Inductive reasoning in simple worlds

Casper Storm Hansen

For you and me, inductive reasoning is complicated. To fully live up to the demands of Bayesianism, we would have to assign prior credences to every conceivable possibility about the future. That's a lot! This complication is an important obstacle when trying to figure out whether it is rational to be an inductivist or a Humean induction skeptic: i.e., whether we ought to increase the credences we assign to future events of a certain kind if we observe earlier events of the same kind, or keep them steady.

However, useful insights about inductive reasoning can be gained by considering simplified models of such reasoning, in the form of agents in worlds with fewer possibilities. That is the working hypothesis of this paper.

To be more precise, the worlds I will describe are *epistemically* simpler, because their inhabitants know a priori (or, if you like, have been informed by a completely trustworthy god) that certain possibilities will not obtain, and are left with a range of epistemic possibilities that lends itself to elegant mathematical treatment. With the aim of improving on Huemer's (2009) contribution to the subject, I will consider a world in which the epistemic agents know that events are governed by constant objective chances. And with the aim of aiding the nomological-explanatory attempt to combat inductive skepticism, I will consider two worlds that are a priori known to be governed by temporally restricted deterministic laws. The paper concludes with a short discussion of the relevance of these simple worlds to our complicated world.

1 Constant objective chances

In World 1, an epistemic agent is suddenly created. At the moment of her creation, she is equipped with ideal rationality, but no empirical knowledge. One of the first things she notices is that some objects in World 1 have a property A and some have a property B. She becomes interested in predicting whether not-yet-observed B things are A things. We stipulate that she knows there is a constant objective chance of B things being A things.

I will follow Huemer (2009) in numbering the Bs according to the order in which they are observed by the agent; in using the notation "A_n" to denote the proposition that the nth B is an A; and in using "U_n" to denote the proposition that all of the first n Bs are As. Thus, U_n implies A_n, but the latter is consistent with an earlier-observed B not being an A. Let P be the agent's credence function at the time of her creation. For a given n, we are interested in whether her credence in A_{n+1} will grow, in between her moment of creation and the instant of time when she observes B number n, if all of the first n Bs are As. Thus, we will also adopt the following definitions of Huemer's: of inductivism as the position that $P(A_{n+1} \,|\, U_n) > P(A_{n+1})$; of inductive skepticism as the position that $P(A_{n+1} \,|\, U_n) = P(A_{n+1})$; and of counter-inductivism as the position that $P(A_{n+1} \,|\, U_n) < P(A_{n+1})$.

Inspired in part by Bayes (1763) and Laplace (1814), Huemer argues that our ideally rational epistemic agent is an inductivist. I will argue for the same conclusion, but show that it can be reached much more easily than it is by Huemer.

The first step of Huemer's argument consists of applying the definition of conditional probability:

$$P(A_{n+1} \,|\, U_n) = \frac{P(A_{n+1} \wedge U_n)}{P(U_n)} = \frac{P(U_{n+1})}{P(U_n)} \tag{1}$$

The second step is to introduce a constant, C, for the objective chance that a B is an A; and a probability density function, ρ, for the agent's credences about its value. There are thus two levels of probability in play: one of objective probabilities and, "on top of that," one of the agent's epistemic probabilities. Only C resides on the former, while the latter is host to P and ρ. Using the continuous version of the theorem of total probability, Huemer can then rewrite the denominator and the numerator of (1) like this:

$$\begin{aligned} P(U_n) &= \int_0^1 \rho(c) \cdot P(U_n \,|\, C = c)\, \mathrm{d}c \\ P(U_{n+1}) &= \int_0^1 \rho(c) \cdot P(U_{n+1} \,|\, C = c)\, \mathrm{d}c \end{aligned} \tag{2}$$

Huemer's third step is to argue that, based on the assumption of the Principle of Indifference, which he supplements with something he calls the "Explanatory Priority Proviso," it is rational for our agent to adopt a uniform probability distribution for the value of C on $[0, 1]$:

$$\rho(c) = 1 \text{ for } c \in [0, 1] \tag{3}$$

These assumptions are unnecessarily strong. However, I will finish the summary before explaining why. The uniform distribution allows Huemer to eliminate the first term of each of the two integrands in (2). His fourth step is to rewrite the two other terms using Lewis' (1980) Principal Principle, according to which the epistemic probability of U_m, conditional on the objective chance of each trial having outcome A being c, is c^m. The results are

$$\begin{aligned} P(U_n) &= \int_0^1 c^n \, \mathrm{d}c = \frac{1}{n+1} \, , \\ P(U_{n+1}) &= \int_0^1 c^{n+1} \, \mathrm{d}c = \frac{1}{n+2} \, . \end{aligned} \tag{4}$$

Plugging (4) back into (1) and reducing, the result is

$$P(A_{n+1} \,|\, U_n) = \frac{n+1}{n+2} \, .$$

Based on the same assumptions, Huemer calculates that $P(A_{n+1}) = \frac{1}{2}$, so $P(A_{n+1} \,|\, U_n)$ is larger whenever n is positive. Therefore, he concludes that the ideally rational agent is an inductivist.

