Indifference to Anti-Humean Chances

J. DMITRI GALLOW †

An indifference principle says that your initial, or *ur-prior*, credences should be distributed uniformly over each of the possibilities you recognise. A chance deference principle says that your ur-prior credences should be aligned with the chances, in the following sense: your ur-prior credence in P, given that the chance of P is x, should be x. My thesis is that, if we are anti-Humeans about chance, then these two principles are incompatible with each other. Anti-Humeans think that it is *possible*—though unlikely—for the actual frequencies to depart from the chances. So long as you recognise possibilities like this, you cannot both invest equal credence in every possibility and defer to the chances. If your ur-prior credences are spread evenly over every possibility, then they will not defer to the chances; and if your ur-prior credences defer to the chances, they will not be spread evenly over every possibility.

In §§1–3 below, I'll introduce anti-Humeanism (AH), a principle of chance deference (CD), and an indifference principle (IP). The principles CD and IP both say something about what your credences should be like in the absence of any evidence—that is to say, they both impose constraints on your initial, or *ur-prior*, credence function. In §4, I'll explain why, if we're anti-Humeans, we cannot satisfy both of these constraints. In §5, I'll consider some anti-Humean responses to this incompatibility. In brief: anti-Humeans may retreat to a weaker indifference principle according to which you should invest equal credence in each *categorical* possibility—where a categorical possibility describes what happens, but does not say what the objective chances are. Alternatively, anti-Humeans could move to an even weaker indifference principle which says only that you should give equal credence to evidentially symmetric propositions, but does not say that, in the absence of evidence, each possibility you recognise is evidentially symmetric with every other. As I'll explain in §5 below, this weaker indifference principle has prominent

advocates—for instance, Keynes (1921) and White (2009). However, the indifference principle which I will call 'IP' is not some straw man. Authors like Pettigrew (2016) argue for exactly this principle.

1 Anti-Humeanism

I'll suppose that you have degrees of belief, or credences, defined over the sentences in some language. In the simplest case, this will be a truth-functional language with a finite number of atomic sentences, $A_1, A_2, \ldots, A_N$. In a language like this, a *state description* is a conjunction of the form $\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N$, where each $\pm A_i$ is either the atomic sentence A_i or its negation. Let 'Ω' be the set of state descriptions, and let $\omega \in \Omega$ be some particular state description. If the language is truth-functional, a state description will settle the truth-value of every other sentence in the language, and every sentence in the language will be equivalent to some disjunction of state descriptions.

Instead of taking your credences to be defined over sentences in a language, we could instead start with a space of *possible worlds*, $\mathcal{W}$, and take your credences to be defined over *propositions* (sets of those possible worlds). Some defenders of indifference principles—Williamson (2010), for instance—use the sentential framework, while others—Pettigrew (2016), for instance—use the propositional framework. But the distinctions between these two frameworks won't make a difference to my discussion here. For it is straightforward to translate between them—I'll leave the details in this footnote.[1] Here, I'll stick to the sentential framework, but the translation scheme allows everything I say about it to carry over to the propositional framework.

In the simplest case, your credences are defined over a simple truth-functional language. However, if you wish to entertain sentences about the chances—sentences

---

1. We may take each state description $\omega \in \Omega$ to correspond to a possible world $w \in \mathcal{W}$. Each sentence is equivalent to some disjunction of state descriptions. So we may associate each sentence with the set of state descriptions in this disjunction, which we may in turn associate with a set of possible worlds, or a proposition. While this translation scheme gives us a surjective function from sentences to propositions (sets of possible worlds), the function is not a bijection. For there will be multiple sentences translated to the same proposition. Even so, any two sentences translated to the same proposition are equivalent. Since your ur-prior is a probability, it assigns equivalent sentences the same probability. Consider the equivalence classes of equivalent sentences. The proposed translation establishes a bijection between propositions and these equivalence classes. So the probability which an ur-prior gives to a proposition (a set of possible worlds) will correspond to the probability which an ur-prior gives to any sentence in the corresponding equivalence class. So we may go back and forth between the two frameworks.

like '$Ch(P) = x$' ('the objective chance of P is x') or '$Ch = ch$' ('ch is the objective chance function')—then we will need to consider a slightly more complicated language. We can generate an appropriately rich language by distinguishing two different kinds of atomic sentences, which I'll call the *atoms* and *chance hypotheses*. The atoms are just the atomic sentences which aren't chance hypotheses. Let the atoms be $A_1, A_2, \ldots, A_N$. The chance hypotheses say what the objective chance function is—that is, the chance hypotheses are sentences of the form '$Ch = ch$', for some probability function ch.[2] We then get the full language by taking the union of the chance hypotheses and the atoms, and closing the resulting set under negation and disjunction. We may recover sentences like '$Ch(P) = x$', since they are equivalent to the disjunction '$\bigvee_{ch:ch(P)=x} Ch = ch$'.

With this richer language, we should change the way we think about a *state description*. If ch and ch' are two distinct probability functions, then both $Ch = ch$ and $Ch = ch'$ will be atoms. If we continue on with our old definition of a state description, there would be a state description which included both of these sentences as conjuncts. These sentences may be known *a priori* to be incompatible with each other. This creates two problems. In the first place, a state description is meant to represent an epistemic *possibility*. But if the state description is *a priori* false, then it is cannot be an epistemic possibility. In the second place, there is a problem for anyone who wishes to endorse an indifference principle like the one I'll introduce in §3. That principle says that every state description should receive the same credence; so, if we used this understanding of 'state description', then this indifference principle would require that you to have a positive credence in something you may know *a priori* to be false.

Once we've got credences defined over chance hypotheses, we must change the way we think about state descriptions. Which change is appropriate may depend upon our metaphysical commitments. Consider the following example: we're going to flip a coin N times, and then destroy it. No other coins have ever or will ever be flipped throughout the history of the universe. Now, suppose that you have

---

2. Here we face a choice point. We could either take the potential chance functions ch to be defined *only* over the sentences in the language generated from the atoms $A_1, A_2, \ldots, A_N$, or we could take them to be defined over *every* sentence in the language. If the objective chances may be uncertain about what the objective chances are, then this second option leads to cardinality worries—since, in general, the space of possible probability distributions over Ω is larger than Ω. However, anti-Humeans should be happy to assume that the objective chances are certain of what the objective chances are, so that, for each potential chance function ch, $ch(Ch = ch) = 1$. (This follows from Lewis's *principal principle*.) This means that, even if we take the objective chances to be defined over sentences like '$Ch = ch$', there will be exactly one such distribution for each function ch, and we avoid cardinality concerns.

credences defined over the atoms $H_1, H_2, \ldots, H_N$, where H_i says that the ith flip landed heads. Now, the question is whether, in this application, we should count $H_1 \wedge H_2 \wedge \cdots \wedge H_N \wedge Ch = ch$ as a state description if $ch(H_i) = 50\%$, for each i. That is: should we have a state description which tells us that the coin landed heads every flip, and that the coin was fair? It's quite natural to think 'yes'—after all, it appears to be *possible* for a fair coin to land heads N times in a row. Sure, it's *unlikely*, but that doesn't make it *impossible*. This is what anti-Humeans think. But Humeans will disagree.

