Prospectus

Stefan Lukits

Contents

1 Abstract 3

2 Literature Review 4

3 Proposal 7

4 Chapter Outline 11

4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

4.2 Affine Constraints . . . . . . . . . . . . . . . . . . . . . . . . 11

4.3 The Principle of Maximum Entropy: Virtues . . . . . . . . . 12

4.4 The Principle of Maximum Entropy: Vices . . . . . . . . . . . 12

4.5 Judy Benjamin . . . . . . . . . . . . . . . . . . . . . . . . . . 13

4.6 The Shimony Objection . . . . . . . . . . . . . . . . . . . . . 13

4.7 The Seidenfeld Objection . . . . . . . . . . . . . . . . . . . . 13

4.8 The Wagner Objection . . . . . . . . . . . . . . . . . . . . . . 14

4.9 Coarsening at Random . . . . . . . . . . . . . . . . . . . . . . 14

4.10 Families . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

4.11 The Two Level Objection . . . . . . . . . . . . . . . . . . . . 15

4.12 Epistemic Entrenchment . . . . . . . . . . . . . . . . . . . . . 16

4.13 Epistemological Implications . . . . . . . . . . . . . . . . . . 16

5 Bibliography 17

1 Abstract

If we assume a prior probability distribution or density, then there may be guidelines for changing it in the light of new evidence (Richard Jeffrey calls the investigation of these guidelines 'probability kinematics'). Standard conditioning is a well-known updating procedure where the | operator provides a posterior probability distribution or density which obeys both basic probability axioms and some intuitions about desired properties of updated probabilities.

There are various situations where standard conditioning is not applicable. E.T. Jaynes has suggested a unified updating procedure which generalizes standard conditioning called the principle of maximum entropy, or pme for short. The pme is based on a subjective interpretation of probabilities as representing the degree of uncertainty, or the lack of information, of the agent holding these probabilities. In this interpretation, probabilities do not represent frequencies or objective probabilities, although there are models of how probabilities relate to them.

The pme is also more narrowly based on a Bayesian interpretation of probabilities, which considers it unavoidable that a prior probability distribution or density precede a meaningful evaluation of evidence in the form of a posterior probability distribution or density. Bayesians also consider standard conditioning to be normative where it applies. Within the Bayesian camp, however, there is disagreement whether there are objective methods of determining probabilities prior to any evidence or observation, for example from some type of principle of indifference, and whether there are objective methods of determining posterior probabilities in case standard conditioning does not apply. My work is concerned with the latter problem, although the pme can also be used to defend objectivism about the former problem.

The scholarly consensus among epistemologists is that the pme is beset with too many conceptual problems and counterexamples to yield a generally valid objective updating procedure. Against the tide of this consensus, my work seeks to establish that the pme as the only candidate for a generally valid, objective updating procedure is also a successful candidate. Both the conceptual problems and the counterexamples are surmountable, as we will show in great detail.

Many of the portrayals of the pme's failings are flawed and motivated by a desire to demonstrate that the labour of the epistemologist in interpreting probability kinematics on a case-by-case basis is indispensable. This 'full employment theorem' of probability kinematics (which has a formally proven equivalent in computer science) is based on the wishful thinking of epistemologists. The pme combines a powerful and simple idea (update your probabilities in accordance with constraints revealed by the evidence without gaining more information than necessary) with a sophisticated formal theory which confirms that the powerful and simple idea consistently works.

Although the role of the pme in probability kinematics as a whole will be the scope of my work, I will pay particular attention to a problem which has stymied its acceptance by epistemologists: conditioning on conditionals. Two counterexamples, Bas van Fraasen's Judy Benjamin and Carl Wagner's Linguist problem, are specifically based on observation of conditionals as well as the pme's alleged failure to update on such observations in keeping with strong intuitions. The pme also faces much of its conceptual criticism on the front of conditioning on conditionals, especially with respect to epistemic entrenchment. Epistemic entrenchment is a fashionable way to look at updating on conditionals which assumes a second tier of commitment to propositions beneath the primary tier of quantitative degrees of uncertainty of belief such as probabilities (or ranks, for example). The pme ignores this second tier and is consequently at odds with the voluminous recent literature on epistemic entrenchment. A large part of my task is to address and defend the pme 's performance with respect to conditionals, both conceptually and with a view to threatening counterexamples.

2 Literature Review

David Hume poses one of the fundamental questions of the philosophy of science, the problem of induction. There is no deductive justification that induction works, as it does not result in necessary truths; and an inductive justification would beg the question. The late 19th and the 20th century bring us two responses to the problem of induction relevant to our project: Bayesian epistemology and a subjective interpretation of probability focus on uncertainty and beliefs rather than measuring frequencies in the world and give prime importance to decision problems; and philosophers of science generally recognize that some difficult-to-nail-down principle (indifference, simplicity, laziness, entropy) justifies entertaining certain hypotheses more seriously than others, even though they are all compatible with experience.

