be certain beyond criticism. Unfortunately, nearly all OR activities involve inductive logic, a much less certain enterprise.

Inductive inference

Induction is usually defined as inference from the particular to the general. In other words, it involves deriving some conclusion from observations or a set of data. Thus, if all the swans I have observed are white, I might inductively conclude that all swans are white. If I observe that absenteeism has over the

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

past 6 months at a particular factory been 20% higher on a Monday than on a Tuesday, I might inductively infer that this is generally the case and absenteeism at the factory on future Mondays might be expected to be higher than on Tuesdays.

The idea that science involves induction can be traced back through Francis Bacon's *Novum Organon* to Aristotle's *Organon*. Isaac Newton's rules for the conduct of science also included induction. But, David Hume formulated an argument that showed that inductive inference lacks justification. However, scientific endeavours have undoubtedly involved inductive inference, and it would be an extreme version of empiricism to hold that nothing can be said beyond the observations made. From the necessarily limited data obtained, scientists want to draw inferences about the state of affairs, past, present and elsewhere, which have not been directly observed. They want to infer laws and theories that give insight into how the world works. They need to bring observations into contact with their theoretical conjectures. As has already been described, one approach, championed by Popper, is to rely on deduction alone to weed out those conjectures that are falsified by the evidence; any inductive reasoning is avoided. The other approach is to allow the use of inductive modes of inference to evaluate conjectures in the light of test evidence despite the lack of justification; statistical methodology and probabilistic calculus formalize the inductive approach. Inductive inference is thus a matter of weighing evidence and judging probability, not proof.

Valid deductive inferences are 100% reliable. If the premises are true, so are the conclusions; it is logically impossible for the conclusion to be false. A deductive inference gives the best guarantee of truth. When a conclusion does not deductively follow from the premises, it is deductively invalid. However, invalid arguments are not without worth as they may provide some support for the conclusion; for example, we might be able to conclude that although the argument is logically invalid, it is *improbable* for the premises to be true and the conclusion false. This formulation suggests that a relation of inductive probability could be defined. Thus, we can formulate the statement $P(C|A)$ to read as 'the (inductive) probability of (conclusion) C given (premises) A'. The premises A serve as evidence for C. P can, therefore, be described as the conditional or posterior probability of C given A. P can be interpreted in terms of probability calculus with $P = 1$ indicating logical consequence; $P = 0$ indicating A contradicts C; and a value between 0 and 1 indicating varying degrees of probabilistic support that A gives C.

Scientific experiments are conducted in the presence of background evidence B. We can say that evidence E increases the confirmation of hypothesis H, if, and only if, $P(H|E\&B) > P(H|B)$. Tentative rules for inductive acceptance (theories are detached and accepted into the corpus of scientific knowledge) can be formulated; for example, if $P(E|H) > 1/2$, the theory is accepted.

We are naturally in the habit of reasoning inductively and attempts have been made to formulate the implicit underlying rules:

The inductive rule for enumerative induction: From information (observational or otherwise) that n instances of P are Q, infer all Ps are Qs (or the next P is a Q).

The rule means that no P was observed not to be Q, and requires that a wide variety of cases of P should be observed. Similarly, we can generalize statistical results:

The inductive rule for statistical generalization 1: From n per cent observed Ps are Qs, infer n per cent of all Ps are Qs.

If n is 100%, then the rule for enumerative induction above is a special case of the statistical rule. The rule is, of course, subject to error; this leads to an improved version of the rule:

The inductive rule for statistical generalization 2: From n per cent observed Ps are Qs, infer $(n \pm \varepsilon)$ per cent of all Ps and Qs (where ε is the error factor).

The determination of the term $\pm\varepsilon$ is addressed by statistical theory.

Although the problem of justifying induction has proved intractable, philosophers have put a good deal of effort into describing our inductive thinking processes. Popular attempts include Hume's idea of induction, namely more of the same, also called conservative induction. However, this does not capture those instances when one is trying to infer that things are going to change. Nor does it cover those instances when one is trying to infer unobservable entities and processes from scientific evidence. It is also too permissive: if we observe only clean-shaven OR consultants, it is not valid to infer that all OR consultants are clean shaven. Despite its limitations, conservative induction does seem to capture our propensity to extend observed patterns, and something like it must play a role in both everyday and scientific inferences (Lipton, 2004, pp 14–15). Another account of inductive support is given by the H-D model (see above) in which a hypothesis is inductively supported when it deductively entails a piece of evidence (datum). The model has the merit that it does seem to reflect scientific practice and is supported by scientists themselves. But, again it is argued by philosophers that the model is both too permissive in some cases and too strict in others (Lipton, 2004, pp 15–16).

Another account of induction focuses on causal induction. As Peter Lipton (2004) says, it is a striking fact about our inductive practice (common sense and scientific) that so many of our inferences depend on inferring from effects to their probable cause; doctors infer from symptoms to disease; car mechanics from engine noises to faults; scientists from data to theoretical explanation. John Stuart Mill suggested that there are two central methods of experimental enquiry: the method of agreement and the method of difference. In the method of agreement, if we find that there is only one antecedent that is shared by all the observed instances of an effect, we infer that it is a cause (each time I drink more than usual, I get a hangover; hangovers are caused by drinking too much).

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

In the method of difference, if we observe two cases where the prior sets of conditions are identical in all respects bar one condition, say *A*, and the results are different in the two cases, we, infer that *A* caused the difference in the results (a recipe for bread is successful, but the next time, when the yeast is left out by mistake, the bread does not rise: yeast causes bread to rise). Of course, these inferences can be challenged on various grounds, but the disputes can at least be clearly formulated. The methods are not in themselves overly permissive and they help bring out the roles in our inferences of competing hypotheses, negative evidence and background knowledge. To be useful, the methods need to be expanded to suggest the principles for selecting from multiple possible agreements or similarities those that are likely to be causes (Lipton, 2004, pp 18–19).

Description of how we draw inferences takes us only so far. Once we have made an inference, we may use it for practical purposes or we may want to learn more about the way the world is; inferences are a source of explanation. An explanation answers the question why; another way of putting it is that explanation addresses the question, what has to be added to knowledge (based on inferences) to yield understanding. A common, but controversial, form of inductive inference is generally known as *inference to the best explanation* (IBE), a term introduced by Gilbert Harman. This holds that other things being equal, if theory T explains more or better than T*, then this is a reason for choosing T over T*. Some hold that better explanation is not just a virtue but is one of the aims, or even the overriding aim, of science. If we accept some facts as premises and we accept that a theory T explains these facts much better than any available rivals, it can be inductively concluded that either T is true; or it is reasonable to believe T is true; or it has greater verisimilitude; or it has greater probability than its rivals: or T ought to be accepted for whatever purpose. A rule such as IBE is opposed by those who think that explanatoriness is a pragmatic virtue of a theory that is not linked to the alethic value of truth. IBE (sometimes previously called abduction) has long been advocated in the sciences (by the pragmatist CS Peirce among others). Today, there is a lot of literature in the fields of artificial intelligence and theories of computation that explores IBE and its application in different areas from medical diagnosis to jury deliberation (Nola and Sankey, 2007, p 119). Lipton (2004) has provided a fuller statement and defence of IBE. The reliability of the approach has been strongly criticized by Laudan (1984) and van Fraassen (1980; 1989).

Musgrave (1988, p 239; 1999, pp 283–286) has made an interesting observation. He first summarizes the IBE approach as:

> F is a fact.
> Hypothesis H satisfactorily explains F.
> No available competing hypothesis explains F better than H does.
> Therefore, it is reasonable to accept H as true.

This line of reasoning, by including such words as satisfactory and reasonable, is inductive rather than deductive. However, the reasoning can be made deductive by including the additional statement:

> It is reasonable to accept a satisfactory explanation of any fact, which is also the best available explanation of the fact, as true.

