Staying Regular?

Alan Hájek

ALI G: *So what is the chances that me will eventually die?*
C. EVERETT KOOP: *That you will die? – 100%. I can guarantee that 100%: you will die.*
ALI G: *You is being a bit of a pessimist…*

–Ali G, interviewing the
Surgeon General, C. Everett Koop

Autobiographical back story

- Over my philosophical career I’ve been interested in various topics, but certain topics have especially gripped me…

Introduction

- I'll discuss the fluctuating fortunes of *regularity:*

If X is possible, then the probability of X is positive.

$$\Diamond X \rightarrow P(X) > 0.$$

Introduction

- I'll give many reasons to care about regularity.
- So it's important to formulate it carefully.
- I'll look at various formulations of it for subjective probability, some implausible, some more plausible.
- I'll offer what I take to be its most plausible version: a constraint that bridges *doxastic* modality and *doxastic* (subjective) probability.
- But even that will fail.

Introduction

- There will be two different ways to violate regularity
  - zero probabilities
  - no probabilities at all (probability gaps).
- Both ways create trouble for pillars of Bayesian orthodoxy:
  - the ratio formula for conditional probability
  - conditionalization, characterized with that formula
  - the multiplication formula for independence
  - expected utility theory

Introduction

- The failure of this seemingly innocuous constraint has ramifications that strike at the heart of probability theory and formal epistemology.

Regularity

If X is possible, then the probability of X is positive.

- We already had the probability axiom:

