SLIDE 1 (below)

Pryor advocates a departure from Bayesianism and towards a belief dynamics that violates finite additivity. He gives four independent types of reasons for this:

1. Probability can't model the difference between ignorance and equal chances.
2. Probability can't model different levels of resilience of a belief.
3. Probability can't model important epistemic priority distinctions.
4. Probability doesn't give Pryor's preferred answer about perceptual justification and skepticism (the Dogmatist position).

I'm going to argue that in 1, 2, and 3, Bayesianism is falsely accused. If I'm right then this means that Pryor's case for the new dynamics rests entirely on 4. 4 has been argued by critics of Pryor's Dogmatist position about skepticism (e.g. Roger White), and I'm going to add further reason to think that position is inconsistent with a standard probability approach and that the standard probability approach is right, under a certain interpretation of the Dogmatist position. In response to 4 Pryor proposes an alternative belief dynamics which violates finite additivity, a move that seems to me more drastic than is motivated. So, the final point I want to make is that if one is sympathetic to the Dogmatist position, then one doesn't need to do something so drastic. This is because there are two alternatives that will have similar effects without giving up finite additivity. One of them is a more faithful interpretation of the Dogmatist position.

SLIDE 2 (below)

Claims 1 and 2 will both seem erroneous to a Bayesian because of the property of **variance**, which every probability distribution or density has. That is, the probability function that the Bayesian imagines as modeling the degrees of belief of a rational subject doesn't merely have arguments and values (propositions and degrees of belief). It has other properties too, like variance, the average distance of the distribution or density from the mean. And the **values** of two probability functions (degrees of belief) can be the **same**, while the **variance** of those functions is **very different**. This is because the "value" of the function (which corresponds to degree of belief) is actually an average, an ***expected*** value. This is a weighted average of the probabilities of the possible values of the random variable in question. It's clear that the average can be the same while the distribution, and thereby the average distance of the distribution from that average (the variance), can be very different. To illustrate: suppose we are interested in the probability of h, that a certain coin will come up heads on the 1,000th toss. (See SLIDE 2 below.) We imagine a random variable *X*, the probability of heads, and it takes as values all the real numbers *x* between 0 and 1. The expected value of our probability density function, which corresponds to our subject's degree of belief, is the weighted average of the density function. P_0 is the density function before we witness any tosses of the coin. Because before any tosses the density function is uniform, the expected value is 1/2.

We compare the prior probability density P_0 with the probability density after 999 coin tosses, P_{999}, and we note that if the coin is fair, the expected value E_{999}, once again the weighted average, is roughly the same as the expected value E_0, namely 1/2. This reflects in the first case the fact that before any tosses we don't know whether or how far the coin is biased, and in the second that we have good reason to believe the coin is fair. The difference between these two states is reflected not in the subject's *degree of belief* that a head will occur on the thousandth toss (the expected values are the same throughout), but in the variance, the average distance of the values of the density function from the mean. For P_0 the variance is very high, for P_{999} much lower. The decrease in variance corresponds exactly to a decrease in the subject's **ignorance** about the true chances of a head on the thousandth toss. The other thing it corresponds to is increasing **resilience** of the subject's degree of belief against changes by future data. Mathematically, one data point, whatever it is, makes far less difference to the expected value of P_{999} than it does to the expected value of P_0. The **upshot** is that probability already is two-dimensional; we don't need to give up finite additivity in order to model the difference between ignorance and equal chances or to model levels of resilience of a degree of belief.

Okay, let's move on to epistemic priority issues.

Varieties of defeating and promoting evidence. Broadly speaking, it seems to me that the concept of **screening off** makes all the important distinctions that Jim's examples contain.

Remainder of slides contain the further remarks.

Undermining and Uncertainty -- Commentary (Roush)

Pryor's argument:

1. Probability can't model the difference between ignorance and equal chances.

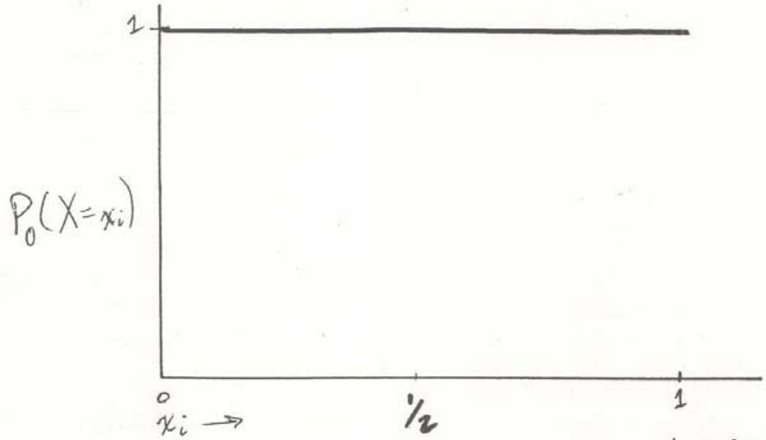
2. Probability can't model different levels of resilience of a belief.