If one simply accepts this epistemic premise, then one can behave as though all IBE argumentation is logical so long as one remembers to declare this all important premise alongside the conclusion to condition it. A similar strategy can be adopted for any mode of inductive thinking. In effect, in Kant's scheme of the things, the would-be inductive analyst adopts the epistemic premise above as a transcendental *a priori* assumption.

Inductive inference and professional practice

Most professionals appeal to inductive inference in some aspect of their activities; criminal lawyers have to interpret forensic evidence; medical practitioners interpret symptoms to make their diagnosis; and engineers conduct tests to establish the strength of the materials they use. Inductive inference, of course, lies at the heart of the statistics profession. More informally, the type of thinking that involves concluding with less than full certainty that one thing causes another, is common in professional thinking.

OR practitioners use inductive thinking frequently in their day-to-day work, for instance, in identifying what to look at, what to analyse. In the case of more formal applications, surveys of OR practice have shown that statistical techniques and forecasting are among the most frequently used of all techniques used (eg, Munro and Mingers, 2004, p 91).

Probabilistic inference

An alternative approach to induction is to express the uncertainty of an outcome in terms of the probability that it will occur in the light of the available evidence. This allows a range of problems to be addressed (eg, in the medical, legal and historical fields), and a wide range of evidence to be introduced. In the form of the subjective Bayesian approach, account can be taken of a person's prior expectations as well as data and other evidence. For example, in the OR domain, forecasts of product demand can combine data as it comes in with a manager's prior expectations. A number of the problems introduced in earlier sections of the article can be addressed by viewing classical deductive and inductive approaches from a probabilistic perspective.

Inductive inference looks to statistical methods and probabilistic analysis to operationalize its approach. However, there is psychological research evidence to show that probabilistic thinking plays a more fundamental role in the way we think, in the way we understand the world and our activities in it. Increasingly, philosophers, psychologists and others have been exploring the possibility that probability provides, not

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

just operational convenience, but an epistemological approach in its own right.

Psychology, human reasoning and probability theory

Since the time of Aristotle and Euclid, thinking rationally has been taken to be synonymous with mathematical and philosophical logic. To be rational is to think and act logically. Such a view was underpinned by Noam Chomsky (1957, 1965), who showed how the syntactic structure of language could be represented as a deductive logical system from which all the grammatical sentences of the language could be generated. Oaksford and Chater (2007) comment:

> Although the implications of this viewpoint for the mental representation of language in the brain have remained contentious, a natural interpretation is that the language processing system involves a logical component, which animates these abstract linguistic rules to construct syntactic structures for specific sentences during language comprehension and production (eg, Shieber *et al*, 1995). Moreover, in the psychology of adult reasoning, this logical conception of mind was again used as the foundation for explaining human thought.

Attempts in psychology to understand and model human reason have generally been based on logical deductive inference. That actual patterns of thinking have been shown to fail to display adherence to deductive logic has been attributed to errors, misunderstanding and human fallibility. The heart of standard cognitive psychology is empirical testing. Psychologists submit their theories and models of how our minds work, of how we reason, to laboratory experiments; people are set carefully selected tasks and their performance measured. The traditional approach to analysing the data is to speculate about the causal mechanisms of the mind, an approach that ranges from cognitive psychology to neuroscience. A second more recent approach is referred to as 'rational explanation'. Rational explanation attempts to explain why the mind works in a particular way by showing this is a 'good', 'rational' way to think from some point of view. Rational explanation includes common-sense folk explanation (eg explaining why a person puts up an umbrella, because she believes that it is going to rain, does not want to get wet and so on), but it also includes evolutionary explanation, rational choice theory in economics and the social sciences, and many aspects of psychological explanation, particularly in the applied domains (eg, cognitive therapy in clinical psychology). '[T]he study of rational explanation has also been intense in disciplines which are fundamentally concerned with specifying normatively justified means of reasoning – including philosophy, logic, probability theory, statistics, operations research and machine learning.' (Oaksford and Chater, 2007, p 284)

In the past, the results of applying the notion of rational explanation to individual behaviour have been disappointing; experimental subjects have proved to be not very good at deductive reasoning, the assumed basis of rational thought. Oaksford and Chater (2007) have revisited these experimental results and tested them against their own probabilistic model. They show that actual choices are much better explained by an assumption of probabilistic reasoning. They conclude that, although humans seem to be no more capable of formal probabilistic reasoning than formal deductive (logical) reasoning (people are generally not good at articulating how they think), their common-sense thinking does seem to reflect probabilistic rather than deductive reasoning. It cannot necessarily be concluded that the mechanisms of the mind are built to work this way, though this is not ruled out (multiple heuristic mechanisms are one possibility). What they have done is provide a new 'rational explanation' based on probabilistic rather than deductive reasoning. On speculating as to how people actually think, they 'assume that the basic process of human knowledge representation and common-sense are likely to use both logical and probabilistic machinery in an integrated way' (p 288). However, the authors observe that almost all cognitive processes, from categorization and memory to learning and motor control, seem to crucially involve dealing with uncertainty. They conclude:

> Almost every aspect of cognition involves uncertainty; and the spectacular success of human reasoning suggests that the brain has developed remarkably effective methods for dealing with uncertainty. Understanding how this is possible requires understanding how human cognition can be provided with a rational analysis: and the only machinery that we have to think effectively about this problem is, we suggest, probability theory (Oaksford and Chater, 2007, p 291).

Common-sense and scientific reasoning are often considered by philosophers as part of the same general project of understanding the world (eg, Goodman, 1951; Quine, 1960; 1990). The biologist TH Huxley famously said, 'Science is nothing but trained and organized commonsense, differing from the latter only as a veteran may differ from a raw recruit: and its methods differ from those of common sense only as far as the guardsman's cut and thrust differ from the manner in which a savage wields his club' (*Collected Essays*, 1893–1894). It should be of no surprise, therefore, that probability theory provides insight into the methods of science as well as common-sense everyday reasoning.

The philosophical basis of probability theory

The mathematical theory of probability is usually taken to begin with a correspondence between Pascal and Fermat in 1654. Jacob Bernoulli's *Ars Conjectandi* (1713) contained most of the ideas of what is now referred to as classical theory, but it was written in Latin. It was, therefore, Laplace's *Essai Philosophique sur les Probabilités* (*Philosophical Essay on Probabilities*), published in French in 1814, that became the standard exposition of the classical interpretation of probability during the 19th century (Gillies, 2000, p 3). Between 1654 (Pascal and Fermat) and 1814 (Laplace), the development of mathematical analysis of probability was primarily

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

motivated by an interest in gambling, and the principal practical success of the theory was in applications to gambling. Attempts were made to apply the new understanding to such subjects as the ratio of male-to-female births, whether it was worth the risk of inoculation against smallpox, and the question of life expectancy and the appropriate value for annuities. However, success in these subject areas was limited and seemed to have little impact on practice (Gillies, 2000, pp 8–10).

The classical theory of the 19th century was developed in the context of the then current views about the natural world. As part of enlightenment thinking and the science of Newton, it was believed that the world was an entirely deterministic system; if one was clever and knowledgeable enough, one could calculate future events from initial conditions according to natural laws (Laplace's demon). If, as humans, we do not know enough about the laws or the initial conditions, we may not be able to make these predictions with certainty and must, therefore, have recourse to probability; the world itself was not taken to exhibit probabilistic features, it was deemed to be deterministic; probability was a measure of human ignorance. In this state of human ignorance, Laplace argued that 'nothing induces us to believe that one of them [events] will occur rather than the others', we must regard the events as equally possible. Thus, the calculus of probabilities can only be applied where we have a number of equally probable cases (eg, an unbiased die). Suppose there are n such cases and m of them are favourable to the outcome A. Then the probability of A is defined to be:

$$\text{Prob}(A) = m/n$$

This is the classical definition of probability based on equal possible cases (Gillies, 2000, p 17). Such an approach has severe and obvious limitations; it cannot even deal effectively with the case of a biased die. Few, if any, today hold to classical theory. The subject of probability has matured such that there is now broad agreement on the mathematical analysis. However, the interpretation of probability, the philosophy, is still hotly disputed. Setting aside classical theory, there are four main philosophical positions: logical theory, subjective theory, frequency theory and propensity theory.