As I'm using the term here, Humeans think that the chance laws supervene upon the distribution of local matters of particular fact. According to Humeans, chances are something like executive summaries of what actually happens in the world. For instance, actual frequentism is a form of Humeanism. The actual frequentist thinks that the chances are just the actual frequencies. In our example, the actual frequentist will say that the chance of the coin landing heads on the ith flip is just the proportion of flips which actually land heads. Thus, if every flip lands heads, the chance of the ith flip landing heads would have to be 100%, and not 50%. For this reason, $H_1 \wedge H_2 \wedge \cdots \wedge H_N \wedge Ch = ch$ will be *impossible* unless $ch(H_i) = 100\%$.

Anti-Humeans disagree. According to them, the chance laws have an independent existence. Chance laws govern outcomes, but they are not reducible to outcomes. A crude analogy: for the anti-Humean, when outcomes are objectively chancy, God is rolling dice. Statements about the chances tell you something about the bias of God's dice, but the chance laws don't necessitate that those dice land any particular way, not even in the long-run. So, for instance, anti-Humeans say that there's no reason that a fair coin couldn't land heads on every flip. More generally, they will say that there's no reason some conjunction of (negations) of atoms $\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N$ couldn't be true even while the chance of it being truth is minuscule. So, if we are anti-Humeans, then we should think that the conjunction $\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N \wedge Ch = ch$ is possible—and we should count it as a state description—even if $ch(\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N)$ is minuscule.

In general, given a language which consists of a set of (non-chance hypothesis) atoms $A_1, A_2, \ldots, A_N$, and potential chance hypotheses $Ch = ch_1, Ch = ch_2, \ldots, Ch = ch_M$, an anti-Humean should be happy to say that a *state description* is any conjunction of the form

$$\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N \wedge \pm Ch$$

where each $\pm A_i$ is either the atom A_i or its negation, and $\pm Ch$ is either one of the chance hypotheses $Ch = ch_i$ or else the negation of every chance hypotheses,

$\bigwedge_{i=1}^{M} Ch \neq ch_i$. In the possible worlds framework, this means that, for every chance hypothesis and every assignment of truth-values to the atoms, there is a possible world at which that chance hypothesis is true and that assignment of truth-values is realised.

For my purposes, I won't need a thorough-going anti-Humeanism. Instead, I will need only the following consequence of it, which I will from here on out refer to as 'anti-Humeanism', or 'AH':

Anti-Humeanism (AH) There is a pair of state descriptions which have the form
$\phi \wedge Ch = ch$ and $\psi \wedge Ch = ch$, where $ch(\phi) \neq ch(\psi)$.

Actual frequentists will want to deny this assumption in the case of Bernoulli processes like coin flips, since any state description with the same chance function will have to have the same frequency of heads landings (or whatever), and the chance function will assign the same chance to state descriptions with the same frequencies. Nonetheless, AH is a very minimal form of anti-Humeanism. Even 'best system' Humeans like Lewis (1994) will be happy to accept AH in many contexts.

2   Chance Deference

The most prominent principle of chance deference is David Lewis (1980)'s *principal principle*.[3] The principal principle says something about a rational initial, or *ur-prior*, credence function, C—the credence function it would be rational to have in the absence of any evidence.[4] In particular, it says: if P is any sentence,[5] t is some future time, $Ch_t(P) = x$ says that the time t chance of P is x, for some real number $x \in [0,1]$, and E is any time t admissible evidence which is compatible with $Ch_t(P) = x$, then

$$C(P \mid Ch_t(P) = x \wedge E) = x$$

The time t won't be important in my discussion, so I'll fix t to be some future time and omit explicit mention of t in the remainder. Likewise, the admissible ev-

---

3. For alternative chance deference principles, see Hall (1994), Ismael (2008), Levinstein (forthcoming), and Dorst *et al.* (2021), for instance.
4. I will assume that C is a probability, by which I mean: 1) $C(P) \geqslant 0$, for every P; 2) if P is *a priori* knowable, then $C(P) = 1$; and 3) if it is *a priori* knowable that no two of $P_1, P_2, \ldots$ are true at once, then $C(P_1 \vee P_2 \vee \ldots) = C(P_1) + C(P_2) + \ldots$.
5. Lewis assumes that the arguments of your credence function are propositions. Since I'm assuming here that the arguments of your credence function are sentences, I've slightly emended his principal principle. Given the translation scheme from §3, the formulation is equivalent.

idence E won't play any important role. I'll assume only that information about the chances is itself admissible at t. If $\mathcal{C}h = ch$ is compatible with $\mathcal{C}h(P) = x$, then $ch(P) = x$. And if $ch(P) = x$, then the conjunction $\mathcal{C}h(P) = x \wedge \mathcal{C}h = ch$ is equivalent to $\mathcal{C}h = ch$. So if we set $E = \mathcal{C}h = ch$ in the principal principle, we get the following:

Chance Deference (CD) Your ur-prior credence in P, given that the objective chance function is ch, should be $ch(P)$.

$$C(P \,|\, \mathcal{C}h = ch) = ch(P)$$

CD governs your *conditional* credences; but I'll suppose that these conditional credences place a constraint on your unconditional credences, *via* the product rule, which says that, for any sentences P and Q, your credence in $P \wedge Q$ is equal to the product of your credence that P *given* Q and your credence that Q. So, if $C(\mathcal{C}h = ch) > 0$, CD implies that

(1) $$\frac{C(P \wedge \mathcal{C}h = ch)}{C(\mathcal{C}h = ch)} = ch(P)$$

CD works well when there are at most a finite number of potential values for x. However, you may want your credences to be defined over uncountably many sentences of the form $\mathcal{C}h(P) = x$—one for each of the uncountably many real numbers, x, between 0 and 1. In that case, CD may have to be generalised. I discuss this generalisation in the appendix, §B.2.

3   Indifference

The indifference principle I'll be interested in here says that, in the absence of evidence, you should give every state description precisely the same probability.

Indifference Principle (IP) For any two state descriptions, ω and ω^*, your ur-prior credence in ω should be equal to your ur-prior credence in ω^*, $C(\omega) = C(\omega^*)$.

In the propositional framework, IP says that, for any two *possible worlds*, w and w^*, your ur-prior credence in w must equal your ur-prior credence in w^*, $C(w) = C(w^*)$.

Let me separate out two different theses which together imply IP:

Symmetry to Indifference (STI) If 'P' and 'Q' are evidentially symmetric, then your credence in 'P' should equal your credence in 'Q'.

State Description Symmetry (SDS) Any two state descriptions are evidentially symmetric.

As I will discuss in §5 below, not every defender of indifference principles has endorsed SDS. So not every defender of indifference principles has endorsed the principle I am here calling 'IP'. However, as I will also discuss in §5, some prominent defenders of indifference principles *have* endorsed the stronger IP. Moreover, prominent arguments for indifference imply the stronger thesis IP, not just the weaker STI.

If there are a finite number of atomic sentences in your language, then there will be finitely many state descriptions in Ω. However, if there are a countable infinity of atomic sentences, there will be uncountably many state descriptions. In this case, IP will be trivially satisfied so long as every state description is given a probability of zero. Nonetheless, there is another form of indifference which we may want to impose in this case. I discuss this stronger indifference principle in the appendix, §B.1.

4   The Incompatibility

In this section, I'll show that AH, CD, and IP are incompatible by assuming all three and deriving a contradiction. This will show that, if we are anti-Humeans, we must choose between indifference and showing deference to the chances.