Pioneers of Bayesian epistemology and subjectivism are Harold Jeffreys (see Jeffreys, 1931, and Jeffreys, 1939) and Bruno de Finetti (see de Finetti, 1931 and 1937). For ideas of simplicity which lead to E.T. Jaynes' articulation of the pme in 1957, there is Richard Avenarius' telling book title Philosophie als Denken der Welt gemaess dem Princip des kleinsten Kraftmasses: Prolegomena zu einer Kritik der reinen Erfahrung (see Avenarius, 1876). Avenarius inspires Ernst Mach, who describes the economy of thought in physics (and the sciences generally) as a result of the desire to "master the sum of one's experience by the smallest possible effort" (see Mach, 1882).

The philosopher who in the beginning stages pursues this project most systematically is Rudolf Carnap in Der logische Aufbau der Welt, whose plan it is "that qualities should be assigned to point-instants in such a way as to achieve the laziest world compatible with our experience ... the principle of least action was to be our guide in constructing a world from experience" (see Quine, 1951). Carnap is one of the first who tried to think systematically about simplicity, preferring already in his dissertation Der Raum 'Einfachheit des Baues' (simplicity of the construct) over 'Einfachheit des Bauens' (simplicity of the construction) (see Carnap, 1922, 82) and then going on to write "Ueber die Aufgabe der Physik und die Anwendung des Grundsatzes der Einfachstheit" (see Carnap, 1923).

When Jaynes introduces the pme (see Jaynes, 1957a and 1957b), however, he is less indebted to the philosophy of science project of giving an account of semantic information (as in Carnap and Bar-Hillel, 1952 and 1953) rather than Claude Shannon's mathematical theory of information and communication. Shannon introduces information entropy in 1948 (see Shannon, 2001), based on work done by Norbert Wiener connecting probability theory to information theory (see Wiener, 1939). Jaynes also traces his work back to Ludwig Boltzmann and Josiah Gibbs, who build the mathematical foundation of information entropy by investigating entropy in statistical mechanics (see Boltzmann, 1877, and Gibbs, 1902). For the further development of the pme in probability kinematics it is important to refer to the work of Richard Jeffrey, who establishes the discipline (see Jeffrey, 1965), and Solomon Kullback, who provides the mathematical foundations of minimum cross-entropy (see Kullback, 1959). It is important to realize that Jaynes' pme is orginally conceived to give objective prior probabilities rather than objective posterior probabilities, but it is easy to generalize the pme for posterior probabilities once Kullback establishes the idea of divergence (first introduced in Kullback and Leibler, 1951). Jaynes himself presents this generalization in 1978 at an MIT conference under the title "Where Do We Stand on Maximum Entropy?" (see Jaynes, 1978), where he explains the Brandeis problem and demonstrates the use of Lagrange multipliers in probability kinematics. Ariel Caticha and Adom Giffin have recently shown, using Lagrange multipliers, that the pme seamlessly generalizes standard conditioning (see Caticha and Giffin, 2006).

Once the pme is formally well-defined and the breadth of its scope established (for the latter, Imre Csiszar's work on affine constraints is important, see Csiszar, 1967), its virtues come to the foreground. While Richard Cox (see Cox, 1946) and E.T. Jaynes defend the idea of probability as a formal system of logic, John Shore and Rodney Johnson use the idea of Cox's theorem to establish the uniqueness of the pme in meeting intuitively compelling axioms (see Shore and Johnson, 1980). By this time, however, an avalanche of criticism against the pme as an objective updating method has been launched. Papers by Abner Shimony (see Friedman and Shimony, 1971, Dias and Shimony, 1981, and Shimony, 1993) convince Brian Skyrms that the pme and its objectivism are not tenable (see Skyrms, 1985, 1986, and 1987). Bas van Fraassen's Judy Benjamin problem (see van Fraassen, 1981) deals another devastating blow to the pme in the literature, motivating Joseph Halpern (who already has reservations against Cox's theorem, see Halpern, 1999) to reject it in his textbook on uncertainty (see Halpern, 2003).

Teddy Seidenfeld leads his own campaign against objective updating methods in articles such as "Why I Am Not an Objective Bayesian" (see Seidenfeld, 1979 and 1986), while Jos Uffink takes issue with Shore and Johnson, casting doubt on the uniqueness claims of the pme (see Uffink, 1995 and 1996). Carl Wagner introduces a counterexample to the pme (see Wagner, 1992), again, as in the Judy Benjamin counterexample (but in much greater generality), involving conditioning on conditionals. In 2009, Igor Douven and Jan-Willem Romeijn write an article on the Judy Benjamin problem (see Douven and Romeijn, 2009) in which they ask probing questions about the compatibility of objective updating methods and epistemic entrenchment.