The logical theory of probability

The logical theory identifies probability with degree of rational belief. It is assumed that given the same evidence, all rational human beings will entertain the same degree of belief in a hypothesis or prediction. In the first few decades of the 20th century, logical theory was developed mainly at Cambridge, but it was also taken up by the Vienna Circle; Carnap supported the theory in the 1950s. The main champion of the (philosophical aspects of) logical theory was John Maynard Keynes (1921), the economist. Keynes's work at Cambridge was associated with the philosophies of Bertrand Russell, GE Moore and the young Wittgenstein. Keynes, Russell and Moore were members of a secret society of the Cambridge intellectual elite known as the Apostles, an elite within an elite (Gillies, 2000, pp 25–49). Gillies describes the issue that arose as follows:

> Russell was working out the principles of deductive logic used in mathematics, but what about the reasoning from evidence to hypotheses and predictions characteristic of science and so many everyday considerations? It could be argued that, as well as a deductive logic, an inductive logic was needed to cover such empirical reasoning. Moreover, this inductive logic would be closely connected to, perhaps identical with, probability theory (Gillies, 2000, p 27).

Keynes took up the challenge, the results being published in his *Treatise on Probability* (Keynes, 1921). Following Russell, Keynes's approach was to start from axioms that were obviously correct to logical intuition, and from these deduce results, which were logically valid, but not necessarily logically intuitive. For Keynes, probability was degree of *rational* belief not simply degree of belief. The word 'rational' was used to indicate that people of a certain intellect would come to the same conclusion when faced with the evidence, the result being a consensus; it meant that probability was not subjective. In working through the consequences of his approach Keynes concluded that for probabilities to be represented as numerical values (between 0 and 1), it was necessary to have a number of *equally* probable alternatives; this is the Principle of Indifference.

Bayesians adopted the Principle of Indifference giving rise to what is referred to as logical Bayesianism. In this logical version of Bayesianism, the Bayesian conditionalization (see Appendix) takes advantage of the Principle of Indifference.

If we have m possible mutually exclusive hypotheses $h_1, h_2, h_3, \ldots, h_m$, then from the Principle of Indifference:

$$P(h_1) = P(h_2) = \ldots P(h_m) = 1/m$$

From probability theory we have:

$$P(e) = \sum_{i=1}^{m} P(e|h_i)P(h_i)$$

As in this case, the probability of each hypothesis is the same, we have:

$$P(e) = P(h) \sum_{i=1}^{m} P(e|h_i)$$

Substituting in Bayes's Theorem (1) we obtain:

$$P(h|e) = \frac{P(e|h)}{\sum_{i=1}^{m} P(e|h_i)}$$

If the likelihoods can be evaluated (often easy to do), the posterior probability of the hypothesis can be calculated. In

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

this way, the logical Bayesians hope to calculate the degrees of rational belief in hypotheses given evidence, or, as it is sometimes put, to construct an inductive logic (Gillies, 2000, p 27).

Keynes openly admitted that the Principle of Indifference runs into some logical difficulties. For example, the *Book Paradox* envisages visiting a library to find a book. We have never visited the library before nor seen the book. We have no more reason to suppose that the book is either red or not red. So according to the Principle of Indifference, we have the probability that the book is red is 1/2. But, it could be similarly argued that the probability that the book is yellow should also be 1/2 and again the probability of blue or green is similarly 1/2; this violates the principle of probability calculus that the sum of mutually exclusive possibilities must be less than or equal to one.

Apart from running into logical difficulties, the Principle of Indifference greatly restricts the applicability of the logic theory of probability; for example, it is not applicable to the continuous case (Gillies, 2000, p 43). Furthermore, it was clear that the foundational assumption of the approach did not correspond to reality; different intellectuals faced with the same evidence did not intuitively arrive at the same conclusions. Altogether, the difficulties inherent in the logical interpretation had by the 1920s reached such a level that the Bayesians needed a new interpretation of probability. Just such a new interpretation arrived in the form of the subjective theory (Gillies, 2000, p 49).

Subjective theory

The subjective theory was discovered independently and at about the same time by Frank Ramsey of Cambridge and Bruno de Finetti in Italy (Gillies, 2000, pp 50–87). The theory identifies probability with the degree of belief of a particular individual. Here, it is no longer assumed that all rational human beings with the same evidence will have the same degree of belief in a hypothesis or prediction; differences of opinion are allowed. The strength of a belief can be measured by examining the character of some action to which it leads; a suitable action for measurement purposes is betting (betting is used to 'operationalize' the theory).

The subjective theory is not an 'anything goes' approach; the probabilities attached to a set of possibilities have to be *coherent*. The Ramsey–De Finetti theorem states that a set of betting quotients is coherent if, and only if, they satisfy the axioms of probability. Coherence means that if someone has to bet on a number of events $E_1, \ldots, E_n$, his betting quotients are said to be coherent if, and only if, someone else cannot choose stakes $S_1, \ldots, S_n$ such that she wins whatever the circumstances (the *Dutch Book* argument). Surprisingly, this condition is both necessary and sufficient for betting quotients to satisfy the three axioms of probability, which are defined by Finetti as follows (these are similar to the famous Kolmogorov axioms).

Let $e, f, \ldots, e_1, \ldots$ stand for events, concerning which we can have some degree of belief whether they will occur, or have occurred. Let Ω denote the certain event, which must occur. The three axioms of probability are:

1. $0 \leqslant P(e) \leqslant 1$ for any e, and $P(\Omega) = 1$.
2. (Addition Law) If $e_1, \ldots, e_n$ are events that are mutually exclusive (ie, no two can both occur) and exhaustive (ie, at least one must occur), then
$$P(e_1) + \cdots + P(e_n) = 1$$
3. (Multiplication Law) for any two events e, f
$$P(e\&f) = P(e|f)P(f)$$

A detailed proof of the Ramsey–De Finetti theorem is given in Gillies (2007, pp 59–64). As Gillies says, the Ramsey–De Finetti theorem is a remarkable achievement, and clearly shows the superiority of the subjective to the logical theory. 'Whereas in the logical theory the axioms of probability could only be justified by a vague and unsatisfactory appeal to intuition, in the subjective theory they can be proved vigorously from the eminently plausible condition of coherence' (Gillies, 2007, p 64). In the subjective theory, there is no need for the unsatisfactory and restrictive Principle of Indifference as there are no unique *a priori* probabilities. Furthermore, Bayes's theorem provides a simple way of adjusting subjective priors in the light of evidence.

Despite its attractions, the subjective Bayesian theory does run into some difficulties. For example, going back for at least 5000 years, it has been recorded that (at appropriate latitudes) the sun rises every single morning. Thus, given that the sun has risen each morning for 1 826 250 days, it can be estimated that the probability of the sun rising tomorrow is approximately 0.9999994. We can be pretty confident that the sun will rise. What happens if, for some extraordinary reason, the sun does not rise tomorrow? What is the probability of it rising the day after tomorrow? According to de Finetti's theory, the new probability is 0.9999989. In other words, it hardly changes at all. Intuitively this is not a reasonable conclusion. Various proposals for solving this problem add complexity to the subjective Bayesian approach. However, it has been one of the most attractive features of the Bayesian approach that it offers a simple mathematical formula for the way in which a rational person should change his or her beliefs in the light of evidence. This appeal is in danger of being lost if complexity is added. The problem would seem to arise when the sequence of tests are not independent. Various ingenious examples have been suggested by the critics of Bayesianism (eg, the game of red–blue, and the chaotic clock). The simplest response would be to first subject the data to some statistical tests for independence. This would be the approach advocated by objectivists. There would, therefore, seem to be a role for objective approaches.