Later in the same paper, Huemer allows for the possibility that the probability-density distribution in (3) may be wrong, and considers some other ways that the Principle of Indifference and the Explanatory Priority Proviso might be used to justify some alternative distributions. He shows that inductivism follows for all the functions he considers.

That concludes my summary. The point I want to make is that the use of these controversial principles to reach the inductivist conclusion is superfluous, as the same conclusion can be achieved on the basis of a much weaker assumption. I would call someone who assigns all of their credence for the value of C to just one value a "dogmatist." This label is justified by the fact that updating by conditionalization does not allow an extreme prior probability of 1 to change to a posterior probability different from 1; so, assigning all the subjective probability to just one value amounts to deciding that no amount of empirical evidence will make you change your mind. It is thus clear (and I will prove mathematically) that someone who does that will be a skeptic in Huemer's sense. But it turns out that, in World 1, being a dogmatist is not merely a sufficient condition for being a skeptic, it is also a necessary condition. That is, merely being open to the possibility of more than one value for the objective probability is enough to allow induction to shift the agent's probability upwards as more confirmations come in. Thus, for the purpose of arguing for inductivism, we can replace the Principle of Indifference and the Explanatory Priority Proviso with this much weaker assumption: it is not rational to be dogmatic about the value of C.

To prove that this assumption suffices, I first need to make the idea of ρ assigning all the probability to a single value of C make sense, even though ρ is a probability-*density* function. That can be achieved by means of the Dirac delta function, which can heuristically be characterized as the function δ such that $\delta(0) = \infty$ and $\delta(x) = 0$ for all $x \neq 0$, and such that the integral of the function over any interval that contains 0 is 1.[1] Each dogmatic distribution can therefore be

[1] See, e.g., Kanwal (1997) for a rigorous definition.

represented by a function ρ such that $\rho(x) = \delta(x - c)$, where c is the value of C that is dogmatically assumed to be correct. Let us call such a function *dogmatic.*

The following theorem is the key to arguing for my claim.[2] To state it, we first need a definition.

Definition. A generalized probability-density function is a function of the form

$$\rho(x) = \sum_{i \in I} a_i \delta(x - b_i) + \left(1 - \sum_{i \in I} a_i\right) \rho_0(x) \ ,$$

where ρ_0 is a regular probability-density function, I is a finite or countable index set, the a_is and their sum belong to the interval $[0, 1]$, and the b_is belong to the domain of the function.

Theorem. *For any generalized probability-density function ρ on $[0, 1]$ and strictly increasing functions f and g on the same interval, it holds that*

$$\int_0^1 \rho(x) f(x) g(x) \, \mathrm{d}x \geq \int_0^1 \rho(x) f(x) \, \mathrm{d}x \int_0^1 \rho(x) g(x) \, \mathrm{d}x \ . \tag{5}$$

If ρ is dogmatic, then equality holds; otherwise strict inequality holds.

Proof. For any $x, y \in [0, 1]$, it follows from the assumptions that $\rho(x)\rho(y) \geq 0$ and $(f(x) - f(y))(g(x) - g(y)) \geq 0$. Hence,

$$\int_0^1 \int_0^1 \rho(x)\rho(y)(f(x) - f(y))(g(x) - g(y)) \, \mathrm{d}x \, \mathrm{d}y \geq 0 \ . \tag{6}$$

If ρ is dogmatic, then $\rho(x)\rho(y) \neq 0$ only if $x = y$; and in that case, the two last factors in the integrand are both equal to 0, so equality holds in (6). Otherwise, there are x_0 and y_0 such that $x_0 \neq y_0$ and the probability of any neighborhood of each is positive. When x_0 and y_0 are different, it also follows that $(f(x) - f(y))(g(x) - g(y)) > 0$ for all x and y in some neighborhood of x_0 and some neighborhood of y_0. Therefore, strict inequality holds in (6).

