A Better Way of Framing Williamson's Coin-Tossing Argument, but It Still Does Not Work

Colin Howson*†

Timothy Williamson claimed to prove with a coin-tossing example that hyperreal probabilities cannot save the principle of regularity. A premise of his argument is that two specified infinitary events must be assigned the same probability because, he claims, they are isomorphic. But as has been pointed out, they are not isomorphic. A way of framing Williamson's argument that does not make it depend on the isomorphism claim is in terms of shifts in Bernoulli processes, the usual mathematical model of sequential coin tossing. But even so framed, the argument still fails.

1. Introduction. The so-called principle of regularity is the doctrine that only necessary truths and falsehoods should have the extreme probabilities 1 and 0, respectively. Within the field of the standard real numbers, outcome spaces of the size of the continuum like the closed unit interval pose an obvious problem for the principle because there can be at most countably many disjoint events of positive (standard) probability. But the hyperreal numbers, a field extension of the real numbers, seem to hold out the promise of overcoming this objection by allowing all those events to take positive infinitesimal probabilities.[1] Indeed a well-known result of Bernstein and Wattenberg (1969) shows that there is a measure that assigns positive infinitesimals over

Received February 2018; revised May 2018.

*To contact the author, please write to: London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK; e-mail: howson.colin@gmail.com.

†For helpful comments I thank the journal's anonymous reviewers and Tim Childers and the other members of the Institute of Philosophy, Czech Academy of Sciences.

1. There is some debate as to whether the hyperreals should be regarded as as 'real' as the reals themselves, which are unique up to isomorphism in the standard model of set theory. There are infinitely many hyperreal extensions, not all of which are order isomorphic unless the continuum hypothesis is assumed. If an inaccessible cardinal κ is as-

Philosophy of Science, 86 (April 2019) pp. 366–374. 0031-8248/2019/8602-0009$10.00
Copyright 2019 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/701957 Published online by Cambridge University Press

the points of the standard unit interval and differs by an infinitesimal from Lebesgue measure on the Lebesgue-measurable sets. The unit interval can be represented by the set of all infinite sequences of zeros and ones, the usual mathematical model of the 'experiment' of tossing a coin infinitely many times.[2] However, in a well-known paper Timothy Williamson (2007) claimed to show, using just that example, that the appeal to hyperreal probabilities cannot save regularity.[3]

Given that a familiar way of defining Lebesgue measure on the measurable subsets of [0, 1] employs the model of infinite independent sequences of tosses of a fair coin, Bernstein and Wattenberg's result described earlier might seem to refute Williamson *avant la lettre*, while in a paper published coincidentally in the same year as Williamson's, Frederik Herzberg (2007) showed how to assign positive infinitesimal probabilities to independent infinite sequences of tosses of a coin with an arbitrary fixed probability of landing heads.[4] But Williamson's argument has garnered some support in the literature (e.g., Hájek 2017), and in what follows I present the argument, argue that what appears to be a crucial premise that two relevant events are isomorphic fails, and show that nevertheless the events must have the same probability. Indeed, this is actually a theorem about a type of Bernoulli process. However, the theorem holds for standard-valued probabilities only, and it is not correct if they can take infinitesimal values. I will conclude by observing that what Williamson's coin-tossing example genuinely does show is that the hyperreals are not merely a device for simplifying and thereby solving standard proofs in mathematics,[5] as is often claimed even by their advocates, but that they can make real discriminations that are invisible when viewed through the lens of the real numbers alone.

---

sumed to exist (as talk of a standard model of set theory seems to presuppose), then there is a κ-saturated extension unique up to isomorphism (see Keisler [2009], 194–97, for explanation and details). Kanovei and Shelah (2004) have shown that a particular countably saturated ultrapower extension is even explicitly definable in ZFC (Zermelo-Fraenkel-plus-Choice) set theory. Nelson's (1977) very different axiomatic treatment, a conservative extension of ZFC, takes the reality of nonstandard numbers as implicit.

2. That 'experiment' might strike some as fanciful, but the 'almost surely' theorems held to be the glory of mathematical probability are of this type.

3. Williamson is not the only one to argue that appealing to hyperreals fails to save regularity. Pruss (2013) has a cardinality-based argument (but see Hofweber 2014a) and another (Pruss 2012) claiming that since infinitesimally valued probabilities are nonconglomerable in some countable partition, one can know beforehand with near certainty what one's conditionalization-updated probability of a suitable event A will be independently of the sample.

