Turning the Tables on Hume

Casper Storm Hansen

ABSTRACT: Certain prior credence distributions concerning the future lead to inductivism and other distributions lead to inductive skepticism. I argue that it is difficult to see how the latter can be considered reasonable. I do not prove that they are not, but at the end of the paper the tables are turned: in line with pre-philosophical intuitions, inductivism has retaken its position as the most reasonable default position while the skeptic is called on to supply a novel argument for his. The reason is as follows. There are certain possibilities concerning the functioning of the world that, if assigned positive credence, support inductivism. Prima facie, one might think that the alternatives to those possibilities, if assigned similar or more credence, cancel out that support. But, on the contrary, I argue that it is plausible that reasonable credence distributions are such that the alternatives instead cancel themselves out and thus leave the support for inductivism in place.

According to Hume (1739; 1748), learning by induction presupposes a question-begging assumption that nature is uniform. To a first approximation, the conclusion of this paper is that it is exactly the other way around: the Huemean skeptic has to make an unreasonable assumption, namely that nature is not governed by any deterministic laws. And the skeptic has to assume this dogmatically, i.e., he must assign it a credence of 0.

One of the reasons this is merely an approximation to the conclusion is the challenge encapsulated by Goodman's (1955) *grue* example. Many have assumed—implicitly, before Goodman, and mostly explicitly after—that they were entitled to take for granted that, e.g., "green" denotes a real property while, e.g., "grue" does not. I will only make use of a weak version of this assumption. A strong version would be that objects that appear green to us have something *mind-independent* in common that is not matched by a similar commonality among objects that "appear grue," and goes beyond mere labeling. In the context of a discussion that throws most of our empirical knowledge into doubt, that is a question-begging assumption that should *not* be assigned credence 1. A weaker version results from removing the term "mind-independent." *That* assumption is a truism: such things have in common their appearance to us. And while the fact of this commonality is a mind-dependent one, it is nevertheless a fact that can enter into explanatory relationships. For example, it *might* be the case that some objects *appear* similar to us *because* they have something mind-independent in common. Assigning positive credence to this possibility arguably suffices for inductivism to be correct. That is a better approximation to the conclusion.

However, I shall not attempt to *prove* this. I doubt that inductivism can be justified through any rigorous deductive argument alone. Instead, the idea is that inductivism can emerge from a process of reasonable estimation and assessment of what priors one should adopt behind a veil of empirical ignorance. That is what will be attempted below. Such an assessment can be more or less detailed, depending on how many possibilities are considered for how the world "works," and how granular they are. I don't know that it can be made perfectly detailed in less than infinitely many pages and therefore I will not be able to conclude definitively that inductivism is correct. But by the end of the paper, I will have turned the tables: in line with prephilosophical intuitions, inductivism will have retaken its position as the most reasonable default position while the skeptic is called on to supply a novel argument for his.

The form of the core of this paper reflects the methodological stance, as it is written as a dialogue, in which two participants—an inductivist, Ingrid, and an induction skeptic, Stefan—engage in such a process of estimation, pushing the assessment in opposite directions. The dialogue shares a characteristic with Plato's: one of the participants, namely Ingrid, is a mouthpiece for me, and is afforded a disproportionate amount of stage time. When I nevertheless chose to use a fictive character to represent me instead of just debating in my own voice against a single fictive character, it is because it is only Ingrid's position *at the end of* the discussion that coincides with my own. Stefan will make some valid points, which will force Ingrid to modify and sharpen her position. By starting out from a position that does not take all these reasonable objections into account, I can introduce some ideas in a simple and accessible way first, and later make the estimation and assessment on a more sophisticated basis. The first assessments, in sections 1 and 2, are too simplistic because they fail to take into account that deterministic laws can have non-uniform effects; yet, they bear important structural similarities to the final one, in section 4. Section 3 prepares the ground for the final assessment by explaining why the bar for inductivism is low, and the bar for skepticism correspondingly high. An additional reason for my use of the Ingrid character is that I am not entirely as confident in her position as she is; I explain this in the final section.

Don't worry: even though the paper is in dialogue form, I will not waste your time with Ingrid and Stefan exchanging pleasantries, talking about the weather, or behaving realistically.

1 Opening moves

INGRID: Let us say that we observe three Bs in a row, and that the first two are both As. I would of course not claim that we then necessarily *know* that the third B will also be an A, but absent any prior empirical knowledge about Bs, I would certainly say that the probability thereof has gone up, as compared to before we learned of the first two; and that, in that sense, we have learned something about the future. How can you think otherwise?

STEFAN: Well, I might assign prior probability $\frac{1}{8}$ to each of the 2^3 possible sequences of A and non-A ("$\overline{A}$") Bs:

