Philosophy 148 — Announcements & Such

- We will have a "homework discussion" before each HW is due. For HW #2 and the Mid-Term, we have: **March 17 @ 6pm @ 223 Dwinelle**.
  - I will go over the relevant parts of Joyce's argument (and the salient bits for the homework). I will also discuss the mid-term as well.
  - **The mid-term is next Thursday (3/20). HW #2 is also due that day.**
- I've posted my LaTeX source for the HW #2 hints handout (see website).
- Today's Agenda:
  - Some final remarks on the Dutch Book Argument.
    - What, precisely, does the Dutch Book Argument show?
    - Some further worries about the Dutch Book Argument.
  - Then: Epistemic Rationality, Revisited
    - The Preface Paradox
      - Logical Consistency and Epistemic Rationality (of *full* beliefs)
- Monday night & Tuesday: Joyce's argument for epistemic probabilism.

What Does the Dutch Book Argument Show?

- Here's a more precise statement of what the Dutch Book Argument shows:
  - ***If*** an agent S satisfies the following conditions:

    1. S views *individual* bets (β on p) within the DBS as *acceptable*, *i.e.*, *fair or favorable* [from an *expected utility* point of view, this *already rules-out* being *non*-probabilistic in *this* way: $q(\sim p) > 1 - q(p)$].
    2. Condition (1) holds *even if* we replace \$'s with *utiles* in the DBS.
    3. S views any *collection* of acceptable bets as *jointly acceptable* [from an *EU* point of view, this is *automatic*, since *EU* is *additive*].

    then, the following is also true of S:

    4. *If and only if* S's betting quotients (in a DBS) q are *non*-probabilistic, *then* there exists a sequence of bets — *which* S *views as both severally and jointly acceptable* — on which S is *guaranteed to lose*.
- Thus, *assuming* (1)–(3), S can avoid the possibility of a sure-loss (in a DBS) by announcing *probabilistic* q's — *and this is the **only** way* to avoid it.

Postscripts to the DBA II: The Value of Money

- We have the caveat about $\$|\mathfrak{s}|$ being "small in comparison to the agent's total wealth" for two reasons. First, if the agent could lose *everything* on a bet, this would undermine the probative value of the argument.
- Also, real agents (and, arguably, also rational agents!) marginally value money in a way that is *non-linear*, especially for larger sums of money.
- The difference in value between \$1 and \$1000 is pretty substantial. But, the difference in value between \$1M + \$1 and \$1M + \$1000 is not.
- This is called the *diminishing marginal utility of money*. Empirical studies show that actual agents have marginal utilities that are close to being linear only for \$ amounts that are "small relative to their total wealth."
- And, this doesn't seem irrational. So, if we want something with *linear* marginal value (for *additivity* purposes — think "package principle"!), money is probably not the best thing to use. This explains (2), above.

Postscripts to the DBA III: The Need for the *Converse* DBT

- The DBT *by itself* cannot secure pragmatic probabilism. All the DBT establishes is that q's coherence entails that q is a probability function. What if the *converse* of this were *false*? Would the DBA still be persuasive?
- If there were some probability functions that were *also* susceptible to Dutch Book, then pragmatic probabilism would *not follow* from the DBA.
- Remember, pragmatic probabilism says *all* probabilistic doxastic states are (pragmatically) better than *all* non-probabilistic doxastic states.
- The mere fact that (DBT) *all non*-probabilistic states are *susceptible* to Dutch Book is *not* sufficient to establish this. One *also* needs to show that (CDBT) *all probabilistic* doxastic states are *immune* from Dutch Book.
- Luckily, the (CDBT) is *true*. We have *not* proven this (and we will not). All we have shown is that *some* Dutch Books are *blocked* by satisfying *some* axioms. This does *not* establish (CDBT). [Note the biconditional in (4).]
- See the Kemeny paper (on website), if you want to see the proof of (CDBT).

Postscripts to the DBA IV: Is the *Converse* DBT *Enough*?

