Indeterminate Probability and Change of View

Isaac Levi

718 Philosophy Hall

Columbia University

New York, NY. 10027 USA

levi@columbia.edu

Abstract

This paper argues that if changes in confirmational commitment (or changes in prior probabilities) are to be defended without begging the question, indeterminacy in probability judgment is required. But no other modification of strict Bayesian doctrine is needed. The resulting qualified version of the Bayesian approach is compared with the way indeterminacy is deployed in the theory of H.E. Kyburg who uses a fixed confirmational commitment but, nonetheless, invokes indeterminacy.

An important motivation for introducing indeterminacy in probability judgment is the need to address the question of how and when probability judgments should be changed. I shall not attempt to address the question of change in probability judgment in detail here but only to explain why the very formulation of the question calls for a consideration of indeterminacy.

X's degree of credal or subjective or personal probability that *h* is true reflects X's judgment at a time of how much to risk on the truth of *h* or X's degree of certainty. But such credal probability judgments are themselves fine grained discriminations between propositions that X judges possibly true where "possible" means consistent with X's state of full belief K.

X's state of full belief constitutes not only X's standard for doxastic possibility. It also constitutes the background information and evidential basis for X's judgments of credal probability.

Changes in credal probability, therefore, are determined by changes in X's state of full belief and by changes in X's standard for assessing the credal state that is warranted by various potential states of full belief. Such a standard is a confirmational commitment (Levi, 1974; 1980, 4.4-4.5).

A *potential confirmational commitment C* is a function from potential states in *K* to sets of conditional credal probability functions *Q*(x/y) satisfying the following requirements:

Probabilistic Coherence: For every K, *C*(K) should be a set of conditional probability functions *Q*(x/y) defined for each proposition x in the algebra and each y in the algebra consistent with K that satisfy the requirements for finitely additive probability and that obey the multiplication theorem.

Consistency: *C*(K) should be nonempty if and only if K is consistent.

Probabilistic Convexity: For each y consistent with K, the set of conditional probability measures in C(K) restricted to $Q_y(x) = Q(x/y)$ is convex.

Confirmational Conditionalization: Suppose agent X who is in state of full belief K considers what X's judgment of credal probability should have been in some weaker state of full belief UK. X's current state of full belief K may be represented as an expansion UK^+_e of UK by adding e. For every probability measure in P(x/ye) in *C*(UK), there is a function $Q_e(x/y) = P(x/ye)$ in *C*(K) and conversely for every $Q_e(x/y)$ in *C*(K) there is a function $P(x/ye) = Q_e(x/y)$.

These constraints constitute minimal requirements for a *weak Bayesian* logic of probability judgment. (Keynes, 1921; Jeffreys, 1957,1961; Carnap 1962; Ramsey, 1990; B. De Finetti 1964, and Savage 1954 subscribe to these requirements.) To be sure, Ramsey, De Finetti and Savage did not explicitly consider confirmational commitments but only credal states. But they tacitly did so. In tackling a given problem in parameter estimation, all of them would consider a prior credal state and a subsequent one where credal probabilities are updated by conditionalizing with the aid of Bayes theorem. Taking the prior credal state to be relative to the state of information or full belief UK given by the model, and letting the new data be e, the posterior credal state is supposed to be obtained according to the following recipe:

Temporal Credal Conditionalization: If K is X's belief state at t_1 and K^+_e is X's (consistent) expansion of K at t_2, and *Q* is a probability function in X's credal state relative to K. Then there is a probability function Q_e in X's credal state relative to K^+_e such that

(*) $Q_e(x/y) = Q(x/ye)$.

Conversely, for every Q_e in X's credal state relative to K^+_e there is a probability function Q in X's credal state relative to K such that (*) holds.

Temporal credal conditionalization holds if and only if the same confirmational commitment is in force both initially and subsequent to the acquisition of e and confirmational conditionalization holds. I have been claiming that deriving posterior probabilities from prior probabilities and data via conditionalization and Bayes theorem as Ramsey, De Finetti and Savage all did is to invoke temporal credal conditionalization. Hence, I contend that these authors may be read as tacitly requiring adoption of confirmational commitments satisfying confirmational conditionalization.