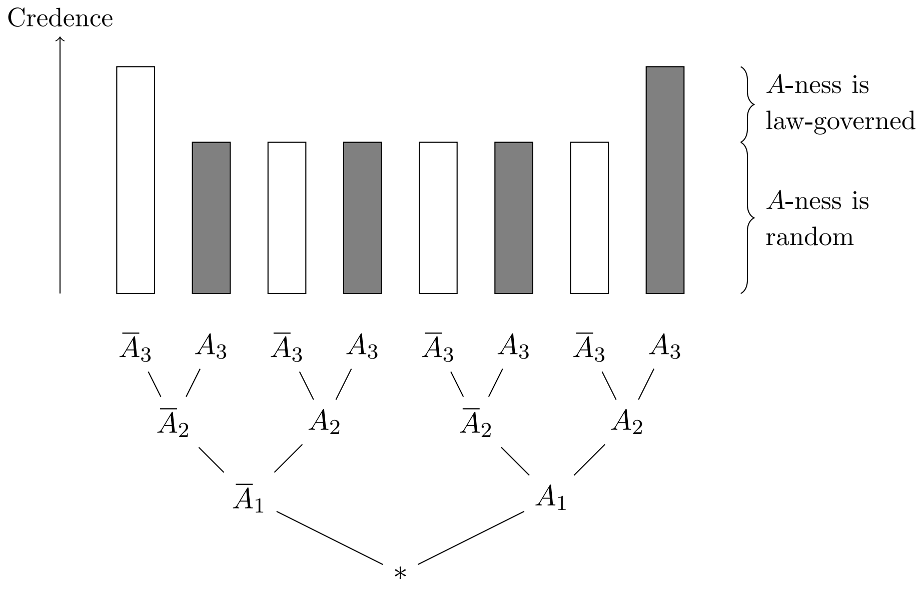
$$\begin{array}{cccc} A_1A_2A_3 & A_1A_2\overline{A}_3 & A_1\overline{A}_2A_3 & A_1\overline{A}_2\overline{A}_3 \\ \overline{A}_1A_2A_3 & \overline{A}_1A_2\overline{A}_3 & \overline{A}_1\overline{A}_2A_3 & \overline{A}_1\overline{A}_2\overline{A}_3 \end{array}$$

Before learning that A_1A_2 (i.e., the first two Bs were A), the probability of A_3 was $4 \cdot \frac{1}{8} = \frac{1}{2}$, because four of those sequences include A_3. But the probability of $A_1A_2A_3$ conditional on the disjunction of $A_1A_2A_3$ and $A_1A_2\overline{A}_3$ is also $\frac{1}{2}$; and that is the probability that the last B is A after it has been learned that the first two Bs were As. So nothing relevant to the probability of A_3 has been learned from the first two.

INGRID: I find fault with that line of reasoning. I know that you do not think that the world is governed by deterministic laws, but you should allow for the possibility that it is—and that one such law governs the properties of Bs. It would be dogmatic of you to completely disregard that option and assign it no probability.

STEFAN: I did no such thing. I included the two uniform sequences that would result if a deterministic law was in effect, namely $A_1A_2A_3$ and $\overline{A}_1\overline{A}_2\overline{A}_3$, and I assigned them a quarter of the probability. My uniform probability distribution is as undogmatic as it could be.

INGRID: You are mistaken. There are two ways that the actual observations could be $A_1A_2A_3$. The first is *because* Bs are subject to a deterministic law that *guarantees* a uniform sequence—what Armstrong (1983; 1991) calls a "strong law," BonJour (1998) calls an "objective regularity," and Foster (1983; 2004) calls an "objective natural necessity." Or, the uniform sequence could be a result of coincidences in the absence of such a law—what Armstrong calls a "mere regularity," BonJour calls a "mere Humean constant conjunction," and Foster calls a "coincidental regularity." Your assignment of a probability of $\frac{1}{8}$ to $A_1A_2A_3$ only takes the possibility of coincidence into account, for that is the probability that results from the assumptions that the probability of A in each observation is $\frac{1}{2}$ and that the observations are stochastically independent. You have, in effect, assigned the possibility of the A-ness of Bs being law-governed a probability of 0, which amounts to a dogmatic assumption that nature is not governed by rules. If you would only allow that there is some probability $\epsilon > 0$ that the observations are governed by a deterministic law (while leaving the remaining probability, $1 - \epsilon$, evenly distributed on the eight possible sequences to reflect the possibility that there is no deterministic law in effect), you would have to agree that the probability of A_3 has increased. That can be seen from this figure:



$P(A_3)$ equals $\frac{1}{2}$ because the fraction of gray column area to the total column area is $\frac{1}{2}$; $P(A_3 \,|\, U_1)$ is higher because the same fraction is higher among the four right-most columns; and $P(A_3 \,|\, U_2)$ is higher yet due to that fraction being even higher among the two right-most columns. So any positive credence assigned to the A-ness of Bs being determined by a deterministic law will skew the distribution in favor of inductivism.

2 Generalization

Stefan: Hume would not agree to assigning positive probability to the existence of a law that is not just a regularity, but instead *responsible for* the regularity. According to him, such a "necessary connection" is not even intelligible, because an idea about it cannot draw its content from impressions alone.

Ingrid: I'm happy to concede that inductive skepticism follows from the rest of Hume's philosophy. Or, to be precise, that it follows from the rest of Hume's philosophy together with the claim that we should assign probability 1 to the world being such that we can conceive of all aspects of it. But is that really the basis for *your* skepticism? Are you committed to Hume's outdated theory of concept formation, and perfectly certain of it?

STEFAN: Well, no. I agree that the possibility of a deterministic law should be assigned positive probability. After all, that the Bs are not subject to a deterministic law is not a mere "relation of ideas," so I do not think that we know it a priori. And even though there are cases where we need to assign probability 0 to propositions we do not know a priori to be false, as pointed out by Howson (2000, 74-75), I agree it would be unreasonable and dogmatic to do so in the case of the proposition that there is *some* deterministic law that governs whether Bs are As.

However, I suspect that the most rational way to assign prior probabilities to sequences would take that into account, and nevertheless end up supporting skepticism.

INGRID: Even in the absence of a specific counter-proposal from you that I can respond to, I think I can prove you wrong, for I have a very general argument.

