6

From Kolmogorov, to Popper, to Rényi: There's No Escaping Humphreys' Paradox (When Generalized)

Aidan Lyon

6.1 Introduction

Humphreys' Paradox is often taken to be a serious challenge for the propensity interpretation of probability; it's one of Eagle's (2004: 21) arguments against propensity analyses of probability.[1] The conclusion that is typically drawn from the paradox is that the propensity interpretation does not satisfy 'the' probability axioms—i.e. Kolmogorov's axioms—and so the propensity interpretation must go (e.g. Salmon, 1984). Humphreys himself, in contrast, thought that it's Kolmogorov's probability axioms that ought to go (Humphreys, 1985: 569–70). This would involve replacing Kolmogorov's axiom system with some other. However, Humphreys offers no replacements, and, as far as I can tell, the subsequent literature has not either.

One might think that a promising alternative would be Popper's axiom system. However, I argue that Popper's axioms fare no better than Kolmorgov's. Interestingly, though, Popper's axioms were inspired by Rényi's (1955) axioms, and it turns out

Thanks to Kenny Easwaran, Alan Hájek, Christopher Hitchcock, Joel Velasco, Alastair Wilson, an anonymous referee, and the audience at the *Chance and Time* conference at Monash University for helpful comments and discussion.

[1] Although there are different versions of the propensity interpretation—some of which some authors have argued do not suffer from the paradox (e.g. Gillies (2000))—I specifically have in mind here the classic, single-case, causal dispositional versions of the propensity interpretation. Moreover, I'm restricting my attention to those propensity interpretations that understand $P(A|B) = x$ as the statement that there is a propensity of strength x for B to produce or bring about A. This is what I take, for example, Popper to have in mind when he writes: 'we can say that the singular event a possesses a probability $p(a, b)$ owing to the fact that it is an event produced, or selected, in accordance with the generating conditions b (Popper, 1959b: 34).

that propensity theorists can avoid Humphreys' Paradox by adopting Rényi's axiom system. This move also allows propensity theorists to avoid some closely related problems that they would otherwise face. Unfortunately, these problems, along with Humphreys' Paradox, are all just special cases of a much more general problem, from which Rényi's axioms provide no safe haven.

I'll begin, in §6.2, by introducing the details of Humphreys' Paradox. At least four versions of the paradox have appeared in the literature and it will be important to distinguish them from one another. In §6.3, I'll argue that all of the versions of the paradox arise in Popper's axiom system, but one of them doesn't arise in Rényi's. In §6.4, I'll introduce Milne's Problem (Milne, 1985), and show that it has at least three variants—all of which are analogous to three of the four versions of Humphreys' Paradox. I'll argue that one of the versions of Milne's Problem also doesn't arise in Rényi's axiom system. Similarly, I'll argue that a problem due to Sober, (2010) can be avoided by adopting Rényi's axioms. However, Sober's problem points the way to a much more general problem, and, in §6.6, I'll argue that this more general problem cannot be solved by adopting Rényi's axioms.

6.2 Humphreys' Paradox

In addition to the original paradox (Humphreys, 1985), two other versions of the paradox have appeared in the literature—and are discussed in Humphreys (2004). In this section, I'll describe these three different versions of the paradox. After doing so, I'll present a fourth version, which has not yet been explicitly identified in the literature.

First, the original paradox. Suppose we have some photons being emitted from a source at time t_0. Some of these photons then impinge on a half-silvered mirror at time t_1. Some of the photons that make it to the mirror are then transmitted through it. Whether a particular photon makes it to the mirror is an indeterministic and probabilistic matter; so too is whether it then gets transmitted or reflected. Let I_{t_1} be the event of a particular photon impinging on the mirror at t_1. Let T_{t_2} be the event of that photon being transmitted through the mirror at the later time t_2. Finally, let B_{t_0} be the background conditions in place at t_0 which include the fact that the photon was emitted from the source at t_0. Thus defined, the events have the following probabilistic constraints:

(i) $Pr_{t_0}(T_{t_2}|I_{t_1}B_{t_0}) = p > 0$
(ii) $1 > Pr_{t_0}(I_{t_1}|B_{t_0}) = q > 0$
(iii) $Pr_{t_0}(T_{t_2}|\neg I_{t_1}B_{t_0}) = 0$

The parameters p and q can take any values within the above constraints and once fixed are to be interpreted as propensity values. It is also assumed that the propensity for the particle to impinge upon the mirror is unaffected by whether the particle is later transmitted or not. Humphreys calls this the *Principle of Conditional Independence* (CI), and formulates it as follows:

$$(\text{CI})\ Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = Pr_{t_0}(I_1 | \neg T_{t_2} B_{t_0}) = Pr_{t_0}(I_{t_1} | B_{t_0})$$