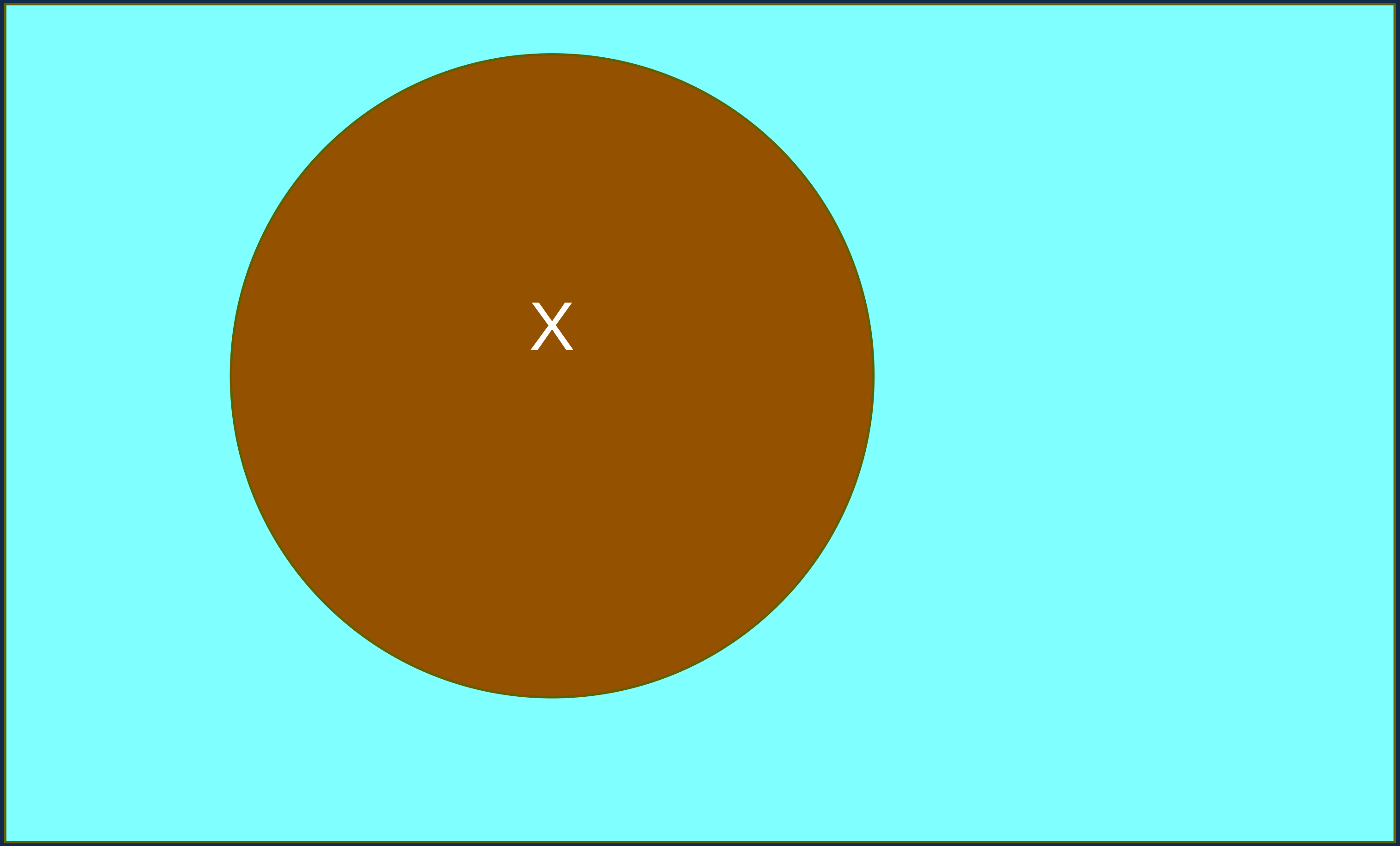
  $$P(X) \geq 0$$

- Now this constraint gets the tiniest strengthening if X is possible; the inequality becomes strict:

  $$P(X) > 0$$

  if X is possible.

- Muddy Venn diagram: no bald spots.



Regularity

- An unmnemonic name, but a commonsensical idea.
- “If it can happen, then it has a chance of happening”…

Advocates of regularity

- Regularity has been suggested or advocated by Jeffreys, Jeffrey, Carnap, Shimony, Kemeny, Edwards, Lindman, Savage, Stalnaker, Lewis, Skyrms, Appiah, Jackson, Hofweber, …

Ten reasons to care about regularity

- Regularity promises a bridge between modality and probability—a bridge that illuminates both.

Ten reasons to care about regularity

- Regularity promises a bridge between probability and truth:

  If X has probability 0, then X is impossible, hence (actually) false.

  If X has probability 1, then X is necessary, hence (actually) true.
- (No assumption of Humean supervenience.)
- If regularity fails, even this is a bridge too far!

Ten reasons to care about regularity

- Regularity may provide a bridge between traditional epistemology and Bayesian epistemology.

Ten reasons to care about regularity

- Regularity promises to illuminate rationality.
- It would provide a much-needed additional constraint on rational credence that goes beyond coherence.

Ten reasons to care about regularity

- Various Bayesian convergence results require regularity.

Ten reasons to care about regularity

- Regularity would allow us to simplify various ‘probability 1’ convergence theorems – for example, the strong law of large numbers.
- The ‘probability 1’ qualification could be removed for any regular probability function, as it would be redundant.

Ten reasons to care about regularity

- Centrepieces of synchronic Bayesian epistemology face problems when regularity fails.

Ten reasons to care about regularity

- The centrepiece of *diachronic* Bayesian epistemology – conditionalisation – faces problems without a version of regularity; yet it also conflicts with regularity.

Ten reasons to care about regularity

- Bayesian decision theory faces problems if regularity fails.
- So failures of regularity pose some of the most important problems for probability theory as a representation of uncertainty.

Ten reasons to care about regularity

- These failures motivate other representations of uncertainty – Popper functions, ranking functions, NAP, comparative probabilities…

Formulating regularity

If X is possible, then the probability of X is positive.

- This is just a schema.
- There are many senses of ‘possible’ in the antecedent…
- There are also many senses of ‘probability’ in the consequent…

Formulating regularity

- Pair them up, and we get many, many regularity conditions.
- Some are interesting, and some are not; some are plausible, and some are not.
- Focus on pairings that are definitely interesting, and somewhat plausible, at least initially.

Formulating regularity

- In the consequent, let's restrict our attention to *rational subjective* probabilities.
- If X is possible, $C(X) > 0$.
- In the antecedent? …

Formulating regularity

- Untenable:

Logical Regularity

If X is LOGICALLY possible, then C(X) > 0.

(Shimony, Skyrms)

Formulating regularity

- Problems: There are all sorts of propositions that are logically possible, but that are a priori knowable to be false, and may rationally be assigned credence 0:
  - ‘Obama is a 3-place relation’
  - ‘Clinton is the number 17’

Formulating regularity

- The probability axioms are not themselves *logically necessary*, so logical regularity curiously would require an agent to give positive credence to their falsehood.