So far, we have only considered your one specific distribution, and the simple family of distributions with the parameter ϵ that I suggested instead. I am not saying that my alternative is correct; I was only suggesting it as an improvement compared to your suggestion. But I can show in general that a distribution supports my position unless it is a trivial variant of your original suggestion. That is, you can replace the number $\frac{1}{2}$ that you choose for the probability of A with some other number between 0 and 1, and inductive skepticism will be the result. But any such choice will be subject to the same criticism. And any alternative that is *not* such a trivial variant will support my case.

To argue for this, I will make use of the framework provided by Huemer (2009), who in turn builds on classical work by Bayes (1763) and Laplace (1814). Let us continue to consider a sequence of Bs, and let "A_i" denote the proposition that B number i is an A. But let us, like Huemer, allow any number of repetitions. Huemer also uses "U_i" to denote the proposition that all of the first i Bs are As. He defines inductivism as the position that our priors ought to be such that $P(A_{i+1} \,|\, U_i) > P(A_{i+1})$, while "inductive skepticism" similarly corresponds to $P(A_{i+1} \,|\, U_i) = P(A_{i+1})$, and "counter-inductivism" to $P(A_{i+1} \,|\, U_i) < P(A_{i+1})$.

Huemer argues for inductivism. He introduces a constant, C, for the objective chance of a B being A and he denotes by "ρ" the probability density function for the subjective prior credences about the value of C held by the person whose general credence function is P.[1] He then argues from

[1] He thus accepts a Lewisian (1980) framework of interaction between subjective and objective types of probability, in opposition to, e.g, de Finetti (1937), who denied that there is such a thing as objective probabilities in addition to subjective ones.

the Principle of Indifference together with a so-called Explanatory Priority Proviso that ρ ought to be uniform on $[0, 1]$. He shows that this implies $P(A_{i+1} \,|\, U_i) > P(A_{i+1})$, i.e., inductivism.

The appeal to these controversial principles can be avoided. Inductivism follows from a much weaker assumption, namely that one ought not be a dogmatist about C. By this I simply mean that one ought not to assign *all* of one's credence to a *single* possible value of C. Doing so implies that one will continue to assign all one's credence to that value no matter on which evidence one later conditionalizes. If one is a dogmatist then one is a skeptic in Huemer's sense because both $P(A_{i+1} \,|\, U_i)$ and $P(A_{i+1})$ equal the dogmatically assumed value of C. In *any* other case, one is an inductivist.

This is a very intuitive consequence of Bayes' Theorem. Consider what happens if you first adopt non-dogmatic priors about C, and then observe a sequence of As. Each A is more likely given higher objective probabilities for As; so because of Bayes' Theorem, your credences will shift away from hypotheses about lower objective probabilities and towards hypotheses about higher objective probabilities. This will happen each time you observe an A. Hence $P(A_{i+1} \,|\, U_i) > P(A_{i+1})$.

I think that leaves you in an untenable position, Stefan: to maintain your skepticism, you must be a dogmatist, and we normally think of "dogmatism" as a position that occupies the opposite extreme on the doxastic scale from skepticism. Of course, you wouldn't exactly be inconsistent, for the dogmatism is in relation to the value of C, while the skepticism is in relation to the next observation; but there is certainly a strong tension. You must be convinced that we have no *ability* to learn anything pertaining to future observations, and combine that with the conviction that there is no *need* to learn anything about the value of C because it is known from the outset. In other words, you must be convinced that we have *no* first-order epistemic powers, but *perfect* second-order ones.

3 Skepticism in the space of probability distributions

Stefan: There is a hidden assumption behind both Huemer's argument and your generalization: namely, that the objective probability C for a B being an A is constant throughout the various observations. That is a premise about regularity in nature, so, as Smithson (2017) points out, it is a framework assumption that is biased in favor of your inductivism.

I realize that our specific suggestions for probability distributions, which we discussed earlier, were both consistent with that assumption. However, I would not justify my suggestion using that assumption. Following Smithson, I think it should be abandoned.

Ingrid: The rejection of that assumption does not seem to me to make much of a difference. It allows you to make independent choices about how you will be dogmatic concerning each i, but to my mind, that still counts as dogmatic. You might go with, say, $P(A_1) = \frac{1}{2}$, $P(A_2) = \frac{1}{3}$, and $P(A_3) = \frac{2}{3}$, instead of $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$. But you still have to assume stochastic independence; so I would complain that you are only covering the option that the series of observations is the result of fluctuating, but known, chances. I would then request that some positive probability be assigned to $A_1A_2A_3$ or $\overline{A}_1\overline{A}_2\overline{A}_3$ being the outcome *because* of a constant (deterministic or probabilistic) law. You conceded something similar earlier.

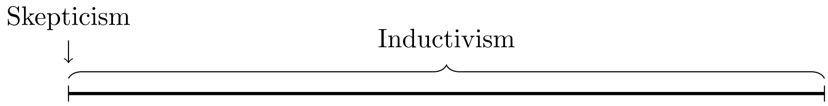
STEFAN: I would not adopt that particular distribution of priors. My point is quite different. Let me approach it indirectly by first reaching back to something from your summary of Huemer's paper. To attain a comprehensive categorization of possible positions, Huemer needed more positions than the two you and I represent. The extra option is that of counter-inductivism. As you mentioned, it is defined by $P(A_{i+1} \,|\, U_i) < P(A_{i+1})$. This option is off the table under the assumption of constant objective probability; but without such an assumption, it needs to be considered.