These authors agreed with Jeffreys and Carnap that credal probability judgment should be determinate so that the following constraint should be satisfied:

Probabilistic Uniqueness: For every consistent K in *K*, *C(*K) is a singleton.

I shall call anyone who endorses probabilistic uniqueness in addition to the principles of minimal weak Bayesian probability logic a *strict Bayesian*. I endorse a weak Bayesian probability logic but I am not a strict Bayesian.

Advocates of the importance of objective probability or chance relative to a kind of trial are often interested in deriving degrees of credal probability from information about chances supplemented by information about the kind of trial that has been implemented on some specific occasion. This is the problem of direct inference. One needs to supplement the principles of probability logic with additional principles of direct inference if objective chance is introduced. De Finetti thought objective chance is meaningless and, as a consequence, had no need to supplement probability logic with principles of direct inference.

Someone who allows for objective probability and requires principles of direct inference in addition to the principles of probability logic just given endorses an *objectivist* probability logic. Those who endorse probabilistic coherence, consistency and convexity and confirmational conditionalization but no principles of direct inference endorse a minimal weak Bayesian probability logic.

No matter what weak Bayesian probability logic X adopts, there is a uniquely weakest confirmational commitment *C** such that *C*(*UK*) consists of all probability distributions satisfying probabilistic coherence and, in the case of objectivist logic, the requirements imposed by principles of direct inference. *C** is the *logical* confirmational commitment.

Necessitarians urge adoption of the logical confirmational commitment *C**. Necessitarian authors like H.Jeffreys and R.Carnap hoped that constraints on probability logic additional to direct inference and conditions of coherence could be defended sufficient to single out a unique standard probability measure on considerations of probability logic alone. Appeal to various principles of insufficient reason or symmetry have been invoked in the hope of securing the desired result.

Ramsey. Jeffreys and Savage were rightly skeptical of this claim but not of the importance of probability logic. On the other hand, they all (with the possible exception of Savage on some occasions) insisted that probability judgment be determinate.

This *personalist* attitude is untenable. By their own admission determinate credal states cannot be sanctioned as mandatory by probability logic. Rational agent X is free to choose any one of many potential confirmational commitments to use in X's deliberations.

In the case of full belief, X may coherently fully believe that *h* where *h* is extralogical. But X should be in a position to change X's full belief in the face of a serious challenge in a manner that begs no question either in favor of or against full belief that *h*. To do this calls for shifting to a position of suspense.

Numerically determinate credal probability judgment is just as opinionated as full belief that extralogical *h* is true. There is nothing wrong with such opinionation *per se* provided that X is prepared to modify X's probability judgment in the face of good non question begging reasons. Thus, it should be possible for an agent to regret his or her prior probability judgment. The trouble is that in the absence of the license to use a logical confirmational commitment, it seems reasonable to provide for the revisability of confirmational commitments. But there does not appear to be any way to see a change from one strictly Bayesian confirmational commitment to another without begging the question.

To avoid question begging, X would have to shift to a credal state that does not rule out any credal probability distribution in contention. That is to say, X's credal state would have to recognize more than one credal probability distribution to be permissible. X's credal state would have to become indeterminate. This is precisely what a strict Bayesian will not allow.

It is as if someone were forbidden to suspend judgment as to the truth of a hypothesis even though no principle of logic or rationality favored adopting either side of the issue. In our case, the personalist idea amounts to insisting that we should adopt a numerically determinate probability judgment even though no principle of probability logic mandates this. Whether we are attending to full belief or probability judgment, the policy is the same. *Do not leave any room for doubt.*

Authors who agree with the personalists that probability logic is not even remotely capable of yielding a strictly Bayesian *logical* confirmational commitment may become skeptics who fully believe only logical truths and refuse to rule out any probability judgment that meets the minimal requirements of probability logic. There is, however, an alternative to such necessitarianism.