Formulating regularity

- More plausible:

<u>Metaphysical Regularity</u>

If X is METAPHYSICALLY possible, then C(X) > 0.

Formulating regularity

- This brings us to Lewis's (1980) formulation of "regularity": "$C(X)$ is zero … only if X is the empty proposition, true at no worlds". (According to Lewis, X is metaphysically possible iff it is true at some world.)
- Lewis regards regularity in this sense as a constraint on "initial" (prior) credence functions of agents as they begin their Bayesian odysseys—Bayesian Superbabies.

Formulating regularity

- A problem for metaphysical regularity as a constraint on Superbabies: it is metaphysically possible for no thinking thing to exist, so by regularity, one must assign positive probability to this.
- But far from being rationally required, this seems to be *irrational*.
- Dutch Book argument.
- It's at least rationally *permissible* to assign probability 0 to no thinking thing existing.

Formulating regularity

- However, *doxastic* possibility seems to be a promising candidate for pairing with subjective probability.
- <u>Doxastic regularity:</u>

  If X is doxastically possible then C(X) > 0.

Formulating regularity

- We might think of a doxastic possibility for an agent as:
  - something that is compatible with what she *believes*;
  - or something that she is not certain is false;
  - or perhaps some other understanding …
  - I will speak of a doxastically *live* possibility—for short, a *live* possibility.

Formulating regularity

- So from now on I will understand regularity as:

  if X is a live possibility then C(X) > 0

- All the better that this can be understood in multiple ways. For I will argue that on any reasonable undertanding of ‘live possibility’, it is false.

Formulating regularity

- If doxastic regularity is violated, then offhand two different attitudes are conflated…
- Not just at 0, but throughout the entire [0, 1] interval.

Formulating regularity

- Doxastic regularity avoids the problems with the previous versions…

Formulating regularity

- And yet doxastic regularity appears to be untenable.

Formulating regularity

- If this version of regularity fails, then various other interesting versions will fail too. E.g.:
- Epistemic regularity:

  If X is epistemically possible, then C(X) > 0.
- This is stronger than doxastic regularity; if it fails, so does this.

Formulating regularity

- <u>Evidential regularity:</u>

  If X is not ruled out by one's evidence, then $C(X) > 0$

Two ways to be irregular

- There are two ways in which an agent’s probability function could fail to be regular:

    1) It assigns zero to some live possibility.

    2) It fails to assign *anything* to a live possibility.

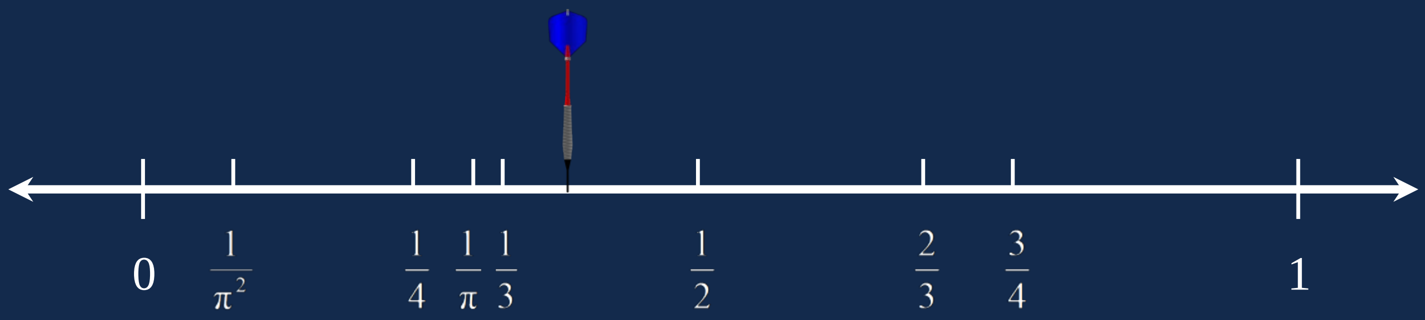
Two ways to be irregular

- Those who regard regularity as a norm of rationality must insist that all instances of 1) *and* all instances of 2) are violations of rationality.
- I will argue that there are rational instances of both 1) and 2).