Epistemic entrenchment figures prominently in the AGM literature on belief revision (for one of its founding documents see Alchourron et al., 1985) and is based on two levels of uncertainty about a proposition: its static inclusion in belief sets on the one hand, and its dynamic behaviour under belief revision on the other hand. Wolfgang Spohn provides an excellent overview of the interplay between Bayesian probability theory, AGM belief revision, and ranking functions (see Spohn, 2012). To what extent the pme is compatible with epistemic entrenchment and a distinction between the static and the dynamic level will be a major topic of our investigation.

3 Proposal

The principle of maximum entropy (pme) is defensible across all counterexamples and conceptual issues raised so far as a generally valid objective updating method in probability kinematics. Subjectivists need to work a lot harder to undermine the validity of the pme, and a fortiori the validity of objectivism.

As it stands, the pme operates on the basis of an astonishingly simple principle: when updating your probabilities, waste no useful information and do not gain information unless the evidence compels you to gain it. The astonishingly simple principle comes with its own formal apparatus (not unlike probability theory itself): Shannon's information entropy, the Kullback-Leibler divergence, the use of Lagrange multipliers, and the sometimes intricate, sometimes quite straightforward relationship between information and probability.

There have been promising and mathematically sophisticated attempts to define probability theory in terms of information theory (see for example Ingarden and Urbanik, 1962, Kolmogorov, 1968, and Kampe de Feriet and Forte, 1967; for a detractor who calls information theory a "chapter of the general theory of probability" see Khinchin, 1957). For the straightforward relationship consider that information is added where probabilities are multiplied, as in the following formula:

-I(X)

P(X) = 2 (1)

While my interest in information theory as a basis for probability theory is historically at the source of my dissertation project, this idea is not at all essential to it. What is at the core of my project is the idea that information theory delivers the unique and across the board successful candidate for an objective updating mechanism in probability kinematics. This idea is unpopular in the literature. Carnap advises great flexibility with respect to formal updating procedures (think of Carnap's [U+03BB]-continuum of inductive methods, of which, as Dias and Shimony report, only [U+03BB] = [U+221E] is consistent with the pme, see Dias and Shimony, 1981). Jeffrey, van Fraassen, Halpern, Skyrms, Persi Diaconis and Sandy Zabell (see Diaconis and Zabell, 1982), Colin Howson and Allan Franklin (see Howson and Franklin, 1994), Douven and Romeijn all follow in his footsteps. In their opinion, probability update is a discipline which requires the epistemologist to study each situation on its own and address it with recourse to a toolkit of updating procedures, each being context-appropriate under different circumstances.

In order to defend our position, we need to address a series of counterexamples which at first glance discredit the pme: Shimony's Lagrange multiplier, van Fraassen's Judy Benjamin case, and Wagner's Linguist. For the latter two, I am confident that we can make a persuasive case for the pme, based on formal features of these problems which favour the pme upon closer examination. For the former (Shimony), Jaynes has written a spirited rebuttal: "This brings us, obviously, to the matter of Shimony. I am not a participant, but, like other readers, only a bewildered onlooker, in the spectacle of his epic struggles with himself" (Jaynes, 1985, 134). According to Jaynes, errors in Shimony's argument have been pointed out five times (see Hobson, 1972, Tribus and Motroni, 1972, Gage and Hestenes, 1973, Jaynes, 1978, and Cyranski, 1979). This does not, however, keep Brian Skyrms, Jos Uffink, and Teddy Seidenfeld from referring again to Shimony's argument in rejecting the pme in the 1980s, so the matter must be sorted out, and we promise to do so.

Once the counterexamples are out of the way, the more serious conceptual issues loom. There is good news and bad news for advocates of the pme. On the one hand, there are powerful conceptual arguments affirming the special status of the pme. Shore and Johnston, who use the axiomatic strategy of Cox's theorem in probability kinematics, show that relatively intuitive axioms only leave us with the pme to the exclusion of all other objective updating methods. Van Fraassen, R.I.G. Hughes, and Gilbert Harman's MUD method, for example, or their maximum transition probability method from quantum mechanics both fulfill their five requirements (see van Fraassen et al., 1986), but do not fulfill Shore and Johnston's axioms. Neither does Uffink's more general class of inference rules, which maximize the so-called R[U+00B4]enyi entropies, but Uffink argues forcefully that Shore and Johnston's axioms rest on unreasonably strong assumptions (see Uffink, 1995). Caticha and Giffin counter that Skilling's method of induction (see Skilling, 1988) and Jaynes' empirical results in statistical mechanics and thermodynamics imply the uniqueness of Shannon's information entropy over rival entropies.