Revisionists in probability judgment provide room for doubt in confirmational commitment. But they do not insist on the radical skepticism of the necessitarian who demands maximal doubt under all circumstances. Because they are prepared sometimes to adopt determinate probability judgments, sometimes maximally indeterminate probability judgments and sometimes probability judgments of intermediate indeterminacy, they must allow for the modifiability of probability judgment through changing confirmational commitments as well as states of full belief.

Perhaps the main advantage of representing probability judgment by means of confirmational commitments rather than by states of credal probability judgment is that confirmational commitments may vary independently of states of full belief and vice versa. This is not true of credal states. Many writers of subjectivist or personalist persuasion take change in credal state as basic and seek to account for changes in credal state without distinguishing between changes due to changes in state of full belief and changes due to changes in confirmational commitment. In so doing, personalists tend to beg the question as to the extent to which probability judgment is or is not independent of changes in states of full belief. By representing credal state as a function of two components, state of full belief and confirmational commitment, the extent of independent variation can be explored without begging the question one way or the other.

Henry Kyburg (Kyburg, 1961,1974), like De Finetti, has no use for objective probability. He thinks relative frequency of target attributes in reference classes can do all the work that objective probability is intended to do. Unlike De Finetti, Kyburg is a necessitarian. To nail down his position, Kyburg has offered an account of direct inference of his own. It derives credal probabilities from information about the frequencies that serve as surrogates for objective probabilities. And it is sufficiently flexible that one can replace the frequencies by statistical probabilities without undermining the salient features of the account of direct inference that emerges. It is one of the most detailed positions available.

Kyburg admits that credal probability judgment may and, in general should go indeterminate. This may be illustrated by case 1.

Case 1: Jones knows that a ball is to be drawn from urn A or from urn B depending upon whether a given fair coin lands heads up or tails up. Urn A contains 90 red balls and 10 blue ones. Urn B contains either (H_1) 90 red balls, one white and 9 blue or (H_2) 80 red, one white and 19 blue. Relative to this information, what probability judgments would Kyburg recommend that Jones make for the predictions that ball drawn is red, is white and is blue?

Since Jones is given no basis for making a probability assignment to H_1 and H_2 conditional on urn B being used because the coin landed tails, we should, according to necessitarian Kyburg, allow any probability from 0 to 1 for H_1. This means that in the "reference class" *T* of trials consisting of tossing a coin and drawing a ball from A if the coin lands heads and from B if the coin lands tails, the probability of red is the interval [0.85,0.9], white is 0.005 and blue is [0.095, 0.195]. Instead of using the three interval valued probabilities, one can take the convex hull of the two distributions <0.9, 0.005, 0.095> and <0,85, 0.005, 0,195>. Kyburg does not follow this practice. Nonetheless, if one has a Boolean algebra of "target" attributes as Kyburg does typically employ, then in case 1, there is no difference between adopting the convex set representation I am employing and supplying interval valued probabilities for each of the target features red, white and blue. This is not always so.

Case 2: Jones knows that trial belongs to the reference class *Ttails* where the coin lands tails up so that ball is selected from urn B.

Here we have statistical information relative to two reference classes: *T* and *Ttails*. The credal state determined by T is already given in case 1. For *Ttails* we may seek to identify the credal state by considering the convex hull of <0.9, 0.01, 0.09> and <0.8, 0.01, 0.19>.

This convex set does not contain and is not contained in the convex set specified for *T*. The kind of experiment is more specific (or the reference class is narrower). It would appear to be perfectly Kyburgian to favor the narrower reference class in such a case and adopt the convex set specified for *Ttails*. *The* interval valued probabilities would be <0.8, 0.9>, <0.01, 0.01>, <0.09, 0.19> for red, white and blue respectively.

Kyburg does not proceed in this fashion. He uses his method for selecting reference classes in direct inference for red, white and blue separately. In those cases Kyburg specifies interval-valued probabilities.

Kyburg's recommendation for red is [0.85, 0.9] - i.e., relative to *T* even though Jones has the more specific information that the trial belongs to *Ttails.* The interval of probabilities [0.8, 0.9] relative to *Ttails* is wider and should according to Kyburg's principles be suppressed. For the same reason, the recommendation for blue is to use the same reference class and obtain the interval [0.095, 0.195].