4. There is a brief discussion of Herzberg's construction in n. 14, below.

5. Their utility from this point of view was given an enormous boost by the seminal work of Loeb (1975). See also Albeverio et al. (2009, 63–67).

https://doi.org/10.1086/701957 Published online by Cambridge University Press

2. Williamson's Argument. A coin is flipped infinitely many times. Suppose H(1. . .) represents the event 'all the outcomes are heads', and H(2. . .) 'all the outcomes after the first are heads' (Williamson 2007, 4; his notation). Then according to Williamson, "H(1. . .) and H(2. . .) are isomorphic events. More precisely, we can map the constituent single-toss events of H(1. . .) one-one onto the constituent single-toss events of H(2. . .) in a natural way that preserves the physical structure of the set-up just by mapping each toss to its successor. H(1. . .) and H(2. . .) are events of exactly the same qualitative type; they differ only in the inconsequential respect that H(2. . .) starts one second after H(1. . .). That H(2. . .) is preceded by another toss is irrelevant, given the independence of the tosses. Thus H(1. . .) and H(2. . .) should have the same probability" (5).

From the last quoted sentence we infer that the common probability of H(1. . .) and H(2. . .) must be 0. For

$$\mathrm{P}(\mathrm{H}(1\ldots)) = \mathrm{P}(\mathrm{H}(2\ldots)|\mathrm{H}(1))\mathrm{P}(\mathrm{H}(1)), \tag{1}$$

where H(1) is the event 'the first toss lands heads'. So by independence

$$\mathrm{P}(\mathrm{H}(1\ldots)) = \mathrm{P}(\mathrm{H}(2\ldots))\mathrm{P}(\mathrm{H}(1)).$$

That is,

$$\mathrm{P}(\mathrm{H}(1\ldots)) = (1/2)\mathrm{P}(\mathrm{H}(2\ldots)).$$

But because of the presumed isomorphism of H(1. . .) and H(2. . .), we have

$$\mathrm{P}(\mathrm{H}(1\ldots)) = \mathrm{P}(\mathrm{H}(2\ldots)),$$

hence

$$2\mathrm{P}(\mathrm{H}(1\ldots)) = \mathrm{P}(\mathrm{H}(1\ldots)),$$

and so, noting that infinitesimals obey the same field rules as the real numbers themselves, we must have $\mathrm{P}(\mathrm{H}(1\ldots)) = 0$, contradicting regularity.

3. Why the Argument Fails. Williamson's argument raises several questions. Most prominently, it is not clear what his claim that the 'events' H(1. . .) and H(2 . . . .) are isomorphic can mean. Events, according to Kolmogorovian orthodoxy, are (measurable) subsets of an outcome space. Since Williamson's experiment consists in observing the outcomes of a coin flipped infinitely many times, the outcome space can be represented by the uncountably infinite set of all infinite sequences of ones and zeroes (1 for heads, 0 for tails). The event H(1. . .), 'all the outcomes are heads', defined in that space is clearly the singleton {x}, where x is the sequence all of which

https://doi.org/10.1086/701957 Published online by Cambridge University Press

are ones, and the event H(2. . .), 'all the outcomes after the first are heads', is equally clearly the pair {x, y}, where y is like x except that its first member is zero. But these two events H(1. . .) and H(2. . .) are two sets without structure, one of which has two members and the other just one, and for unstructured sets isomorphism reduces to merely having the same cardinality. Howson (2017) pointed this out and concluded that no appeal to isomorphism can justify inferring the equiprobability of H(1. . .) and H(2. . .).