Dart example

Throw a dart at random at the [0, 1] interval of the reals …

Dart example



Dart example

- Various non-empty subsets get assigned probability 0:
  - All the singletons
  - Indeed, all the finite subsets
  - Indeed, all the countable subsets
  - Even various uncountable subsets (e.g. Cantor's 'ternary set')

Dart example

- Examples like this pose a threat to regularity as a norm of rationality.
- Any landing point in [0, 1] is a live possibility for our ideal agent.

Arguments against regularity

- In order for P to be regular, there has to be a certain harmony between the *cardinalities* of P's sample space and its range.
- If the sample space is too large relative to P, regularity will be violated.

Arguments against regularity

Kolmogorov's axiomatization requires *P* to be *real*-valued. This means that any uncountable probability space is automatically irregular. (Hájek 2003).

Arguments against regularity

- It is curious that this axiomatization is restrictive on the *range* of all probability functions: the real numbers in [0,1], and not a richer set;
- yet it is almost completely permissive about their *domains:* Ω (the sample space) can be any set you like, however large, and *F* (the set of subsets that get assigned probabilities) can be any field on Ω, however large.

Arguments against regularity

- *We can apparently make the set of* <u>*contents*</u> *of an agent's thoughts as big as we like.*
- *But we limit the* <u>*attitudes*</u> *that she can bear to those contents*—the attitudes can only achieve a certain fineness of grain.
- Put a rich set of contents together with a relatively impoverished set of attitudes, and you violate regularity.

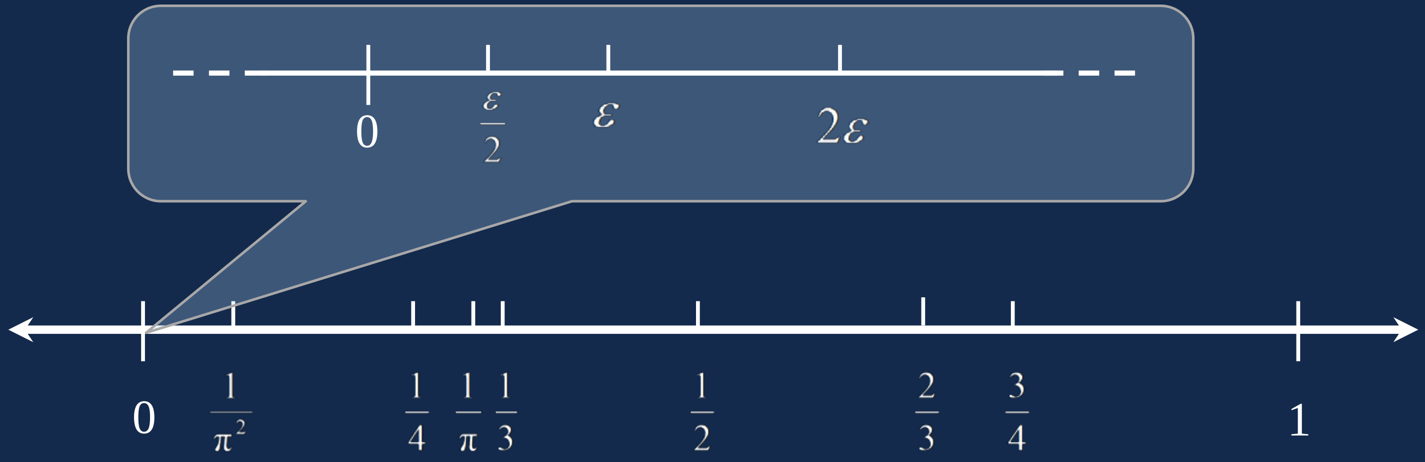
Infinitesimals to the rescue?

The friend of regularity replies: if you're going to have a rich *domain* of the probability function, you'd better have a rich *range.*

Lewis:

"You may protest that there are too many alternative possible worlds to permit regularity. But that is so only if we suppose, as I do not, that the values of the function *C* are restricted to the standard reals. Many propositions must have infinitesimal *C*-values ... (See Bernstein and Wattenberg (1969).)"

Infinitesimals to the rescue?



Infinitesimals to the rescue?