- *Even if* we use both DBT and CDBT (which *are* both theorems), does the Dutch Book *Argument* establish pragmatic probabilism? Maybe not.
- Alan Hájek thinks there is still a gap in the DBA, *even with* both directions of the DBT, *and even if* we grant the "package principle" [*i.e.*, (1)–(3)].
- As Hájek explains in his paper "Scotching Dutch Books" (now on website), the structure of the DBA is more-or-less something like the following:
  1. S's q is coherent $\Longleftrightarrow$ S's q is a probability function. [DBT + CDBT]
  2. S is susceptible to Dutch Book $\Longleftrightarrow$ S's q is incoherent. [definition]
  3. Susceptibility to DB is "bad" and immunity from DB is "good". [ass.]
  4. $\therefore$ S's q is *better* if it is probabilistic than if it is *non*-Pr. [P-probabilism]
- Does (4) follow from (1)–(3)? Clearly, (1)–(3) entail that there is *some sense* in which q is *guaranteed* to be "better" *in virtue of* being probabilistic.
- What if there is *also* some sense in which q is *guaranteed* to be "*worse*" *in virtue of* being probabilistic. Is there a symmetric "Good Book" argument?

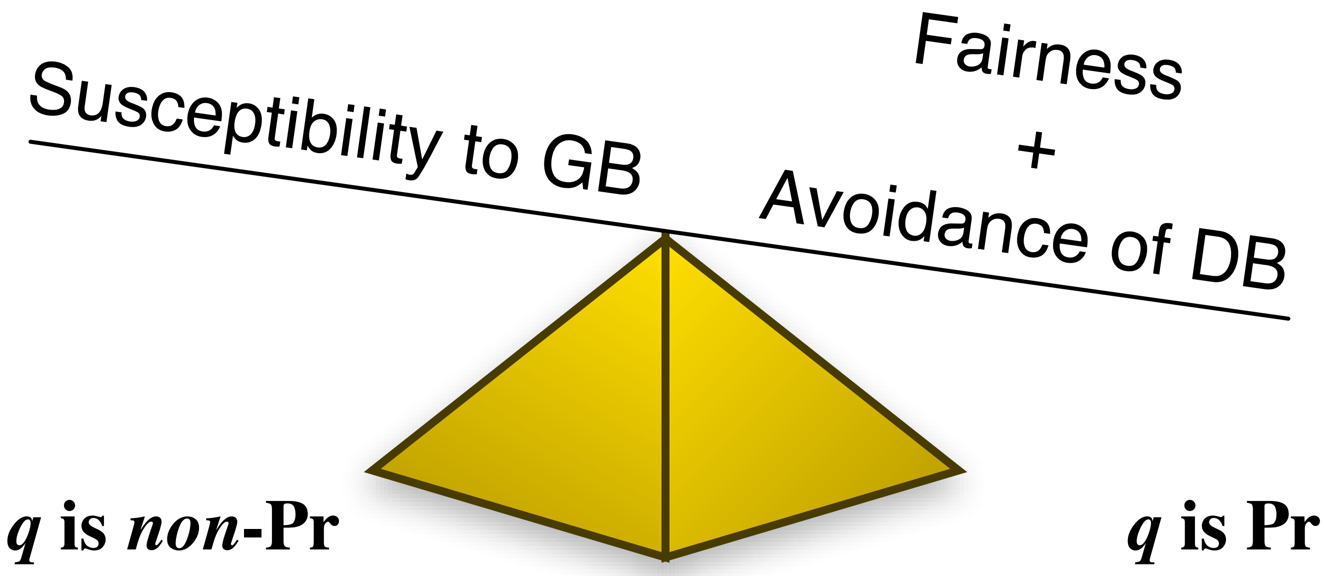
Postscripts to the DBA V: Hájek's "Good Book" Argument

- A *Good Book* is a sequence of bets that *wins* money *come what may*. Betting quotients q are *good* if they are susceptible to a Good Book.
- It turns out that the following as *also* a theorem (we won't prove it, but it shouldn't be surprising, given the Dutch Book theorem and it Converse):

  Good Book Theorem (GBT). q is good $\iff$ q is *not* a probability function.
- So, why not the following, symmetric argument *against* P-probabilism?
  1. S's q is good $\iff$ S's q is not a probability function. [GBT]
  2. S is susceptible to Good Book $\iff$ S's q is good. [definition]
  3. Susceptibility to GB is "good" and immunity from GB is "bad". [ass.]
  4. $\therefore$ S's q is *worse* if it is probabilistic than if it is *non*-Pr. [~P-probabilism]
- The question now becomes: Is the DBA more compelling than the GBA? Which is pragmatically *better*: immunity from DB or susceptibility to GB?
- And, are there yet *other dimensions* of "goodness" we have overlooked?