By multiplying into the parentheses and exploiting the additivity of integration, the double integral can be rewritten as the sum of four double integrals. The first two are

$$\begin{aligned}
\int_0^1 \int_0^1 \rho(x)\rho(y) f(x) g(x) \, \mathrm{d}x \, \mathrm{d}y &= \int_0^1 \rho(y) \int_0^1 \rho(x) f(x) g(x) \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_0^1 \rho(x) f(x) g(x) \, \mathrm{d}x \int_0^1 \rho(y) \, \mathrm{d}y \\
&= \int_0^1 \rho(x) f(x) g(x) \, \mathrm{d}x
\end{aligned}$$

[2] It is a version of the Fortuin-Kasteleyn-Ginibre Inequality (Fortuin, Kasteleyn, and Ginibre 1971).

and

$$
\begin{aligned}
-\int_0^1 \int_0^1 \rho(x)\rho(y)f(x)g(y)\,\mathrm{d}x\,\mathrm{d}y &= -\int_0^1 \rho(y)g(y)\int_0^1 \rho(x)f(x)\,\mathrm{d}x\,\mathrm{d}y \\
&= -\int_0^1 \rho(x)f(x)\,\mathrm{d}x \int_0^1 \rho(y)g(y)\,\mathrm{d}y \\
&= -\int_0^1 \rho(x)f(x)\,\mathrm{d}x \int_0^1 \rho(x)g(x)\,\mathrm{d}x \ .
\end{aligned}
$$

By renaming variables, the third and fourth addends can be shown to be equal to the first and second, respectively; so, division by 2 eliminates them. Adding the second addend on both sides of (6) (and its $=$ and $>$ versions) then produces (5) (and its $=$ and $>$ versions). □

Let us apply the theorem. Go back to (2), skip what I referred to as "Huemer's third step," and just take the fourth step. That results in these equations:

$$
\begin{aligned}
P(U_n) &= \int_0^1 \rho(c)\cdot c^n\,\mathrm{d}c \\
P(U_{n+1}) &= \int_0^1 \rho(c)\cdot c^{n+1}\,\mathrm{d}c
\end{aligned}
$$

Similarly, we have

$$
P(A_{n+1}) = \int_0^1 \rho(c)\cdot c\,\mathrm{d}c \ ,
$$

so it follows from the theorem that

$$
P(U_{n+1}) \geq P(U_n)\cdot P(A_{n+1}) \ .
$$

Assuming that $P(U_n) > 0$, it thus follows from (1) that

$$
P(A_{n+1}\,|\,U_n) \geq P(A_{n+1}) \ ,
$$

which is the disjunction of inductivism and skepticism. The last part of the theorem allows us to distinguish between the two: if ρ is a dogmatic function, then skepticism follows; if not, then inductivism follows.[3]

This is a complete characterization of generalized probability-density functions, dividing them into those that lead to inductivism and those that lead to skepticism. And we can see that the latter class of functions is small indeed: only the absence of any doubt whatsoever will prevent the agent from learning through induction. I claim that the irrationality of such dogmatism is immediately obvious, independently of whether it can be reduced to more fundamental principles. Thus, in World 1, Humean inductive skepticism can be soundly rejected.

[3]The assumption that $P(U_n) > 0$ is only false if ρ is the Dirac delta function itself. This is a dogmatic function, but it leads to neither inductivism, skepticism, nor counter-inductivism, as defined by Huemer, for $P(A_{n+1}\,|\,U_n)$ is not defined. I will set this limit case aside.

2 Expiring deterministic laws

According to the nomological-explanatory solution to the problem of induction advocated by Armstrong (1983, 1991), BonJour (1998), and Foster (1983, 2004), there are many cases in which observation of a good number of Bs that are all As warrants an inference to the best explanation; which, according to them, is that there is a universal, deterministic law that implies that *all* Bs are As *by necessity*. Subsequent to making such inference, a rational agent can *deduce* that the next B will be an A.

The defenders of this solution believe that the possibility of temporally restricted laws is a major challenge for them. That is, the observation of a bunch of Bs that are all As could potentially also be explained by the existence of a law according to which all Bs *so far* are As by necessity; and they think that this would undermine inductivism. So, they have dedicated considerable energy to arguing that universal laws make for better (i.e., a priori more likely) explanations. The point of this section is to argue that that assumption is not necessary: just like with Huemer, Bayes, and Laplace above, it is easier to argue for inductivism than Armstrong, BonJour, and Foster realize.

Consider World 2A in which an epistemic agent is created knowing that at the moment of creation, there is a deterministic law in effect necessitating that Bs be As. She also *knows* that the law is temporally restricted, i.e., that it will expire after some unknown number, N, of B observations. Since we saw in the previous section that known, constant, objective chances are bad for inductivism, let us also stipulate that after the law has expired, each B has an objective chance, C, of being A for some constant $C \in [0, 1)$ that is known to our agent. Even in these apparently terrible circumstances, the agent will be an inductivist, unless she assigns priors in an extremely dogmatic way.