Assume AH, CD, and IP. By AH, there is a pair of state descriptions which have the form $\phi \wedge Ch = ch$ and $\psi \wedge Ch = ch$, where $ch(\phi) \neq ch(\psi)$. Then,

$$\frac{C(\phi \wedge Ch = ch)}{C(\psi \wedge Ch = ch)} = \frac{C(\phi \wedge Ch = ch)/C(Ch = ch)}{C(\psi \wedge Ch = ch)/C(Ch = ch)} = \frac{ch(\phi)}{ch(\psi)} \neq 1$$

The final equality follows from (1), with 'ϕ' and 'ψ' substituted in for 'P'. As we saw in §2, (1) follows from CD. It therefore follows from CD that

(2) $$C(\phi \wedge Ch = ch) \neq C(\psi \wedge Ch = ch)$$

On the other hand, since IP requires that every state description get the same credence, it implies that

(3) $$C(\phi \wedge Ch = ch) = C(\psi \wedge Ch = ch)$$

Contradiction.

So, if we assume AH, CD, and IP, we arrive at a contradiction. Assuming we are anti-Humeans, then, we face a choice between CD and IP. We cannot both be indifferent and show deference to the chances.

Perhaps this incompatibility only arises because we considered a *finite* number of chance hypotheses? In appendix B, I show that a similar incompatibility arises even if there are uncountably many chance hypotheses.

5   Further Discussion

One kind of reaction to this incompatibility is to reject one of the principles and leave it at that. For instance, Humeans may say: so much the worse for anti-Humeanism! (As an anti-Humean myself, I am more inclined to see the foregoing as a reason to reject IP, though I won't insist upon that here.) I won't have anything further to say about this kind of reaction. However, there is a more moderate re-action which is worth discussing: anti-Humeans may wish to defer to the chances and still endorse *some* form of indifference principle. They may achieve this by weakening the principle IP.

As I formulated IP, it says that your credence in any state description must be equal to your credence in any other state description. And as I've understood it, a *state description* specifies all of the things your language is able to tell you about the world. It describes matters in as precise a detail as your language will permit. If ω is a state description, then any other description of the world (in your language) is either entailed by ω or incompatible with ω. However, we might weaken IP by having it say that your credence in any *categorical* state description is the same as your credence in any other—where a *categorical* state description is logically weaker than a full state description. It specifies what happens, but fails to say what the objective chances of those happenings are. Such a weakening of IP need not conflict with CD and AH.

Again, suppose that you have some collection of atomic sentences in your lan-guage: the 'atoms' $A_1, A_2, \ldots, A_N$, and the 'chance hypotheses', $Ch = ch_1, Ch = ch_2, \ldots, Ch = ch_M$. Say that a *categorical state description* describes the world in as rich a detail as the atoms (excluding the chance hypotheses) permit. That is, it is a conjunction of the form $\pm A_1 \wedge \pm A_2 \wedge \cdots \wedge \pm A_N$, where each $\pm A_i$ is either A_i or $\neg A_i$. Then, consider the weakened principle WIP.

Weak Indifference Principle (WIP) For any two categorical state descriptions, ω

and ω^*, your ur-prior credence in ω should be equal to your ur-prior credence in ω^*, $C(\omega) = C(\omega^*)$.

There is no conflict between AH, CD, and WIP. So why not simply restrict IP in this way so as to make it consistent with AH and CD?

We could certainly do so. However, I personally have a hard time seeing the philosophical motivation for accepting WIP while rejecting IP. By way of explanation, let me say something about what kind of constraint IP imposes on an ur-prior credence, and why its defenders have thought you should satisfy this constraint when you lack evidence. In general, a credence function will encode relations of evidential relevance. If your credence in P given Q is greater than your credence in P, this encodes the fact that you take Q to be evidence for P. The IP imposes a rather demanding constraint on what kinds of evidential relevance relations you're permitted to recognise in the absence of evidence. It forbids taking any atomic sentence of your language to be evidence for any other atomic sentence of your language in the absence of evidence.

Williamson (2010) justifies the ur-prior recommended by IP on the grounds that it is leads to maximally cautious actions: it "is on average the more cautious policy when it comes to risky decisions", in the sense that it "minimises worst-case expected loss".[6] Similarly, Pettigrew (2016) argues for IP on the grounds that it *epistemically* cautious: it minimises the worst case with respect to the *accuracy* of your beliefs. According to Pettigrew, "what is wrong with assigning greater credence to one possibility over another in the absence of evidence is that by doing so you risk greater inaccuracy than you need to risk. [If you violate IP, then] there is an alternative [ur-prior] credence function, namely the uniform distribution…that has lower inaccuracy in its worst-case scenario than you have in yours."[7] Neither of these arguments depend in any way upon assumptions about the *content* of the atomic sentences in your language, nor whether they are about chance hypotheses. So I have a hard time seeing why we should find those arguments any less compelling when some of the atomic sentences in your language are chance hypotheses. Moreover, if we grant an exemption for the chance hypotheses, one wants to know why a similar exemption cannot be granted for *other* atomic sentences.

Suppose your language contains only the atomic sentences $B_1, B_2, \ldots, B_N$, where B_i says that the ith raven is black. Some of us think that, even before receiving evidence, you should take the first k ravens being black to be evidence for the

---

6. Williamson (2010, pp. 62 & 65). See Williamson (2010, §3.4.4) for more.
7. Pettigrew (2016, p. 164). See Pettigrew (2016, part III) for more.

$k + 1$st raven being black. Some of us say that—even with this simple language, and even in the absence of evidence—your credence in B_N, given $B_1 \wedge B_2 \wedge \cdots \wedge B_{N-1}$, should be greater than your unconditional credence in B_N. Both IP and WIP disagree. They say that, with this simple language, before you have any evidence, you must *not* take the fact that the first $N - 1$ ravens are black to be evidentially relevant to whether the Nth raven is black. They say that your credence in the state description $B_1 \wedge B_2 \wedge \cdots \wedge B_{N-1} \wedge B_N$ (every raven is black) must be the same as your credence in the state description $B_1 \wedge B_2 \wedge \cdots \wedge B_{N-1} \wedge \neg B_N$ (every raven is black except for the last one). And if that's so, then your credence that the Nth raven is black, given that the first $N - 1$ ravens are black, will be 1/2, which will be the same as your unconditional credence that the Nth raven is black. (Exactly half of the state descriptions include 'B_N', and exactly half contain its negation.) So, if you satisfy either IP or WIP, then you won't see the blackness of the first $N - 1$ ravens as evidence for the Nth raven being black.

More generally, IP requires that—in the absence of evidence—every atomic sentence is given a credence of 1/2, and every atomic sentence is probabilistically independent of every other. So it forbids recognising evidential relations between atomic sentences, unless you have *evidence* supporting those evidential relations. This imposes a kind of *a priori* inductive skepticism. It forbids an ur-prior from recognising many evidential relations typically recognised by inductive methods. It says that, in the absence of evidence, it is irrational to take 'John testifies that P' or 'It appears that P' to be evidence for 'P'.