It seems that Williamson is simply overlooking his own description of H(2. . .) as the event 'all tosses *after the first* land heads', defined in the sample space just described, and hence as a disjunctive event in that space and taking it, incorrectly, to be a (singleton) sequence of events in its own right.[6] Such a view of H(2. . .) does seem implicit in his claim that "we can map the constituent single-toss events of H(1. . .) one-one onto *the constituent single-toss events of H(2. . .)* in a natural way that preserves the physical structure of the set-up just by mapping each toss to its successor" (Williamson 2007, 5, my emphasis). But in that case how does Williamson reconcile a view of H(2. . .) as a singleton sequence with the step $P(H(1...)) = (1/2)P(H(2 ...))$? According to Benci, Horsten, and Wenmackers (2018, 528–30), Williamson is simply conflating two distinct experiments, or 'models', as they call them, one of which (their model A) records the outcomes of an infinite sequence of coin tosses starting at some given moment, and the other (model B) records as its *n*th outcome the $n + 1$th outcome as recorded by A. In model B, H(2. . .) is indeed now a singleton sequence containing ones only, thereby identical to H(1. . .) in model A and so justifying the identity $P_A(H(1...)) = P_B(H(2 ...))$. But, Benci et al. continue, "[Williamson] glosses this as $Prob(H(1...)) = Prob(H(2 ...))$, thus turning the probabilities involved into evaluations within the same model" (529).

The problem with Benci et al.'s account is that Williamson actually does distinguish their model A and model B, even if not in those precise terms, and still insists that $P(H(1...)) = P(H(2 ...))$. For in an alternative version of his argument, he considers two spatiotemporally distinct sequences of tosses of two 'qualitatively identical' coins, both yielding the outcomes 'all heads'. One of those sequences (clearly corresponding to Benci et al.'s model B) starts at the second toss of the other coin (corresponding to their model A):

6. As does Easwaran (2014). Weintraub and Hofweber agree with Williamson that H(1. . .) and H(2. . .) are isomorphic (although Hofweber puts the word in quotes), but each denies that they must have identical probabilities—Weintraub, because the supposed isomorphism does not preserve what she claims is a 'global' feature of the situation, that H(2. . .) is a proper subsequence of H(1. . .) (2008, 249), and Hofweber, simply because he thinks, in the light of Cantor's theory of transfinite cardinals, such a relation is not decisive (2014b, 12). But there are not two sequences of events but two events, and H(2. . .) is not a proper subsequence of H(1. . .).

https://doi.org/10.1086/701957 Published online by Cambridge University Press

"suppose that another fair coin, qualitatively identical with the first, will also be tossed infinitely many times at one second intervals, starting at the same time as the second toss of the first coin, all tosses being independent. Let $H^*(1...)$ be the event that every toss of the second coin comes up heads, and $H^*(2...)$ the event that every toss after the first of the second coin comes up heads" (Williamson 2007, 6).

Thus, in this version of Williamson's argument Benci et al.'s $P_B(H(2...))$ becomes Williamson's $P(H^*(1...))$. And while Williamson does indeed 'exploit the intuition that $Prob_A(H(1...)) = Prob_B(H(2...))$', he does not 'gloss' it as $Prob(H(1...)) = Prob(H(2...))$. Instead, he renders it in his own notation as "$H(1...)$ and $H^*(1...)$ should be equiprobable, because the probability that a coin comes up heads on every toss does not depend on when one starts tossing, and there is no qualitative difference between the coins" (Williamson 2007, 5).[7] Williamson's equating $Prob(H(1...))$ and $Prob(H(2...))$ arises not from any unconscious conflation of two models, therefore, but from a belief, as the following quotation makes quite clear, that $H(2...)$ is an infinite sequence in the same sample space as $H(1...)$, structurally identical to $H^*(1...)$ and hence—given his immediately preceding observation—to $H(1...)$: "But for the same reason $H^*(1...)$ and $H(2...)$ should also be equiprobable. *These two infinite sequences of tosses proceed in parallel*, synchronically, and there is no qualitative difference between the coins; in particular, that the first coin will be tossed once before the $H(2...)$ sequence begins is irrelevant.[8] By transitivity, $H(1...)$ and $H(2...)$ should be equiprobable: [hence] $Prob(H(1...)) = Prob(H(2...))$" (2007, 5, my emphasis).

But surely, it might be said in defense of Williamson, $H(2...)$ does have a straightforward representation as a single sequence of events, one sanctioned by formal probability theory itself. For where X_i is a two-valued random variable defined on $X = \{0, 1\}^N$ (usually written 2^N), and $N = \{1, 2, 3, ...\}$, with $X_i(x) = x_i$, $i = 1, 2, ...$, $H(2...)$ is the sequence $\langle X_2 = 1, X_3 = 1, ..., X_{n+1} = 1, ... \rangle$, a sequence moreover that clearly is a subsequence of $\langle X_1 = 1, X_2 = 1, ..., X_n = 1, ... \rangle$, that is, of $H(1...)$. Thus, the successor function $f : N \to N$, $f(i) = i + 1$, which sends '$X_i = 1$' to '$X_{f(i)} = 1$', really does appear to induce an isomorphic embedding of $H(2...)$ into $H(1...)$. So is Williamson right after all?