Postscripts to the DBA V: Hájek's "Good Book" Argument II

- Perhaps Hájek's "Good Book" argument can be blocked, by noting that the only way to ensure *fairness* is to have *probabilistic* q's. [See Kemeny.]
- In other words, maybe "Good Books" are just as *good* as "Dutch books" are *bad*, but *fairness* "tips the scales" in favor of *probabilism*:



- This does require us to assume that it is a *good thing* to be able to *ensure fairness* (or *breaking-even*) in the Dutch Book setup. This seems OK to me.

Postscripts to the DBA VI: States *vs*. Processes Again 1

- The Dutch Book Argument we have seen only aims to establish that rationality requires an agent's doxastic state *at a particular time* to be representable as a probability model. This is a *state* requirement.
- These kinds of requirements are also sometimes called requirements of *synchronic* rationality. One might wonder: what about *processes* that lead us from one probability model to another — *diachronic* rationality?
- Traditional subjective probabilists (like Ramsey and de Finetti) didn't think there were any process requirements in this sense. So long as you are "coherent" at each time, that's all there is to pragmatic rationality.
- Contemporary subjective probabilists (often called "Bayesians") do offer *diachronic* Dutch Book arguments in support of what I called norm (3), which is sometimes called the "rule of conditionalization" (ROC).
- We won't look at diachronic Dutch Book arguments in this course. But, there are various arguments of this kind in the literature.

Postscripts to the DBA VI: States *vs*. Processes Again 2

- [The "accuracy arguments" we will look at next in our *epistemic* probability unit do not seem to have any diachronic analogues. This is an interesting asymmetry in the literature on subjective probability.]
- There are other kinds of process requirements in the literature on pragmatic subjective probability. One of these is (roughly) as follows:
  - When an agent goes from one doxastic state to another, they should do so in a way that constitutes a "minimal change"— the new state should be "closest" to their old one, subject to some "constraints".
- This rough idea can be made more precise, and it can lead to answers that diverge from the (ROC), depending on how one precisifies the terms "learning", "minimal change", "closest", and "constraints".
- *E.g.*, if "learning" does not require *assigning probability 1* to what is learned, then this sort of "minimal change" approach leads to a more general form of conditionalization known as *Jeffrey Conditionalization*. [We may touch on this later, when we discuss Bayesian confirmation.]

Postscripts to the DBA VII: Logical Omniscience

- If the DBA is sound, then *pragmatically* rational agents are *logically omniscient* — they can demarcate the logical truths from the non-logical-truths in their doxastic state $\mathcal{B}^t$. This is a strong requirement!
- Several authors (*e.g.*, Hacking, Harman) have objected that this is simply too strong a requirement to place on a pragmatically rational agent.
- One could try to weaken this requirement in some way. Hacking's "Slightly More Realistic Personal Probability" (website) is a good example.
- It is not so easy to weaken this requirement and still ensure that the resulting doxastic states are *probability models*. Consider two proposals:
  - S is required to have $\Pr(p) = 1$ *only* for p's S *knows* to be logical $\top$'s.
  - S is required to have $\Pr(p) = 1$ *only* for p's that are expressible (say, in some language S uses) with *at most* some degree k of *complexity*.
- Problem: how will we ensure that $\mathcal{B}^t$ is a *Boolean algebra* on these proposals? We'll return to this issue in the Bayesian confirmation unit.

Digression: The Preface Paradox and Epistemic Rationality I

- Before discussing Joyce's argument for *epistemic probabilism* (*i.e.*, for the claim that *probabilism* is good from an *epistemic* point of view), I want to digress briefly to discuss the following epistemic norm for belief:
  1. Logically *consistent* belief states are better than inconsistent states.
     - Why? Inconsistent sets must contain some falsehoods.
- Given that truth is epistemically valuable, it seems *epistemically good* to avoid states that are *guaranteed* (by logic) to contain some false beliefs.
- But, intuitively, this must be read "other things being equal". There seem to be cases in which this epistemic norm (call it "the consistency norm") comes into conflict with other epistemic norms. *E.g.*, *Prefaces* like these:

  > To those of my colleagues and students who have given me encouragement and stimulation, I wish to express sincere thanks. I am especially grateful to ... The errors and shortcomings to be found herein are not their fault, and are present only in spite of their wise counsel.

- Such Prefaces sound reasonable — even *epistemically responsible*. *But* ...