Weakening IP to WIP makes an exception to the general rule of not recognising evidential relations between atomic sentences. Such an exception could, of course, be granted. But the reasons provided for IP by defenders like Jaynes (1957), Williamson (2010), and Pettigrew (2016) do not seem to motivate such an exemption. Take an anti-Humean ur-prior which satisfies CD by being more confident in state descriptions in which $A \wedge Ch(A) = 0.6$ than it is in state descriptions in which $\neg A \wedge Ch(A) = 0.6$. This ur-prior builds in more information, and so has greater entropy, than one which satisfies IP by spreading its credence equally over all state descriptions. If we should minimise prior information about whether nature is uniform, whether testifiers are trustworthy, and whether appearances are deceiving, then why shouldn't we *also* minimise information about whether the chances are accurate? If the outcome of a risky action depends upon whether $A \wedge Ch(A) = 0.6$, the ur-prior which satisfies CD will lead to less cautious actions than the one which satisfies IP. If we shouldn't take incautious actions when it comes to whether nature is uniform, whether testifiers are trustworthy, and whether appearances are deceiving, then why should we take incautious actions when it comes to whether

the chances are accurate? And, if we should minimise worst-case epistemic risk when it comes to whether nature is uniform, testifiers are trustworthy, and appearances are deceiving, why shouldn't we also minimise worst-case epistemic risk when it comes to whether chance is accurate? I am not contending that there is no reason for a selective *a priori* inductive skepticism, according to which we have *a priori* grounds to trust in chance, but no *a priori* grounds to trust in regularities, testifiers, or our senses. I am contending that, to my knowledge, no such reason has been given.

Of course, we could weaken WIP further by allowing an ur-prior to build in assumptions about the uniformity of nature, as well as the reliability of testifiers and appearances. More generally, we could allow in any number of *a priori* rationality constraints, and say only that you should spread your ur-prior credences as evenly as possible *subject to* these constraints. That is: your ur-prior credences should be spread evenly, except when this conflicts with some other *a priori* norm of rationality.

Let me make four observations about a principle like this. Firstly, *some* authors who have defended indifference principles have a principle like this in mind. For instance, White (2009) defends the following, which he calls "the principle of indifference":

> If '*P*' and '*Q*' are evidentially symmetrical, then your credence in '*P*' should equal your credence in '*Q*'

In §3, I called this principle 'Symmetry to Indifference' (STI). If we combine STI with the assumption that, in the absence of evidence, any two state descriptions are evidentially symmetrical (SDS), then we get back the principle IP. However, White is not committed to the evidential symmetry of state descriptions. When explaining what it takes for '*P*' and '*Q*' to be *evidentially symmetrical*, he makes it clear that this can include *a priori* reasons to think '*P*' is more likely than '*Q*'. He writes: "I mean to understand evidence very broadly here to encompass whatever we have to go on in forming an opinion about the matter. This can include non-empirical evidence or reasons, if there are such."[8]

Secondly, several *other* authors who have defended indifference principles have the stronger thesis IP in mind. For instance, Pettigrew (2016, §12.1) explicitly rejects White's thesis, and, in its place, advocates the following stronger formulation:

---

8. White (2009, p. 161–2)

Suppose that $\mathcal{F}$ is a finite, rank-complete set of propositions. If an agent has an initial credence function c_0 defined on $\mathcal{F}$, then rationality requires that c_0 is the uniform distribution on $\mathcal{F}$...[where the uniform distribution] assigns to each proposition the proportion of the possible worlds at which it is true.[9]

Pettigrew formulates this principle in a framework where the arguments of your credence function are sets of possible worlds. But, as I explained in §3 above, we may translate between a framework like this and a framework where the arguments of your credence function are sentences. Translating between the two frameworks, his requirement that $\mathcal{F}$ be finite is analogous to requiring that there are finitely many atomic sentences. (The notion of a *rank complete* set is a slightly technical notion which is needed for Pettigrew's theorem, but which isn't relevant to our discussion here. Just note that, given our translation scheme, this condition will be satisfied so long as your language is closed under negation and conjunction and you have a credence in every sentence in your language.[10]) Given the translation, Pettigrew's principle says exactly what IP does: your ur-prior should give every state description the same credence.[11]

Thirdly, while there may be good reason for an anti-Humean to endorse STI while rejecting stronger principles like WIP and IP, the arguments of Williamson (2010) and Pettigrew (2016) do not support this more moderate position. An anti-Humean ur-prior which satisfies CD will lead to less cautious actions than one which satisfies IP. So adopting the weaker principles does not minimise worst-case expected loss. Since Williamson's justification of IP appeals to a principle of minimising worst-case expected loss, that justification cannot be used to support the moderate position. Similarly, an anti-Humean ur-prior which satisfies CD will lead to less *epistemic* caution than one which satisfies IP. As Pettigrew taught us, if you satisfy CD, then there is an alternative ur-prior—namely, the uniform ur-prior—which has a lower inaccuracy in its worst-case scenario than you have in yours (assuming AH). Since Pettigrew's justification of IP appeals to a principle

---

9. Pettigrew (2016, p. 164)

10. For the curious: this is what it is for $\mathcal{F}$ to be rank-complete: if there is a proposition $P \in \mathcal{F}$ which contains N possible worlds, then every other set of N worlds is also included in $\mathcal{F}$.

11. I believe that Williamson (2010) also endorses the principle I've called 'IP', though this is more difficult to establish exegetically, since it hinges upon whether Williamson understands 'evidence' to include *a priori* knowledge, and the text says very little about evidence. In any case, Williamson rejects CD, accepting instead a *diachronic* norm which says that, upon *learning* that the chance of P is x, your credence in P should be x. This means, by the way, that Williamson rejects the diachronic norm of conditionalisation.

which says that in the absence of evidence, you must minimise your worst-case in-accuracy, that justification cannot be used to support accepting STI while rejecting SDS, either.

Finally, depending upon how exacting the other *a priori* norms of rationality are, there may be little to no work left over for STI to do. For instance, suppose that the other *a priori* norms of rationality pin down a precise rational credence in every state description. Then, STI would be vacuously satisfied—which is to say, it would impose no constraint at all. There would be no difference between it and a norm which says to spread your credence as *unevenly* as possible, given the (other) *a priori* rational norms.

In closing, it's worth noting that there is another, less conservative, reaction to the incompatibility from §4. Anti-Humeans may decide to abandon the framework which represents your degrees of confidence with a precise real-valued credence function, C. In its place, they may wish to move to a framework in which your degrees of confidence are represented with a comparative confidence ordering, or a framework in which they are represented with an *imprecise* probability distribution. For the interested reader, I discuss these alternative reactions in appendix A. In brief: a similar incompatibility arises in both of these alternative frameworks.

6   In Summation

In sum: anti-Humeans cannot accept both CD and IP. If they wish to spread their ur-prior credences evenly over each possibility they recognise, then they must not defer to the chances; if they wish to defer to the chances, they cannot spread their ur-prior credences evenly over each of the possibilities they recognise. We could slightly weaken IP to render it compatible with AH and CD, though I personally have a hard time seeing the philosophical motivation for this weakening. There is also an even weaker indifference principle anti-Humeans could satisfy while defer-ring to the chances. This principle allows an ur-prior credence distribution to be uneven so long as this unevenness is required by some other *a priori* requirement of rationality. It says merely that your ur-prior credences should be as even as the other requirements of rationality allow them to be. This principle does not conflict with AH and CD. While there may be good reason to endorse this weaker indiffer-ence principle, it is not supported by the arguments of Williamson and Pettigrew.