The intersubjective approach, suggested by Gillies, argues that beliefs about probabilities are often the shared outcome of

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

a group rather than purely personal subjective views. Members of a group are forced to agree on betting probabilities or else a third party could make a *Dutch Book* against them. Thus, their views need to be collectively coherent.

Frequency theory

The frequency theory defines the probability of an outcome as the limiting frequency with which the outcome appears in a long series of similar events (Gillies, 2000, pp 88–112). In the 19th and early 20th centuries, the tendency developed of using frequency as the basis of probability theory. Robert Ellis and John Venn of Cambridge, the Americam pragmatist CS Peirce, Karl Pearson of London and others embarked on such an empirical definition. However, it was the flowering of empiricism brought about by the Vienna Circle that led to renewed interest and further development by Richard von Mises and Hans Reichenbach. In 1928, Richard von Mises published his most famous work on the subject, *Probability, Statistics and Chance*, providing the conceptual clarity and completeness that best exemplifies frequency theory (von Mises, 1928). In von Mises's approach, probability theory is considered as a scientific theory in mathematical form like mechanics or thermodynamics. Its subjects are *mass phenomena* or *repeatable events* as they appear in games of chance, in insurance problems, in heredity theory and in the ever-growing domain of applications in physics.

In his examples, von Mises examines a set of repetitive events or mass phenomenon, which have an attribute that varies from one element to another. A set of attributes is regarded as *a priori* possible; these form what von Mises refers to as an attribute space. The attribute space is usually depicted as Ω (also sometimes referred to somewhat misleadingly as the sample space). Ω is a set of possible outcomes. von Mises introduced the technical term *collective* to describe repetitive events or mass phenomenon of this type. He claimed that we can only introduce probabilities in a mathematical or qualitative sense when there is a large set of uniform events; we must observe his maxim, 'first the collective—then the probability'.

von Mises observed that in the game of dice and all the other mass phenomena observed, the relative frequencies of certain attributes become more and more stable; in the limit, as the number of events is increased, the average of the attributes observed converges to some value. For von Mises, the relative frequency of the repetition is the measure of 'probability', just as the length of a column of mercury is the measure of 'temperature'. Thus, the mathematical notions are given operational definition in terms of observables. The problem von Mises addressed was how to move from empirical observations (the empirical collective) to the axioms of mathematical theory. Once the mathematical theory has been set up, consequences can be logically deduced and predictions made.

von Mises's definition (which he insisted should be strictly interpreted) excluded many common language uses of the term probability. Thus, 'the probability of winning a battle' had no place in his theory because we cannot think of a collective to which it belongs. Frequency theory did make sense when it was first propounded, but in the early thirties, the new subjectivist approach, as we have seen, gave a method for measuring probabilities (as degrees of belief) in such a way that the axioms of probability could be derived from the plausible condition of coherence. This showed that it was possible to extend quantitative probabilities to many cases where no collective is involved. Thus, although it was not shown that frequency theorem was wrong, it was shown that it was possible to extend the mathematical calculus to examples that lie outside its scope (Gillies, 2000, p 99). Another related criticism (made by De Finetti) is that the frequency theory does not deal with the role of probability theory in induction and confirmation.

Propensity theory

Popper initially adopted the frequency theory as the objective theory of probability. However, he identified a problem. He noted that frequency theory could not provide objective probabilities for single events; for quantum theory, he argued that these were necessary. He proposed instead, what he called, the propensity theory. The propensity theory, or at least one of its versions, takes probability as a propensity inherent in a set of repeatable conditions. To say that the probability of a particular outcome is p is to claim that the repeatable conditions have a propensity such that, if they were to be repeated a large number of times, they would produce a frequency of the outcome close to p (Gillies, 2000, pp 113–168).

The emphasis is on *repeatable conditions*, but there are various interpretations as to what this entails. In a later version, Popper (1990) and Miller (1994, pp 185–186), suggest that strictly speaking every propensity must be referred to the complete situation of the universe at the time. Thus, propensity seems to have been changed from a scientific to a metaphysical concept that cannot be tested. On the other hand, Fetzer (2001) proposes that propensities depend on a *complete* set of *relevant* conditions; these, of course, may also be difficult to define and test (Gillies, 2000, p 128). Gillies suggests that long-run propensities should be taken to depend on a set of relevant conditions, but not necessarily a complete set. This makes assignment of probabilities testable, but it means that they cannot be extended to single cases; Gillies suggests using subjective probabilities in this instance (Gillies, 2000, p 129).

One of the achievements of the propensity theory is that it can be used to derive and make more specific the theories of frequency. The key is an issue raised by Popper. As a result of his interest in testing scientific hypotheses, he was faced with the problem of how to falsify scientific statements couched in probabilistic terms; in principle, such statements are impervious to strict falsification (any outcome is in theory consistent with such a statement). In answer to this difficulty,

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

Popper appealed to the notion of methodological falsifiability. This was an issue that had already been independently addressed by the statisticians Pearson (the chi-square test), Gosset (the student t-test) and Fisher (the F-test and analysis of variance). By combining a falsifying rule with the axioms of probability, it is possible to derive theoretically the empirical laws of frequency theory. Furthermore, not only can the empirical law (that as the number of trials increases towards infinity, the long run frequency tends to a fixed value of probability—the Law of the Stability of Statistical Frequencies) be derived, but also the rate at which this convergence occurs can be predicted; this prediction can be empirically tested. The empirical testing of the convergence rates confirms the theory.

A pluralistic approach to probabilistic inference

It would seem that there is no single theory that is superior to all others and each has its problems. Gillies classifies the logical, subjective and intersubjective theories collectively as epistemological (or epistemic) theories; the frequency and propensity theories are classified as objective. For the epistemological theories, probabilities measure degree of knowledge, degree of rational belief or degree of belief, or words to that effect. Objective interpretations of probability take probability to be a feature of the objective material world, having nothing to do with human knowledge or belief. (Gillies, 2000, p 2). Gillies concludes that no one approach can satisfy all requirements. He suggests that the propensity theory is the most convincing objective theory and is appropriate for science; the subjective (or its intersubjective variant) is the dominant epistemological theory and is appropriate for social science.

As we have seen, Bayes's theory is associated with both the logical and subjective approaches to probability. A growing number of researchers in a variety of fields now advocate that the whole subject matter of inference should be viewed from a Bayesian perspective; they advocate a Bayesian rationality. Thus, the Bayesian perspective itself provides the framework for a pluralistic approach. They have been encouraged by the light that such a probabilistic, subjective Bayesian approach has been found to shed on some of the problems faced by the deductive and inductive approaches.

Bayesian confirmation

Bayes's theorem addresses the issue of confirmation; by one definition, h is confirmed by e if $P(h|e) > P(h)$; h is disconfirmed by e if $P(h|e) < P(h)$. Thus, using the first version of the theorem, Equation (A.1) in the Appendix, confirmation is given by whether the Bayesian multiplier (the ratio of the likelihood of h on e over the probability of e) is greater than one. If the background information b is included, as in Equation (A.2) of the Appendix, the definitions can be modified to include the idea of conformation or disconformation of h by e relative to b. Other definitions of confirmation are possible; for example, h is confirmed if $P(h|e) > P(h|\neg e)$. From qualitative notions of confirmation, various quantitative measures have been suggested; for example, the simplest, but not necessarily the best, is the arithmetical difference between the posterior and the prior probabilities of h.