Digression: The Preface Paradox and Epistemic Rationality II

- Let's think a bit harder about what underlies these sorts of Prefaces.
- There are two things going on with Prefaces for long (non-fiction) books:
  1. Assuming the author S is epistemically responsible, then S will be justified in believing each individual claim in the book: $p_1, \ldots, p_n$.
     - Presumably, a responsible author wouldn't assert a claim p_i, if they didn't have sufficient grounds/justification for doing so.
  2. But, it also seems epistemically reasonable for S to believe that *any* long book (written by humans) will contain *some* errors/*false claims*.
     - We all know that human beings are imperfect. Thus, it is reasonable to believe that *any* long book is bound to contain *some* errors. [It also seems reasonable to believe the same of *one's own* (long) book.]
- **Preface Paradox.** (1) and (2) entail that S's belief set $\mathcal{B}$ is *inconsistent*.
  - Why? $\mathcal{B}$ *contains* $p_1, \ldots, p_n$, but $\mathcal{B}$ also *entails* the proposition that *at least one* of the claims $p_1, \ldots, p_n$ is *false* — *i.e.*, $\sim(p_1 \,\&\, \cdots \,\&\, p_n)$.

Digression: The Preface Paradox and Epistemic Rationality III

- Here's a more general paradox. Let $\mathcal{B}$ be your current (total) belief set.
- By assumption, you *believe* each of the elements $p_1, \ldots, p_n$ of $\mathcal{B}$.
- But, you know $\mathcal{B}$ is a very large and complex set of beliefs, and you also know that (all) human believers are epistemically imperfect. So, it also seems reasonable for you to believe that *some* element(s) of $\mathcal{B}$ are false.
- However, *that* "self-humility" claim (H) is *inconsistent with* the members of $\mathcal{B}$ ($p_1, \ldots, p_n$). So, if you *also* believe H, then your belief set $\mathcal{B}$ — which then contains $p_1, \ldots, p_n$ *and* H — will be inconsistent. General pattern:
  - *A fortiori*, you believe each element of your belief set $\mathcal{B}$.
  - You know your belief set $\mathcal{B}$ is very large and complex.
  - You know that all human beings are prone to epistemic error, and $\therefore$ (reasonably) believe that all large belief sets contain some falsehoods.
  - On this basis, you (reasonably) believe $\mathcal{B}$ contains some falsehoods.
  - But, now, $\mathcal{B}$ is *inconsistent*, and you $\therefore$ violate the consistency norm!

Epistemic Probability: Joyce's Argument for Epistemic Probabilism I

- Joyce's argument is based on the following assumptions and set-up:
  - Each agent S has a finite (or countable) set of propositions $\mathcal{B}$ (it need not be a Boolean algebra) that they are capable of entertaining and assigning some degree of credence (*i.e.*, epistemically rational d.o.b) to.
  - We are concerned with the *accuracy* of an agent's credence function q. We use the notation $I(q, w)$ to denote the *in*accuracy of q in world w.
  - $I(q, w) = \sum_{X \in \mathcal{B}} \rho(q(X), w(X))$, where $q(X)$ is S's credence in X, $w(X)$ is the truth-value of X at w (0 if X is false in w, 1 if X is true in w), and $\rho(q(X), w(X))$ is a measure of "distance" between $q(X)$ and $w(X)$.
  - Note two things about I. First, I is defined *piecewise*. Second, I is an *additive* function of the "distances from the truth" of each $p \in \mathcal{B}$.
  - The **Norm of Gradational Accuracy** (NGA) says that an epistemically rational agent strives to have credences that are *as close to the truth as possible*, *i.e.*, to *minimize* $I(q, w)$ in some suitable sense of "minimize".

Epistemic Probability: Joyce's Argument for Epistemic Probabilism II

- **Theorem.** Given four assumptions $\mathcal{A}$ (to be discussed below) about I, if q is not a probability function on $\mathcal{B}$, then there exists a probabilistic credence function q', which is strictly more accurate than q, in every possible world w (*i.e.*, *no matter what the truth-values of the p's in $\mathcal{B}$ turn out to be*). *I.e.*, $\exists$ a Pr-function q' such that $I(q', w) < I(q, w)$, for all w.
- This is a nice *epistemic* result. It purports to show that non-probabilistic credence functions are *epistemically* inferior to probabilistic ones — in terms of their *accuracy*. Of course, the devil is in the details: $\mathcal{A}$.
- Next, I'll discuss each of the assumptions in $\mathcal{A}$ (*i.e.*, the pre-conditions of the above Theorem). As we will see, the first two of these seem very plausble. The second pair of assumptions is more controversial.
- We will discuss at some length some objections to the second pair advanced by Patrick Maher. One assumption in the first pair also been called into question by Gibbard (see his paper). We won't discuss that.