In the case of white, the recommendation is to use the more specific reference class *Ttails* according to which the interval is [0.01, 0,01]. The resulting network of intervals does *not* agree with the intervals obtained by appealing first to the families of distributions licensed according to *T* and according to *Ttails.*

But even if one modifies Kyburg's proposals so that they are used to derive sets of probability distributions over target predicates belonging to some privileged partition and all Boolean combinations of these, Kyburg's proposals for selecting reference classes violate weak Bayesian strictures. According to weak Bayesian requirements, an inquirer should always base direct inference on the narrowest (or strongest) reference class under which the experiment is known to fall. Kyburg's proposals allow for violations of this requirement whether we use Kyburg's own version of his recommendations or my modification. In particular, when the set of probability distributions relative to the narrowest reference class includes the set of distributions relative to the broader one, the set relative to the broader reference class prevails.

Case 3: Urn C contains 90 black balls and 10 white. Urn D contains 10 black balls and 90 white. Trial *T** is selecting a ball at random blindfolded from one of the urns. Trial *T*black* is a trial of kind *T** where the color of ball drawn is black. An outcome has the target attribute of being a winner iff it is a black ball selected from C or a white ball selected from D. It is a loser otherwise. The objective statistical probability of a winner is 0.9 on a trial of kind *T**. The statistical probability of a winner on a trial of kind *T*black* is the interval [0,1] according to Kyburg. Smith knows that the trial is of kind *T*black* and, hence, of kind *T*.* Kyburg would recommend adopting as one's credal probability 0.9 for a winner.

This recommendation has a remarkable property. Prior to sampling, the prior probability that the urn is C is the interval [0,1]. Upon finding out that the ball selected is black, the credal probability changes to 0.9. Starting with maximal ignorance, a very determinate probability judgment is obtained from the data. This is a clear violation of confirmational conditionalization. So Kyburg's theory violates the minimal requirements for a weak Bayesian probability logic.

Does that refute Kyburg's theory? Anyone who is prepared to abandon deductive closure as a condition on rationally coherent states of belief (as Kyburg is prepared to do) is not likely to be daunted by this point. And many who endorse the importance of indeterminacy in probability judgment are unlikely to criticize allies in the resistance against the tyranny of strict Bayesian probabilism. But I, for one, think that resistance to strict Bayesianism is not enough. One should consider the basis on which one's resistance is founded.

My own view has been that strict Bayesianism would not be so bad were it not so strict. I am a fan, for example, of the Sure Thing Principle and maintain that alleged failures of experimental subjects to obey it are often a reflection of indeterminacy in either probability or utility judgment (Levi 1986, 1997). Now the Sure Thing Principle lies at the cornerstone of the interpretation of conditional probabilities in terms of called off bets where a possible state is rendered irrelevant by featuring a constant payoff to all the available options. This definition presupposes the Sure Thing Principle according to which the preferences among options remain the same when the payoffs to all options in a given state are the same regardless of the magnitude of the constant payoff.

In called-off bets, the constant column remains nonetheless a serious possibility consistent with the agent's state of full belief. But consider the initial state of belief modified by ruling out that state as a serious possibility. If this transformation does not alter the evaluation of the options, the new probabilities are a conditionalization of the old. The idea behind confirmational conditionalization is that when the confirmational commitment is the same, the evaluation of the options should not vary (Levi 1980, 10.4).

Anyone who finds this line of thinking compelling, as I do, must reject Kyburg's view of direct inference. I am strongly opposed to the dogmatism of strict Bayesianism in all its forms. But I am an unreconstructed weak Bayesian. I think that abandoning confirmational conditionalization is throwing out the baby with the bathwater. Kyburg, on the other hand, has been a persistent champion of an alternative to weak Bayesianism accounts of rational probability judgment.

Weak Bayesian revisionists need to furnish accounts of how confirmational commitments are changed. Kyburg's necessitarianism relieves him of this obligation. It may seem more congenial to the proceduralism defended by many computer scientists when they explore questions in epistemology. Changing confirmational commitments, like changing states of full belief, is often a matter of decision based on deliberation relative to the goals of the deliberating agent. One program does not fit all. Nor do a small number of programs. It is understandable why those sympathetic with proceduralism might sympathize with Kyburg.