3. Probability can't model important epistemic priority distinctions.

4. Probability doesn't give my preferred answer about perceptual justification and skepticism (the Dogmatist position).

Therefore, 5., we need a new belief dynamics.

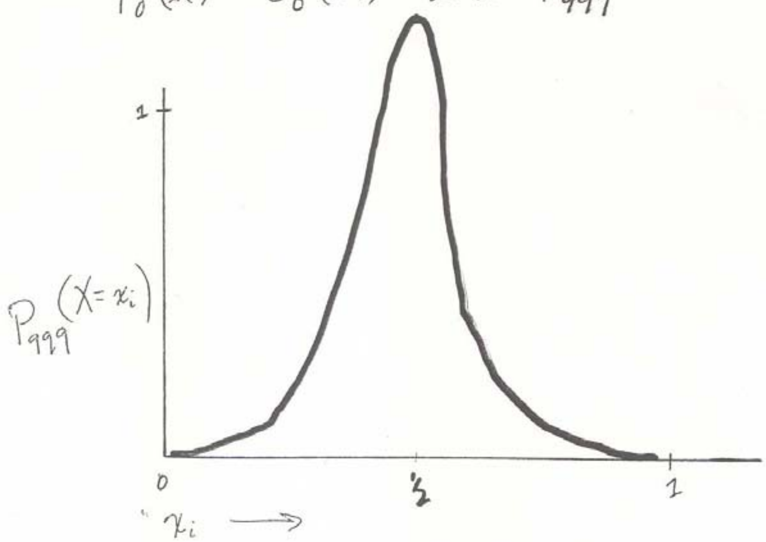
Variance: Ignorance & Resilience



$E_0(h) = \frac{1}{2}$

X is the probability of heads, h = A head will occur on 1,000th toss

$P_0(h) = E_0(h)$ and $P_{999}(h) = E_{999}(h)$



$E_{999}(h) \approx \frac{1}{2}$

p *screens off* e from q iff $P(q/e.p) = P(q/p)$ and $P(q/e) \neq P(q)$.

We can think of this as full neutralization by p of e's power to affect q. This phenomenon also comes in degrees, of course.

Varieties of Defeating Evidence:

e = Jim's brother says his landlord is shifty-looking.
h = Landlord is dishonest.

Say, $P(h/e) > P(h)$, so e is evidence for h.

Consider:
E_1 = Roommate of brother says landlord is not shifty-looking.
(This evidence "opposes e.")

E_2 = Landlord is active churchgoer, etc. (This evidence "narrows reference class.")

E_3 = Jim says brother has never met landlord, is prejudiced against him. (This evidence "undermines justification for h given by e.")

E_4 = Roommate of brother says their landlord has a kind of palsy that makes him look shifty involuntarily. (Something else again.)

All of these are cases of defeating evidence. What's similar about them is this: while $P(h/e) > P(h)$, $P(h/e.E_n) < P(h/e)$.

Pryor says probability can't model what's different about them. Here goes.

e = Jim's brother says his landlord is shifty-looking.
h = Landlord is dishonest.

Say, $P(h/e) > P(h)$, so e is evidence for h.

Consider:
E_1 = Roommate of brother says landlord is not shifty-looking.

E_2 = Landlord is active churchgoer, etc.

E_3 = Jim says brother has never met landlord, is prejudiced against him.

E_4 = Roommate of brother says their landlord has a kind of palsy that makes him look shifty involuntarily.

Let $b(x)$ = Brother says x, q = Landlord is shifty-looking.
Then, $e = (b(x) \text{ and } x = q)$. I.e., $e = b(q)$.
Now, $P(h/e) > P(h)$, as we are imagining the case, only if

1) $\mathbf{P(q/b(q)) > P(q)}$ and
2) $\mathbf{P(h/q) > P(h)}$.

That is, the brother's saying that the landlord is shifty-looking is evidence that the landlord is dishonest only if 1) the brother's saying this is evidence that the landlord is shifty-looking and 2) the landlord's being shifty-looking is evidence he is dishonest.

-- E_1 and E_3 block 1). I.e., $P(q/b(q).E_1) \approx P(q) < P(q/e)$ and
$P(q/b(q).E_3) \approx P(q) < P(q/e)$

-- E_2 and E_4 block 2). I.e., $P(h/q.E_4) \approx P(h) < P(h/e)$, which implies that $P(h/e.E_4) \approx P(h)$ since e can't give any more to h than q has.

These are two different ways of defeating e's support for h and probability makes the distinction. I don't see further distinctions to make here.

Epistemic priority, cont'd

Challenge: distinguish cases

1) where your justification for believing (A or B) comes from your justification to believe the more specific B, from cases

2) where your justification for believing B comes from evidence for the more general (A or B), interacting with other background information you have.

Evidence of the first kind, e, satisfies the following condition:

P((A or B)/B) = P((A or B)/B.e),

i.e. B screens off e from (A or B).

Evidence of the second kind, E, satisfies a different condition:

P(B/(A or B).E) = P(B/(A or B)),

i.e., (A or B) screens off E from B.

Neither type of evidence (necessarily) satisfies the condition for the other. These probabilistic conditions characterize the difference between the two types of evidence.