From these premises and the Kolmogorov axioms of probability, a contradiction can be derived. Using Bayes' Theorem:[2]

$$\begin{aligned} Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) &= \frac{Pr_{t_0}(T_{t_2} | I_{t_1} B_{t_0}) Pr_{t_0}(I_{t_1} | B_{t_0})}{Pr_{t_0}(T_{t_2} | I_{t_1} B_{t_0}) Pr_{t_0}(I_{t_1} | B_{t_0}) + Pr_{t_0}(T_{t_2} | \neg I_{t_1} B_{t_0}) Pr_{t_0}(\neg I_{t_1} | B_{t_0})} \\ &= \frac{pq}{pq+0} \\ &= 1 \end{aligned}$$

But from CI:

$$Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = q < 1$$

which contradicts the value of $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0})$ obtained from Bayes' Theorem.

That is the original paradox. Two other closely related paradoxes have been discussed in the literature. They arise when CI is replaced with alternative principles. The first alternative principle has been called the *Fixity Principle* (F) (Humphreys, 2004: 670; first presented by Milne, 1985), and states that:

$$(F)\, Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 1 \text{ or } Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 0$$

The idea behind this principle is that once t_2 has occurred, I_{t_1} has already occurred or not occurred—and since the past is fixed, this matter cannot change. The principle as formulated as above doesn't quite capture this idea, though. Presumably part of the intuition is that the value of $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0})$ depends on whether I_{t_1} actually occurs. So, a better way of capturing the relevant intuition might be:

$$(F')\text{ If } I_{t_1}, \text{ then } Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 1 \text{ and if } \neg I_{t_1}, \text{ then } Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 0$$

If I_{t_1} doesn't actually occur, then $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 0$, which contradicts the assignment obtained from Bayes' Theorem (which applies regardless of whether I_{t_1} actually occurs).

The second alternative principle, which has been called the *Zero Influence Principle* (ZI) (Humphreys, 2004: 670), states that:

$$(\text{ZI})\ Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 0 \tag{6.1}$$

This assignment is clearly inconsistent with the assignment obtained from Bayes' Theorem. The idea behind ZI is that the event T_{t_2} has no propensity to produce I_{t_1} since if it occurs, it occurs after I_{t_1}. However, 'no propensity' can be ambiguous between 'zero propensity' and literally '*no* propensity'—i.e., the absence of a

[2] McCurdy (1996: p. 116) argues that Bayes' Theorem, or any other inversion theorem, is not needed to arrive at this value, since $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0}) = 1$ can be derived from the arrangement of the system. However, Humphreys (2004: 674) appears to deny this point. This detail does not matter for my purposes here, for there is a violation of Kolmogorov's axioms either way.

propensity of any strength. This suggests a fourth principle: the No Propensity principle (NP):[3]

$$\text{(NP)}\ Pr_{t_0}(I_{t_1} \mid T_{t_2} B_{t_0}) \text{ is undefined} \tag{6.2}$$

NP is clearly inconsistent with the assignment obtained from Bayes' Theorem, since NP says there is no assignment and Bayes' Theorem says there is. Salmon's presentation of the problem—one of the first—seems to be along these lines:

'If numerical values are given, we can calculate the propensity of [a] factory to produce defective corkscrews. So far, so good. Now, suppose an inspector picks one corkscrew from the day's output and finds it defective. Using Bayes' theorem we can calculate the probability that the defective corkscrew was produced by the new machine, but it would hardly be reasonable to speak of the propensity of that corkscrew to have been produced by the new machine.' (Salmon, 1984: 88).

So we have four paradoxes, each stemming from one of the principles CI, F/F′, ZI, and NP combined with (i)–(iii) and Bayes' Theorem.[4]

Before moving on, it's worth noting that CI, F', ZI, and NP are pairwise inconsistent when combined with (i)–(iii)—and in some cases, even when not combined with (i)–(iii). NP clearly contradicts CI, F′, and ZI, since it says $Pr_{t_0}(I_{t_1} \mid T_{t_2} B_{t_0})$ is not defined and they all assume it is. ZI contradicts F' since ZI says that $Pr_{t_0}(I_{t_1} \mid T_{t_2} B_{t_0}) = 0$ regardless of whether I_{t_1} occurs. (ZI is not inconsistent with F, but it does seem to be inconsistent with the motivation behind F.) ZI also contradicts CI combined with (ii) since from ZI we have $Pr_{t_0}(I_{t_1} \mid T_{t_2} B_{t_0}) = 0$ and putting that into CI we get $Pr_{t_0}(I_{t_1} \mid B_{t_0}) = 0$, which contradicts (ii). Similarly, both F and F' contradict CI combined with (ii).