A Indifference with Comparative Confidence and Imprecision

I have been taking for granted a traditional Bayesian framework in which your degrees of confidence get represented with a precise probability function. However, there are other frameworks available, and these other frameworks afford us different ways of thinking about what's involved in being 'indifferent' between state descriptions, and what's involved in deferring to the chances. In this appendix, I'll look at two alternative approaches: an approach which represents rational degrees of confidence with a comparative confidence ordering (§A.1), and an approach which represents rational degrees of confidence with an *imprecise* probability function (§A.2).

A.1   Comparative Confidence

In this section, I'll introduce comparative confidence orderings. Representing rational doxastic states with these orderings allows us to formulate indifference principles which avoid the familiar objections to IP$^\infty$ from §B.1.1. However, even with these orderings, we face an analogue of the incompatibility from §4.

The most general kind of comparative confidence ordering is a *conditional* comparative confidence ordering. This is a binary relation between pairs of sentences, which we may write '$[A \mid E] \geqslant [B \mid F]$', and give the interpretation that you are at least as confident in A, given E, as you are in B, given F. From this ordering, we may recover an unconditional comparative confidence ordering by setting the 'conditioning' sentences equal to a tautology. That is, we assume that you think A is not less likely than B exactly when $[A \mid \top] \geqslant [B \mid \top]$, which I will abbreviate with '$A \geqslant B$'.[12] As usual, we may stipulate that $[A \mid E] > [B \mid F]$ iff $[A \mid E] \geqslant [B \mid F]$ and it's not the case that $[B \mid F] \geqslant [A \mid E]$. And we may stipulate that $[A \mid E] \approx [B \mid F]$ iff $[A \mid E] \geqslant [B \mid F]$ and $[B \mid F] \geqslant [A \mid E]$. For my purposes, I'll only need to assume that $\geqslant$ is reflexive and transitive, and that $A \wedge E \geqslant B \wedge E$ whenever $[A \mid E] \geqslant [B \mid E]$.[13]

In this framework, the natural analogue of CD is this:

Comparative Chance Deference (CCD) In the absence of any evidence, for any
sentences P, Q, and any *ch*: if $ch(P) > ch(Q)$, then, given that the chance

---

12. See Fine (1973, ch. 2), and the references contained therein.
13. This follows from Fine's QCC7 (see p. 30).

function is *ch*, you should be more confident in P than Q,

$$[P \mid Ch = ch] \; > \; [Q \mid Ch = ch]$$

There are multiple ways we might try to formulate an indifference principle in this framework. Adapting a proposal from Norton (2008), we could say that you should be as confident in any one state description as you are in any other.

Comparative Indifference Principle (CIP) For any two state descriptions, ω and ω^*, in the absence of evidence, you should be as confident in ω as you are in ω^*,

$$\omega \; \approx \; \omega^*$$

CIP is incompatible with CCD whenever there is a pair of state descriptions $\phi \wedge Ch = ch$ and $\psi \wedge Ch = ch$ such that $ch(\phi) > ch(\psi)$. For CCD tells us that

(4) $$[\phi \mid Ch = ch] \; > \; [\psi \mid Ch = ch]$$

Given our assumptions, it follows from (4) that

$$\phi \wedge Ch = ch \; > \; \psi \wedge Ch = ch$$

But, by CIP, we have that

$$\phi \wedge Ch = ch \; \approx \; \psi \wedge Ch = ch$$

Contradiction.

Eva (2019) proposes another way of showing indifference. His suggestion is that, in the absence of evidence, you shouldn't have *any* comparative judgements about one state description being more or less likely than another. That is: rather than saying that any two states descriptions are equally likely, you should say that any two state descriptions are *incomparable*. Abbreviate '$(A \not\succcurlyeq B) \wedge (B \not\succcurlyeq A)$' with '$A \odot B$'. Then, Eva's proposal is this:

Comparative Indifference Principle′ (CIP′) For any two state descriptions, ω and ω^*, in the absence of evidence, you should not make any comparative confidence judgements about ω and ω^*,

$$\omega \odot \omega^*$$

CIP′ is also incompatible with CCD, and for the same reason that CIP is. Again, consider a pair a state descriptions, $\phi \wedge Ch = ch$ and $\psi \wedge Ch = ch$, where $ch(\phi) > ch(\psi)$. Again, CCD implies that

$$\phi \wedge Ch = ch \;>\; \psi \wedge Ch = ch$$

whereas CIP′ implies that

$$\phi \wedge Ch = ch \;\odot\; \psi \wedge Ch = ch$$

Contradiction.

A.2 Imprecision

In this section, I'll introduce imprecise credence functions. Representing rational doxastic states with imprecise credences allows us to formulate indifference principles which avoid the familiar objections to IP$^\infty$ from §B.1.1. But imprecise credences will nonetheless give rise to an analogue of the incompatibility from §4.

An imprecise credence function, $\mathbb{C}$, is just a set of precise credence functions, with the interpretation that your doxastic state has all and only the features shared by every credence function in $\mathbb{C}$. A helpful metaphor: each credence function $C \in \mathbb{C}$ is a 'committee member' who gets a vote in determining your doxastic state. Your doxastic state has a property iff the committee passes a motion saying that it has the property. Each committee member $C \in \mathbb{C}$ votes in favour of a motion saying your doxastic state has a certain property exactly when C has that property. And the committee only passes a motion when the vote is unanimous. For instance, suppose that, for every real number x between 1/3 and 2/3, there is a committee member whose credence in P is x. Then, the committee unanimously agrees that your confidence in P is between 1/3 and 2/3, though, when it comes to your confidence in P, it does not agree on anything stronger than this.[14]

In this framework, it's natural to impose a principle of chance deference by demanding that *every* committee member defers to the chances.

Imprecise Chance Deference (ICD) For every sentence P, and every committee member $C \in \mathbb{C}$, C's credence in P, conditional on the chance function being

---

14. For more comprehensive and thorough introductions to imprecise credences, see van Fraassen (1990, 2006), Walley (1991), Seidenfeld & Wasserman (1993), Joyce (2010), Schoenfield (2017), and Moss (2020), a.o.

ch, is $ch(P)$.

$$C(P \mid Ch = ch) = ch(P)$$

Corresponding to the set $\mathbb{C}$ is a set-valued function which we can write '$\mathbb{C}(P)$', and which is defined to be the set of all real numbers x such that, for some $C \in \mathbb{C}$, $C(P) = x$. Likewise, we can let $\mathbb{C}(P \mid Q)$ be the set of real numbers x such that, for some $C \in \mathbb{C}$, $C(P \mid Q) = x$. In these terms, the principle ICD requires that $\mathbb{C}(P \mid Ch = ch) = \{ch(P)\}$.

Within this framework, your attitudes are maximally undecided exactly when they are maximally imprecise. That is: what it is for you to assume nothing at all about whether P is for your committee members to agree on nothing at all about your attitude towards P, other than that it lies somewhere between 0 and 1. That is: what it is for you to be maximally undecided about P is for $\mathbb{C}(P)$ to be the unit interval.

In general, indifference principles say that, in the absence of evidence, your doxastic state should build in as little information as possible about which state description is true. So, in the imprecise framework, it is natural to formulate an indifference principle by saying that, in the absence of evidence, you should be maximally undecided about every state description.

