Mushy Credences II

Topics in Epistemology
Oxford, 7/6/12

Daniel Rothschild
daniel.rothschild@philosophy.ox.ac.uk

1 Dilation

Even if we can convince ourselves that unsharp credences can be combined with a decision theory that does not issue in palpable irrationalities, we might still reject the representation. This is because of a very strange property of mushy credences known as DILATION in the literature.[1]

1.1 van Fraassen's exam

van Fraassen [2006] considers a variation of a case in Seidenfeld and Wasserman [1993]. Here is another variation. Suppose there is an exam next Tuesday. You think you have a 90% chance of passing the exam. You also think there is a 50% chance of rain. You have no beliefs about the relation between the exam and the rain (i.e. for all you know they are correlated or they are independent. One presentation of this state is the set $C = \{p : p(\text{'rain'}) = .5 \text{ and } p(\text{'pass'} = .90\}$. Suppose you learn that it will rain, how does that affect your credence that it will pass? If you conditionalize every member of C, with 'rain' or 'not rain' you will get a C that assigns the interval from .8 to 1 to 'pass'.
Variations: problems get even worse if we make the initial belief in rain a large interval. (See van Fraassen for this example).
Another case: two coin flips with 50/50 chance each, but which may or may not be independent to any degree. After first flip our expectations for second flip expand.
Note: in all of these problems we *assumed* total ignorance of independence. Maybe this itself is what's problematic?

1.2 White's coin

White [2010] gives perhaps the most compelling case where dilation leads to what seem to be absurd consequences. His case can be described as follows:

> Take some past proposition, p for which you are entirely clueless. Say, for instance, that the price of oil was linked to the price of cotton in 1885. Now suppose your friend Bill knows the truth of this proposition. He takes a fair coin and effaces both the head and tails side, so they are no longer visible. He then writes 'p' on heads if the proposition is true, and 'p' on tails if the proposition is false. Then he flips the coin, and you observe, let us say, a p.

[1] Seidenfeld and Wasserman [1993] give a rather technical account of dilation, and seem to accept it as a basic feature of these kinds of representation of probability.

So once you know that p shows, you know that the coin landed heads iff p. It seems like then you should assign the same credence to the coin landing heads and p showing but how does this work?
Let us label the following events: 'Heads', 'p shows', 'p'. Initially we describe your credal set C includes all probability functions that assign 50/50 to 'Heads' while assigning any probability to p. We know that 'p shows' iff ('Heads' and p) or ('not Heads' and $\neg p$). One probability in this C assigns 1 to 'p', so if we conditionalize on 'p shows' we then assign 1 to 'Heads'. Another assigns 0 to 'p', so if we conditionalize on 'p shows' we then assign 0 to 'Heads'. By convexity we thus must assign the whole range from [0,1] (assuming convexity is preserved under conditionalizing, which it is).[2]
Formally, then when we update by conditionalizing each member of C on 'p shows' we *dilate* our probability of heads from .5 to [0,1]. This, White argues, is weird. We could also *narrow* our probability in p in response to the coin flip, which would also be strange.
Some aspects of the strangeness:

Betting behavior On pretty much any dilation view it should be permissible to accept bets with bad odds for the coin flip *after* seeing the answer. (This worry seems somewhat independent of Elga's as solutions to Elga's don't seem to help here.)

Failures of reflection A variation on van Fraassen's REFLECTION principle states that if you now know that you will get new credences in response to evidence that will come, then you should change your credences already. Either narrowing or dilation violates this. While the Reflection principle has many known counterexamples (such as cases where you know where you will forget things) it seems like the failure here is not explicable in such terms since nothing epistemically goes wrong, it appears.

Breakdown of statistical reasoning Consider a million events with mushy credences. A million coins are labelled and flipped. What beliefs should you have about the coin flips afterwards? Statistical laws get quite a lot of bite here, but once you have seen the ps and blanks, your credences should be entirely open. Like White, I take this as the most severe worry, as it just seems to violate basic norms of rationality to not believe that about half the coins will be heads after the flip.