So let us consider another probability distribution for the sequence of three trials, which fails to satisfy Huemer's assumption. For simplicity, we can stick with a shared assumption, namely that $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$, i.e., that our credence, for each of the three trials, of an A outcome is $\frac{1}{2}$ before any empirical evidence is gathered. However, our imagined counter-inductivist will interpret A_1A_2 as definitive evidence for $\overline{A}_3$, i.e., he assigns $P(A_1A_2A_3) = 0$.

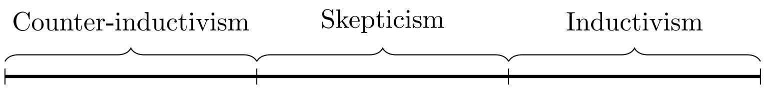
Obviously, I do not endorse such a distribution. But I consider it no less reasonable than that of an (also imagined) inductivist who interprets A_1A_2 as definitive evidence for A_3 and therefore has $P(A_1A_2A_3) = P(A_1A_2)$. For every inductivist position regarding, say, $i = 3$ (note that Huemer's categorization of positions is really relative to the value of i), there is a mirror-image counter-inductivist position: if P_I is the inductivist distribution, then just define the counter-inductivist distribution P_C by $P_C(A_1A_2A_3) = P_I(A_1A_2\overline{A}_3)$, $P_C(A_1A_2\overline{A}_3) = P_I(A_1A_2A_3)$, and P_C otherwise as having the same value as P_I.

That changes the picture quite radically. Using Huemer's assumption, you made my position look like an extreme limit case in the space of possible probability distributions, and you called me dogmatic. But actually, I occupy the moderate center ground, with dogmatic inductivists on one side and dogmatic counter-inductivists on the other. So I believe that the distribution I first suggested is correct, not because I am dogmatic about one specific objective chance, but because it is the average of all the dogmatic positions.

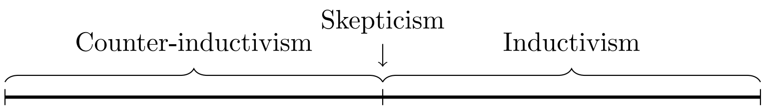
INGRID: That is a more challenging objection. I agree that without the assumption of a constant objective chance, it is much more accurate to picture skeptics like you as placed in the middle instead of off to the side. However, it is also part of the picture that your middle position is extremely narrow. Let me actually draw some pictures. We agree that this is wrong:



But so is this:



As long as we stick to simple pictures, this is the most accurate:



I think that the narrowness of the position implies that it is very likely wrong. But I will first concentrate on justifying the accuracy of the picture. I claim that as a skeptic you are left with a very narrow space of probability functions. In the picture it is shown as a point on a line segment, that is, as a subspace of one dimension less than the full space. But actually, the subspace of skeptical distributions is even "smaller" relative to the space of all distributions than that, when the sequence of observations is longer than 2.

I need to set the stage before I can explain that point. The outcome space we have implicitly been discussing is $\{A, \overline{A}\}^n$ for some natural number n. The full space of probability distributions on this outcome space is the set of all mappings P from $\{A, \overline{A}\}^n$ to $[0, 1]$ such that $\sum_{x \in \{A, \overline{A}\}^n} P(x) = 1$.

Every ρ function in the framework I used earlier induces such a P function. For instance, the dogmatic function that assigns all the credence to the possibility that $C = \frac{1}{2}$ induces the constant P function that maps every element of $\{A, \overline{A}\}^n$ to 2^{-n}. But P functions that are *not* induced by a ρ function also exist: for example, the function that maps the alternating sequence $A\overline{A}A\overline{A}\ldots$ of length n to 1 and every other sequence to 0.

Your thesis that we cannot learn about the future from the past can now be expressed as the thesis that for all $i \in \{1, \ldots, n-1\}$ and all sequences $X_1 X_2 \ldots X_i$, it is the case that $P(A_{i+1} \mid X_1 X_2 \ldots X_i) = P(A_{i+1})$. Let me compare the number of "degrees of freedom" that this leaves you against the number of degrees of freedom in general. In general, there is one degree of freedom consisting in the choice of the value of $P(A_1)$, and that is the same for you. And in general there are two further degrees of freedom corresponding to the values of $P(A_2 \mid A_1)$ and $P(A_2 \mid \overline{A}_1)$. But you are only afforded one degree of freedom for $i = 2$, consisting in the choice of $P(A_2)$, for you are under the constraint that $P(A_2 \mid A_1) = P(A_2 \mid \overline{A}_1) = P(A_2)$. For $i = 3$ there are in general four degrees of freedom, as the values of $P(A_3 \mid A_1 A_2)$, $P(A_3 \mid A_1 \overline{A}_2)$, $P(A_3 \mid \overline{A}_1 A_2)$, and $P(A_3 \mid \overline{A}_1 \overline{A}_2)$ can be chosen independently; whereas you only have one degree of freedom at this stage as well. And so on. The result is that while there are $2^n - 1$ degrees of freedom in general, you are limited to n.

Because you are occupying a tiny n-dimensional subspace of a much larger $(2^n - 1)$-dimensional probability space, it is extremely unlikely for a typical series of observations that the most rational distribution is skeptical. For instance, if we assume that there is a case in which it would be rational to assign, say, probability .8 to there being a constant objective chance, distributed in some particular way by an inductivist probability function P_1; probability .1 to something that is itself distributed in some particular way by a counter-inductivist probability function P_2; and probability .1 to neither of those being the case, distributed in some particular way by a probability function P_3. Since the skeptical subspace is so small, it is extremely unlikely that $.8P_1 + .1P_2 + .1P_3$ is a skeptical distribution.