Imprecise Indifference Principle (IIP) For every state description ω, in the absence of evidence, your credence in ω should be maximally imprecise,

$$\mathbb{C}(\omega) = [0, 1]$$

Assuming there is at least one state description $\phi \wedge Ch = ch$ such that $ch(\phi) = x < 1$, ICD and IIP are incompatible. For ICD requires *every* committee member $C \in \mathbb{C}$ to give a credence of x to ϕ, conditional on $Ch = ch$. This means that the greatest credence *any* committee member could give to $\phi \wedge Ch = ch$ is x—any greater, and its credence in ϕ, conditional on $Ch = ch$, would be greater than x, in violation of ICD.[15] So, if you satisfy ICD, you'll have

$$\mathbb{C}(\phi \wedge Ch = ch) \subseteq [0, x]$$

15. More carefully: $C(\phi \mid Ch = ch)$ is the ratio $C(\phi \wedge Ch = ch)/C(Ch = ch)$. Since ICD says this ratio must be equal to x, we have that $C(\phi \wedge Ch = ch) = x \cdot C(Ch = ch)$. By setting $C(Ch = ch)$ equal to 1, we may set $C(\phi \wedge Ch = ch)$ equal to x, but if $C(Ch = ch)$ is any lower than 1, $C(\phi \wedge Ch = ch)$ will be less than x.

However, IIP requires that

$$\mathbb{C}(\phi \wedge Ch = ch) = [0, 1]$$

Since $x < 1$, there's no way to satisfy both of these principles at once.

Joyce (2010, p. 289–290) proposes another, weaker, way of understanding 'indifference' in the imprecise framework. He puts forward an imprecise analogue of the principle I called 'Symmetry to Indifference' (STI) in §3. According to this principle, given any partition[16] of evidentially symmetric sentences, your attitude towards these sentences should be symmetric, in the sense that any committee member who deviates from a uniform distribution is 'balanced out' by committee members who deviate from the uniform distribution to the same degrees, but in different ways. More carefully:

Imprecise Symmetry to Indifference If $\mathscr{E} = \{E_1, E_2, \ldots, E_N\}$ is a partition such that, for every $E_i, E_j \in \mathscr{E}$, E_i and E_j are evidentially symmetric, then for any $C \in \mathbb{C}$, and any permutation p of $\mathscr{E}$, there is some $C^* \in \mathbb{C}$ such that, for each $E \in \mathscr{E}$, $C^*(p(E)) = C(E)$.

If we combine this principle with the assumption that any two state descriptions are evidentially symmetric (SDS), we get the following.

Imprecise Indifference Principle$'$ (IIP$'$) For any $C \in \mathbb{C}$, and any permutation of state descriptions, p, there is some $C^* \in \mathbb{C}$ such that, for every state description ω, $C^*(p(\omega)) = C(\omega)$.

Assume that there is a pair of state descriptions, $\omega_\phi \overset{\text{def}}{=} \phi \wedge Ch = ch$ and $\omega_\psi \overset{\text{def}}{=} \psi \wedge Ch = ch$, such that $ch(\phi) = z \cdot ch(\psi)$, for some $z \neq 1$. Then, IIP$'$ will be incompatible with ICD. Without loss of generality, suppose that $z > 1$, so that $ch(\phi) > ch(\psi)$. By ICD, for every $C \in \mathbb{C}$, $C(\omega_\phi) = z \cdot C(\omega_\psi)$. So *every* committee member gives a higher credence to ω_ϕ than they do to ω_ψ.

(5) $$\forall C \in \mathbb{C} \qquad C(\omega_\phi) > C(\omega_\psi)$$

Now, consider a permutation p which swaps ω_ϕ with ω_ψ, but maps every other state description to itself. Since there's some $C \in \mathbb{C}$ such that $C(\omega_\phi) > C(\omega_\psi)$

---

16. For our purposes, we can take a partition to be a set of sentences such that no sentence in the set is knowable *a priori* to be false, and such that it is knowable *a priori* that the set contains exactly one truth.

(by 5), IIP′ requires that there's another $C^* \in \mathbb{C}$ such that $C^*(\omega_\psi) = C^*(p(\omega_\phi)) = C(\omega_\phi)$ and $C^*(\omega_\phi) = C^*(p(\omega_\psi)) = C(\omega_\psi)$. So IIP′ requires that

(6) $\qquad\qquad\qquad \exists C^* \in \mathbb{C} \qquad C^*(\omega_\psi) > C^*(\omega_\phi)$

But (5) and (6) contradict each other.

B Infinite Indifference to Anti-Humean Chances

In this appendix, I will discuss whether the incompatibility between deference to the chances, indifference, and anti-Humeanism extends to contexts in which there are infinitely many state descriptions (or possible worlds). In §B.1, I introduce an infinitary analogue of IP, IP^∞. In §B.2, I introduce an infinitary analogue of CD, CD^∞. Then, in §B.3, I introduce an infinitary analogue of AH, AH^∞, and I argue that accepting IP^∞, CD^∞, and AH^∞ leads to a contradiction.

B.1 Infinite Indifference

If there are countably many atomic sentences, then a state description will be an infinitary conjunction, and there will be continuum-many state descriptions. If there are uncountably many state descriptions, then any non-trivial indifference principle will require us to impose additional structure on the set Ω. We find some random variable, V, which maps every state description $\omega \in \Omega$ to some real number, $V(\omega) \in \mathbb{R}$. Then, we can assign to each value v in the range of the variable V a *credence density*, $\rho_V(v)$. This density function *doesn't* say what your credence that $V = v$ is.[17] If you abide by IP, your credence that V takes on any particular value, v, will have to be zero. Instead, $\rho_V(v)$ says how *dense* your credence is at $V = v$. Think about it like this: for any narrow interval $[v, v + \epsilon]$, the ratio $C(V \in [v, v + \epsilon])/\epsilon$ is the density of your credence over the interval $[v, v + \epsilon]$. By taking the limit of this ratio as ϵ goes to zero, we get the density of your credence at the *point* $V = v$, $\rho_V(v)$.

With a credence density function, ρ_V, we can determine your credence distribution by integrating over ρ_V. For instance, your credence that V is between a and b will be given by $\int_a^b \rho_V(v)\, dv$. And, in general, for any measurable set of values $\mathbf{v}$, your credence that V is within $\mathbf{v}$ is given by $\int_{\mathbf{v}} \rho_V(v)\, dv$.[18] Then, indifference

17. Notation: '$V = v$' is the disjunction of state descriptions which V maps to v, $V = v \overset{\text{def}}{=} \bigvee_{\omega \in \Omega \,:\, V(\omega) = v} \omega$.

18. In general, we could characterise the possibilities in Ω with any finite number of real-valued vari-
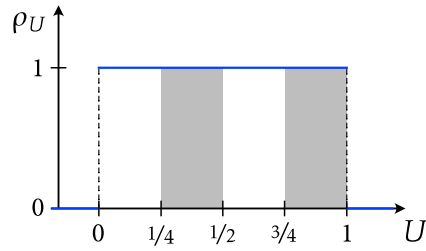
FIGURE 1: The uniform credence density over U. Your credence that U lies in the set $\mathbf{u} = [1/4, 1/2] \cup [3/4, 1]$ is given by the integral $\int_{\mathbf{u}} \rho_U(u)\, \mathrm{d}u$, which is the area under the curve $\rho_U(u)$ shown in grey.

may be implemented by saying that your credences should have a *uniform density*. That is: every value of v should have exactly the same credence *density*.