Consider the significance of new evidence. If we take Equation (A.2) from the Appendix and rearrange it, we have:

$$\frac{P(h|e\&b)}{P(h|b)} = \frac{P(e|h\&b)}{P(e|b)}, \text{ where } P(e|b) \neq 0 \qquad (9)$$

Taking background b as earlier evidence and e as new evidence, the ratio on the left-hand side of the equation tells us the degree to which the evidence confirms the hypothesis h. The confirmation ratio is equal to the factors on the right-hand side. The numerator is a likelihood that we can assume to be close to 1. Then, the confirmation of h varies inversely with the denominator. The denominator is the expectedness of e against a background b of earlier evidence. Clearly, if the evidence is more of the same kind of evidence found in b, the denominator will be high and hence the confirmation is low (only slightly above 1). If, however, e is a different variety of evidence, or is surprising evidence given b, then the denominator on the right-hand side will be low and the confirmation of h high. In this way, the common idea that variety of evidence, or the surprising characteristic of evidence, boosts confirmation, has a rationale within Bayesianism (Nola and Sankey, 2007, pp 231–232).

The problems of Bayesian priors and old evidence

The probabilistic, Bayesian perspective provides a way of understanding, and in some cases alleviating, problems with both the H-D and the inductive methods (see below). However, some new problems are introduced. The first of these is the problem of choosing priors at the initiation of the analysis. A charge that is often made against pure Bayesianism is that 'anything goes'; this refers to the choice of priors that are open to personal choice. One response is to argue that it is no worse a position than that of deductive logic; after all, deductive logic places no restrictions on the premises. A second response is to point out that the priors will soon be 'swamped' or 'washed out' by the evidence; however, this is not the case when the priors are very close to one or zero (as they often are from the perspective of committed advocates of one theory or another). A third response is to add something to pure Bayesianism to constrain or guide the choice of priors. Various approaches have been advocated proposing that priors should: be well calibrated; or be tempered; or draw on analogies; or maximize entropy; or display open-mindedness; or reflect empirical examination of previous choices; or reflect likely future views or the views of experts in the field; or simply be reasonable.

More generally, Bayesianism is accused of making science a subjective affair. How does it explain the widespread agreement that science is correctly opposed to superstition in its

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

claims on our credence because it is based on objectively justifiable canons of inference, not on what people actually do believe and the extent to which they believe it? Howson and Urbach (1989) argue that this objection rests on a confusion, and that a Bayesian reconstruction of the procedures of inductive inference poses no threat whatsoever to the objectivity of the scientific enterprise (p 258). They argue that it is the subject matter, degrees of belief, that is subjective not the Bayesian rules of consistency imposed on them (p 290).

Another problem is that of old evidence. An example is the resolution of the anomalous behaviour of the precession of the perihelion of the planet Mercury; this could not be explained by Newtonian mechanics. Einstein's general theory of relativity provided an account of the anomaly. Given that the data describing the behaviour of Mercury is old, $P(e) = 1$. It follows from Bayes's theorem that the data gives no confirmation to Einstein's theory. But, scientists clearly believed that it did; the new theory had resolved a long-standing anomaly. The problem has attracted considerable attention in the literature and a number of solutions have been proposed.

There are other objections to subjective Bayesianism, largely by Popper and his followers, an account of which (together with rebuttals) can be found in Howson and Urbach (1989, pp 257–296).

Probabilistic inference and professional practice

Probability plays a major role in both the education and practice of OR. It is a subject in its own right and provides the basis for statistical and forecasting theory. It is also used in areas as diverse as queuing theory, inventory control, simulation, decision trees, decision theory and maintenance analysis. Some of these applications create models of, what the analysts take to be, objective features of objectively real processes. In other words, the probability that a machine (of a particular type, operated and maintained in a particular way) will break down (at a particular time) is taken to be a feature of the particular machine in its context. It is thus the equivalent of a natural process observed by scientists to exhibit probabilistic behaviour. There are all sorts of philosophical objections to the idea that these processes are in fact real and that they can be said to exhibit objective probabilistic behaviour. For example, many scientists recognize that their descriptions of unobservable sub-atomic entities (particles, waves, strings and so on) are just metaphors to help human comprehension of abstract mathematical models. The very concept of probability is considered by some philosophers as man-made; after all we do not observe probabilities directly, we infer them from observed events; we apply our human notion of probability to observed behaviour. However, there seems to be no harm in assuming objectivity in our practice in these areas. In many other modelling applications, such as dynamic programming, decision trees and decision analysis, the probabilities represent the subjective views of the decision maker and are recognized as such by the modeller.

Students of OR are taught both probability theory and statistics. This generally will include Bayes's theorem. Surveys of OR practitioners indicate that statistical methods and forecasting are among the most frequently used techniques. Thus, OR practitioners have to be well versed in the application of probabilistic thinking and, of course, statistics. It is difficult to tell the degree to which Bayes's theorem is used in practice. A scan of *JORS* indices and article titles over the past 5 years (2003–2007) for mention of Bayes revealed only three articles (Choi *et al*, 2003; Stavrulaki *et al*, 2003; Rajagopal and del Castillo, 2007); all three are theoretical articles. Certainly, a Bayesian approach is sometimes favoured in practice for forecasting, particularly where the limited amount of information that needs to be carried forward is an advantage (eg, retailing where many stocked items are involved and ordering is frequent). However, it seems likely that standard non-Bayesian statistical methods still dominate OR practice.

From a Bayesian perspective

A number of issues raised in the debate about scientific inference are now addressed from a Bayesian perspective.

Popper's severity of a test and corroboration

One of the problems that Popper struggled with was how to indicate, for practical purposes, the reliability of a particular hypothesis: how well has it been established? Popper is anti-inductivist but not an anti-probabilist (however, he is opposed to a subjective interpretation of probabilities). While holding that no test can positively prove a theory, he argues that a hypotheses that has withstood a severe test is to be preferred. He defines his idea of the severity of a test in probabilistic terms as proportional to either the difference between $P(e|h\&b)$ and $P(e|b)$, or their ratio. In the ratio form, it is the same as the right-hand side of the Equation (9) above (a rearranged form of Bayes's theorem); thus, it is equal to the Bayesian confirmation, the left-hand side of the equation. In other words, Popper's measure of severity tells us in Bayesian terms about the conformation that evidence e provides. However, the approaches do differ: instead of using confirmation to indicate the strength of a hypothesis, Popper proposes his notion of corroboration that includes severity in its difference form. Corroboration of a hypothesis h by evidence e against a background b is defined as:

$$Cor(h|e\&b) = [P(e|h\&b) - P(e|b)]/[P(e|h\&b) - P(e\&h|b) + P(e|b)]$$

This formulation cannot be justified in Bayesian terms.

Bayesianism and induction

Equations (A.3) and (A.4) in the Appendix can be rewritten in terms of h and its negation $\neg h$ rather than h_1 and h_2.

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

Rearranging and defining the likelihood ratio LR as $P(e|\neg h)/P(e|h)$, then we have:

$$P(h|e) = \frac{P(h)}{P(h) \times (1 - LR) + LR}$$

This equation tells us that the confirmation of h by e varies directly with the prior $P(h)$ and inversely with the likelihood ratio $P(e|\neg h)/P(e|h)$ or directly with the likelihood ratio $P(e|h)/P(e|\neg h)$. This intuitively feels right in the light of induction. Given some initial priors, Bayesians can explain inductive inference in terms of serial conditionalization. However, no method of choosing the initial priors can be shown to be uniquely correct. Although this may seem a problem to inductionists, for Bayesians it is not. In general, Bayesians do not believe that induction can be justified. Like Popper they seek to by-pass it, in their case with their Bayesian approach.