Counter-inductivists have something important in common with inductivists: they think that we can use evidence about the past to learn about the future. So the position that we cannot learn about the future depends on such a precarious balance that even the smallest asymmetry will destroy it.

STEFAN: I will have to grant you that skepticism is a narrow position. But you cannot prove that the most rational probability distribution does not, nevertheless, belong there. That is one of the main points made by Howson (2000), and it is argued convincingly.

INGRID: No, I cannot *prove* it. But proof should not be required as a prerequisite to granting that it is rational to be an inductivist. We are discussing which probability distribution in a space with a continuum of options is the most rational in the case of typical properties B and A. Whether something is proved, on the other hand, is a binary matter: either it is, or it isn't. There is no reason to consider the continuous epistemic space as subject to a binary epistemic meta-level. I should not be forced to become a skeptic if I only consider it 99.9% likely that inductivism is correct. If you are unsure what the most reasonable distribution of priors is, then you cannot do better than taking a weighted average over distributions of priors according to the most reasonable distribution of probabilities over distributions of priors that you can think of. It may be—to borrow a phrase from Jeffrey (1992, 11)—probabilities all the way down![2]

[2]Atkinson and Peijnenburg (2017) have proved that infinitely descending hierarchies

I don't know if Carnap was right to assume that there is a uniquely rational distribution of priors, but I do think we can reject some dogmatic distributions as irrational, even if that claim cannot be reduced to more fundamental principles through proof.[3]

4 Skewing the probability distribution away from skepticism

STEFAN: Well, I don't even think you can make it *plausible* that the most rational probability distribution is inductivist. You claim that the possible existence of a deterministic law helps you in that regard, but I don't think it does, because it is also possible that such a law is temporally restricted. And I am not impressed with the attempts to argue that "temporally unrestricted" makes for a better explanation for an observed uniform sequence than "temporally restricted" (Foster 2004, Lecture 4).

Let us say that we have observed a string of n Bs that were all As. Let D_p be the proposition that such sequence has been governed by a deterministic law, at least so far. Then, assuming a suitable outcome space that is more fine-grained than the one you specified, we can express our probability distribution $P(\cdot\,|\,U_n)$ as $P(D_p)\cdot P(\cdot\,|\,U_n \wedge D_p) + P(\overline{D}_p)\cdot P(\cdot\,|\,U_n \wedge \overline{D}_p)$. If the possibility of D_p is supposed to save you, you should be willing to accept for the sake of argument that $P(\cdot\,|\,U_n \wedge \overline{D}_p)$ is a skeptical distribution.

INGRID: I am.

STEFAN: Good. Then to continue with our shared, simplifying assumption, both $P(A_{n+1})$ and $P(A_{n+1}\,|\,U_n \wedge \overline{D}_p)$ equal $\frac{1}{2}$. Let D_f be the proposition that the deterministic law will continue in effect. Then $P(A_{n+1}\,|\,U_n \wedge D_p \wedge D_f)$ equals 1, but $P(A_{n+1}\,|\,U_n \wedge D_p \wedge \overline{D}_f)$, i.e., the probability that the next B will be an A, on condition of the deterministic law having been in effect until now and *only* until now, equals 0. If we also assume $P(D_p \wedge D_f) = P(D_p \wedge \overline{D}_f)$, which seems reasonable, then the skeptical result that $P(A_{n+1}\,|\,U_n) = \frac{1}{2}$ follows.

INGRID: No, $P(D_p \wedge D_f) = P(D_p \wedge \overline{D}_f)$ is not reasonable at all. $D_p \wedge \overline{D}_f$ only covers the option that the law ceases to in effect exactly after n A observations, while $D_p \wedge D_f$ covers the option that the law ceases to be in effect after $n+1$ such observations, the option that the law ceases to in effect after $n+2$ such observations, etc.; in addition to the option that the law is temporally unrestricted. Why would it be rational to assign those two the same credence?

Anyway, there is another mistake that is more fruitful to point out. This

of probabilities are—with the exception of a few limit cases—not only well-defined but surprisingly well-behaved.

[3] See Huemer (2017, section 5) for a similar view.

bigger mistake is that you let $P(A_{n+1} \,|\, U_n \wedge D_p \wedge \overline{D}_f)$ be equal to 0. The condition here is that there has been a law that guaranteed Bs were As until now, and that such law now ceases to be in effect. But that possibility includes the possibility that the next B will *happen to be* an A, that is, that the determinism ends but the uniformity continues. Hence, for the purpose of predicting whether the next B will be an A, $D_p \wedge \overline{D}_f$ resembles $\overline{D}_p$, so $P(A_{n+1} \,|\, U_n \wedge D_p \wedge \overline{D}_f) = \frac{1}{2}$ is more reasonable.

Notice how this mistake is similar to the one you made at the beginning of our discussion. You thought that the possibility of a deterministic law had been taken into account by assigning a quarter of the probability to $A_1A_2A_3$ and $\overline{A}_1\overline{A}_2\overline{A}_3$, and you thought that the opposite possibility balanced it out, whereas in fact, the opposite possibility merely balanced *itself* out. That is confirmed now, for we see that if you split the possibility of a deterministic law into the possibilities of a constant, deterministic law and a non-constant deterministic law, the latter at most balances itself out; it does not balance out the possibility of a constant law.