1. **Dominance.** If $q_1(y) = q_2(y)$, for all $y \neq x$, then $I(q_1, w) > I(q_2, w)$ iff $\rho(q_1(x), w(x)) > \rho(q_2(x), w(x))$. [Nobody objects to this one!]
   – This ensures the appropriateness of the linear, piecewise way of thinking about $I(q, w) = \sum_{X \in \mathcal{B}} \rho(q(X), w(X))$. *No Holism*!
2. **Normality**. If, for all x, $|w(x) - q_1(x)| = |w'(x) - q_2(x)|$, then $I(q_1, w) = I(q_2, w')$. [Allan Gibbard objects to this. We won't discuss.]
   – *Over*estimating a proposition's p's truth-value by ϵ is counted as exactly as inaccurate as *under*estimating its truth-value by ϵ.

---

3. **Weak Convexity** (WC). If $I(q_1, w) = I(q_2, w)$, then $I(q_1, w) \geq I(\frac{1}{2}q_1 + \frac{1}{2}q_2, w)$ with identity *only if* $q_1 = q_2$.
   – *Averaging* two equally inaccurate q's [q_1, q_2] can't yield a *strictly more inaccurate* $\bar{q}$. And, $\{\bar{q}, q_1, q_2\}$ are *equally inaccurate* ***only if*** $q_1 = q_2$.
4. **Symmetry** (S). If $I(q_1, w) = I(q_2, w)$ then, for any $\lambda \in [0, 1]$, $I(\lambda \cdot q_1 + (1 - \lambda) \cdot q_2, w) = I((1 - \lambda) \cdot q_1 + \lambda \cdot q_2, w)$.
   – Symmetric weighted averages of equally inaccurate q's remain so.

- (WC) and (S) are the really controversial (and non-obvious) ones. To wit ...

Epistemic Probability: Joyce's Argument for Epistemic Probabilism III

- (WC) and (S) place constraints on which underlying distance measures ρ are admissible. There are *many* measures of "distance" ρ which lead to inaccuracies I satisfying (WC) & (S). *E.g.*, the *square* difference measure $\rho^{\dagger}$:

$$\rho^{\dagger}(q(X), w(X)) \overset{\text{def}}{=} (q(X) - w(X))^2$$

  leads to an I satisfying (WC) and (S). Distance measures like $\rho^{\dagger}$ are called *proper scoring rules* (PSRs) [see below]. Are all measures of distance PSRs?

- Patrick Maher points out that the *absolute value* difference measure ρ^*:

$$\rho^*(q(X), w(X)) \overset{\text{def}}{=} |q(X) - w(X)|$$

  leads to an I which *violates* (WC) & (S). And, Maher shows that ρ^* can*not* undergird Joyce's Theorem, since ρ^* allows non-probabilistic credence functions to be more accurate than probabilistic ones (at some w's)!

- Maher argues that ρ^* is an adequate measure of distance. If Maher is right, then this is a serious problem for any approach like Joyce's.

Epistemic Probability: Joyce's Argument for Epistemic Probabilism IV

- ρ^* leads to an I that violates both (WC) and (S). To find a counterexample to (WC) for ρ^*, one needs at least 2 propositions in $\mathcal{B}$ (HW!). Here's one:
  - Assume A is true, and B is false, in w. So, $w(A) = 1$ and $w(B) = 0$.
  - Let $q_1(A) = q_1(B) = 1$, and $q_2(A) = q_2(B) = 0$.
  - Therefore, $q_3(A) = \dfrac{q_1(A) + q_2(A)}{2} = \dfrac{1+0}{2} = \dfrac{1}{2}$.
  - And, $q_3(B) = \dfrac{q_1(B) + q_2(B)}{2} = \dfrac{1+0}{2} = \dfrac{1}{2}$.
  - Thus, $I(q_1, w) = \rho^*(q_1(A), w) + \rho^*(q_1(B), w) = |1-1| + |1-0| = 1$.
  - And, $I(q_2, w) = \rho^*(q_2(A), w) + \rho^*(q_2(B), w) = |1-0| + |0-0| = 1$.
  - And, $I(q_3, w) = \rho^*(q_3(A), w) + \rho^*(q_3(B), w) = |\frac{1}{2} - 1| + |\frac{1}{2} - 0| = 1$.
  - So, we have $I(q_1, w) = I(q_2, w) = I(q_3, w)$, but $q_1 \neq q_2$, violating (WC).
  - Note: $\rho^\dagger(q_3(A), w) = \rho^\dagger(q_3(B), w) = (\frac{1}{2} - 1)^2 = \frac{1}{4}$, which blocks the counterexample. [HW: there are *no* 2-proposition counterexamples for $\rho^\dagger$.]