The H-D method from a Bayesian perspective (Nola and Sankey, 2007, pp 202–204)

The basic idea of the H-D method is that a hypothesis h under test, in conjunction with whatever auxiliaries a, entails a set of test consequences c, which we can suppose is identical to some evidence e. Applying the version of Bayes's theorem shown in Equation (2) we have

$$P(h|e\&a) = \frac{P(e|h\&a)P(h)}{P(e|a)}$$

However, in the case of the H-D method, the evidence has been anticipated by deduction for the hypothesis under test. Therefore, $P(e|h\&a) = 1$. We, therefore, have:

$$P(h|e\&a) = \frac{P(h)}{P(e|a)}$$

This says there is confirmation of h by e with respect to the background of auxiliary theory a. Thus, by setting the H-D method in the Bayesian context, it is automatically equipped with a theory of confirmation it would otherwise lack. Furthermore, Baysians claim that the 'tacking problem' can be addressed in Bayesian terms (though in my view not entirely satisfactorily). Impressively Bayesians do seem to have found a way (actually several ways) of dealing with the Quine–Duhem thesis. Dorling, who wrote one of the early, original articles on the thesis (Dorling, 1979) used the historical example of the secular speeding up of the moon in its orbit as analysed by Laplace using Newtonian physics. Dorling concludes that although the results can be derived without recourse to the actual practice of scientists or to the past successes of such practice, the results seem to be fully in line with the intuition of scientists (Jeffrey, 2004, pp 35–44). Dorling challenges the social scientists critics of science thus:

> The Bayesian analysis indeed vindicates the rationality of experienced scientists' reactions in many cases where those reactions were superficially paradoxical and where the relevant scientists themselves must have puzzled over the correctness of their own intuitive reactions to the evidence. It is clear that in many such complex situations many less experienced commentators and critics have sometimes drawn incorrect conclusions and have mistakenly attributed the correct conclusions of the experts to scientific dogmatism. Recent philosophical and sociological commentators have sometimes generalised this mistaken reaction into a full-scale attack on the rationality of men of science, and as a result have mistakenly looked for purely sociological explanations for many changes [in] scientists' beliefs, or the absence of such changes, which were, in fact, as we now see, rationally de rigueur (Dorling quoted in Jeffrey, 2004, pp 43–44).

Dorling concludes that for the first time the subjective, probabilistic (Bayesian) approach provides scientists with an adequate theory of how to change their beliefs in the light of evidence. Subjective probability theory provides the simplest account available of the logical relations between degrees of belief which fall short of certainty (Jeffrey, 2004, pp 42–43).

Bayesianism and IBE

IBE compares hypotheses and asks which provides the best explanation. If we make a Bayesian comparison of two hypotheses with a common background b, we have Equation (A.6) from the Appendix:

$$\frac{P(h_1|e\&b)}{P(h_2|e\&b)} = \frac{P(h_1|b)}{P(h_2|b)} \times \frac{P(e|h_1\&b)}{P(e|h_2\&b)}$$

Wesley Salmon calls this the *Bayesian algorithm for theory preference*. The expressions $P(h_1|b)$ and $P(h_2|b)$ on the right-hand side reflect the initial plausibility of the two hypotheses against b, the background knowledge of science. Nola and Sankey (2007, p 240) provide an example to illustrate this. The fact to be explained is how Michelangelo painted the Sistine Chapel ceiling, which is high above the ground. The two rival hypotheses are (i) that he used scaffolding to climb and (ii) that he levitated. But, the hypothesis of levitation has, from what we know about science and ordinary life, a low probability (what Lipton calls likeliness) and so is implausible. The ratio on the right-hand side $P(h_1|b)/P(h_2|b)$ will reflect this relative implausibility. The expressions $P(e|h_1\&b)$ and $P(e|h_2\&b)$ on the right-hand sides are likelihoods. What is the probability of e given other kinds of varying information? Nola and Sankey say:

> A number of factors become important in assessing this. [Earlier] we also noted that issues to do with consilience of facts [a theory is consilient if it brings together two different classes of fact under its scope: the more consilient it is, the better], and the importance of facts, have a role in determining how good an explanation is; and we also noted other contributing factors such as simplicity, scope and fertility of the explanatory hypotheses (Nola and Sankey, 2007, p 241).

When taken together these explanatory considerations provide an account of the overall credibility of an

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

explanation captured in the right-hand expression in the above Equation, (A.6) in the Appendix. Applied to all the hypotheses, this procedure can lead to a uniquely best explanation. The Bayesian framework has provided a way to allow explanatory considerations to be taken into account when determining the priors and likelihoods. The left-hand side is not couched in terms of truth, but in terms of probabilities. Explanationalists would want to go further and draw a conclusion about the truth of the statement. On the other hand, strict Bayesians would not want to be constrained by explanatory considerations and would certainly want to maintain coherence in the choice of probabilities. Lipton suggests that though explanationist considerations cannot fix priors and likelihoods precisely they can help fix them within certain bounds: 'explanatory considerations provide a central heuristic we use to follow the process of conditionalisation, a heuristic we need because we are not very good at making probabilistic calculations directly' (Lipton, 2004, p 107).

Bayesianism and Khunian values

Kuhn suggests that when choosing between an established theory and an upstart competitor the characteristics accuracy, consistency, scope, simplicity and fruitfulness (and perhaps others) are brought to bear. Nola and Sankey have shown that, at least on some interpretations of what they mean, Kuhn's characteristics can be reflected in the Bayesian algorithm for theory preference above (A.6 from the Appendix). For example, if a theory is internally inconsistent, $P(h|b)$ will be 0 no matter what the evidence is and what the other probabilities are; thus, $P(h|e\&b)$ will be 0. Kuhn argued that the characteristics function as criteria subjected to weights, which differ from scientist to scientist. He concludes that subjective weightings undermine the idea of an algorithm for theory choice. Nola and Sankey say:

> For Kuhn this was an important feature of scientific method that is overlooked by those who claim that there is an objectivity and rationality to science; for example, supporters of the H-D method would reject Kuhn's claim that 'paradigm choice can never be unequivocally settled by logic and experiment alone' (Kuhn 1970, p 94). What Kuhn resists is the idea that there can be a 'logic-based algorithm' (like the H-D method) that, in a choice situation in which several theories are being considered, churns out one and only one theory as the right one to choose. Such a 'logic-based algorithm' goes along with a strong account of the objectivity and 'rationality' of methods of choice in science. The probability calculus, including Bayes's theorem, might seem to be such a 'logic-based algorithm'; in this respect Bayesians differ from Kuhn (Nola and Sankey, 2007, p 237).

However, the probabilities of the Bayesian approach are to be understood as coherent degrees of personal belief. As in Kuhn's approach, they are subjective, but Bayesians can make use of their convergence of opinion theorems. According to Nola and Sankey: 'This restores a certain objectivity, and even rationality, to the Bayesian position not found in Kuhn's [position]' (p 237).

The implications for OR practice

As has been described above psychological research has shown that in everyday life people use deductive, inductive and probabilistic inference. The same is true of scientists in general. University OR course syllabi, surveys of OR techniques used in practice, and the canon of OR publications show that all three approaches to inference are employed in OR both formally and informally. Deploying any one of the approaches is, in effect, to buy into an intellectual framework; each approach comes with a proposition:

Proposition for deductive inference: (a) If you can identify a question that can be formulated as a deductive puzzle, then deductive logic and mathematics can be deployed. But, any conclusions must be accompanied by a full exposition of the assumptions made (the premises) including the structure assumed (boundaries, linearity and so on), the parameters used and the objectives assumed by the model. (b) If you wish to draw conclusions from evidence using deductive methods, you have to accept the arguments of Popper that this is possible (and desirable). Any conclusions must be accompanied by an assessment of whether the background assumptions themselves are reliable.

Proposition for inductive inference: If you wish to draw conclusions from evidence and you consider inductive inference is reasonable and you are confident that you are a good judge of what is reasonable in a particular situation, then you can apply inductive logic and the supporting statistical methods with reasonable confidence. But, any conclusions must be conditioned by a statement that includes the reasonableness assumption together with an estimation of error.

Proposition for probabilistic inference: (a) If it is believed that the logic of the situation can be represented in terms of probability (understood as objective or subjective), then the calculus of probability can be deployed (b) If you wish to draw conclusions from evidence and believe that you or your clients can suggest sensible priors, then a Bayesian approach to inferences can be taken. Any conclusions must indicate the dependence on the priors; have they been swamped? In both cases, any conclusions reached must be stated in terms of their probability.