To continue with the good news, the pme seamlessly generalizes standard conditioning and Jeffrey's rule where they are applicable (see Caticha and Giffin, 2006). It underlies the entropy concentration phenomenon described in Jaynes' standard work Probability Theory: the Logic of Science, which also contains a sustained conceptual defence of the pme and its underlying logical interpretation of probabilities. Entropy concentration refers to the unique property of the pme solution to have other distributions which obey the affine constraint cluster around it. When used to make predictions whose quality is measured by a logarithmic score functions, posterior probabilities provided by the pme result in minimax optimal decisions (see Tops[U+00F8]e, 1979, Walley, 1991, and Gruenwald, 2000). Under a logarithmic scoring rule these posterior probabilities are in some sense optimal.

On the other hand, now turning to the bad news, the belief revision literature has in the last twenty years mostly turned its attention to the AGM paradigm (named after Carlos Alchourron, Peter Gaerdenfors, and David Makinson), which operates on the basis of fallible beliefs and their logical relationships. These are really at this point two different epistemological worlds: the one where doxastic states are cashed out in terms of fallible beliefs which move in and out of belief sets; the other the world of probabilities where 'beliefs' are vague labels for a more deeply rooted, graded notion of uncertainty. Richard Jeffrey with his radical probabilism pursues a project of epistemological monism (see Jeffrey, 1965) which would reduce beliefs to probabilities, while Spohn seeks reconciliation between the two worlds, showing how their formal structure reveals many shared features so that in the end they have more in common than what separates them (see Spohn, 2012, 201).

In the end, our project, like Spohn's, is not about the semantics of doxastic states. We do not argue the eliminativism of beliefs in favour of probabilities; on the contrary, the belief revision literature has opened an important door for inquiry in the Bayesian world with its concept of epistemic entrenchment. This is a good example for the kind of cross-fertilization between the two different worlds that Spohn had in mind, mostly in terms of formal analogies and with little worry about semantics. Pioneering papers in probabilistic versions of epistemic entrenchment are recent (see Bradley, 2005, and Douven and Romeijn, 2009).

The guiding idea behind epistemic entrenchment is that once an agent is apprised of a conditional (indicative or material), she has a choice of either adjusting her credence in the antecedent or the consequent. Often, the credence in the antecedent remains constant and only the credence in the consequent is adjusted (Bradley calls this 'Adams conditioning'). Douven and Romeijn's give an example where the opposite is possible and only the credence in the consequent is left constant. Douven and Romeijn speculate that an agent could theoretically also take any position in between, and they use Hellinger's distance to represent these intermediary positions formally.

Even though Bradley, Douven, and Romeijn are in the world of probabilities, they are using a notion frequently used and introduced by the AGM literature to capture formally analogous structures in probability theory. The question is how compatible the use of epistemic entrenchment in probabilistic belief revision (probability kinematics) is with the pme. The pme appears to assign probability distributions or densities to events without any heed to epistemic entrenchment. The Judy Benjamin problem is a case in point. pme's posterior probabilities are somewhere in between epistemic entrenchments, as though mediating between them. The pme emphatically does not accept two levels of epistemic commitment: the static and the dynamic. AGM belief revision theory, Spohn's ranking functions, and epistemic entrenchments according to Bradley, Douven, and Romeijn suppose that beneath our credences (static probabilities), believers entertain a second set of dynamic probabilities which are determinative of the kinematics once doxastic states are subject to change. I cautiously hypothesize that the pme can only understand this second dynamic set of graded commitments as information. In other words, entrenchments are evidential, not epistemic.

Despite the potholes in the historical development of the pme, on account of its unifying features, its simple and intuitive foundations, and its formal success it deserves more attention in the field of belief revision and probability kinematics, definitely more attention than the many competing ad hoc methods which patch one problem while flagrantly tearing up many more somewhere else. The pme is not only the single principle which can hold things together over vast stretches of epistemological terrain (intuitions, formal consistency, axiomatization, case management), it also puts in question the scholarly consensus that such a principle is not needed.

4 Chapter Outline

4.1 Introduction

The introduction introduces the problem of probability kinematics for affine constraints, underlines the philosophical relevance, and embeds the discussion in a wider epistemological context.

4.2 Affine Constraints

Affine constraints come in three forms: observation and certainty of an event; complete re-partitioning of the event space; and reassessment of expectation. This chapter shows how these forms are related. The third form is a generalization of the second form, and the second form is a generalization of the first form. Consequently, a method for solving the problem of probability kinematics for the third form automatically solves this problem for the other forms. This chapter introduces standard conditioning and reviews the arguments why standard conditioning is widely accepted as a solution for affine constraints of the first form. This chapter introduces Jeffrey conditioning and reviews the arguments why Jeffrey conditioning is less widely accepted as a solution for affine constraints of the second form. The thesis of the dissertation is conditional on the acceptance of Jeffrey conditioning as a solution to the problem of probability kinematics for the second form, although the dissertation provides a brief overview of counterarguments in this chapter. Then this chapter introduces affine constraints which are not covered by Jeffrey conditioning. Let us call these strictly affine constraints. Strictly affine constraints exist, and they complete the list of forms that we need to consider for the problem of probability kinematics. There is a formal basis for these two claims. My work provides counterarguments to the claim that the pme can be subsumed under standard conditioning (for example in Domotor, 1985 and Skyrms, 1985, or retrospective conditioning, summary and problems in Diaconis and Zabell, 1982, 822, in action for the Judy Benjamin problem see Grove and Halpern, 1997).