Epistemic Probability: Joyce's Argument for Epistemic Probabilism V

- To find a counterexample to (S) for ρ^*, one also needs ≥ 2 propositions:
  - Assume A is true, and B is true, in w. So, $w(A) = 1$ and $w(B) = 1$.
  - Let $q_1(A) = 1$ and $q_1(B) = 3$. And, let $q_2(A) = q_2(B) = 0$.
  - Thus, $I(q_1, w) = \rho^*(q_1(A), w) + \rho^*(q_1(B), w) = |1 - 1| + |1 - 3| = 2$.
  - And, $I(q_2, w) = \rho^*(q_2(A), w) + \rho^*(q_2(B), w) = |1 - 0| + |1 - 0| = 2$.
  - $I(q_1, w) = I(q_2, w)$. So, (S) $\Rightarrow I(q_3, w) = I(q_4, w)$, where $\lambda = \frac{1}{4}$; and:

$$q_3 = \lambda \cdot q_1 + (1 - \lambda) \cdot q_2 = \tfrac{1}{4} \cdot q_1 + \tfrac{3}{4} \cdot q_2$$
$$q_4 = (1 - \lambda) \cdot q_1 + \lambda \cdot q_2 = \tfrac{3}{4} \cdot q_1 + \tfrac{1}{4} \cdot q_2$$

But,

$$\begin{aligned} I(q_3, w) &= \rho^*(q_3(A), w) + \rho^*(q_3(B), w) \\ &= |1 - (\frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 0)| + |1 - (\frac{1}{4} \cdot 3 + \frac{3}{4} \cdot 0)| = \frac{3}{4} + \frac{1}{4} = 1 \\ I(q_4, w) &= \rho^*(q_4(A), w) + \rho^*(q_4(B), w) \\ &= |1 - (\frac{3}{4} \cdot 1 + \frac{1}{4} \cdot 0)| + |1 - (\frac{3}{4} \cdot 3 + \frac{1}{4} \cdot 0)| = \frac{1}{4} + \frac{5}{4} = \frac{3}{2} \end{aligned}$$