The three modes of thinking are not mutually exclusive in the practical situation. In deductive mode, when formulating the assumptions (the premises), it may be necessary to operate in inductive or probabilistic mode to establish the parameters. In inductive mode, probability calculus is used in the statistical methods; chains of deductive reasoning may well also be involved. In probabilistic mode, some additional inductive reasoning may be required; once formulated in probabilistic terms, the problem becomes a puzzle to be analysed using

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

probability calculus (the deductive logic and mathematics of probability based on the axioms of probability).

A practitioner's choice of approach will be determined by his or her understanding of the logic of the situation and his or her personal preferences and habits (his or her strand of practice). Some analysts will seek out opportunities to apply deductive logic (mathematical modelling, for example); more generally, consultants will concentrate on meeting the concerns of a client, deploying deductive, inductive or probabilistic thinking as appropriate. OR is undoubtedly a pluralistic activity in terms of the approach taken to inference. For example, consider CPA or PERT analysis for project management. At their core, the models are based on deductive logic. On the other hand, the estimates of time to complete each task and so on are subjective estimates generally induced from past experience. Very often probabilities will be included to explore the range of possibilities. Take another example, DEA analysis. DEA is essentially an exercise in deductive logic, but it is usually preceded by a statistical analysis to identify the most important factors to be included in the model. Forecasting, another example, can either be based on classical statistical estimating or Bayesian conditioning. Business models (eg, energy models) usually combine logical relationships (eg, supply of energy type A equals the demand for energy type A plus the change in stocks) with behavioural parameters (for instance, energy demand price elasticities) based on inductive statistical analysis. Choosing and combing various methods, techniques and modes of inference is ubiquitous throughout OR practice and is a crucial craft skill.

The choices that practitioners make can be thought of as being an application of values. OR practitioners (and professionals of other disciplines) will undoubtedly weight potential values rather differently from pure scientists. After all, the OR consultant is only interested in understanding and truth in so far as they inform choice and action. The understanding that is reached does not have to be universal; local and *ad hoc* may be more useful. Practitioners are more likely in circumstances they deem appropriate to put greater weight on pragmatic values. It is not that the epistemic values can be ignored, clearly in some cases they will be all important. It is just that there will be many cases where pragmatic values are influential. For example, for supermarkets forecasting the demand for a particular product line at a particular store (an activity required for thousands of product lines every day), ease of computability will be an important pragmatic value that will be weighed against forecasting accuracy. Robinson (2008a), in an article about simulation, suggests that there are four requirements for conceptual models for simulation modelling: (i) validity, (ii) credibility, (iii) utility and (iv) feasibility. In drawing up this list, he cites instances in the literature where the requirements for a conceptual model are suggested. These articles use a variety of terms including completeness, aptness for client's problem, understandable, transparency, execution speed, ease of modification, maintainability, portability, timely, elegance and time and cost to build. Robinson suggests that the overarching requirement is simplicity. Reporting on an application in the Ford Motor Company, he provides an illustration in the OR domain of Kuhn's observation that the weights applied to the various values (requirements or assessment criteria) will vary from person to person:

> A second reason why there is no right conceptual model is because the model is an agreement between more than one person (the modeller, clients and domain experts). Each has his/her own preferences for and perceptions of what is required. These preferences and perceptions are expressed through the four requirements of a conceptual model. ... The conceptual model is, therefore, some compromise between alternative preferences and perceptions of the world (Robinson, 2008b, p 303).

Robinson cites other sources of assessment criteria in the literature, mainly, he says, in the domain of large-scale military and public policy models. In the late 1970s, when the US government was funding a large number of such policy models (often in the US referred to as systems analysis), guidelines were laid down by the US Government Accountability Office (GAO) for model evaluation. The guidelines suggested that the model documentation should be evaluated and the models themselves should be assessed against four criteria: (i) (theoretical, data, and operational) validity, (ii) (computer model) verification, (iii) maintainability and (iv) usability (US GAO, 1979). The last two criteria are clearly pragmatic in nature. Thus, in the three domains discussed, forecasting, simulation and policy modelling, pragmatic as well as epistemic values are well to the fore. These three domains cover a good deal of OR practice and it can reasonably concluded that OR is generally in the habit of balancing epistemic and pragmatic considerations.

Philosophers and sociologists of science raise another issue related to values. They argue that facts (established by inference) and values cannot be separated. The OR response is generally to separate out facts and values, with OR providing advice on how the two can be combined in the decision-making process. OR also advises how facts can be established. However, it is the decision maker (singly or collectively) who determines the values; in other words the objectives to be pursued. This approach is labelled by the critics 'decisionism'. Both the fact–value duality and decisionism have come under sustained attack (eg, Putman, 2002). Roughly speaking, they argue that looked at closely all activities are value-laden including the activities of those engaged in the establishment of 'facts'. Moreover, by separating technical work and moral choice, the way is open for unscrupulous people to use science and technology for undesirable (undemocratic, say) ends. It should also be noted that the term 'rational' used in the article title would also be challenged as claiming too much. There are issues to be addressed here, but it is beyond the scope of this paper.

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

The implications for OR research

The application of the Bayesian approach to a wide range of issues has taken great strides in recent years. It is a programme that is still energetic and expanding; it is attracting resources. We can, therefore, expect more insights to become available (and some no doubt to be discarded). However, the account given in this article is only a partial account and there are other Bayesian and non-Bayesian programmes underway (eg, formal learning theory and theory of error statistics). Particularly interesting for OR are those programmes that look beyond problems of inference, those that address choice between options with different consequences (rational choice in Figure 1). There are, therefore, opportunities to conduct research into the progress of Bayesian and other programmes and consider their implications for OR: broadening the use of the Bayesian perspective as an epistemological lens. There are also a number of other research opportunities. One could engage in empirical research to examine the use of different modes of inference-making in practice: a micro-epistemological investigation. At a higher level of abstraction, one could examine the justification of OR activities from a Bayesian perspective: a macro-epistemological examination. Another avenue would be to look at soft OR and problem structuring methods from a Bayesian perspective. What inferences are being attempted? How are they made? One of the attributes of the Bayesian perspective promoted by its proponents is that it can be used in a wide variety of fields.

The implication for teaching OR

If one concludes that the probabilistic Bayesian perspective is a superior viewpoint from which to understand some OR activities, then the obvious thing to do is to build it into courses as an integrating framework for the variety of methods being taught. However, before advocating such a course of action, I would want to see progress on the research agenda in the previous section. At a more prosaic level, course content could be adjusted to take account of some of the recent advances in the understanding of inference making outlined in this article.

Concluding remarks

It is much easier when addressing practical problems to choose between deductive/inductive, deterministic/probabilistic, objective/subjective thinking than it is to describe the way we do it, to formulate the rules we employ and to justify our habits. The article has reviewed the thinking behind three practical approaches to making inferences, namely deductive, inductive and probabilistic. Each has a part to play in OR and professional practice. However, it is the probabilistic view in the form of the subjective version of Bayes's theorem that gives us a new perspective on the various choices; it provides a new epistemological approach. Just as we have been able to draw on deductive mathematics, inductive statistical methods and probabilistic calculus in our modelling efforts, we can now draw on the Bayesian perspective to describe and formalize our practice habits; it gets us closer to understanding a major component of our own craft skills. Current research shows that probabilistic thinking is reflected in everyday and scientific habits; there is no reason to think that OR and other professionals and their clients are different. This would seem to be reason enough to believe that a probabilistic perspective is likely to be productive and produce insights into the conduct of practice. There are plenty of possible avenues for research to take this programme forward in the domain of professional practice in general and OR in particular.

References

Bayes T (1763). An essay towards solving a problem in the doctrine of chances. *Phil Trans Royal Soc* **53**: 370–418.

Choi T-M, Li D and Yan H (2003). Optimal two-stage ordering policy with Bayesian information updating. *J Opl Res Soc* **54**: 846–860.

Chomsky N (1957). *Syntactic Structures*. Mouton: The Hague.