Infinitary Indifference Principle (IP$^\infty$) Your credence *density* should be uniform.

For instance: consider a random variable, U, which tells us what percentage of space is *u*noccupied. U can take on values between 0 and 1. Then, indifference requires that the density of your credence should be uniform over these values. This uniform credence density is shown in figure 1.

B.1.1. Familiar Objections to IP $^\infty$. In this subsection, I'll briefly review some familiar objections to IP$^\infty$ which are orthogonal to my interests here. The uninterested reader should skip ahead to §B.2.

If Ω is infinite, then there will inevitably be more than one way of parametrising the state descriptions in Ω. For instance, consider the variable R, which gives the *ratio* of unoccupied space to occupied space. Like U, R maps each state description to some real number.[19] However, unlike U, R raises two pressing issues for IP$^\infty$. The first issue is that, unlike U, the potential values of R are *unbounded*—R could take on any value from 0 to ∞. So, if we demand that the density of your credence is *uniform* over R, then we will run into a conflict with normalisation. For the uniform density over R will either be positive for each r or else it will be zero for each r. If positive, $\rho_R(r) = \alpha > 0$, then your credence in Ω will be $\int_0^\infty \alpha = \infty$. If zero, then your credence in Ω will be $\int_0^\infty 0 = 0$. Either way, you will violate normalisation. In response, defenders of IP$^\infty$ could allow that, even if a perfectly

---

ables, $V_1, V_2, \ldots, V_N$. Then, instead of having a density function on $\mathbb{R}$, we'd have a density function on $\mathbb{R}^N$. However, we won't require these additional complications here.

19. More carefully, it maps each possibility to an *extended* real number—if no space is occupied, then we will stipulate that $R = \infty$.
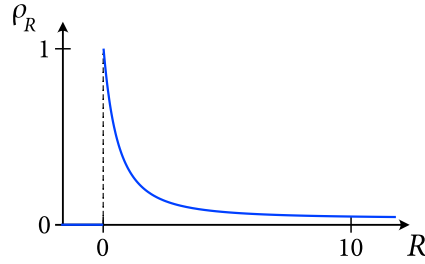
FIGURE 2: A uniform credence density over U induces a non-uniform credence density over $R = U/(1 - U)$.

uniform credence density is impossible, the density of your credences should still be *sufficiently* uniform.[20]

The second issue: once it has pronounced on your credence density over U, IP$^\infty$ has *already* pronounced on your credence density over R. For there is a logical relationship between the values of U and R: necessarily, $R = U/(1 - U)$. But this means that a uniform credence density over U induces the following credence density over R: for each $r \geqslant 0$, $\rho_R(r) = (1 + r)^{-2}$ (if $r < 0$, then $\rho_R(r) = 0$). This credence density is shown in figure 2. The second problem is just that *this* density is *far* from uniform. With this density function, your credence that R is between 0 and n is given by $n/(n + 1)$. So, you will be 90% confident that R is between 0 and 9 and 99% confident that R is between 0 and 99.[21]

Either IP$^\infty$ should be applied to *multiple* parametrisations, or else there is one privileged parametrisation to which it should be applied. In the first case, IP$^\infty$ is outright inconsistent. In the second case, the principle is either language-dependent or arbitrary. Arbitrariness and language-dependence are better than contradiction, so I'll suppose that defenders of IP$^\infty$ think that there is some privileged parametrisation,[22] or that the requirements of rationality are language-dependent.[23]

B.2 Infinite Chances

Suppose you want your credences to be defined over uncountably many sentences of the form $\mathcal{C}h(P) = x$—one for each of the uncountably many real numbers, x, between 0 and 1. Then, so long as your credences are real valued, you'll have to assign a credence of zero to uncountably many of the sentences $\mathcal{C}h(P) = x$. If your

20. Cf. Williamson (2010).
21. Versions of this problem appear in Bertrand (1889). For more recent philosophical discussion, see van Fraassen (1989).
22. This is the route taken by White (2009)—though White does not endorse IP$^\infty$ (see §5 above).
23. This is the route taken by Williamson (2010).

credence in $Ch(P) = x$ is zero, then the product rule will not impose any constraint on the relationship between $C(P \mid Ch(P) = x)$ and $C(P \wedge Ch(P) = x)$. Lewis was not concerned with this, because he allowed rational credences to take on infinitesimal values.[24] So he thought that, even when you're spreading your credences over uncountably many state descriptions, you needn't give a credence of zero to any of them. If we agree with him about this, then perhaps CD is already general enough. But I've been persuaded that Lewis was wrong to rely upon infinitesimals.[25] If, like me, you want your credences to be real-valued, then you should be looking for a natural generalisation of CD for the case where you have credences over uncountably many chance sentences.

Even if your credence that the chance of P is x will be zero for any particular choice of x, your credence that the chance of P lies within an *interval* of values $[x, x + \epsilon]$ (with $\epsilon > 0$) can be non-zero, no matter how small the interval $[x, x + \epsilon]$. So a natural generalisation of CD says that a rational ur-prior credence in P, given that the chance of P lies in some interval $[x, x + \epsilon]$, is within the interval $[x, x + \epsilon]$:

Infinitary Chance Deference (CD$^\infty$) Your credence that P, given that the chance of P is between x and $x + \epsilon$, should be between x and $x + \epsilon$ (for any $\epsilon > 0$).

$$x \leqslant C(P \mid Ch(P) \in [x, x + \epsilon]) \leqslant x + \epsilon$$

If your credence in P, given $Ch(P) \in [x, x+\epsilon]$, is in the interval $[x, x+\epsilon]$, then your credence in $\neg P$, given $Ch(P) \in [x, x + \epsilon]$, is within the interval $[1 - x - \epsilon, 1 - x]$:

$$1 - x - \epsilon \leqslant C(\neg P \mid Ch(P) \in [x, x + \epsilon]) \leqslant 1 - x$$

So long as $C(Ch(P) \in [x, x + \epsilon]) > 0$, for any positive ϵ, no matter how small, it then follows from the product rule that, for any $\epsilon > 0$,

$$C(P \wedge Ch(P) \in [x, x + \epsilon]) \leqslant \frac{x + \epsilon}{1 - x - \epsilon} \cdot C(\neg P \wedge Ch(P) \in [x, x + \epsilon])$$
$$\text{and } C(P \wedge Ch(P) \in [x, x + \epsilon]) \geqslant \frac{x}{1 - x} \cdot C(\neg P \wedge Ch(P) \in [x, x + \epsilon])$$

Divide both sides of these inequalities by ϵ, and take the limit as ϵ goes to zero. Thereby, we get that the *density* of your credence in the conjunction $P \wedge Ch(P) = x$ must be $x/(1 - x)$ times the *density* of your credence in the conjunction $\neg P \wedge$

---

24. See Lewis (1980, pp. 267–8).
25. See Williamson (2007), Easwaran (2014), and Hájek (ms, §7).

$$\mathcal{Ch}(P) = x,$$

$$(7) \qquad \rho(P \wedge \mathcal{Ch}(P) = x) = \frac{x}{1-x} \cdot \rho(\neg P \wedge \mathcal{Ch}(P) = x)$$

Equation 7 follows from CD^∞. It will be important in §B.3 below.