1.3 A Variation

Can we answer the statical worry: well one thing to say is that having *these particular* thick probabilities is just not rational, even if dilation is acceptable in general. Nothing after all in the picture implies that *any* set of thick credences is rationally permissible. White's case has a feature that you might think is significant: the symmetry of the example, so that unsharply considered proposition is as likely to be wrong as right.[3] Consider first a case where we believe p with (precise) .6 credence. We then label the

[2] Note we did not specify that C makes the coin toss and p independent, but that specification will not change this case, so independence (or lack thereof) does not play a key role as it did in van Fraasssen's case.

[3] This seems particularly important because White discusses centers his paper around the principle of indifference which would require .5 probability in these cases, but says nothing about unbalanced cases.

coin and flip it as in the White scenario. In this case if p shows, it seems like our probability of heads should shift to .6.

Sturgeon [2010] discusses a variety of cases like this and argues (inter alia) that in non-centered cases with mushy credences we *should* shift our credences that way. Sturgeon writes:

> In each of these unbalanced cases is it pre-theoretically obvious that your new take on pup should be shaped, somehow, by your old take on p, your old take on pup, your initial view that the truth-values of p and *pup* have nothing to do with one another, and so forth. In a nutshell: it is pre-theoretically clear, in the unbalanced cases above, that your new take on *pup* should be shaped by your old take on the set up. Yet that clarity is not itself based on theory or fidelity to model. It is just common-sense. In the unbalanced cases above, therefore, it is pre-theoretically plausible that your new view of *pup* should inherit some thickness from your old take on p.

Consider an arbitrary proposition, U, for which you have a thick, but non-symmetric credence in with range $[p-\epsilon, p+\epsilon]$. We can think of U as a random variable that is either equal to 0 or 1 (false or true). Then imagine a fair coin, F, which is another random variable with probability .5 of being 0 or 1 (tails or heads) with a probability independent of U. Then consider a third random variable $V = 1$ if $U = F$, and 0 otherwise. Note that the while V is dependent on U, it has a sharp probability itself, namely .5.[4] However, once we conditionalize on the value of U we will get an unsharp probability for the proposition that $V = 1$, particularly $p(V = 1|U = 1)$ is the whole interval $[p-\epsilon, p+\epsilon]$, whereas $p(V = 1|U = 0)$ is $[1-(p+\epsilon), 1-(p-\epsilon)]$. In either case, something unsharp.

Why is this problematic? Characterizing a reflection failure here is somewhat hard (?) but statistical problems will arise. Here's a sketch of an argument: take 1,000,000 events/flips (each independent and each of each other, and each of the events with a thickness of at least ϵ). One should have very high confidence that the value of V will be within a certain percentage of 50% of 1,000,000. However, once you see the value of the coin flips, your confidence will erode with the size of ϵ. Even a small ϵ will be problematic with a sufficient number of flips.[5]

1.4 Another variation: a puzzle without explicit conditioning

Let us again consider a large number of (this time, say, 1000), propositions for which you have *even*, mushy credences. Take a set of 1000 random variables U_i with values 0 or 1 (i.e. true or false), $0 \leqslant i < 1000$. Here are two things that seem plausible to say about ones' credences about this set of events:

Statistical ignorance You can't have, say, 90% confidence that about 500 over them will turn out to be true.

Irrelevance of negation considering the set of events which in which some subset form the original are replaced by their negations (instead of the original proposition) shouldn't affect one's statical beliefs about them (since one's views are symmetric).

These two beliefs are probably not maintainable together. For if you decided to arbitrarily negate some of the propositions, choosing each time with a coin flip, then the expected number of events that will be true (with those negations) should have a binomial distribution (at least this follows from probability theory for any probability distribution over all the original events). The argument for this is simple: there are really three random variables here: U_i, the fair coin F_i and then the variable representing whether the possibly negated proposition is true which is V_i as above. For any mushy probability of U_i we will still have a sharp .5 probability for V_i. Even if the U_i are not independent all the V_i's will be, thus the some of them has a binomial distribution (i.e. it is like the sum of a lot of coin flips).