Since (i)–(iii) are not up for debate, and the probability axioms are typically not up for debate, it seems that the blame for all of these contradictions ought to be placed on the principles. Not all of them can be true, but it's not entirely clear which ones are false, and all four principles have some intuitive force behind them.

6.3 Changing the Axioms

The literature typically sees Humphreys' Paradox as a problem for propensity interpretations of probability (see e.g. Eagle, 2004: 402). However, Humphreys himself thought that the paradox spells trouble for the standard probability axioms:

'It is time, I believe, to give up the criterion of admissibility [the criterion that an interpretation of probability should satisfy the standard probability calculus]. We have seen that it places an unreasonable demand upon one plausible construal of propensities. Add to this the facts that

[3] Humphreys notes that there is a difference between a propensity not existing and having a value of zero (2004: 671 n. 7), but nevertheless does not discuss NP.

[4] Actually, since the principles are inconsistent with each other and different intuitions motivate them, it's probably better to say that we have four arguments, each of which is a paradox only for those that have the relevant intuitions.

limiting relative frequencies violate the axiom of countable additivity and that their probability spaces are not sigma–fields unless further constraints are added; that rational degrees of belief, according to some accounts, are not and cannot sensibly be required to be countably additive; and that there is serious doubt as to whether the traditional theory of probability is the correct account for use in quantum theory. Then the project of constraining semantics by syntax begins to look quite implausible in this area.' (Humphreys, 1985: 569–70).

Humphreys' conclusion is that we should not be beholden to Kolmogorov's axiom system, and that which axiom system we adopt should be sensitive to our interpretation of probability. Propensity theorists would be well advised to find an axiom system that is suited to their interpretation.

Interestingly, Popper, who was a propensity theorist, developed his own axiom system as an alternative to Kolmorogov's. Popper had several motivations for doing so, but one of them was that he felt that an axiom system should not rule out possible probability interpretations:[5]

'In view of the fact that a probability statement such as '$p(a, b) = r$' can be interpreted in many ways, it appeared to me desirable to construct a purely 'formal' or 'abstract' or 'autonomous' system, in the sense that its 'elements' ... can be interpreted in many ways, so that we are not bound to any particular one of these interpretations. ...

There are three main characteristics which distinguish a theory of this kind from others. (i) It is formal; that is to say, it does not assume any particular interpretation, although allowing for at least all known interpretations. ...' (Popper, 1959a: 329–30).

One of the defects of Kolmogorov's system, Popper argued, was that it ruled out some interpretations of probability. Since Popper was a propensity theorist and also clearly sympathetic to Humphreys' proposal that probability syntax shouldn't constraint probability semantics, we might expect that Popper's propensity interpretation would be a possible interpretation of his own axiom system.

Somewhat ironically, this is not the case. As Humphreys notes, his paradox can be generated within Popper's axiom system (Humphreys, 1985: 559–60). Indeed, this is true not just for the version of the paradox based on CI, but also for the versions based on F/F′, ZI, and NP. The reason why it is true for the versions of the paradox based on F/F′ and ZI is pretty straightforward: the inversion theorems of Popper's axiom system can determine probability values that contradict F/F′ and ZI.

The reason why the NP-version of the paradox arises is the same as before and straightforward too: Popper's axiom system guarantees the existence of 'inverse' conditional probabilities. However, the reason why Popper's axiom system has this property is more interesting. In addition to an axiom system not ruling out possible probability interpretations, Popper also thought that it was a virtue for an axiom system to be what he called *symmetrical*. In continuation of the previous quote, he writes:

5 He expresses the same sentiment in many other places—e.g. Popper, 1938: 275; 1955: 51.

(iii) It is symmetrical; that is to say, it is so constructed that whenever there is a probability $p(b, a)$—i.e. a probability of b given a—then there is always a probability $p(a, b)$ also—even when the absolute probability of b, $p(b)$, equals zero; that is, even when $p(b) = p(b, a\bar{a}) = 0$. (Popper, 1959a: 330).

It is this condition of being symmetrical that guarantees that Popper's axiom system suffers from the NP-version of Humphreys' Paradox.

Popper seems to have been inspired by the development of another alternative axiom system, due to Rényi (1955):

I have received considerable encouragement from reading A. Rényi's most interesting paper 'On a New Axiomatic Theory of Probability', *Acta Mathematica Acad. Scient. Hungaricae* 6, 1955, pp. 286–335. Although I had realized for years that Kolmogorov's system ought to be relativized, and although I had on several occasions pointed out some of the mathematical advantages of a relativized system, I only learned from Rényi's paper how fertile this relativization could be. The relative systems published by me since 1955 are more general still than Rényi's system which, like Kolmogorov's, is set-theoretical, and non–symmetrical ...' (Popper, 1959a: 352).