- Bernstein and Wattenberg’s article does not substantiate Lewis’ strong claim that there are too many possible worlds to permit regularity *only if C*’s values are restricted to the reals.
- Bernstein and Wattenberg show that using infinitesimals, one can give a regular probability assignment to the landing points of our fair dart throw.

Infinitesimals to the rescue?

- But that's a very specific case, with a specific cardinality!
- Lewis himself thinks that the cardinality of the set of possible worlds is greater than that (at least beth-2).
- We need a similar result that holds if the set of possibilities has higher cardinality than that of the real interval [0, 1].
- Indeed, the set of doxastic possibilities may well be *a proper class! …*

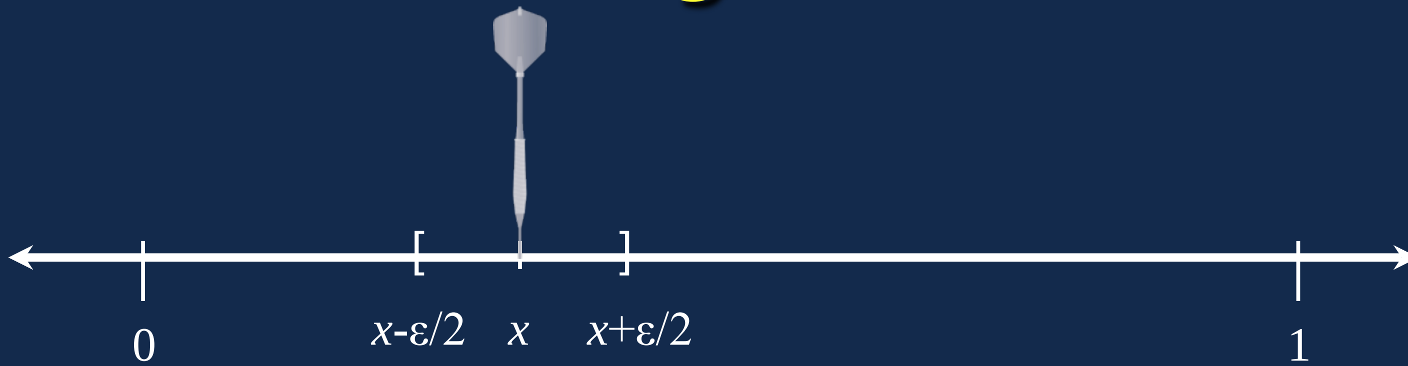
Arguments against regularity, even allowing infinitesimals

- I conjectured that a version of the cardinality problem would always arise.
- Pruss proved it: if the cardinality of Ω is greater than that of the range of P, then regularity fails.
- No symmetry assumption is needed – cardinalities do all the work.

Arguments against regularity, even allowing infinitesimals

- But if we add a symmetry assumption, we have another, more intuitive argument.
- We can scotch regularity even for a hyperreal-valued probability function by correspondingly enriching the space of possibilities…
- The dart is thrown at the [0, 1] interval of the *hyperreals*.

Arguments against regularity, even allowing infinitesimals



Not to scale!

- x is strictly contained within nested intervals of width ε, for each infinitesimal ε. The probability of each interval is its width, ε. (This assumption can be somewhat weakened.)
- So the point's probability is bounded above by *all* these ε, and thus it must be smaller than all of them—i.e. 0.

Arguments against regularity, even allowing infinitesimals

I envisage a kind of arms race:

- We scotched regularity for real-valued probability functions with sufficiently large domains (uncountable).
- The friends of regularity fought back, enriching their ranges: making them hyperreal-valued.
- The enemy of regularity counters by enriching the domain.
- And so it goes.
- By Pruss's result, the enemy can always win (for anything that looks like Kolmogorov's probability theory).

Arguments against regularity, even allowing infinitesimals

- Could we tailor the range of the probability function to the domain, for each particular application? (Like the general of a defense force …)
- The trouble is that in a Kolmogorov-style axiomatization the commitment to the range of P comes *first*…
- On the tailoring approach, a probability function is a mapping from F to …—well, to *what*?
- What will the additivity axiom look like?
- In any case, this 'wait and see' approach is quite a departure from Kolmogorov.

Arguments against regularity, even allowing infinitesimals

- On a Kolmogorov-style approach, there will always be an Ω that will have non-empty subsets assigned probability 0.