7. Benci et al. do note Williamson's second version of his argument, but the only significance they see in it is that "this allows even more freedom in choosing the sample space and embedding the events in it" (2018, 530).

8. Easwaran denies that intuition justifies equating the probabilities (2014, 20), but Williamson's argument was not based on intuition but on a specific claim of isomorphism ("But $H(1...)$ and $H(2...)$ are isomorphic events. . . . Thus [they] should have the same probability"; Williamson 2007, 5), which Easwaran does not address.

https://doi.org/10.1086/701957 Published online by Cambridge University Press

No. $\langle X_1 = 1, X_2 = 1, \ldots, X_n = 1, \ldots \rangle$ is not a sequence but a conjunction of formulas '$X_i = 1$', signified by the fact that in the set-theoretical language of measure theory it is represented as the set-theoretical intersection of all sets of the form $\{x : X_i(x) = 1\}$.[9] Correspondingly, the formal-logical representation of H(2. . .) is a disjunction of conjunctions of atomic sentences '$X_i = 1$', i = 1, 2, ..., in the minimal language of the experiment: it is '$(X_1 = 1 \& Y) \vee (X_1 = 0 \& Y)$', where Y is '$X_2 = 1 \& X_3 = 1 \& \ldots \& X_{n+1} = 1 \& \ldots$'. By contrast, H(1. . .) is a simple conjunction of atomic items. "But there are no such things as infinite conjunctions" I hear someone say. In fact there are, and they are well-behaved items of so-called infinitary languages extending first-order logic, obeying the de Morgan's laws, and so on.[10] The ordinary language of set theory makes these languages redundant, at least for measure-theoretic applications, which is why they are rarely if ever employed in those contexts; instead, countably infinite intersections and unions do routine service in defining infinitary events.

4. The Argument Redux. At the end of their discussion of Williamson's argument, Benci et al. remark that it "crucially turns on translation symmetry" (2018, 530). I assume (they do not elaborate) that they are referring to Williamson's seeming beliefs that H(2. . .) is just the sequence H(1. . .) shifted, or translated, forward one place and that such a shift must therefore preserve probabilities. Williamson's argument fails because H(2. . .) is not the sequence H(1. . .) translated, but there is something to it nonetheless. As we will now see there is a translation relating the two events, which can be shown to preserve probabilities without appealing to any spurious assumption of isomorphism and without conflating distinct probability models. On the contrary, only one is involved, called a *one-sided Bernoulli process with probability p*, typically seen as modeling an infinite sequence of coin tosses under the assumptions that (i) the tosses are independent and (ii), one might say in Williamsonese, 'qualitatively identical' in that there is exactly the same probability p of a head at each toss however far along the sequence one goes (the conditions i and ii are often abbreviated to 'iid', for 'independent, identically distributed', and they are exactly the formal conditions appealed to by Williamson in his argument). In the current case the parameter p is of course equal to 1/2.

9. This is essentially just basic model theory: negations homomorphically map to complements, and disjunctions to unions.

10. The closest to a first-order language is the infinitary language $L(\omega_1, \omega)$, which allows denumerably infinite conjunctions and disjunctions (hence ω_1) and finite strings of quantifiers (hence ω). It can be endowed with a set of logical axioms and an infinitary proof schema that are provably complete (for countable sets of formulas; see Bell 2016).