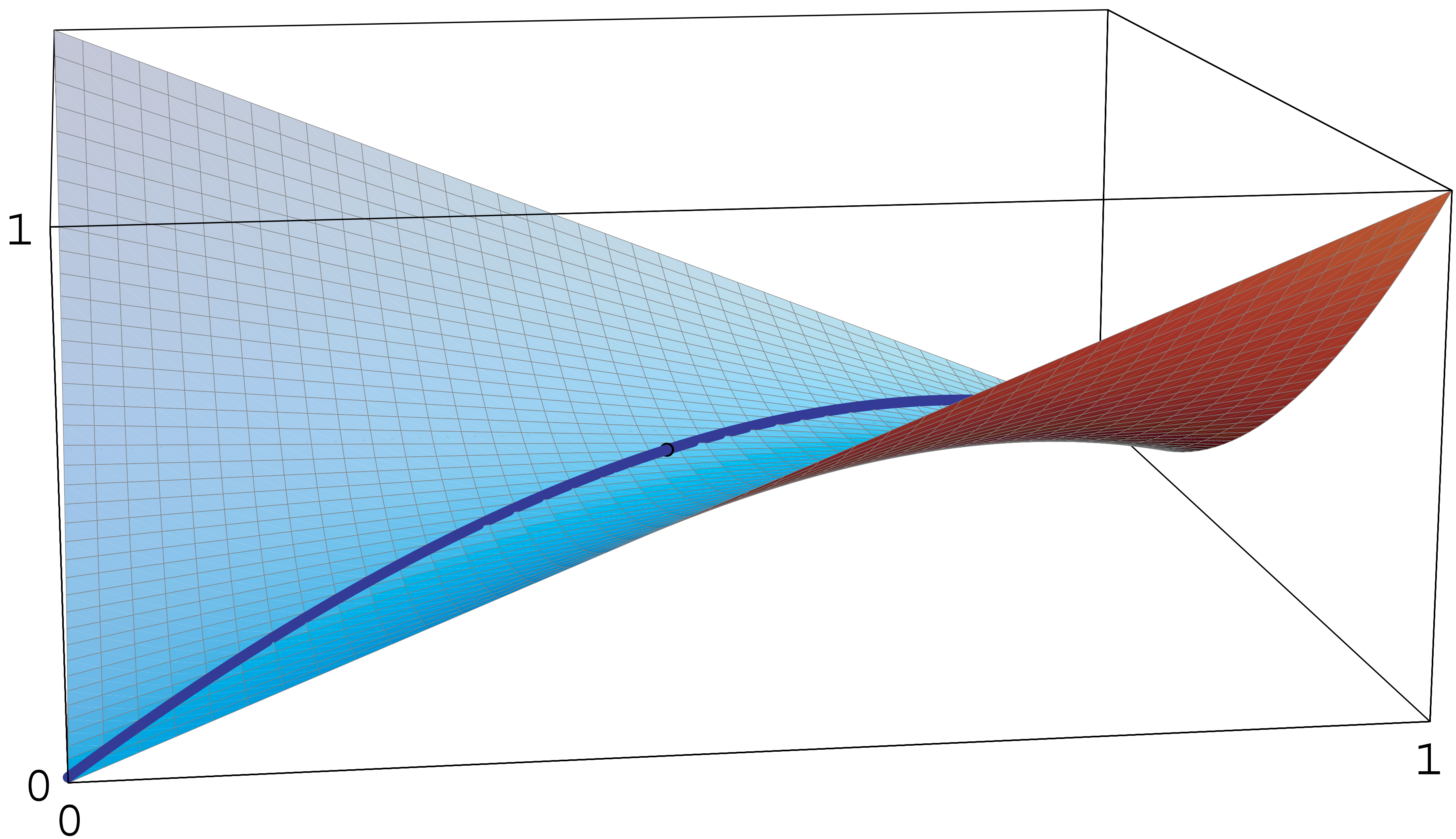
Epistemic Probability: Joyce's Argument for Epistemic Probabilism VI

- **Joyce's Reply**: define *expected distance of* q' *(according to* q*) from the truth concerning* X*, relative to* ρ*,* $EDT(q', q, X, \rho)$, in the following way:

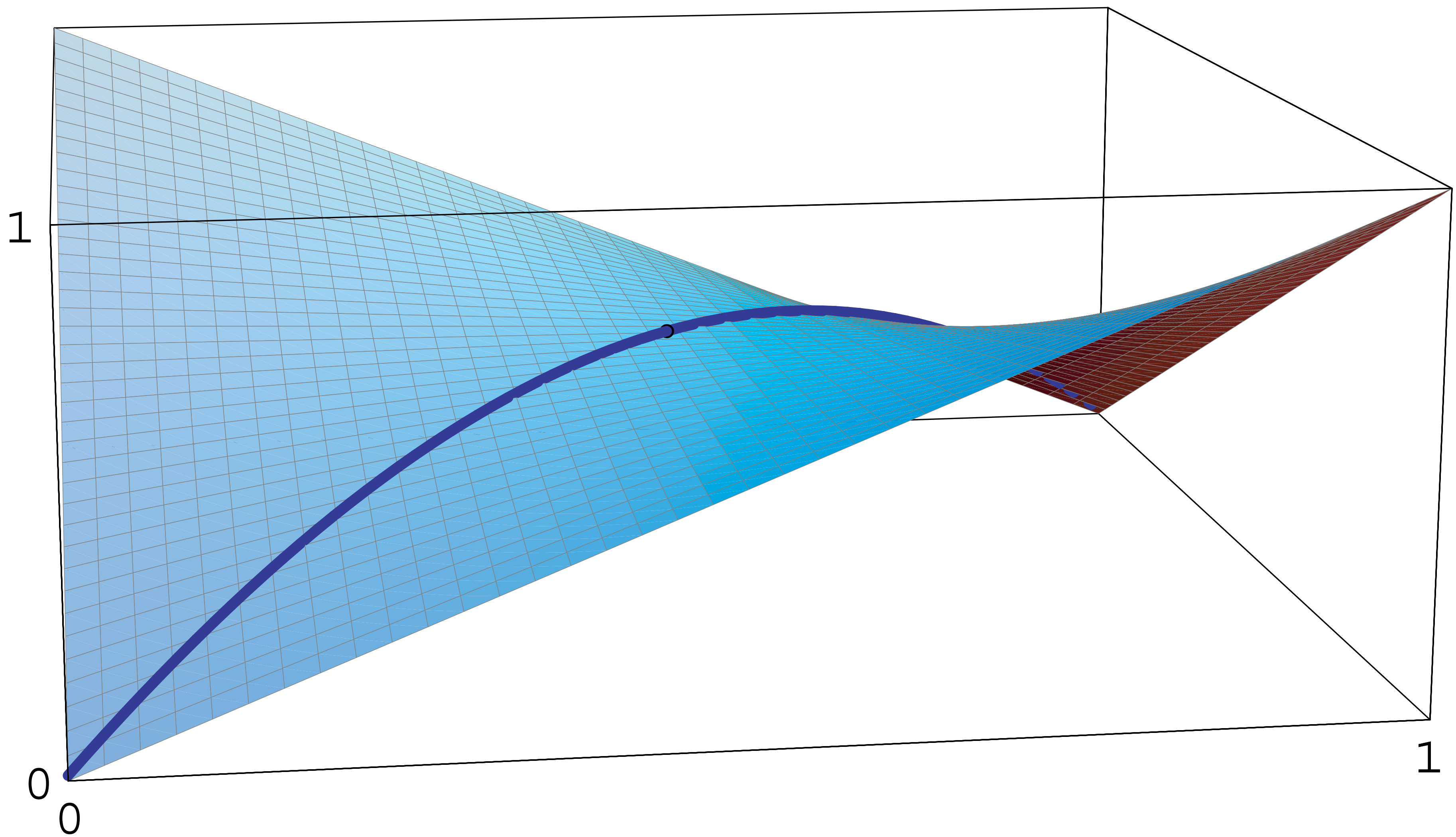
$$\begin{aligned} EDT(q', q, X, \rho) &= q(X) \cdot \rho(q'(X), 1) + (1 - q(X)) \cdot \rho(q'(X), 0) \\ &= q \cdot \rho(q', 1) + (1 - q) \cdot \rho(q', 0) \end{aligned}$$