4.3 The Principle of Maximum Entropy: Virtues

This chapter introduces the pme and highlights its virtues. The pme provides a unique solution for all affine constraints and is well-immunized against transformation variance (i.e. changing the angle on a problem does not change the results, as in Bertrand's paradox, see Jaynes, 1973). This solution uniquely fulfills several desiderata, and thus we are able to provide an axiomatic basis for its use. The pme accords with standard conditioning if the affine constraint is of the first form; and with Jeffrey conditioning if the affine constraint is of the second form. While it is highly contested as an objective updating method for strictly affine constraints, it exhibits important virtues as such a method. It underlies the entropy concentration phenomenon. Entropy concentration refers to the unique property of the pme solution which it has in relation to distributions which obey the affine constraint cluster around it. When used to make predictions whose quality is measured by a logarithmic score functions, posterior probabilities provided by the pme result in minimax optimal decisions. Under a logarithmic scoring rule these posterior probabilities are in some sense optimal.

4.4 The Principle of Maximum Entropy: Vices

A large majority of philosophers of science and epistemologists, however, rejects the idea that the pme is normative as an objective updating procedure given strictly affine constraints. Instead, they adhere to a type of 'full employment theorem,' according to which affine constraints must be submitted to the individualized attention of an expert before the problem can be considered properly investigated. I will call this camp 'opponents,' not because I necessarily believe that their claims are false, but because I believe that the arguments on which their claims rest are either faulty or incomplete. They have not made a convincing case that the pme lacks generality. Their case rests on the vices of the pme, which we will try to address in a comprehensive manner in this dissertation. This chapter portrays these vices in their strongest form possible: the Judy Benjamin problem; the Shimony objection; the Seidenfeld objection; and the Wagner objection. All of these objections derive counterintuitive results from the pme and claim that advocates of the pme find themselves in the awkward position of having to assent to something to which they would intuitively withhold assent.

4.5 Judy Benjamin

Opponents prominently cite van Fraassen's Judy Benjamin case to undermine the generality of the pme. This chapter shows that an intuitive approach to Judy Benjamin's case supports the pme. This is surprising because based on independence assumptions the anticipated result is that it would support the opponents. The chapter also demonstrates that opponents improperly apply independence assumptions to the problem. Analogous to the (1/2)er camp and the (1/3)er camp in the Sleeping Beauty case, this chapter gives rigorous arguments for the (1/3)er camp in the Judy Benjamin case.

4.6 The Shimony Objection

In a series of papers, Abner Shimony has highlighted cases in which the pme results in the counterintuitive claim that the observation (which is supposed to be informative) is expected with certainty before it is learned. This chapter introduces Shimony's objection and mounts a defence of pme against it.

4.7 The Seidenfeld Objection

Teddy Seidenfeld claims that the pme as a rule to represent partial information (an affine constraint of the third form) is unacceptable. It leads to precise probabilities that are excessively aprioristic, containing more information than the evidence generating them allows. This is a common objection to pme among philosophers. This chapter shows how this charge cannot be leveled against the pme without being leveled at Bayesian epistemology as a whole. It provides also a brief overview of reasons why within appropriate contexts it is advisable to accept the normative claims of Bayesian epistemology and therefore, in these contexts, the normative claims of the pme.

Besides this more general objection, Teddy Seidenfeld develops several formal lines of argument against the pme, for example that under certain circumstances involving noise factors the pme will inappropriately provide more information based on less evidence. Another example involves an alleged incompatibility between Bayes' Theorem and the pme. Seidenfeld also explains and expands the Shimony objection in a new light. This chapter looks in detail at available lines of defence for the pme against Seidenfeld's objections.

4.8 The Wagner Objection

Carl Wagner is a typical representative of full employment, who in his own words values the hard work of judging by warring against the temptations of mechanical updating. Other representatives of full employment with similar sentiments are Peter Gr[U+00A8]unwald, Joseph Halpern, Richard Bradley, Persi Diaconis, and Sandy Zabell. E.T. Jaynes, the most prominent proponent of the pme, sometimes speaks derisively about the relationship of the statistician (or formal epistemologist) and the client as one between a doctor and his patient, which I have characterized as full employment in allusion to the full employment theorem in computer science.

Wagner's formal attack against the pme concerns constraints that are a generalization of Jeffrey conditioning, for which Wagner suggests a revision procedure which violates the pme. Wagner's application of the principle of maximum entropy, however, is incorrect. A correct application agrees with his intuition. I present a formal proof that the principle of maximum entropy seamlessly and elegantly generalizes Wagner's generalization of Jeffrey conditioning, with the added benefit that it is not merely an ad hoc rule such as Wagner's, but integrated in a unified approach to probability kinematics that also makes intuitive sense.