Popper thought that Rényi's axiom system was a great step forward (away from Kolmogorov), but that it still had the defects of being set-theoretical and non-symmetrical. Since it is the symmetry of Popper's system that results in the NP-version of Humphreys' Paradox, we might expect, then, that the asymmetry of Rényi's axiom system allows it to *avoid* this version of the paradox.

This is, in fact, the case. Rényi's axiom system allows for $Pr(A|B)$ to be defined and $Pr(B|A)$ to be undefined. This is because Pr is defined over an algebra $\mathcal{A}$ crossed with a subset $\mathcal{B}$, i.e., $Pr : \mathcal{A} \times \mathcal{B} \longrightarrow \mathbb{R}$. Since $\mathcal{B}$ only has to be a subset of $\mathcal{A}$, some elements of $\mathcal{A}$ need not appear in $\mathcal{B}$. So if A and B are in $\mathcal{A}$, but only B is in $\mathcal{B}$, then $Pr(A|B)$ gets a value and $Pr(B|A)$ doesn't. This feature of the axiom system blocks the NP-version of Humphreys' Paradox, since B_{t_0}, I_{t_1}, T_{t_2} can all be in $\mathcal{A}$, but only B_{t_0} and I_{t_1} need be in $\mathcal{B}$, which means, for example, $Pr_{t_0}(T_{t_2} | I_{t_1} B_{t_0})$ can be defined without $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0})$ having to be defined—in agreement with NP.

However, the paradoxes based on CI, F/F′, and ZI still go through. This is because the statements of CI, F/F′, and ZI all involve the inverse conditional probabilitiy $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0})$, and so for them to be true, $\mathcal{B}$ must include $T_{t_2} B_{t_0}$ as an element. If $T_{t_2} B_{t_0}$ is in $\mathcal{B}$, then there are instances of the Multiplication Rule and Bayes' Theorem that involve $Pr_{t_0}(I_{t_1} | T_{t_2} B_{t_0})$, and the derivation of the paradoxes based on CI, F/F′ and ZI go through.

The debate over which of CI, F/F′, ZI, and NP is true now looms large. We have a reasonably standard probability axiom system with independent motivation that doesn't result in a paradox if NP is true. It seems, therefore, that a propensity theorist would be well advised to accept NP (thereby rejecting CI, F/F′, and ZI) and Rényi's axiom system. For what it's worth, it seems that NP is something a propensity theorist should say anyway: propensities are causal dispositions and there are no 'backwards' causal dispositions.

6.4 Milne's Problem

I have argued that the propensity theorist can solve Humphreys' Paradox by adopting Renyi's axiomatization of probability and the idea that there are no backwards propensities. I'll now argue that this move (when slightly tweaked) also solves a closely related problem for the propensity theorist: Milne's Problem. However, just as with Humphreys' Paradox, there is more than one version of Milne's Problem.

Consider the following example:

Let us consider an unbiased die in an indeterministic universe in which the real single-case probabilities have their familiar values. If *a* denotes the outcome '6'–uppermost, and *b* denotes the event 'even number'–uppermost, then $p(a) = 1/6$, $p(b) = 1/2$ and, by definition, $p(a|b) = 1/3$. How is $p(a|b)$ to be interpreted? It is certainly not the probability that the outcome *a* is realised given that the outcome *b* has been realised, for if *b* has been realised exactly one of the events '2'–uppermost, '4'–uppermost, or '6'–uppermost has occurred. In the first two cases *a*'s occurrence is impossible, in the third it is certain. The event *b* is realised by the occurrence of *a* or of an event incompatible with *a*. It is the realisation of *a* or one of these other events which constitutes the occurrence of *b*. In terms of real single–case probabilities, when *b* occurs there is no longer any matter of chance, no indeterminacy, about *a*'s occurrence, it is fully determinate. (Milne, 1985: 130).

In our notation, we can describe the example as follows. A fair die is rolled at t_0. Let B_{t_0} be the background conditions at t_0, which include the event of the beginning of this roll, and also suppose that the roll of the die has various propensities to produce events at t_1. For example, the propensity for the role to produce 'six' is:

$$Pr_{t_0}(Six_{t_1}|B_{t_0}) = 1/6$$

Similarly, the five other possible outcomes have propensities of strength $1/6$ to occur. This probability distribution and the (Kolmogorov) probability axioms together entail that:

$$Pr_{t_0}(Six_{t_1}|Even_{t_1}B_{t_0}) = 1/3$$

where *Even* is equivalent to *Two* or *Four* or *Six*. However, this conditional probability is inconsistent with the following principle:

$$(\mathrm{F}^*) Pr_{t_0}(Six_{t_1}|Even_{t_1}B_{t_0}) = 1 \text{ or } Pr_{t_0}(Six_{t_1}|Even_{t_1}B_{t_0}) = 0,$$

which we may call the *fixity of simultaneous events.*

Of course, stating F* explicitly like this makes it clear that there is an analogy here with F—indeed I have extracted them both from Milne's paper, and they seem to be motivated by the more general principle:

[T]he occurrence of the conditioning event does not determine the occurrence or otherwise of the conditioned event. What makes the probabilities 0 or 1 is that the occurrence or otherwise of the conditioned event is determinate before or concurrently with the occurrence of the condition.' (Milne, 1985: 131).