Advocates of a coherentist or objectivist version of Bayesian probability logic can be as resolute in their resistance to strict Bayesianism as can advocates of Kyburg's brand of fiducial reasoning. But formally the shift from strict Bayesianism to coherentist or objectivist weak Bayesianism is relatively tame. Those who look for raw meat would do well to sample from Kyburg's menu.

Nonetheless, Kyburg's proposals like those of Fisher and the different proposal of Dempster have the dubious merit of promising something for nothing. We are supposed to be able to distill something from data even when our prior probability judgment is as indeterminate as probability logic will allow. This is *creatio ex nihilo* of the sort one comes to believe when situated on the big rock candy mountain. As I see it, there is no free lunch. The revisionist version of weak Bayesianism I am advocating sees the choice of prior probabilities in deliberation and inquiry as a problem that needs to be addressed within the context of the issue being addressed.

To be sure, appealing to context can be the refuge of the philosophical scoundrel. To do so responsibly, one should try, as much as possible, to specify how. That is the larger and very serious issue that cannot be addressed here.

There are, of course, several alternative approaches to indeterminate probabilities besides Kyburg's and mine. T. Fine (Fine 1973) is an old veteran. His student P. Walley (Walley 1991) has written a magisterial survey of the subject. And, of course, mention should be made of the fundamental contributions of Kadane, Schervish and Seidenfeld (Kadane *et al*. 1999) and Seidenfeld and Wasserman (Seidenfeld *et al.* 1993).

My aim here has been to argue for two claims: (a) Accounting for the revisability of confirmational commitments in a nonquestion begging way requires acknowledgment of indeteterminacy in probability judgment. (b) *Contra* Kyburg among others, it is desirable to do so while preserving the core requirements of the classical strict Bayesian doctrine other than numerical determinateness. A weak Bayesian probability logic supplemented perhaps by rules of direct inference for objective probability should provide an adequate probability logic.

Supplementing this approach with an account of how states of full belief and confirmational commitments ought to be changed should yield all that is required for the purpose of applying probability judgments in scientific inquiry and in practical contexts.

References

Carnap, R. 1962. *Logical Foundations of Probability* 2nd ed. Chicago: Chicago University Press.

De Finetti, B. 1964. Foresight: Its Logical Laws, Its Subjective Sources. In *Studies in Subjective Probability* ed. by H.E. Kyburg and H.E. Smokler, New York: Wiley.

Fine, T.L. 1973. *Theories of Probability*. New York: Academic Press.

Jeffreys, H. 1957. *Scientific Inference* 2nd ed. Cambridge: Cambridge Unversity Press.

Jeffreys, H. 1961. *Theory of Probability* 3rd ed. London: Oxford University Press.

Kadane, J.B., Schervish, M.J. and Seidenfeld, T. 1999. *Rethinking the Foundations of Statistics*, Cambridge: Cambridge University Press.

Keynes, J.M. 1921. *A Treatise on Probability*. London: MacMillan.

Kyburg, H.E. 1961. *Probability and the Logic of Rational Belief*, Middletown, Conn.: Wesleyan University Press

Kyburg, H.E. 1974. *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.

Levi, I. 1974. On Indeterminate Probabilities, *The Journal of Philosophy*.

Levi, I. 1980. *The Enterprise of Knowledge*, Cambridge, Mass.: MIT Press, paper, 1983.

Levi, I. 1986. *Hard Choices*, Cambridge: Cambridge University Press.

Levi, I. 1997. *The Covenant of Reason*, Cambridge: Cambridge University Press.

Ramsey, F. 1990. *Philosophical Papers*. D.H. Mellor ed. Cambridge: Cambridge University Press.

Savage, L.J. 1954. *The Foundations of Statistics.* New York: Wiley.

Seidenfeld, T. and Wasserman, L. 1993. Dilation for Sets of Probabilities. *Annals of Statistics* 21, 1139-54.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hill.