- **Theorem.** If $\rho = \rho^*$, then $EDT(q', q, X, \rho)$ can be minimized *only* by assigning $q'(X) = 0$, $q'(X) = \frac{1}{2}$, or $q'(X) = 1$. If $\rho = \rho^\dagger$, $EDT(q', q, X, \rho)$ will be minimized when $q'(X) = q(X)$, which can lie *anywhere* on $[0, 1]$.
- If you aim to minimize your *expected inaccuracy* of your own q, then you shouldn't use ρ^*, since this will lead you to view your current degrees of belief as inferior to one that assigns 0, 1 or $\frac{1}{2}$ to *every* X in $\mathcal{B}$.
- But, it's OK to use $\rho^\dagger$, since this leads to a "*stable*" q. See plots, below.
- Questions: *why* minimize $\boldsymbol{EDT}$? Wasn't the goal to minimize DT? Also, why do we get to *assume* here that $q({\sim}X) = 1 - q(X)$? Question begging?

Plot of $EDT(q', q, X, \rho^{\dagger})$ — the curve $q' = q$ is a *stable* "saddle".



Plot of $EDT(q', q, X, \rho^*) - q' = q$ is *un*stable (except at $\{0, \frac{1}{2}, 1\}$).



Review of Joyce's Epistemic Argument for Probabilism

- Joyce aims to show that, unless q is a probability function, there will exist a q' which is closer to the truth – *no matter what the truth turns out to be*.
- The "closeness to the truth" of q in world w is based on some measure ρ of the distance between $q(X)$ and the truth-value ($w(X) = 0$ or 1) of X in w, summed over all of the X's in S's set of entertainable propositions $\mathcal{B}$.
- Joyce's argument rests on four assumptions $\mathcal{A}$ about the inaccuracy measure I. $\mathcal{A}$ is consistent with $\rho^{\dagger}(q(X), w(X)) = |q(X) - w(X)|^2$, but not $\rho^{*}(q(X), w(X)) = |q(X) - w(X)|$, or various other distance measures.
- Maher shows that ρ^{*} violates two of Joyce's axioms: (WC) and (S), and he argues that ρ^{*} is nonetheless a reasonable measure of "distance from truth". Maher also shows that Joyce's Theorem *fails* if one uses ρ^{*}.
- Joyce's response appeals to "Expected Distance from Truth". But, this undermines one of the nice things about Joyce's original argument — that it made no appeal to Expected Utility Theory (which seems *pragmatic*).