As discussed earlier, the constant law does not even have to be deterministic; it may also be probabilistic, as long as this possibility is taken into account in a way that is not dogmatic concerning the value of the objective probability.

STEFAN: I still think you are wrong, but for a reason I hadn't thought about earlier. Take a seemingly random sequence like $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$. In spite of the "randomness", it is possible that this sequence results *because* there is a constant, deterministic law in effect that guarantees exactly that sequence. We should also take that possibility into account. The same is the case for any other sequence as well. We should assign those possibilities the same probability, if the prior probability of A outcomes is $\frac{1}{2}$. Similar considerations apply to any other reason to assign credence to $A_1A_2A_3A_4A_5A_6A_7A_8A_9$. Thus, all sequences should be treated on a par with one another, and when you do that, you end up with a skeptical distribution.

INGRID: Okay, I do need to amend my position. But the essence of my argument survives, because you are making the same kind of mistake once again, although this version of the mistake is more subtle. The general form of my argument is that there is a possibility that supports induction, and that it is plausible that the alternatives to that possibility at worst balance each other out. Your mistake is to assume that the alternatives instead balance out the "induction possibility."

However, I failed to identify the correct "induction possibility." It is not that there is a constant, deterministic law (or a constant probabilistic law—but let us leave that complication aside for now). Rather, it is that there is a constant, deterministic law that guarantees a uniform sequence *because* Bs *must* be As or Bs *must* be $\overline{A}$s. And some of the alternatives consist in there being a deterministic law that guarantees some sequence, but not *because* Bs *must* be As or Bs *must* be $\overline{A}$s. Let me explain.

I acknowledge that there could be a deterministic law that produces the sequence $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$, and that we should assign it some (very small) positive probability. But if a deterministic law effects those outcomes, the Bs are not all As or all $\overline{A}$s, so *a fortiori* the deterministic law does not effect the outcomes $\overline{A}_1$, $\overline{A}_2$, A_3, $\overline{A}_4$, A_5, A_6, A_7, A_8, and $\overline{A}_9$ because Bs must be As or Bs must be $\overline{A}$s. It is for some other reason (or, perhaps it is a brute fact that the law is like that). In particular, the explanation (if there is one) of why A_3 is guaranteed by the law is not (simply) that it is a B and therefore must be an A. For instance, there could be a deterministic law that produces the sequence $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$ because Bs must be As iff they are Cs, and precisely Bs number 3, 5, 6, 7, and 8 happen to be Cs.

Similarly, there could be a deterministic law that produces the sequence $A_1A_2A_3A_4A_5A_6A_7A_8A_9$, without that being *because* Bs must be As. In particular, the explanation (if there is one) for why A_3 is guaranteed by the law might not simply be that it is a B and therefore must be an A. For instance, there could be a deterministic law that produces the sequence $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ because Bs must be As iff they are Cs, and all of the first nine Bs happen to be Cs.

So if you assign some probability to $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$ happening because of a deterministic law that does not necessitate all Bs to be As, it is also plausible that it is reasonable, absent specific reasons to the contrary, to assign at least as much[4] to $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ happening because of a deterministic law that does not necessitate all Bs to be As; and similarly for the other sequences. That is, those probabilities at worst balance each other out. In addition, I submit that you should assign positive probability to $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ happening because of a deterministic law that *does* necessitate that all Bs are As. That will again break the skeptical balance and result in an inductivist distribution.

And, even if you were to deny that at least the same probability should be assigned to $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ as to $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$ happening because of a deterministic law that does not necessitate all Bs to be As, you would have a difficult time arguing that the delicate skeptical balance can be precisely restored, once it has been broken by the asymmetry.

STEFAN: I'm not sure I got that.

INGRID: Okay, I will try to explain it in a different way. Assume that all Bs observed so far were As, so that [necessarily, Bs are As] remains an epistemic

[4] "At least as much" applies to the simplest case where $P(A_i) = P(\overline{A}_i) = \frac{1}{2}$. In the general case, it should instead be "at least $\frac{P(A_i)^9}{P(A_i)^5 \cdot P(\overline{A}_i)^4}$ times as much."

possibility. For simplicity, assume further that at least one such observation has been made, so that [necessarily, Bs are $\overline{A}$s] is not an epistemic possibility. And hold on to the simplifying assumption that the prior probability of A and $\overline{A}$ observations were both $\frac{1}{2}$.

I claim that the possibility [necessarily, Bs are As] is a symmetry-breaker, because there is a natural partition of the space of epistemic possibilities such that, for any possibility π_1 that implies that the next B will be a $\overline{A}$, there is another possibility π_2 that implies that the next B will be an A, and is a more direct counterpart of π_1 than [necessarily, Bs are As] is; and it is rational to assign a probability to π_2 that is larger than or equal to the probability assigned to π_1.

Now, I don't have a proof of this claim, nor can I give a precise definition of "direct counterpart." But I think some examples will suffice to make the term intelligible, and the claim plausible. Here are three:

- [It is a brute fact that the next B is an A] is the direct counterpart of [it is a brute fact that the next B is a $\overline{A}$].
- [It was a law until now that all Bs are As, and the next B will randomly be an A] is the direct counterpart of [it was a law until now that all Bs are As, and the next B will randomly be a $\overline{A}$].
- [Necessarily, Bs are As iff they are Cs; and the next B is a C] is the direct counterpart of [necessarily, Bs are As iff they are Cs; and the next B is a $\overline{C}$].