4.9 Coarsening at Random

Coarsening at random (CAR) involves using more naive (or coarse) event spaces in order to arrive at solutions to probability updating problems. Gr[U+00A8]unwald and Halpern show that updating on the naive space rather than the sophisticated space is legitimate for event type observations when the set of observations is pairwise disjoint or when the CAR condition (as defined by them) holds. For Jeffrey type observations, there is a generalized CAR condition which applies likewise. For strictly affine constraints, the pme essentially never gives the right results, according to the authors. Gr[U+00A8]unwald and Halpern mention the Judy Benjamin problem as their prime example for a case in which the pme delivers the wrong result. It does so in analogy to the evidently wrong result in the Monty Hall or the Three Prisoners problem by naive conditioning. This chapter shows how the analogy is misguided and the authors' claim misleading that the pme essentially always gives the wrong results.

4.10 Families

Jos Uffink targets especially Shore and Johnson's assumptions when they identify the pme as the unique method for determining updated probability distributions, given certain types of rationality constraints. Uffink shows how a more reasonable restatement of Shore and Johnson's assumptions results in a whole class of updating procedures, the so-called R[U+00B4]enyi entropies. Van Fraassen, with collaborators R.I.G. Hughes and Gilbert Harman, also concludes that there is a family of candidates which fulfills five requirements or principles that he establishes with the Judy Benjamin problem in view. If the five requirements are complemented by two performance criteria, the pme is not the best updating method with respect to any of the performance criteria. The objection that instead of a unique rational updating method there is a family of such methods is again common among philosophers, especially because it accords with Rudolf Carnap's continuum of inductive methods in a different context. This chapter presents and expands on Ariel Caticha and Adom Giffin's arguments against families of updating methods.

4.11 The Two Level Objection

For people like Spohn, you can only have meaningful updating methods, either for rankings or for probabilities, if behind a first (static) layer of probabilities there is a second (dynamic) layer which determines how the first layer behaves when changes occur. The principle of maximum entropy, by contrast, only operates on one level and must incorporate the second level as information (see also epistemic entrenchment and Douven/Romeijn's Hellinger's distance). Degrees of belief about degrees of belief may be coherent and useful (see Jeffrey, 1974; Skyrms, 1980; and Domotor 1980, 1981). Our contention is that they can be subsumed under evidence-handling and are thus of a completely different source and nature than first-order degrees of belief.

4.12 Epistemic Entrenchment

Upon learning the truth of a conditional, degree of belief in the antecedent and the consequent may be re-evaluated. AGM belief revision theory has at its heart the idea of epistemic entrenchment, the degree to which an agent wants to maintain belief in a proposition when she is informed of its antecedent role in conditionals. Igor Douven and Jan-Willem Romeijn have applied epistemic entrenchment to the Judy Benjamin problem and evaluate it in terms of 'Adams conditioning,' the kind of conditioning where the degree of belief in the antecedent remains unaltered. This solution, which receives support in a number of other papers (for example, by Joseph Halpern), in contrast to the pme contradicts van Fraassen, Hughes, and Harman's five requirements.

Epistemic entrenchment and the pme are incompatible at first glance. Wolfgang Spohn's work in ranking theory illustrates a strong commitment to the idea that a rational agent needs both a quantitative assessment (either in terms of rankings or in terms of probability) as well as parameters of epistemic entrenchment in order to make sense out of updating in light of new observation or evidence. The pme, however, updates by operating on probability distributions alone, only in conjunction with an objective principle of minimum information gain. This chapter will seek reconciliation between the strong intuitive appeal of both the pme and epistemic entrenchment.

4.13 Epistemological Implications

In the concluding chapter, I want to look at the wider epistemological implications of a more affirmative stance towards the pme as an objective updating method. Most importantly, information becomes a basic notion to which philosophers will increasingly pay attention, possibly at the expense of probability. Probability theory is easily expressible in the terms of information theory, whereas the reverse is not true. Questions in epistemology, the philosophy of science, and perhaps also in the philosophy of cognition may be rearticulated and answered in different ways if information is more widely used as a common currency between these fields. In physics, information already plays a vital role, whereas the philosophy of information in many respects appears to be in its infancy.

5 Bibliography

Alchourr[U+00B4]on, Carlos E, Peter G[U+00A8]ardenfors, and David Makinson. "On the

Logic of Theory Change: Partial Meet Contraction and Revision

Functions." Journal of Symbolic Logic 510--530.

Avenarius, Richard. Philosophie als Denken der Welt gem[U+00A8]a[U+00DF] dem Prinzip

des kleinsten Kraftma[U+00DF]es. Leipzig: Reisland, 1876. ID: 254997873.