However, there are also analogues of ZI and NP that may seem equally or more intuitive and which also result in probability assignments that are inconsistent with $Pr_{t_0}(Six_{t_1}|Even_{t_1}B_{t_0}) = 1/3$.

For example, one may reason that even if Six_{t_1} does occur, it is not because of any causal efficacy of $Even_{t_1}$: $Even_{t_1}$ has zero dispositional strength in bringing Six_{t_1} about. Put another way, $Even_{t_1}$ has 'zero influence' on Six_{t_1}:

$$(\mathrm{ZI}^*) Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0}) = 0.$$

regardless of whether Six_{t_1} occurs. This, of course, is no help to the propensity theorist, for ZI* is also inconsistent with $Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0}) = 1/3$.

Another intuition one may have is that since Six_{t_1} and $Even_{t_1}$ must be simultaneous if they both occur, then there can be *no* causal relation between them, and so *no* propensity for either to bring the other about. This is an analogue of NP:

$$(\mathrm{NP}^*)\ Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0}) \text{ is undefined.}$$

That is, there simply isn't a propensity for $Even_{t_1}$ to produce Six_{t_1}, not even one of zero strength. NP* is also inconsistent with $Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0}) = 1/3$, so the propensity theorist with the NP* intuition is also in trouble.

However, it is only this last principle, when coupled with Renyi's axiomatization of probability that allows the propensity theorist to escape unscathed. This is because Renyi's axiom system allows $Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0})$ to not be defined, even when conditional probabilities such as $Pr_{t_0}(Six_{t_1} | B_{t_0})$ and $Pr_{t_0}(Even_{t_1} | B_{t_0})$ are.

At this point, one may protest: although Renyi's axiom system does not require that $Pr_{t_0}(Six_{t_1} | Even_{t_1} B_{t_0})$ be defined, it nevertheless should be, for we so clearly and intuitively know it's value: 1/3. It seems that there is a deeper problem with the propensity interpretation of probability. Forget worries about whether it satisfies this or that axiom system, the real problem is that the interpretation doesn't satisfy a basic platitude about the die: the probability of *six* given *even* is 1/3, not 'undefined'.

Is this bad news for the propensity interpretation of probability? No; I think it's bad news for probability interpretation *monism*: the view that there is one interpretation for all probability statements. A small degree of pluralism—a degree of pluralism that almost all philosophers of probability already have—can solve the problem.

Most philosophers of probability agree that along with the notion of objective probability, there is the notion of subjective probability—a.k.a. degree of belief, or credence. For example, Lewis said that both concepts are required for the proper understanding of science:

> We subjectivists conceive of probability as the measure of reasonable partial belief. But we need not make war against other conceptions of probability, declaring that where subjective credence leaves off, there nonsense begins. Along with subjective credence we should believe also in objective chance. The practice and the analysis of science require both concepts. Neither can replace the other. Among the propositions that deserve our credece we find, for instance, the proposition that ... any tritium atom that now exists has a certain chance of decaying within a year. Why should we subjectivists be less able than other folk to make sense of that? (Lewis, 1986: 83).

Indeed, Lewis sought to connect the two via what he called the *Principal Principle* (which I'll explain shortly). It this connection between chance and credence that solves the problem for the propensity theorist.

Roughly speaking, the intuition behind the Principal Principle is that if you know the chance of some proposition and you have no other information relevant to that proposition, then your credence in the proposition should be equal to the chance of the proposition. Formally, if Cr is a reasonable initial credence function, and $Ch_t(A) = x$ is the proposition that the chance of A at time t is x, then:

$$Cr(A|Ch_t(A) = x \wedge E) = x$$

where E is any admissible proposition (roughly, any proposition irrelevant to A). We can use the Principal Principle to understand 'the probability' of six given on a fair roll of a die: it is a credence that is grounded in certain objective chances, via the Principal Principle.