Chomsky N (1965). *Aspects of the Theory of Syntax*. MIT Press: Cambridge, MA.

Dorling J (1979). Bayesian personalism, the methodology of scientific research programmes, and Duhem's problem. *Stud Hist Philos Sci* **10**: 177–187.

Duhem P (1954). *The Aim and Structure of Physical Theory*. Princeton University Press: Princeton, NJ.

Fetzer J (ed) (2001). *The Philosophy of Carl G Hempel: Studies in Science, Explanation and Rationality*. Oxford University Press: Oxford.

Gillies D (2000). *Philosophical Theories of Probability*. Routledge: London.

Goodman N (1951). *The Structure of Appearance*. Harvard University Press: Cambridge, MA.

Howson C and Urbach P (1989). *Scientific Reasoning: The Bayesian Approach*. Open Court: La Salle, IL.

Jeffrey R (2004). *Subjective Probability*. Cambridge University Press: Cambridge.

Keynes JM (1921). *Treatise on Probability*. Macmillan: London.

King P and Shapiro S (1995). History of Logic. In: Honderich T (ed). *The Oxford Companion to Philosophy*. Oxford University Press: Oxford, pp 496–500.

Kuhn TS (1962). *The Structure of Scientific Revolutions*. University of Chicago Press: Chicago, IL. (Page numbers refer to the 1996 third edition).

Kuhn TS (1970). Reflections on my critics. In: Lakatos I and Musgrove A (eds). *Criticism and the Growth of Knowledge*. Cambridge University Press: Cambridge, pp 231–278.

Laudan L (1984). *Science and Values: The Aims of Science and their Role in Scientific Debate*. University of California Press: Berkley, CA.

Lipton P (2004). *Inference to the Best Explanation*, 2nd edn. Routledge: London.

Miller D (1994). *Critical Rationalism: A Restatement and Defence*. Open Court: Chicago, IL.

Munro I and Mingers J (2004). Response to Richard Ormerod. *J Opl Res Soc* **55**: 90–93.

Musgrave A (1988). The ultimate argument for scientific realism. In: Nola R (ed). *Relativism and Realism in Science*. Kluwer: Dordrecht, pp 229–252.

Musgrave A (1999). *Essays on Realism and Rationalism*. Rodopi: Amsterdam.

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms

Nola R and Sankey H (2007). *Theories of Scientific Method*. Acumen: Stocksfield.
Oaksford M and Chater N (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press: Oxford.
Ormerod RJ (2002). On the nature of OR: taking stock. *J Opl Res Soc* **53**: 475–491.
Ormerod RJ (2009). The history and ideas of critical rationalism: the philosophy of Karl Popper and its implications for OR. *J Opl Res Soc* **60**: 441–460.
Popper KR (1934). *Logik der Forschung*. Springer: Vienna.
Popper KR (1959). *The Logic of Scientific Discovery*. Hutchinson: London.
Popper KR (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford University Press: London.
Popper KR (1990). *A World of Propensities*. Thoemmes: Bristol.
Putman H (2003). *The Collapse of the Fact/Value Dichotomy*. Harvard Press: Cambridge, MA.
Quine WVO (1960). *Word and Object*. Harvard University Press: Cambridge MA.
Quine WVO (1990). *Pursuit of Truth*. Harvard University Press: Cambridge, MA.
Quine WVO and Ullian J (1978). *The Web of Belief*, 2nd edn. Random House: New York.
Rajagopal R and del Castillo E (2007). A Bayesian approach for multiple criteria decision making with applications in Design for Six Sigma. *J Opl Res Soc* **58**: 779–790.
Robinson S (2008). Conceptual modelling for simulation Part I. *J Opl Res Soc* **59**: 278–290.
Robinson S (2008). Conceptual modelling for simulation Part II. *J Opl Res Soc* **59**: 291–304.
Shieber SM, Schabes Y and Pereira FCN (1995). Principles and implementation of deductive parsing. *J Logic Programming* **24**: 3–36.
Stavrulaki E, Fong DKH and Lin DKJ (2003). Two-resourse stochastic capacity planning employing a Bayesian methodology. *J Opl Res Soc* **54**: 1198–1208.
van Fraassen B (1980). *The Scientific Image*. Oxford University Press: Oxford.
van Fraassen B (1989). *Laws and Symmetry*. Clarendon: Oxford.
von Mises R (1928). *Probability, Statistics and Chance*, 2nd revised English edn (1961) Allen & Unwin: London.
US GAO (1979). *Guidelines for Model Evaluation*. PAD-79-17. US GAO: Washington, DC.
Williamson C (1995). Syllogism. In: Honderich T (ed). *The Oxford Companion to Philosophy*. Oxford University Press: Oxford, p 862.

Appendix

Bayes's theorem

In its simplest form Bayes's theory is formulated as follows:

$$P(h|e) = \frac{P(e|h)P(h)}{P(e)} \tag{A.1}$$

$P(h|e)$ is known as the posterior probability of h given e.
$P(e|h)$ is known as the likelihood.
$P(h)$ is known as the prior probability of h.
$P(e)$ is the prior probability of e.

The change from $P(h)$ to $P(h|e)$ is known as *Bayesian conditionalization*. The likelihood divided by the prior probability of e, which conditions the prior probability of h, is referred to as the *Bayesian multiplier*. If we include a further variable b, we can account for any background assumptions, perhaps a background theory or some old evidence:

$$P(h|e\&b) = \frac{P(e|h\&b)P(h|b)}{P(e|b)} \tag{A.2}$$

If two hypotheses are being compared on the basis of the same evidence e, we have:

$$P(h_1|e) = \frac{P(e|h_1)P(h_1)}{P(e|h_1)p(h_1) + P(e|h_2)P(h_2)} \tag{A.3}$$

and

$$P(h_2|e) = \frac{P(e|h_2)P(h_2)}{P(e|h_1)p(h_1) + P(e|h_2)P(h_2)} \tag{A.4}$$

We can divide these two equations to give:

$$\frac{P(h_1|e)}{P(h_2|e)} = \frac{P(h_1)}{P(h_2)} \times \frac{P(e|h_1)}{P(e|h_2)} \tag{A.5}$$

That is, the ratio of the degree of confirmation of the rival hypotheses by the same evidence e is given by the ratio of the prior probabilities of the rivals multiplied by the likelihood of the rivals, or in other words the ratio of the degree to which each rival predicts the evidence. We can expand (A.5) to include background assumptions:

$$\frac{P(h_1|e\&b)}{P(h_2|e\&b)} = \frac{P(h_1)}{P(h_2)} \times \frac{P(e|h_1\&b)}{P(e|h_2\&b)} \tag{A.6}$$

This formulation of Bayes's theorem yields an understanding of crucial experiments. However, more generally, each hypothesis will have its own background assumptions. We, therefore, have:

$$\frac{P(h_1|e\&b_1)}{P(h_2|e\&b_2)} = \frac{P(h_1)}{P(h_2)} \times \frac{P(e|h_1\&b_1)}{P(e|h_2\&b_2)} \tag{A.7}$$

For n rival hypotheses, we have:

$$P(h_k|e) = \frac{P(e|h_k)p(h_k)}{\sum_{i=1}^{n} P(e_i|h_i)p(h_i)} \tag{A.8}$$

Thomas Bayes's (1763) theory was published posthumously. It seems likely that it was developed in response to David Hume's identification of 'the problem of induction'. As a statement of logic, the theorem is not generally considered controversial. However, it is the use of the theory with subjective probabilities and as a metaphor for addressing epistemological questions, which is of primary importance for this paper. These usages are more controversial, but have recently received widespread attention in a variety of fields. A fuller account of Bayes's theorem can be found in Howson and Urbach (1989).

Received April 2008;
accepted June 2009 after three revisions

This content downloaded from 89.206.115.10 on Tue, 05 Nov 2019 22:10:44 UTC
All use subject to https://about.jstor.org/terms