Bar-Hillel, Yehoshua, and Rudolf Carnap. "Semantic information." The

British Journal for the Philosophy of Science 4, 14: (1953) 147--157.

Boltzmann, Ludwig. On the Nature of Gas Molecules. Taylor and Francis,

1877.

Bradley, Richard. "Radical Probabilism and Bayesian Conditioning."

Philosophy of Science 72, 2: (2005) 342--364.

Carnap, Rudolf. "Der Raum: Ein Beitrag zur Wissenschaftslehre.", 1922.

Kant-Studien 1978, Erg[U+00A8]anzungsheft 56. Berlin: Reuther und Reichard.

Translated by Michael Friedman and Peter Heath as Space. A

Contribution to the Theory of Science. Unpublished.

. "[U+00A8]Uber die Aufgabe der Physik und die Anwendung des

Grundsatzes der Einfachstheit." Kant-Studien 28: (1923) 90--107.

Carnap, Rudolf, and Yehoshua Bar-Hillel. An Outline of a Theory of

Semantic Information. Cambridge, MA: MIT, 1952.

Caticha, Ariel, and Adom Giffin. "Updating Probabilities." In MaxEnt

2006, the 26th International Workshop on Bayesian Inference and

Maximum Entropy Methods. 2006.

Cox, Richard. "Probability, Frequency and Reasonable Expectation."

American Journal of Physics 14: (1946) 1.

17

Csisz[U+00B4]ar, Imre. "Information-Type Measures of Difference of Probability

Distributions and Indirect Observations." Studia Scientiarum

Mathematicarum Hungarica 2: (1967) 299--318.

Cyranski, John F. "Measurement, Theory, and Information." Information

and Control 41, 3: (1979) 275--304.

De Finetti, Bruno. "Sul significato soggettivo della probabilit`a."

Fundamenta mathematicae 17: (1931) 298--329.

. "La pr[U+00B4]evision: ses lois logiques, ses sources subjectives." In

Annales de l'institut Henri Poincar[U+00B4]e. Presses universitaires de France,

1937, volume 7, 1--68.

Diaconis, Persi, and Sandy L. Zabell. "Updating Subjective Probability."

Journal of the American Statistical Association 822--830.

Dias, P., and A. Shimony. "A Critique of Jaynes' Maximum Entropy

Principle." Advances in Applied Mathematics 2: (1981) 172--211.

Domotor, Zoltan. "Probability Kinematics and Representation of Belief

Change." Philosophy of Science 384--403.

. "Higher Order Probabilities." Philosophical Studies 40, 1: (1981)

31--46.

. "Probability Kinematics, Conditionals, and Entropy Principles."

Synthese 63, 1: (1985) 75--114.

Douven, I., and J.W. Romeijn. "A New Resolution of the Judy Benjamin

Problem." CPNSS Working Paper 5, 7: (2009) 1--22.

Kamp[U+00B4]e de F[U+00B4]eriet, J., and B. Forte. "Information et probabilit[U+00B4]e." Comptes

rendus de l'Acad[U+00B4]emie des sciences A 265: (1967) 110--114.

Friedman, Kenneth, and Abner Shimony. "Jaynes's Maximum Entropy

Prescription and Probability Theory." Journal of Statistical Physics 3,

4: (1971) 381--384.

Gage, Douglas W, and David Hestenes. "Comment on the Paper Jaynes's

Maximum Entropy Prescription and Probability Theory." Journal of

Statistical Physics 7, 1: (1973) 89--90.

Gibbs, Josiah Willard. Elementary Principles in Statistical Physics. New

Haven, CT: Yale University, 1902.

18

Grove, A., and J. Halpern. "Probability Update: Conditioning Vs.

Cross-Entropy." In Proceedings of the Thirteenth Conference on

Uncertainty in Artificial Intelligence. Providence, Rhode Island:

Citeseer, 1997.

Gr[U+00A8]unwald, Peter. "Maximum Entropy and the Glasses You Are Looking

Through." In Proceedings of the Sixteenth conference on Uncertainty in

artificial intelligence. Morgan Kaufmann Publishers, 2000, 238--246.

Halpern, Joseph. "A Counterexample to Theorems of Cox and Fine."

Journal of Artifcial Intelligence Research 10: (1999) 67--85.

. Reasoning About Uncertainty. Cambridge, MA: MIT, 2003.

Hobson, Arthur. "The Interpretation of Inductive Probabilities." Journal

of Statistical Physics 6, 2: (1972) 189--193.

Howson, C., and A. Franklin. "Bayesian Conditionalization and

Probability Kinematics." The British Journal for the Philosophy of

Science 45, 2: (1994) 451--466.

Ingarden, R. S., and K. Urbanik. "Information Without Probability."

Colloquium Mathematicum 9: (1962) 131--150.

Jaynes, E.T. "Information Theory and Statistical Mechanics." Physical

Review 106, 4: (1957a) 620--630.