This can be seen from the following. The objective probability of A_{n+1} is equal to C if the law has expired by B number $n+1$, i.e., if $n+1 > N$. Otherwise, it is equal to 1. Hence, the subjective probability of A_{n+1}, before U_n is known, can be expressed as

$$P(A_{n+1}) = \sum_{i=1}^{n} P(N = i) \cdot C + \sum_{i=n+1}^{\infty} P(N = i) \cdot 1 \, .$$

Similarly, when U_n *is* known, the subjective probability of A_{n+1} is

$$P(A_{n+1} \,|\, U_n) = \sum_{i=1}^{n} P(N = i \,|\, U_n) \cdot C + \sum_{i=n+1}^{\infty} P(N = i \,|\, U_n) \cdot 1 \, .$$

At that point, the subjective probability for the proposition that the law covers exactly i Bs is as follows, according to Bayes' Theorem:

$$P(N = i \mid U_n) = \frac{P(U_n \mid N = i) \cdot P(N = i)}{P(U_n)} = \begin{cases} \dfrac{C^{n-i} \cdot P(N = i)}{P(U_n)} & \text{if } i \leq n \\ \dfrac{1 \cdot P(N = i)}{P(U_n)} & \text{if } i \geq n \end{cases}$$

So, both $P(A_{n+1})$ and $P(A_{n+1} \mid U_n)$ are weighted averages of some 1s and some Cs, which are smaller than 1. In the latter average, the 1s have greater weights than in the former average or, in some limit cases, the same weight. Hence, $P(A_{n+1} \mid U_n) \geq P(A_{n+1})$. There are two limit cases. The first is when $P(N = 1)$, $\ldots, P(N = n-1) = 0$, or equivalently when $P(U_n) = 1$. The second is when $P(N = n+1), \ldots = 0$. Both of these cases must be characterized as extremely dogmatic.

I want to emphasize with an example how remarkable, in its counter-intuitiveness, this simple mathematical result is. Let us say that the agent is considering whether B number ten will have the A property, before she has observed any Bs. In this example, she is almost *certain* that the temporally restricted law ensures that exactly the first nine Bs will be As, and only assigns a tiny amount of credence to the possibility that it might expire earlier and to the possibility that it might expire later. She is also *certain* that when it expires, all subsequent Bs will be non-As. According to the intuition that drives Armstrong, BonJour, and Foster, this should be a awful case for inductivism. Yet, our Bayesian epistemic agent realizes that her credence in A_{10} will be higher if she observes U_9 than it currently is. That is, she is an inductivist with respect to A_{10}. Observation of U_9 will make her credence in A_{10} go up, because the only hypotheses disconfirmed by this evidence are some according to which the objective probability of A_{10} is 0.

When discussing World 2A with colleagues, I have learned that some have the intuition that I am still being too nice to the inductivist, because the assumption of a known, constant, objective probability seems too much like "regularity in nature." According to them, we should instead assume that the expiration of the deterministic law is followed by the opposite entirely: total chaos! That should undermine inductivism! But what, in terms of a probability distribution, does such an assumption look like? The best answer I can think of is a distribution that assigns the same probability to every post-law sequence. That is complete randomness; every possibility is treated the same; it is the opposite of a deterministic law! However, that is the same probability distribution as the one that results from the assumption of a constant objective probability, with that probability being $\frac{1}{2}$. So, no; the result is again—perhaps counter-intuitively—inductivism.

Since I cannot cover every option in a short paper—indeed, in any finite paper—I will instead leave this exercise for the nomological-explanatory reader: come up with a credence distribution for what comes after the law's expiration that implies non-inductivism, and then convince yourself that it is the most reasonable distribution, and that it would be adopted by an ideally rational agent. You will rejoice in your failure, I think.

Is it possible to be a counter-inductivist in some world different from 2A where the inhabitants also know that it is governed by an expiring deterministic law? Yes, as it turns out. However, one has to go a bit out of one's way to design a world and an agent specifically with that goal in mind. I will describe one such contrived example. In World 2B, there is also a temporally restricted deterministic law that lasts for N rounds, and its expiration is also followed by a constant, objective probability of $C \in (0,1)$; except, that is, for the *first* round after the law has expired, where the objective probability is some lower value $C_0 \in [0, C)$. Again, the epistemic agent in this world knows this, including the values of C and C_0. The *temporary* drop in objective probability means that some (but only some) priors are counter-inductivist. Since it both makes counter-inductivism easier to "achieve," and greatly simplifies the mathematics, let us say that $C_0 = 0$.