That leaves the possibility [necessarily, Bs are As] completely free to skew what might otherwise be a skeptical distribution into an inductivist one, given that we undogmatically assign it positive probability.

STEFAN: I see. But that amounts to giving the $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ sequence special treatment compared to, e.g., $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$, and I don't think that's justified. Under a different way of categorizing outcomes, $\overline{A}_2$ might be similar to A_3, and we might describe the $\overline{A}_1\overline{A}_2A_3\overline{A}_4A_5A_6A_7A_8\overline{A}_9$ sequence as an "$A'_1A'_2A'_3A'_4A'_5A'_6A'_7A'_8A'_9$ sequence." So if we should give $A_1A_2A_3A_4A_5A_6A_7A_8A_9$ special treatment, we should give all sequences special treatment, which amounts to giving no sequences special treatment. And the possibility [necessarily, Bs are As] would be balanced out by [necessarily, Bs are A's].

You have a problem in Goodman's (1955, Chapter III) so-called new riddle of induction. Let me summarize it. As an example, he stipulates that all emeralds examined up until now, time t, have been green. Naively, that would seem to raise the probability that the next emerald to be dug out of a mine will also be green. However, Goodman defines the predicate "grue" to mean "green and first examined before t, or blue and first examined after t," and points out that we might also take the evidence at hand—that all emeralds examined up until now have been grue—as confirmation of the thesis that the next emerald will be grue. But that is the same as concluding that the probability of the next emerald being blue has gone up. If we assume for the sake of argument that we know for certain that all emeralds are either green or blue,[5] then we need to both raise and lower the probability that the next emerald will be green. Instead, I say that we should not change the probability at all.

Goodman points out that there is no simple syntactical criterion to distinguish "good" predicates from "bad" ones. We cannot determine that sequences of green events should be assigned higher probability than sequences of grue events simply on the basis of "grue" being a complex predicate, having been defined from two other color predicates and by reference to temporal properties, while "green" is simple. That is because such a measure of simplicity is relative to which predicates happen to be available in everyday English. Had "grue" and its dual "bleen" been available instead of "green" and "blue," then the latter two could be defined in terms of the former two, and would appear to be complex.

I think it is most fair-minded to be equally open to deterministic laws that can be described in a simple way in English, and to deterministic laws that cannot. Hence, the probability that we assign to there being a deterministic law in effect should, just like the probability that we assign to the outcomes being objectively random, be distributed uniformly across sequences.

Ingrid: Well, of course the "A" label might be arbitrary—it could, for instance, mean "grue." I have assumed it was not arbitrary, i.e., that it designated a genuine property. If it is a genuine property, as *green* is and *grue* isn't, my point stands: there is a possibility that, if only it is afforded a minimum of positive probability, will skew the probability distribution into an inductivist one, because the alternatives balance each other out.

If I am not allowed to take for granted that the labels are non-arbitrary, I just have to formulate the "induction possibility" or "symmetry-breaker" more carefully: there is a constant, deterministic law that guarantees some sequence because Bs must be As, or must be $\overline{A}$s, *simply by virtue of what it is to have the B property and the A property.* And an alternative that balances itself out is that there is a constant, deterministic law that guarantees some sequence, but not *because* Bs must be As, or must be $\overline{A}$s, simply by virtue of what it is to have the B property and the A property. There could be a deterministic law that guarantees a grue sequence of emeralds, but not because emeralds must be grue simply by virtue of what it is to have the properties *emerald* and *grue.* There could also be a deterministic law that

[5] Alternatively, "non-green" could replace "blue" in the definition of "grue."

guarantees a green sequence of emeralds, but not because emeralds must be green simply by virtue of what it is to have the properties *emerald* and *green*. In addition, there could be a deterministic law that guarantees a green sequence of emeralds because emeralds must be green simply by virtue of what it is to have the properties *emerald* and *green*.

In contrast, there couldn't be a deterministic law that guarantees a grue sequence of emeralds because emeralds must be grue simply by virtue of what it is to have the properties *emerald* and *grue*; for *grue* is not a real property. The reason the deterministic law guarantees a grue sequence would have to be more complex. For instance, it might be that emeralds must be green when there is chromium present on Earth and must be blue when there is none, and that chromium is bound to disappear from Earth at time t. And that possibility belongs to the category that is balanced out. In this case it would be balanced out by, among others, the possibility that emeralds must be green when there is chromium present on Earth and must be blue when there is none, and that chromium is bound to never disappear from Earth.

STEFAN: I don't think we are entitled to the assumption that *grue* is not a real property. Remember that, even though you take yourself to have lots of induction-based knowledge, you must for the purpose of this discussion imagine yourself as a rational subject about to receive, but so far innocent of, empirical information.

INGRID: That's true. In that situation, the agent has to start out from "subjective properties" determined by a subjective sense of similarity. Since such an agent might not initially be able to identify emeralds, I will explain what I mean using a simpler example. Say that the agent is thrown into an extremely simple world in which his experiences are only of cubes and spheres that are either green or blue. Let "Object$_0$" denote the first object he observes. Is it rational for him to increase his credence in new objects that appear similar to Object$_0$ in shape also appearing similar to Object$_0$ in color, as he observes more and more of the former that are all the latter? I say that it is. The subjective sense of similarity gives the sets of similar-looking objects a privileged status over grue-like alternative sets, and this suffices to create the slight asymmetry required.