https://doi.org/10.1086/701957 Published online by Cambridge University Press

Formally, we are talking about a single Kolmogorovian probability space (X, F, P), where $X = 2^N = \{0, 1\}^N$, the set of all denumerably infinite sequences of zeroes and ones; F is the class of events defined in X, a σ-algebra of subsets of X generated by the cylinder sets (so all singleton sequences are events); and P is a product measure.[11] The distinctive additional feature of such a Bernoulli process is a map T from X to X, the *shift*, or *left shift*, which moves each individual outcome in any sequence $x \in X$ one place to the left.[12] It is easily proved that the shift operation is probability preserving: that is, for any event C in F, $P(C) = P(T^{-1}C)$. But this implies that P(H(1...)) = P(H(2 ...)). For T(0, 1, 1, . . .), that is, the shift of the sequence zero followed by all ones, is the sequence (1, 1, 1, . . ., 1, . . .), that is, the singleton sequence in the extension of H(1. . .), and the shift of that sequence is itself. Hence, T^{-1}(1, 1, ... , 1, . . .) is the pair {(0, 1, 1, ... , 1, . . .), (1, 1, 1, ... , 1, . . .)}, that is, H(2. . .). Hence, by the measure-preservation property of T we have P(H(2 ...)) = P(H(1...)).[13] So we seem to have secured Williamson's premise that P(H(1...)) = P(H(2 ...)) as a theorem about a single probability space, appealing only to the iid properties of Bernoulli processes, with no appeal to any spurious assumption that H(1. . .) and H(2. . .) are isomorphic.

QED for Williamson? Not at all. What his argument reveals is that measure invariance is an artifact of standard probabilities where infinitary events like H(1. . .) and H(2. . .) are concerned, since all infinite sequences with probability 1/2 of heads have limiting standard probability 0, with that of H(2. . .) being the sum of two zeros. By contrast, if P can take values in the hyperreal unit interval, as finer measures of possibility than the reals, then those probabilities can quite consistently be infinitesimal and positive. Indeed, the sequence $(2^{-n})_{0\leq n}$ itself, factored by some nonprincipal ultrafilter on N, would be as good as any other infinitesimal value of the probability of H(1. . .). But while there may be indeterminacy of specific values for P(H(1. . .)) and P(H(2. . .), the relation between them is not at all indeterminate.[14] For Williamson's own proof

11. These are sets of sequences that agree on finitely many indices. Thus, $\{x \in X: x_{i(1)} = a_1, \ldots, x_{i(k)} = a_k\}$, for any $k \in N$, is a cylinder set.

12. So $T(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)$.

13. This might seem to go some way to answering Hofweber's question: "Why is the chance of H(1. . .) supposed to be the same as that of H(2. . .)? This is, of course, very plausible for finite events, i.e., events that have only finitely many constitutive parts. But once we deal with infinite events, like an infinite sequence of coin tosses that has infinitely many constitutive events as parts, i.e., the individual coin tosses, this is far from clear" (2014b, 11).

14. Herzberg uses the transfer principle to pass, *salva veritate*, between the standard Bernoulli distribution of the number of heads in a finite sample from the population of all standard infinite sequences of heads and tails (ones and zeroes) and an internal Bernoulli distribution on the internal algebra generated by the sequences of some arbi-

https://doi.org/10.1086/701957 Published online by Cambridge University Press

tells us that whatever infinitesimal probability value H(2. . .) is assigned, it has to be twice that of H(1. . .), and if we like we can even construe this as actually measuring the infinitesimal difference between the extensions of H(1. . .) and H(2. . .), that is, between the singleton sequence in X consisting of ones only and the pair of sequences one of which is the sequence consisting of ones only and the other is the same sequence but starting with a zero instead of a one.

But allowing infinitesimal-valued probabilities does seem to entail violations of what have hitherto been thought of as fundamental symmetry principles,[15] not just measure invariance under shifts in Bernoulli processes (although of course all finite sequences remain invariant) but also translation invariance of Lebesgue measure on the measurable subsets of [0, 1], rotational invariance on the circle, and so on. If one accepts hyperreals as legitimate entities in their own right, however, and not merely Leibnizian 'ideal elements' simplifying standard definitions and proofs,[16] then those classical results do become an artifact of restricting values to the real number system. But, to the extent to which the real and complex numbers remain the working tool of mathematical physics, those principles are unlikely to be dethroned.