The analogues to the above equations are as follows:

$$P(A_{n+1}) = \sum_{i=1}^{n-1} P(N=i) \cdot C + P(N=n) \cdot C_0 + \sum_{i=n+1}^{\infty} P(N=i) \cdot 1$$

$$P(A_{n+1} \,|\, U_n) = \sum_{i=1}^{n-1} P(N=i \,|\, U_n) \cdot C + P(N=n \,|\, U_n) \cdot C_0 + \sum_{i=n+1}^{\infty} P(N=i \,|\, U_n) \cdot 1$$

$$P(N=i \,|\, U_n) = \frac{P(U_n \,|\, N=i) \cdot P(N=i)}{P(U_n)} = \begin{cases} \dfrac{C_0 \cdot C^{n-1-i} \cdot P(N=i)}{P(U_n)} & \text{if } i < n \\ \dfrac{1 \cdot P(N=i)}{P(U_n)} & \text{if } i \geq n \end{cases}$$

With $C_0 = 0$, we therefore have

$$P(A_{n+1}) = \sum_{i=1}^{n-1} P(N=i) \cdot C + P(N=n) \cdot 0 + \sum_{i=n+1}^{\infty} P(N=i) \cdot 1 \,,$$

$$P(A_{n+1} \,|\, U_n) = \sum_{i=1}^{n-1} 0 \cdot C + \frac{P(N=n)}{P(U_n)} \cdot 0 + \sum_{i=n+1}^{\infty} \frac{P(N=i)}{P(U_n)} \cdot 1 \,.$$

Thus, before learning U_n, the credence in A_{n+1} is an average of some Cs, a 0, and some 1s; afterwards, it is just an average of a 0 and some 1s, with the relative weights of the 0 and the 1s staying the same. Hence, the credence decreases iff the weight of the 1s divided by the weights of the 0 and the 1s is smaller than C. That is,

$$P(A_{n+1} \,|\, U_n) < P(A_{n+1}) \;\leftrightarrow\; \frac{P(N > n)}{P(N \geq n)} < C \,.$$

This means that unless C is close to 1, the agent, in order to be a counter-inductivist, has to assign a relatively large credence to the possibility that the law expires after exactly n rounds, as opposed to at *any* later time.

3 Relevance to our world

Why is any of this relevant to our complicated world? We do not have the same a priori knowledge as any of the three agents we have considered, so even if they should be inductivists, how does that prove that we should? Well, it does not *prove* it. But the conclusions about the simple worlds make it more plausible. Here is why. Let $\{\pi_i\}_{i \in I}$ be a partition of all the possibilities that you and I, in our complicated world, would have to take into consideration if we did not yet have any empirical evidence. Using the theorem of total probability, we can then rewrite the condition for inductivism, $P(A_{n+1} \,|\, U_n) > P(A_{n+1})$, as

$$\sum_{i \in I} P(A_{n+1} \,|\, U_n \wedge \pi_i) \cdot P(\pi_i) > P(A_{n+1}) \ .$$

That is, it is a weighted average over the elements of the partition that must be larger than $P(A_{n+1})$. If some elements of the partition with positive weights are, by themselves, inductivist—i.e., $P(A_{n+1} \,|\, U_n \wedge \pi_i) > P(A_{n+1})$—then the overall probability distribution is also inductivist, unless other elements of the partition are counter-inductivist, and sufficiently counter-inductivist and with enough weight to fully counterbalance the inductivist elements.

Such counterbalancing becomes more difficult when one realizes that the *entirety* of the possibility that there is a constant objective chance of Bs being As would make for an inductivist part of a partition. That is the relevance of World 1. And it becomes more difficult yet when it is demonstrated that the possibility of laws only being in effect for a limited time does not obviously make for a counter-inductivist part, as assumed by the advocates of the nomological-explanatory solution; and that there are in fact sub-possibilities thereof that definitely (World 2A) or plausibly (World 2B) make for inductivist parts.

These considerations show that the inductivists are holding better cards than they themselves realize. However, they are not sufficient by themselves for a full argument for inductivism—for such an argument would also have to take account of Goodman's (1955) *grue* challenge and Smithson's (2017) reply to Huemer, among other things—but I believe they can be part of one.

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.

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.

Fortuin, C. M., P. W. Kasteleyn, and J. Ginibre (1971). Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics 22*, 89–103.

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.

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

Kanwal, R. P. (1997). *Generalized Functions: Theory and Technique* (Second ed.). Birkhäuser.

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.