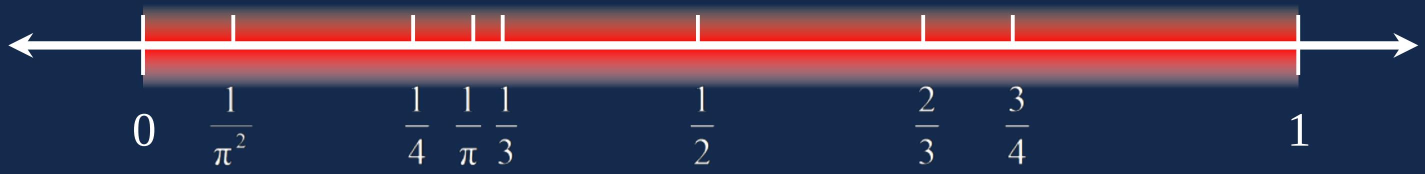
Doxastically possible credence gaps

- I will argue that you can rationally have credence gaps.

Examples of doxastically possible credence gaps

- Non-measurable sets

Dart example



- Certain subsets of Ω—so-called *non-measurable* sets—get no probability assignments whatsoever.

Examples of doxastically possible credence gaps

- Chance gaps
- The Principal Principle says (roughly!!):

  your credence in X, *conditional* on it having chance x, should be x:

  $$C(X \mid \text{chance}(X) = x) = x.$$

Examples of doxastically possible credence gaps

- A relative of the Principal Principle? Roughly:

  your credence in *X*, *conditional* on it being a chance gap, should be gappy:

  $C(X \mid \text{chance}(X) \text{ is } \textit{undefined}) \text{ is } \textit{undefined}.$

- All I need is that rationality *sometimes permits* your credence to be gappy for a hypothesized chance gap.

Examples of doxastically possible credence gaps

- There are arguably various examples of chance gaps:
  - Chance statements themselves
  - Cases of indeterminism without chances: Earman’s space invaders, Norton’s dome (Eagle)

Examples of doxastically possible credence gaps

- One’s own free choices
- Kyburg, Gilboa, Spohn, Levi, Briggs, Liu and Price contend that when I am making a choice, I must regard it as free. In doing so, I *cannot* assign probabilities to my acting in one way rather than another (even though onlookers may be able to do so).
- “Deliberation crowds out prediction”—or better, it crowds out probability.

Examples of doxastically possible credence gaps

- To be sure, these cases of probability gaps are controversial.
- But these authors are committed to there being credence gaps, and thus violations of regularity.
- All I need is that it is *permissible* for them to be credence gaps.

Ramifications of irregularity for Bayesian epistemology and decision theory

- I have argued for two kinds of counterexamples to regularity: rational assignments of zero credences, and rational credence gaps, for doxastic possibilities.
- I now want to explore some of the unwelcome consequences these failures of regularity have for traditional Bayesian epistemology and decision theory.

Problems for the conditional probability ratio formula

- The ratio analysis of conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

... provided P(B) > 0

Problems for the conditional probability ratio formula

- What is the probability that the dart lands on ½, *given* that it lands on ½?
- 1.
- But the ratio formula cannot deliver that result, because P(dart lands on ½) = 0.

Problems for the conditional probability ratio formula

- Gaps create similar problems.
- Take your favorite probability gap, G.
- The probability of G, given G, is 1.
- But the ratio formula cannot deliver that result, because

  P(G) is undefined.

Problems for the conditional probability ratio formula

- We need a more sophisticated account of conditional probability.
- I advocate taking conditional probability as primitive (in the style of Popper and Rényi).

Problems for conditionalization

- The zero-probability problem for the conditional probability formula quickly becomes a problem for the updating rule of *conditionalization*, which is defined in terms of it.
- Suppose the agent learns evidence *E.*

$$P_{new}(X) = P_{old}(X \mid E) \quad \text{(provided } P_{old}(E) > 0\text{)}$$

Problems for conditionalization

- Suppose you *learn* that the dart lands on ½. What should be your *new* probability that the dart lands on ½?
- 1.
- But

  $$P_{old}(\text{dart lands on } ½ \mid \text{dart lands on } ½)$$

  is undefined, so conditionalization (so defined) cannot give you this advice.

