

Probability and Randomness

Antony Eagle

Random outcomes, in ordinary parlance, are those that occur haphazardly, unpredictably, or by chance. Even without further clarification, these glosses suggest an interesting connection between randomness and **probability**, in some of its guises. But we need to be more precise, about both probability and randomness, to understand the relationship between the two subjects of our title.

It is a commonplace that there are many sorts of probability; and of each, we may ask after its connection with randomness. There is little systematic to say about randomness and **credence**; even rational degrees of belief may be certain about a random outcome, and uncertain about a non-random one. More of substance can be said about connections between randomness and **evidential probability** – particularly according to Solomonoff's version of objective Bayesianism, which uses Kolmogorov complexity (§2.2) to define a privileged prior **algorithmic probability** (Solomonoff, 1964; see also Li and Vitanyi, 2008; Rathmanner and Hutter, 2011).

For reasons both of concision and focus, in the present chapter I set those issues aside, to concentrate on randomness and **physical probability**, or **chance**: probability as a physical feature of certain worldly processes.¹ A number of philosophers have proposed an intimate connection between randomness and chance, perhaps even amounting to a reduction of one to the other. I explore, with mostly negative results, the prospects for such views; and discuss some weaker but still interesting ways in which randomness bears on chance. I begin by clarifying and distinguishing a number of kinds of randomness.

¹ For more on the nature of chance, see the entries by Frigg, Gillies, La Caze, and Schwarz in the present volume.

1. Different kinds of randomness: product and process

The paradigm sort of case that might involve randomness is a series of tosses of a fair coin, or similar chance device. A **typical** example of such a series will be a disorderly and patternless sequence of outcomes (Heads and Tails, or 0s and 1s), because it will have been produced by a genuinely chancy process.

However, it is at least conceivable that a chancy process could produce an orderly series of outcomes – a lucky series of coin tosses may land all heads. It is similarly conceivable that a disorderly series of outcomes could have been produced by a non-chancy process. It seems arbitrary to believe that, in the more typical case, it is really just one of these features that is responsible for the presence of randomness. Hence we might find it useful to regiment our terminology, distinguishing different sorts of randomness in virtue of the different features that can prompt ascriptions of randomness.

A sequence of 100 consecutive Heads is *prima facie* not random, even if produced by fair tosses of a fair coin. Yet a sequence of 100 mixed Heads and Tails from the same fair coin with no discernable pattern is *prima facie* random. The difference here in the *prima facie* appearances cannot be grounded in the underlying chances, which are the same in both cases (each has chance $1/2^{100}$). The appearance is grounded rather in the intrinsic disorderliness (or otherwise) of the outcome sequences themselves. Let's say that a **random product** is a series of outcomes of some repeated process that is disorderly and irregular, regardless of how it was produced. Moreover, let's say that a **random process** is one involving genuine chance, in which some of the possible outcomes of the process have non-trivial objective chances of coming to pass. There is precedent for this regimentation, assuming that 'probabilistic laws' support objective chances:

I group *random* with stochastic or chancy, taking a random process to be one which does not operate wholly capriciously or haphazardly but in accord with stochastic or probabilistic laws. (Earman, 1986: 137; my italics)

This proposed regimentation diverges to a certain extent from intuition, in some cases where the chances are nearly 1 or nearly 0. An exercise of a quite reliable skill, like catching a ball, can be a random process, if there is some genuine chance of failing to catch the ball. It is awkward to characterize the occurrence of the – entirely expected – outcome of such a process as random. In the case of human action, perhaps our judgments are being confounded by the fact that, while my catching a ball is partly a matter of chance, it is not solely due to chance. So instead we should say: a process

is random if at least some of its outcomes are *purely* a matter of chance – they happen not just ‘in accord with’ probabilistic laws, but are adequately explained by citing probabilistic laws.

Even refined in this way, the regimentation classifies some highly probable outcomes as random – any event which is best explained by citing a chancy law of nature. Suppose we consider the outcomes of this repeated chance process: roll two fair dice, record ‘1’ if double six comes up, and ‘0’ otherwise. In the long run, the frequency of 1s in the outcome sequence will be (close to) $1/36$. Because of this, the outcome sequence will be quite orderly: it will be almost all 0s, with a few 1s scattered here and there. Yet this process is purely chancy (assuming that dice rolls are), and so is random. Ordinary intuition appears to be split on such cases. On the one hand, the outcomes are quite regular, and even predictable (if you always bet on ‘0’, you’d be almost always right). On the other, the best opinion to have about a given outcome is simply to set your credence in line with the chances, and the best explanation of why some outcome came up in a particular case just cites the chances. Once we make the distinction between product and process randomness, we can vindicate both these intuitive judgments: this scenario involves a random process, but produces something non-random. Those who would refuse to call biased chance processes ‘random’ are – I venture – letting views about the randomness of the product drive their opinion of randomness of the process. We have a more useful taxonomy if we keep these two categories apart.