B.3   Incompatibility

To keep matters simple, let's suppose that there is just a single atom, A. Then, we may have, for each $x \in [0,1]$, a chance hypothesis $Ch = ch_x$, where ch_x is a probability function defined over the sentences we get by taking the set $\{A\}$ and closing it under negation and disjunction. Every such sentence will be equivalent to one of the following four: 1) $A \wedge \neg A$, 2) $A \vee \neg A$, 3) $\neg A$, and 4) A. Since chance is a probability function, we must have $ch_x(A \wedge \neg A) = 0$, $ch_x(A \vee \neg A) = 1$, and $ch_x(\neg A) = 1 - ch_x(A)$. So we may characterise each potential chance function ch_x with a single parameter, x, which is the probability ch_x assigns to the atom A.

In this context, I will take 'anti-Humeanism' to be the following thesis:

Infinitary Anti-Humeanism (AH^∞) For each $x \in [0,1]$, there are two corresponding state descriptions: $A \wedge Ch = ch_x$ and $\neg A \wedge Ch = ch_x$.

In that case, there are uncountably many state descriptions in Ω. To apply IP, then, we must first parametrise these state descriptions by using an appropriate random variable from Ω to $\mathbb{R}$. We can encode the information of which chance hypothesis is true with a variable Ch_A, which maps a state description $\omega \in \Omega$ to x iff the chance hypothesis $Ch = ch_x$ is included in ω. But this variable on its own doesn't tell us everything. Besides the chance of A, we also need to know whether A is true or false. I will encode this information with a variable 2_A, which maps a state description $\omega \in \Omega$ to the value 2 iff A is included in ω, and maps ω to 0 if $\neg A$ is included in ω. We can then put these two pieces of information together with a variable $V = Ch_A + 2_A$. V tells us everything there is to tell about both the chance of A is and whether A is true or false. If V is between 0 and 1, then A is false and the chance of A is the value of V. If V is between 2 and 3, then A is true and the chance of A is $V - 2$.

What IP^∞ says will depend upon how we parametrise the state descriptions. (See the discussion from §B.1.1.) The parametrisation I've chosen here in terms of V is meant to be as natural as possible. It cleanly gives us exactly the information of whether A is true and what A's chance is, and the uniform distribution over that chance corresponds to the standard Lebesgue measure. Applying IP^∞ to this very
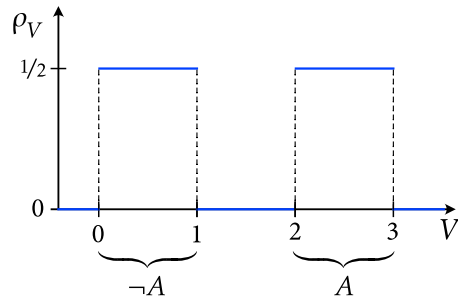
FIGURE 3: The uniform density over $V = Ch_A + 2_A$

natural parametrisation, it tells you to have the uniform credence density shown in figure 3.

But this is incompatible with CD^∞. For CD^∞ requires that, for any v between 0 and 1,

(8) $$\rho_V(v+2) = \frac{v}{1-v} \cdot \rho_V(v)$$

(Equation 8 follows from from equation 7, which itself follows from CD^∞, as we saw in §B.2.) But the uniform credence density shown in figure 3 sets $\rho_V(v+2) = \rho_V(v) = 1/2$ for every value of v between 0 and 1. So the uniform credence density will violate equation 8 for every value of v other than $v = 1/2$. So the uniform credence density violates CD^∞.

References

Bertrand, Joseph. 1889. *Calcul des Probabilités*. Paris: Gauthier-Villars. [viii]

Dorst, Kevin, Levinstein, Benjamin A., Salow, Bernhard, Husic, Brooke E., & Fitelson, Branden. 2021. "Deference Done Better." In *Philosophical Perspectives*, **35** (1): 99–150. [5]

Easwaran, Kenny. 2014. "Regularity and Hyperreal Credences." In *The Philosophical Review*, **123** (1): 1–41. [ix]

Eva, Benjamin. 2019. "Principles of Indifference." In *The Journal of Philosophy*, **116** (7): 390–411. [ii]

Fine, Terrence L. 1973. *Theories of Probability: An Examination of Foundations*. New York: Academic Press. [i]

Hájek, Alan. ms. "Staying Regular?" Available at `http://hplms.berkeley.edu/HajekStayingRegular.pdf`. [ix]

Hall, Ned. 1994. "Correcting the Guide to Objective Chance." In *Mind*, **103** (412): 505–517. [5]

Ismael, Jenann. 2008. "Raid! Dissolving the Big, Bad Bug." In *Noûs*, **42** (2): 292–307. [5]

Jaynes, E. T. 1957. "Information Theory and Statistical Mechanics." In *Physical Review*, **106** (4): 620–630. [10]

Joyce, James M. 2010. "A Defense of Imprecise Credences in Inference and Decision Making." In *Philosophical Perspectives*, **24** (1): 281–323. [iii], [v]

Keynes, John Maynard. 1921. *A Treatise on Probability*. New York: Harper & Row. [2]

Levinstein, Ben. forthcoming. "Accuracy, Deference, and Chance." In *The Philosophical Review*. [5]

Lewis, David K. 1980. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey, Berkeley: University of California Press, volume II, 263–293. [5], [ix]

Lewis, David K. 1994. "Humean Supervenience Debugged." In *Mind*, **103** (412): 473–490. [5]

Moss, Sarah. 2020. "Global Constraints on Imprecise Credences: Solving Reflection Violations, Belief Inertia, and Other Puzzles." In *Philosophy and Phenomenological Research*. doi:{10.1111/phpr.12703}. [iii]

Norton, John D. 2008. "Ignorance and Indifference." In *Philosophy of Science*, **75** (1): 45–68. [ii]

Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford: Oxford University Press. [2], [9], [10], [11], [12], [13]

Schoenfield, Miriam. 2017. "The Accuracy and Rationality of Imprecise Credences." In *Noûs*, **51** (4): 667–685. [iii]

Seidenfeld, Teddy & Wasserman, Larry. 1993. "Dilation for sets of probabilities." In *Annals of Statistics*, **21**: 1139–1154. [iii]

van Fraassen, Bas C. 1989. *Laws and Symmetry*. Oxford: Oxford University Press. [viii]

van Fraassen, Bas C. 1990. "Figures in a Probability Landscape." In *Truth or Consequences: Essays in Honor of Nuel Belnap*, edited by J. Michael Dunn & Anil Gupta, Dordrecht: Kluwer Academic Publishers, 345–356. [iii]

van Fraassen, Bas C. 2006. "Vague Expectation Value Loss." In *Philosophical Studies*, **127**: 483–491. [iii]

Walley, Peter. 1991. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman & Hall. [iii]

White, Roger. 2009. "Evidential Symmetry and Mushy Credence." In *Oxford Studies in Epistemology*, edited by Tamar Szabo Gendler & John Hawthorne, Oxford University Press, 161–186. [2], [11], [viii]

Williamson, Jon. 2010. *In Defense of Objective Bayesianism*. Oxford: Oxford University Press. [2], [9], [10], [12], [13], [viii]

Williamson, Timothy. 2007. "How probable is an infinite sequence of heads?" In *Analysis*, **67** (3): 173–180. [ix]