Problems for conditionalization

- Gaps create similar problems.
- Suppose you *learn* that G. What should be your *new* probability for G?
- 1.
- But

  $$P_{\text{old}}(G \mid G)$$

  is undefined, so conditionalization cannot give you this advice.

Problems for conditionalization

- We need a more sophisticated account of conditionalization.
- Primitive conditional probabilities to the rescue!

Problems for independence

- We want to capture the idea of A being probabilistically uninformative about B.
- A and B are said to be *independent* just in case

$$P(A \cap B) = P(A)\, P(B).$$

Problems for independence

- According to this account of probabilistic independence, anything with probability 0 is independent of *itself*:

  If $P(X) = 0$, then $P(X \cap X) = 0 = P(X)P(X)$.

- But identity is the ultimate case of (probabilistic) *d*ependence.

Problems for independence

- Suppose you are wondering whether the dart landed on ½. *Nothing* could be more informative than your learning: the dart landed on ½.
- But according to this account of independence, the dart landing on ½ is independent of the dart landing on ½!

Problems for independence

- Gaps create similar problems.
- Suppose you are wondering whether G. *Nothing* could be more informative than your learning: G.
- But there is no verdict from this account of independence.

Problems for independence

- We need a more sophisticated account of independence – e.g. using primitive conditional probabilities.
- Branden Fitelson and I have been working on this!

Problems for expected utility theory

- Arguably the two most important foundations of decision theory are the notion of *expected utility*, and *dominance reasoning*.

Problems for expected utility theory

- And yet probability 0 propositions apparently show that expected utility theory and dominance reasoning can give conflicting verdicts.

Problems for expected utility theory

- Suppose that two options yield the same utility except on a proposition of probability 0; but if that proposition is true, option 1 is far superior to option 2.

Problems for expected utility theory

- You can choose between these two options:
  - Option 1: If the dart lands on 1/2, you get a million dollars; otherwise you get nothing.
  - Option 2: You get nothing.

Problems for expected utility theory

- Expected utility theory apparently says that these options are equally good: they both have an expected utility of 0.
- But dominance reasoning says that option 1 is strictly better than option 2. Which is it to be?
- I say that option 1 is better.
- I think that this is a counterexample to expected utility theory as it is usually interpreted.
- Both evidential and causal.
- (To be sure, there are replies …)

Problems for expected utility theory

- Gaps create similar problems.
- You can choose between these two options:
  - Option 1: If G, you get a million dollars; otherwise you get nothing.
  - Option 2: You get nothing.

Problems for expected utility theory

- Expected utility theory goes silent.
- I say that option 1 is better.
- We need a more sophisticated decision theory.

Conclusion

- Irregularity makes things go bad for the orthodox Bayesian; that is a reason to insist on regularity.
- The trouble is that regularity appears to be untenable.
- I focused on doxastic regularity, but other interesting regularities will meet similar downfalls.
- I think, then, that irregularity is a reason for the orthodox Bayesian to become unorthodox.

Conclusion

- I have advocated replacing the orthodox theory of conditional probability, conditionalization, and independence with alternatives based on Popper/Rényi functions. Expected utility theory appears to be similarly in need of revision.

Conclusion

- And then there are some possibilities that *really should* be assigned zero probability …

Thanks especially to Rachael Briggs, David Chalmers, John Cusbert, Kenny Easwaran, Branden Fitelson, Renée Hájek, Thomas Hofweber, Leon Leontyev, Hanti Lin, Aidan Lyon, John Maier, Daniel Nolan, Alexander Pruss, Wolfgang Schwarz, Mike Smithson, Weng Hong Tang, Peter Vranas, Clas Weber, and Sylvia Wenmackers for very helpful comments that led to improvements; to audiences at Stirling, the ANU, the AAP, UBC, Alberta, Rutgers, NYU, Berkeley, Miami, the Lofotens Epistemology conference; to Carl Brusse and Elle Benjamin for help with the slides; and to Tilly.

Ten reasons to care about regularity

- Regularity may provide a bridge between logic and probability.
- Failure of regularity is a thorn in the side of probabilistic semantics for logic. Probabilistic notions of entailment, incompatibility are poor surrogates for their logical counterparts.