5. Conclusion. Not only can regularity be consistently maintained in the context of Williamson's example, but we see that the hyperreals can provide a finer analysis of what happens at standard limit points, where the standard lens fails to reveal fine enough differences of structure. That H(2. . .) is a composite event in the domain of P while H(1. . .) is a singleton sequence is a real, if infinitesimal, difference between them, which the hyperreals capture and which the standard reals cannot, blurring as they do the probabilities of all the singleton sequences as the zero limit of (2^{-n}).[17]

---

trary hyperfinite length h. Under the restriction map, the infinitesimal measure of a sequence x in $\{0, 1\}^h$ becomes the probability of the corresponding sequence in $\{0, 1\}^N$, maximized among $p \in [0, 1]$ by the standard part of the internal relative frequency of tails in x (Herzberg 2007, app. B). Herzberg's object was to define a maximum-likelihood ranking for 'probability laws'—in this case values of p in [0, 1]—which all assign standard probability 0 to the infinite sequences of outcomes (the problem had been raised initially by David Lewis), but in the Williamson example there is no concern to find a 'best-fit' solution for the infinitesimal probabilities, and in principle any could be assigned to H(1. . .) (for an explanation of the transfer principle and the difference between internal and external sets and structures, see Albeverio et al. [2009], 15–22).

15. See Barrett (2010) for a discussion of this type of objection.

16. To take a stock example, compare the standard ε, δ definition of continuity at a point x with the nonstandard one: f is continuous at x if, for every y infinitely close to x, *f(y) is infinitely close to *f(x).

17. Compare Hofweber: "the situation preserves our prima facie judgment that on the one hand the chances are the same (up to a small infinitesimal difference), but on the other hand they are different (since there is a small, infinitesimal difference)" (2014b, 10).

https://doi.org/10.1086/701957 Published online by Cambridge University Press

REFERENCES

Albeverio, S., J. E. Fenstad, R. Høegh-Krohn, and T. Lindstrom. 2009. *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*. New York: Dover.
Barrett, M. 2010. "The Possibility of Infinitesimal Chances." In *The Place of Probability in Science: in Honor of Ellery Eells (1953–2006)*, ed. E. Eells and J. H. Fetzer, 65–79. Dordrecht: Springer.
Bell, John L. 2016. "Infinitary Logic." In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. https://plato.stanford.edu/archives/win2016/entries/logic-infinitary/.
Benci, V., L. Horsten, and S. Wenmackers. 2018. "Infinitesimal Probabilities." *British Journal for the Philosophy of Science* 69:509–22.
Bernstein, A. R., and F. Wattenberg. 1969. "Nonstandard Measure Theory." In *Applications of Model Theory to Algebra, Analysis and Probability*, ed. W. A. J. Luxemburg, 171–86. New York: Holt, Rinehart & Winston.
Easwaran, K. 2014. "Regularity and Hyperreal Credences." *Philosophical Review* 123:1–41.
Hájek, A. 2017. "Staying Regular." Unpublished manuscript, Australian National University.
Herzberg, F. 2007. "Internal Laws of Probability, Generalized Likelihoods and Lewis' Infinitesimal Chances: A Response to Adam Elga." *British Journal for the Philosophy of Science* 58:25–43.
Hofweber, T. 2014a. "Cardinality Arguments against Regular Probability Measures." *Thought: A Journal of Philosophy* 3:166–75.
———. 2014b. "Infinitesimal Chances." *Philosopher's Imprint* 14:1–34.
Howson, C. 2017. "Regularity and Infinitely Tossed Coins." *European Journal for the Philosophy of Science* 7:97–100.
Kanovei, V., and S. Shelah. 2004. "A Definable Non-standard Model of the Reals." *Journal of Symbolic Logic* 69:159–64.
Keisler, H. J. 2009. *Foundations of Infinitesimal Calculus*. Madison: University of Wisconsin Press.
Loeb, P. A. 1975. "Conversion from Nonstandard to Standard Measure Spaces and Applications in Probability Theory." *Transactions of the American Mathematical Society* 211:113–22.
Nelson, E. 1977. "Internal Set Theory: A New Approach to Nonstandard Analysis." *Bulletin of the American Mathematical Society* 83:1165–98.
Pruss, A. R. 2012. "Infinite Lotteries, Perfectly Thin Darts and Infinitesimals." *Thought: A Journal of Philosophy* 1:81–89.
———. 2013. "Probability, Regularity, and Cardinality." *Philosophy of Science* 80:231–40.
Weintraub, R. 2008. "How Probable Is an Infinite Sequence of Heads?" *Analysis* 68:247–50.
Williamson, T. 2007. "How Probable Is an Infinite Sequence of Heads?" *Analysis* 67:173–80.

https://doi.org/10.1086/701957 Published online by Cambridge University Press