So if you take probability theory seriously (and the statistical inferences it licenses) then **Irrelevance of negation** for your confidences is not compatible with **Statistical ignorance**. Of course, this reasoning depends on taking an arbitrary *sharp* probability function to represent the probability of the U_is, but since our credence in them is mushy we could argue that that is not a legitimate move. However, I suspect that *as a matter of fact* most of the time when you take a large set of propositions arbitrarily negate some of them by a coin flip, then count the number of true ones, that number will be around 50% of the cardinality of the original set. So maintaining **Statistical ignorance** seems impossible if you accept **Irrelevance of negation**. We could thus try to reject **Irrelevance**. But on what grounds? If we really don't know much of anything the likelihood of a set of propositions, why should our beliefs change if we consider the negation of some them?

Here's a more precise case that illustrates the parallels to White's coin case:

Consider first 4 events for which you have even, mushy credences, but which you are certain are independent probabilistically. We'll assume for each your confidence ranges from .1 to .9 (and all combinations of credences for all 4 are live possibilities). What should your confidence none of these events happen. Well it should range from the lowest probability ($.1^4 = .0001$) to the highest ($.9^4 = .6561$). Now consider instead if we think of the probability that the first is false and the other three are true, this will give the same answer. Indeed you will have the same answer for *any* assignment of truth values to the four propositions. Now suppose you decide to flip 4 fair coins to determine which truth values you are interested in: if the first is heads you are interested in whether the first proposition is true, if tails than false, and so on for the other three coins. So you might be interested in the following proposition: will the truth values come out exactly as the 4 four coin flips land? Regardless of the coin flips we already have a confidence in this though, namely the interval .0001 to .6561 since that's our confidence interval for *any distribution of truth and falsity across these four propositions*. However, it's also true that *whatever the actual distribution of truth and falsity is for these propositions* the probability the coin toss will match the distribution should be $.5^4 = .0625$, by simple probabilistic reasoning.

[4] For any $p \in C$:

$$p(V = 1) = p(U = F = 1) + p(U = F = 0) = .5 * p(U = 1) + .5(1 - p(U = 1)) = .5.$$

[5] This would need to be worked out more carefully: it's just a sketch of the consideration.

2 Against conditioning sets

As Halpern [2003] notes it is not at all obvious (independently of dilation phenomenon) that conditioning each member is a good way to update sets of probability in light of new information. Suppose you have a coin of uncertain bias in the range of [.3, .6]. You are going to flip it hundred times. What should you probability be that it lands heads on the 100th flip after it lands heads on each of the prior 99 flips. It will still be [.3, .6]: after all when updating any individual function in the set there is no effect of past flips on probability of future flips. Again it seems like a better representation to have some form of distribution *over* biases rather than a mushy representation.

This criticism of course, cannot apply directly to one-off events (but it might apply indirectly).[6] However, it should make us wary of using conditionalizing each member as a way of updating mushy sets.

Should we then advocate sharpening as the right response to White's coin puzzle (at least in some cases)? After all it seems that sharpening *is* sometimes an appropriate epistemic response to new evidence (independently of the dilation cases).

However, note that hear each member of the credal set has certainty in the bias of the coin: this seems wrong. So the issue might just be a poor choice of thick credences.[7]

References

Joseph Y. Halpern. *Reasoning about Uncertainty*. MIT Press, 2003.

Sarah Moss. Epistemology formalized. *Philosophical Review*, forthcoming.

Teddy Seidenfeld and Larry Wasserman. Dilation for sets of probabilities. *The Annals of Statistics*, 21:1139–1154, 1993.

Scott Sturgeon. Confidence and course-grained attitudes. In Tamar S Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology*, volume 3. OUP, 2010.

Eric Swanson. *Interactions with Context*. PhD thesis, MIT, 2006.

Bas van Fraassen. Vague expectation value loss. *Philosophical Studies*, 127:483–491, 2006.

Roger White. Evidential symmetry and mushy credences. In Tamar S Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology*, volume 3. Oxford University Press, 2010.

Seth Yalcin. Epistemic modals. *Mind*, 116:983–1026, 2007.

[6] Consider an expert who designs a coin with a bias corresponding to his sharp credence in the one-off event.

[7] Thanks to Sarah Moss for pointing this out to me.