STEFAN: But unless you are suddenly assuming idealism, your previous proposal for an "induction possibility" does not work. Objects do not appear similar to a prototypical green thing *because* they appear similar to a prototypical cube, simply by virtue of what it is to appear similar to those things.

INGRID: Well, strictly speaking, I don't think that could be ruled out in the *tabula rasa* situation considered, so it could serve as *an* induction possibility. But if we set that aside—and I'm happy to do so—you are right that I once again have to amend my formulation, as you have robbed me of further simplifying assumptions. I claim that the following is an induction possibility: Objects that appear similar to Object_0 w.r.t. shape also appear similar to Object_0 w.r.t. color because

- there is an objective property Y and a constant, deterministic law that guarantees that all and only Ys appear similar to Object_0 w.r.t. shape simply by virtue of what it is to have the Y property and the property of appearing similar to Object_0 w.r.t. shape;
- there is an objective property X and a constant, deterministic law that guarantees that all and only Xs appear similar to Object_0 w.r.t. color simply by virtue of what it is to have the X property and the property of appearing similar to Object_0 w.r.t. color; and
- there is a constant, deterministic law that guarantees that Ys are Xs simply by virtue of what it is to have the Y property and the X property.

One among many alternatives to this possibility is that there are two different objective properties, Y_1 and Y_2, that make their bearers appear similar to Object_0 w.r.t. shape; that Y_1 always makes the object appear similar to Object_0 w.r.t. color, while Y_2 always makes the object appear different from Object_0 w.r.t. color; and that it is by mere chance that only Y_1 objects have been observed so far. This alternative is plausibly balanced out by the possibility that is the same, except that Y_2 also always makes the object appear similar to Object_0 w.r.t. color.

We don't have to take *for granted* that there are objective properties corresponding to subjective experiences of similarity. We don't even have to consider it a priori more likely than any alternatives. Again, the mere possibility, together with non-dogmatism about that possibility, can be enough to skew a probability distribution into one that allows for induction.

The regress that you have, to your credit, forced me into, Stefan, comes to an end with the fact that some sequences of experiences *appear* to be uniform, and stand out as privileged in that respect. Appearing-to-be-green is a real (subjective) property that could possibly be due to an objective property (such as emitting light at a specific range of frequencies), while cubic objects appearing-to-be-green-before-t-or-appearing-to-be-blue-after-t could not possibly be explained *simply* by cubic objects all having some real, objective property in common.

5 Epilogue

I am not *sure* that Ingrid's final position—the claim about a natural partition in the middle of section 4 combined with the induction possibility from the end of section 4—is true. But at the very least I think it is a good approximation to an important fact about the priors that an ideally rational person would arrive at if they had infinitely much time to consider every metaphysical possibility in turn, and assign each an a priori probability. I doubt that those priors can be derived from some elegant principle, based on indifference or otherwise. It may be very messy. Hence, having neither ideal rationality nor infinite time at my disposal, I don't know how to become sure that Ingrid is right.

Stefan might cease on this admission that inductivism has still not been proved to claim that we must continue to be skeptics until the epistemic situation has been improved. And much of the philosophical tradition since Hume would be in support of that sentiment. Skepticism has become the default position from which only a rigorous proof will be accepted as sufficient to dislodge us. I, on the other hand, believe that the burden of proof should be equally shared. Hence, if we agree that we must always have *some* credences—we should not suspend judgment in the radical sense of refusing to adopt any—a plausibility argument is sufficient to actually make it rational for those of us with bounded rationality to adopt inductivist credences. And Ingrid has delivered such a plausibility argument. The onus is now on Stefan and his allies to convince us that there are considerations that justify a credence shift back to the delicate skeptical balance.

References

Armstrong, D. M. (1983). *What is a Law of Nature?* Cambridge University Press.

Armstrong, D. M. (1991). What makes induction rational? *Dialogue 30*, 503–512.

Atkinson, D. and J. Peijnenburg (2017). *Fading Foundations: Probability and the Regress Problem*. Springer.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society 53*, 370–418.

BonJour, L. (1998). *In Defense of Pure Reason*. Cambridge University Press.

de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré 17*, 1–68. Translated as "Foresight: Its logical laws, its subjective sources" in H. E. Kyburg and H. E. Smokler (eds), *Studies in Subjective Probability*, Wiley, 1964.

Foster, J. (1983). Induction, explanation and natural necessity. *Proceedings of the Aristotelian Society, New Series 83*, 87–101.

Foster, J. (2004). *The Divine Lawmaker*. Oxford University Press.

Goodman, N. (1955). *Fact, Fiction, and Forecast*. Harvard University Press.

Howson, C. (2000). *Hume's Problem: Induction and the Justification of Belief*. Oxford University Press.

Huemer, M. (2009). Explanationist aid for the theory of inductive logic. *The British Journal for the Philosophy of Science 60*, 345–375.

Huemer, M. (2017). There is no pure empirical reasoning. *Philosophy and Phenomenological Research 95*, 592–613.

Hume, D. (1739). *A Treatise of Human Nature*. Oxford University Press.

Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Oxford University Press.

Jeffrey, R. (1992). *Probability and the Art of Judgement*. Cambridge University Press.

Laplace, P.-S. (1814). *Essai Philosophique sur les Probabilités*. Paris. Translated as "Philosophical Essay on Probabilities", Springer, 1995.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability*, Volume II. University of California Press.

Smithson, R. (2017). The principle of indifference and inductive scepticism. *The British Journal for the Philosophy of Science 68*, 253–272.