A propensity theorist understands the chance of A at time t as the propensity of A to be produced by some set of background conditions at time t. That is:

$$Ch_t(A) := Pr_t(A|B_t)$$

From Milne's example of the roll of the fair die, we have:

$$Pr_{t_0}(Six_{t_1}|B_{t_0}) = 1/6$$
$$Pr_{t_0}(Even_{t_1}|B_{t_0}) = 1/2.$$

If we plug these into the Principal Principle, we get:

$$Cr(Six_{t_1}|Pr_{t_0}(Six_{t_1}|B_{t_0}) = 1/6 \wedge Pr_{t_0}(Event_1|B_{t_0}) = 1/2) = 1/6$$
$$Cr(Even_{t_1}|Pr_{t_0}(Six_{t_1}|B_{t_0}) = 1/6 \wedge Pr_{t_0}(Event_1|B_{t_0}) = 1/2) = 1/2$$

From here on, I'll abbreviate '$Pr_{t_0}(Six_{t_1}|B_{t_0}) = 1/6 \wedge Pr_{t_0}(Even_{t_1}|B_{t_0}) = 1/2) = 1/6$' as simply K. So we have $Cr(Six_{t_1}|K) = 1/6$ and $Cr(Even_{t_1}|K) = 1/2$.

The Principal Principle is meant to be a rationality constraint on credences. Another is Bayesian Conditionalization, which (roughly) says that when we learn some proposition K, our new credence function should be our old credence function conditional on K:

$$Cr_K(-) = Cr(-|K)$$

In reasoning about the fair die, we therefore ought to conditionalize on our knowledge of the chances associated with the die. So, by Bayesian Conditionalization, we get:

$$Cr_K(Six_{t_1}) = 1/6$$
$$Cr_K(Even_{t_1}) = 1/2$$

Another plausible rationality constraint on our credences is one on our conditional credences:

$$Cr(A|B) = Cr(A \wedge B)/Cr(B), \text{ if } Cr(B) > 0$$

So from this principle, we have:

$$Cr_K(Six_{t_1}|Even_{t_1}) = Cr_K(Six_{t_1} \wedge Even_{t_1})/Cr_K(Even_{t_1})$$

One final intuitive rationality constraint on credences: if A entails B then $Cr(A \wedge B) = Cr(A)$. From this, we get:

$$\begin{aligned} Cr_K(Six_{t_1}|Even_{t_1}) &= Cr_K(Six_{t_1} \wedge Even_{t_1})/Cr_K(Even_{t_1}) \\ &= (1/6)/(1/2) \\ &= 1/3 \end{aligned}$$

So from our knowledge of the chances/propensities of the die, and some intuitive rationality constraints on credences, we see that *our* probability of six given even ought to be one-third.

I submit that, in Milne's example, propensity theorists should understand 'the probability' of *six* given *even* as a rationally constrained credence function of an agent who understands the set-up of the roll. This allows us to use the full resources of standard axiomatizations of probability without thereby committing ourselves (as propensity theorists) to understanding all such probabilities as propensities. In short, the probability axioms that capture the rationality constraints on credences are different to the probability axioms that circumscribe propensities, but nevertheless there is a connection between credences and propensities (e.g. the Principal Principle).[6]

One common obstacle to understanding a probability statement as a statement of credence is that, if the probability statement has some objectivity to it, then it's hard to account for that objectivity with a subjective credence function. For example, attempts to understand the probabilities of statistical mechanics as credences quickly run into the problem of determining a unique and objective prior that all rational agents must have (see e.g. Jaynes, 2003). However, my proposal for understanding 'the probability' of six given even as a rationally constrained credence function of an agent who understands the set-up of the roll doesn't suffer from this problem. My proposal does not require a unique objective prior. It only requires some very plausible rationality constraints on credences and knowledge of the set-up of roll of the die by some (possibly hypothetical) rational agent.

6.5 Sober's Problem

Humphreys' Paradox is said to a serious problem for causal propensity interpretations of probability, e.g.:

> This problem is devastating for views that take propensities to involve weakened or intermittent causation. This is because causation fails simple inversion theorems of the probability calculus. (Eagle, 2004: 36).

[6] Similar reasoning can be applied to other examples, such as the one in Humphreys' Paradox.

'Humphreys's paradox' is ... the basis for one of the most fundamental criticisms of the propensity interpretation of probability. (McCurdy, 1996: 105).

And the root of the problem is often attributed to a mismatch between the temporal asymmetry of propensities and symmetry of probabilities:

[T]here is an asymmetry in propensities as causes that is not present in probability; so probabilities cannot be propensities.

The point is simple: the interpretation of probability should not require actual backwards causation for every well defined inverse probability! (Eagle, 2004: 37).