. "Information Theory and Statistical Mechanics II." Physical

Review 108, 2: (1957b) 171.

. "The Well-Posed Problem." Foundations of Physics 3, 4: (1973)

477--492.

. "Where Do We Stand on Maximum Entropy." In The Maximum

Entropy Formalism, edited by R.D. Levine, and M. Tribus, Cambridge,

MA: MIT, 1978, 15--118.

. "Some Random Observations." Synthese 63, 1: (1985) pp.

115--138.

Jeffrey, Richard. The Logic of Decision. New York, NY: McGraw-Hill,

1965.

. "Preference Among Preferences." The Journal of Philosophy 71,

13: (1974) 377--391.

19

Jeffreys, Harold. Scientific Inference. Cambridge University, 1931.

. The Theory of Probability. Cambridge University, 1939.

Khinchin, A.I. Mathematical Foundations of Information Theory. New

York, NY: Dover, 1957.

Kolmogorov, A.N. "Logical Basis for Information Theory and Probability

Theory." IEEE Transactions on Information Theory 14, 5: (1968)

662--664.

Kullback, Solomon. Information Theory and Statistics. Dover

Publications, 1959.

Kullback, Solomon, and Richard Leibler. "On Information and

Sufficiency." The Annals of Mathematical Statistics 22, 1: (1951) 79--86.

Mach, Ernst. Die [U+00A8]okonomische Natur der physikalischen Forschung.

Vienna, Austria: Kaiserliche und K[U+00A8]onigliche Hof- und Staatsdruckerei,

1882.

Quine, W. V. Two Dogmas of Empiricism. New York: Longmans, 1951.

Seidenfeld, Teddy. "Why I Am Not an Objective Bayesian; Some

Reflections Prompted by Rosenkrantz." Theory and Decision 11, 4:

(1979) 413--440.

. "Entropy and Uncertainty." In Advances in the Statistical

Sciences: Foundations of Statistical Inference, Springer, 1986, 259--287.

Shannon, Claude Elwood. "A Mathematical Theory of Communication."

ACM SIGMOBILE Mobile Computing and Communications Review 5,

1: (2001) 3--55.

Shimony, Abner. The Search for a Naturalistic World View. Cambridge

University Press, 1993.

Shore, J., and R.W. Johnson. "Axiomatic Derivation of the Principle of

Maximum Entropy and the Principle of Minimum Cross-Entropy."

IEEE Transactions on Information Theory 26, 1: (1980) 26--37.

Skilling, John. "The Axioms of Maximum Entropy." In Maximum-Entropy

and Bayesian Methods in Science and Engineering, edited by G.J.

Erickson, and C.R. Smith, Dordrecht, Holland: Springer, 1988, 173--187.

20

Skyrms, Brian. "Higher Order Degrees of Belief." In Prospects for

Pragmatism: Essays in Memory of F.P. Ramsey, edited by David

Mellor, Cambridge University, 1980, 109--137.

. "Maximum Entropy Inference as a Special Case of

Conditionalization." Synthese 63, 1: (1985) 55--74.

. "Dynamic Coherence." In Advances in the Statistical Sciences:

Foundations of Statistical Inference, Springer, 1986, 233--243.

. "Updating, Supposing, and Maxent." Theory and Decision 22, 3:

(1987) 225--246.

Spohn, Wolfgang. The Laws of Belief: Ranking Theory and Its

Philosophical Applications. New York, NY: Oxford University, 2012.

Tops[U+00F8]e, F. "Information-Theoretical Optimization Techniques."

Kybernetika 15, 1: (1979) 8--27.

Tribus, Myron, and Hector Motroni. "Comments on the Paper Jaynes's

Maximum Entropy Prescription and Probability Theory." Journal of

Statistical Physics 4, 2: (1972) 227--228.

Uffink, Jos. "Can the Maximum Entropy Principle be Explained as a

Consistency Requirement?" Studies in History and Philosophy of

Science 26, 3: (1995) 223--261.

. "The Constraint Rule of the Maximum Entropy Principle."

Studies In History and Philosophy of Science 27, 1: (1996) 47--79.

Van Fraassen, Bas. "A Problem for Relative Information Minimizers in

Probability Kinematics." The British Journal for the Philosophy of

Science 32, 4: (1981) 375--379.

Van Fraassen, Bas, R.I.G. Hughes, and Gilbert Harman. "A Problem for

Relative Information Minimizers, Continued." The British Journal for

the Philosophy of Science 37, 4: (1986) 453--463.

Wagner, Carl G. "Generalized Probability Kinematics." Erkenntnis 36, 2:

(1992) 245--257.

Walley, P. Statistical Reasoning with Imprecise Probabilities. Chapman

and Hall London, 1991.

Wiener, Norbert. "The Ergodic Theorem." Duke Mathematical Journal 5,

1: (1939) 1--18.

21