➔➔ Probability has a lot of resources.

The notion of screening off is also important for understanding independent evidence. See: Sober, "Independent Evidence for a Common Cause," *Philosophy of Science* (1989), Roush, "Testability and the Unity of Science," *Journal of Philosophy* (Nov. 2004), and Roush, "Testability and Candor," *Synthese* (July 2005)

Note also: A screens off B from C iff B and C are independent relative to A. Conditional independence is the key notion of Bayes Nets forms of confirmation theory. See e.g. Bovens and Hartmann, *Bayesian Epistemology*, OUP 2003.

Transmission of warrant *vs.* Closure of warrant under known implication

Closure: Having warrant to believe p suffices for you to have warrant to believe q, assuming you know that p implies q.

Transmission: Your warrant to believe p is part of what makes you warranted in believing q, assuming you know that p implies q.

The difference is illustrated by cases of transmission failure (Crispin Wright):

p = A soccer goal was just scored.
q = A soccer game is taking place (and not a filming of a movie scene).

p implies q, and (if you know this then) if you are warranted in believing p you are also warranted in believing q (closure). However, arguably, your warrant for believing p doesn't *transmit* to q. Instead, you had to have warrant for believing q as a precondition for having warrant, in the way you did, for believing p.

It seems to me this point is easy to capture in Bayesian terms. Let
e = I have an experience as of a soccer goal being scored.
e is evidence for p, in the Bayesian sense, since P(p/e) > P(p). However, **e is not evidence for -m**, where m = I am watching the filming of a movie scene, because

P(-m/e) = P(-m)

e doesn't discriminate between this being a movie scene and a real soccer game, and that accounts for our sense that warrant isn't transmitted, since q implies -m.

Note:
This transmission failure is a case of failure of **transitivity of positive relevance**: P(p/e) > P(p), P(-m/p) > P(-m), but not P(-m/e) > P(-m).

This transmission failure is also an illustration of the fact that Bayesian confirmation theory does not obey the **special consequence condition**.

Also, -q **screens off** e from p. That is, P(p/e.-q) = P(p/-q). Given -q, e makes no difference to the probability of p. This corresponds to Wright's sense that you had to have antecedent justification for q in order even for e to *be evidence for* p.

Note the similarity in these respects of the Wright case to the skeptical case:

e = I have an experience as of a hand
p = I have a hand.
q = I am not a handless brain in a vat.

P(p/e) > P(p), e is evidence for p,
P(q/p) > P(q), since p implies q,
but not P(q/e) > P(q), e is not evidence for q, because

P(e/q) = P(e/-q), i.e., e does not discriminate between q and -q.
(Thus, the Bayesian has something that corresponds to the epistemologist's intuition that **we have no (observational) evidence for or against skepticism**.)

So, we have once again a failure of transitivity of positive relevance.

In this case, Jim wants to say that one's justification for p (e, or rather the experience e describes) *does transmit* to q.

Question for Jim: why should we think warrant transmits here if we aren't supposed to think that in the Wright case?

We have once again here that -q screens off e from p, i.e. P(p/e.-q) = P(p/-q). More generally, if your antecedent belief in q is low, then the appearance of a hand is that much less evidence that you have a hand. This seems perfectly intuitive.

I will come back to this.

Finite Additivity

$P(A_1 \vee A_2 \vee \ldots \vee A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$, $\quad A_1, \ldots, A_n$ exclusive

A lovely property, but is it a requirement for rational credences?
Dutch Book arguments say 'yes'. Opponents argue these arguments are circular. Also, that superadditivity explains the Ellsburg paradox. I am moved by some of these opposing arguments, but I am not much moved by Jim's reasons for superadditivity.

Two Alternatives

1. Intuitionist Bayesianism (http://brian.weatherson.net/conprob.pdf or *Notre Dame Journal of Formal Logic*, 2003)

P(A) + P(-A) need not sum to 1, on this account. But finite additivity is preserved because instances of excluded middle are not assigned probability 1 by the axioms. Instances of excluded middle get their probabilities raised in light of evidence just as their component propositions do.

It seems to me that this will do for Jim just what superadditivity does with biased evidence and its behavior toward underminers, because what he wants there is that something can rationally raise the credence of not-B without taking away from the credence of B. This can be so if $P(B \vee -B)$ does not start at 1.

2. View the justification of p (I have a hand) not as the *proposition* e but as the *experience* as of a hand. Then Dogmatism is a claim about the transition from experience as of a hand to belief in p (leading to conditionalization on p).

Why this makes a difference: only the relationship between propositions p (I have a hand) and q (I am not a brain in a vat) are constrained by the probability axioms. The rest of what Jim wants to say is a supplement that tells us something about the transition from **experience** as of p to **belief** in p.

Unlike before, the axioms do not dictate that in order for the experience to justify belief in p I have to **already have justified belief in q**.

This is because on this account -q does not screen off the justification of p from p; it can't do so because the justification in question is not a proposition.

By the principle of minimal mutilation, Option 2 is the best way to incorporate Jim's Dogmatism into a formal account of evidence.