The essence of the issue can easily be conveyed. Suppose some conditional propensity exists, the propensity for D to occur conditional on C, $Pr(D|C)$. ... Standard theories of conditional probability require that when $Pr(D|C)$ exists, so does the inverse conditional probability $Pr(C|D)$.... Yet the inverse propensity ..., $Pr(C|D)$, ... is not related to $Pr(D|C)$ in any simple way, if indeed it is mathematically dependent at all. One might even doubt whether such an inverse propensity exists. (Humphreys, 2004: 668).

If propensities are causal tendencies—that is, if $Pr(Y \text{ at } t_2|X \text{ at } t_1)$ represents the causal tendency of X at t_1 to produce Y at t_2—then the propensity interpretation cannot make sense of the 'backwards probabilities' ... that have the form $Pr(X \text{ at } t_1|Y \text{ at } t_2)$, at least not if cause must precede effect (this objection is due to Paul Humphreys; see Salmon (1984: 205). (Sober, 2010: 149).

When construed as causal tendency, probability cannot satisfy standard axioms because of the temporal asymmetry between cause and effect. (Milne, 1987: 330).

It is true: there is a mismatch between propensities and probabilities. However, the mismatch is not only because of the time–asymmetry of propensities and the non–time–asymmetry of probabilities.[7] There are more general properties of propensities that make them unable to be (standard) probabilities.

To see this consider the following example, inspired by Sober (2010: n. 32). Suppose we roll a pair of dice fairly, and let B_{t_0} be the background conditions that include that the dice were rolled at t_0. Suppose that one die lands and finishes rolling before the other. Let t_1 and t_2 be the two times that the dice finish rolling ($t_1 < t_2$), and let One_{t_1} be that the die that finishes at t_1 lands with the face with one dot facing up, and similarly for other possible outcomes of the roll. There are all sorts of unproblematic propensities associated with this setup:

$$Pr(One_{t_1}|B_{t_0}) = 1/6, Pr(One_{t_2}|B_{t_0}) = 1/6, \text{ etc.}$$

But there are also plenty of problematic ones. Consider:

$$Pr(One_{t_2}|Four_{t_1}B_{t_0}) = 1/6$$

If we understand the formalism as we have so far, then this is the propensity for the first die landing *four* to result in the second die landing *one*. But, intuitively, there is no such propensity—which, again, is ambiguous between there being a propensity

[7] Why 'non–time–asymmetry' and not simply 'time–symmetry'? Because time does not figure in the (standard) probability axioms, and also because they are not, strictly speaking, symmetric (in Popper's sense), anyway.

but with zero strength (another analogue of ZI), or there literally being no propensity of any strength (another analogue of NP). B_{t_0} is a so-called common cause for One_{t_2} and $Four_{t_1}$ (if they happen), and there is no causal relation between One_{t_2} and $Four_{t_1}$ themselves.

We therefore have a problem that is similar to Humphreys' Paradox and Milne's Problem: we're forced, by the probability axioms, to have a propensity where we don't think there is one—or forced to have one with a strength which we don't think is correct. Fortunately, Rényi's axiom system can come to the rescue again, in much the same way as it came to rescue in the case of Milne's Problem. One can simply insist that the domain of the probability function that represents the propensities of the situation is $A \times B$, and that B only contains B_{t_0}, and so problematic conditional probabilities such as $Pr(One_{t_2}|Four_{t_1}B_{t_0})$ remained undefined.

6.6 The General Problem

Unfortunately, this victory is short-lived. Sober's problem is similar to Humphreys' Paradox and Milne's Problem, but it is also different in an important respect: there is no reliance on details concerning backward propensities or downward (synchronic) propensities. As noted earlier, Humphreys' Paradox is often attributed to the temporal asymmetry of propensities which is not shared by probabilities. However, Sober's problem shows that there is a general feature of the probability axioms causing the trouble: it forces there to be propensities of certain strengths where we think there are no propensities. The temporal asymmetry of propensities is just one kind of way in which this happen. There are other ways, and some of them are ones on which appealing to Rényi's axioms is of no help.

We can capture the three problems that we have seen so far with three slogans. Humphreys' Paradox: there are no backwards propensities from later events to earlier events. Milne's Problem: sometimes there are no propensities between synchronic events. Sober's Problem: sometimes there are no propensities between events brought about by a common cause. But there are more slogans, and so there are more problems. Whenever A and B are causally independent, there will be a problem with interpreting $P(A|B)$ (or $P(B|A)$) as a propensity.

For example: sometimes there are no propensities between events in independent causal chains. Suppose that two dice are rolled at t_0, but one is rolled here on Earth and the other is rolled really far away—say, outside of our light cone. Let B_{t_0} be the background conditions at t_0 that only include the details of the roll on Earth, and let B'_{t_0} be the background conditions at t_0 that only include the details of the roll that happens outside of our light cone. Similarly, let One_{t_1} be the event of the Earth die resulting in 'one' at t_1 and One'_{t_2} be the distant die resulting in 'one' at t_2 (t_1 can be before, after or the same time as t_2, it doesn't matter—but both are after t_0). Again, various unproblematic propensities exist, e.g.:

$$Pr(One_{t_1}|B_{t_0}) = 1/6$$

$$Pr(One'_{t_1} | B'_{t_0}) = 1/6$$

But there are some problematic ones too, e.g.:

$$Pr(One_{t_1} | B'_{t_0}) = ?$$

The '?' is there because it isn't clear what the value should be, but it *is* clear that Kolmogorov's axioms require that there be one ($Pr_{t_0}(B'_{t_0})$ will be greater than 0 if $Pr(One'_{t_1} | B'_{t_0})$ is defined). Similarly, Popper's axioms requires that there be a value too, for any Popper function is defined over $\mathcal{A} \times \mathcal{A}$. One might hope that Rényi's axioms fare better, but they don't. Since $Pr(One_{t_1} | B_{t_0}) = 1/6$ and $Pr(One'_{t_1} | B'_{t_0})$ are assumed to be defined, this means that $One_{t_1}, One'_{t_1} \in \mathcal{A}$ and $B_{t_0}, B'_{t_0} \in \mathcal{B}$, and since the domain of any Rényi function is $\mathcal{A} \times \mathcal{B}$, this means that it has to assign a value to $Pr(One_{t_1} | B'_{t_0})$.

This, I believe, illustrates the most general form of the problem for propensity interpretations: there are all sorts of pairs of events that have no propensity relations between them, and all three axiom systems—Kolmogorov's, Popper's, and Rényi's—will force there to be conditional probabilities between some of them. That is bad news for propensity theorists if conditional probabilities are meant to represent propensity relations.

6.7 Conclusion

As mentioned earlier, the typical reaction to Humphreys' Paradox is that it is a problem for the propensity interpretation. Humphreys thought otherwise, and concluded that it's a problem for the Kolmogorovian probability axiom system, and that 'the project of constraining semantics by syntax begins to look quite implausible in this area' (Humphreys, 1985: 570). In this spirit, I have argued that one can solve Humphreys' Paradox, Milne's Problem, and Sober's problem, if one rejects Kolmogorov's axioms as the correct probability axioms for propensities, adopts Rényi's axioms for propensities instead, accepts the NP and NP* principles, and allows that there can be different probability axiom systems for different probability interpretations.

However, Sober's problem points the way to a more general problem for the propensity interpretation. There are all sorts of pairs of events that have no propensity relations between them, and all three axiom systems—Kolmogorov's, Popper's, and Rényi's—will sometimes force there to be conditional probabilities between them. This is not an argument that there is no alternative axiom system that propensity theorists can adopt, but it is an argument that the three main contenders are not viable.

References

Eagle, A. (2004). Twenty-One Arguments Against Propensity Analyses of Probability. *Erkenntnis*, 60, 371–416.

Gillies, D. (2000). *Philosophical Theories of Probability.* Routledge: New York.

Humphreys, P. (1985). Why Propensities Cannot Be Probabilities. *Philosophical Review*, 94, 557–70.

—— (2004). Some Considerations on Conditional Chances. *British Journal for the Philosophy of Science*, 55, 667–80.

Jaynes, E. (2003). *Probability Theory: The Logic of Science.* Cambridge: Cambridge University Press.

Lewis, D. K. (1986). A Subjectivist's Guide to Objective Chance. In *Philosophical Papers*, (vol. 2, pp. 83–132). Oxford: Oxford University Press.

McCurdy, C. (1996). Humphrey's Paradox and the Interpretation of Inverse Conditional Propensities. *Synthese*, 108(1), 105–25.

Milne, P. (1985). Can there be a Realist Single-Case Interpretation of Probability? *Erkenntnis*, 25, 129–32.

Popper, K. R. (1938). A Set of Independent Axioms for Probability. *Mind*, 47(186), 275–77.

—— (1955). Two Autonomous Axiom Systems for the Calculus of Probabilities. *The British Journal for the Philosophy of Science*, 6(21), 51–7.

—— (1959a). *The Logic of Scientific Discovery.* New York: Basic Books.

—— (1959b). The Propensity Interpretation of Probability. *British Journal for the Philosophy of Science*, 10(37), 25–42.

Rényi, A. (1955). On a New Axiomatic Theory of Probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 286–335.

Salmon, W. (1984). *Four Decades of Scientific Explanation.* Minneapolis: University of Minnesota Press.

Sober, E. (2010). Evolutionary Theory and the Reality of Macro Probabilities. In E. Eells and J. Fetzer (eds), *Probability in Science* (pp. 133–62). La Salle, Ill.: Open Court.