On my suggested regimentation, there is a definitional link between process randomness and physical probability. The relationship between product randomness and physical probability is less clear. The typical outcome of a repeated random process will be a random product; but it is conceivable both that some random products are not produced by chance, and, that some repeated trials of random processes don’t yield random products. Having regimented our vocabulary so as to be able to make such distinctions, it is clear that we can neither **define** randomness of a product in terms of randomness of the process generating it, nor *vice versa*.

But for all that, it may be that we can **reduce** one to the other, finding some metaphysical account which either grounds the randomness of a sequence in the randomness of some process, or conversely. This converse reduction of process randomness to product randomness is the characteristic idea behind the **frequentist** account of chance (von Mises, 1957; La Caze, this volume). Von Mises claims, in effect, that a process is chancy iff it would produce a random sequence of outcomes which exhibit stable frequencies; those frequencies are to be identified with the chances of those outcomes. Frequentists often proposed a definitional link between

chance and frequencies. They needn't have; they need only claim that frequencies in random sequences **actually realize** the chances.

I begin the rest of this chapter by exploring proposals in the vicinity of frequentism in this respect: proposals that attempt to ground process randomness – at least in part – in facts about product randomness. That task involves giving a theory of product randomness that makes no reference to chance or probability, a theory that characterises disorderliness of a sequence directly. In §2, I discuss first von Mises' attempt to provide such a theory, then related theories due to Martin-Löf and Kolmogorov. In §3, I examine whether there is any viable prospect of using randomness of sequences to ground facts about chance. My conclusions are largely negative. At the end of §3, I discuss the weaker claim that product random sequences can be good **evidence** for chances. Above, I noted our ability to distinguish orderly and disorderly sequences, independently of the randomness or otherwise of the underlying process. If we can come up with a sensible account of product randomness, we may be able to use the existence of a product random sequence as evidence for the existence of a random process. It seems that if the outcome sequence is product random, it will be unpredictable; and accordingly it will be epistemically irrational to seek certainty in advance concerning future outcomes. In such a situation, probabilistic theories look particularly attractive, providing reliable information about the future when certainty cannot be had. Given familiar difficulties concerning the epistemology of chance, the possibility that product randomness is evidentially significant for chance is a good reason to explore product randomness more deeply.

* * *

The distinction between process and product randomness doesn't exhaust the theoretically useful categories in this vicinity. Typically, random processes yield **unpredictable outcomes**, in the sense that gathering additional information concerning already observed outcomes doesn't generally improve predictive accuracy compared to estimates based solely on the underlying chances. Likewise, a random product involves unpredictable outcomes, in that information about earlier outcomes in the sequence doesn't provide information about later outcomes in the sequence – for if it did, that would be a pattern (perhaps somewhat subtle) in what is supposed to be a disorderly sequence. (We might be mystified by the process involved, but a sequence of coin tosses which reliably came up heads every prime-numbered toss would be at least partially predictable.) So both product and process randomness

tend to lead to unpredictability.² But the converse isn't true. A process may be unpredictable because we can't figure out how to model it, even though it is perfectly deterministic and non-chancy. (Some examples like this are discussed in §3.2 below.) Likewise, a product may be unpredictable, even though there is an underlying pattern or order, because that pattern is too difficult to discern.

Elsewhere, I've argued that it would be theoretically elegant to identify randomness with unpredictability, on the grounds that only such a liberal identification would capture all cases of randomness (Eagle, 2005). My basic argument was that product and process randomness are each too narrow to apply to everything characterised as random by the various special sciences. The notion of unpredictability I offered was, however, explicated just in terms of credence, and didn't say much about how these 'random' phenomena justify our subjective probability assignments. In focussing on the connections between randomness and physical probability in this chapter, I will discuss some outcomes that are unpredictable because they are process and/or product random. But I will neglect the class of unpredictable outcomes more generally, because there is little both systematic and true that can be said about the relationship between unpredictability and physical probability.

I also largely set aside another interesting notion in the vicinity: **pseudorandomness**. Like product randomness, pseudorandomness is a property of sequences. Informally, a sequence is pseudorandom if it *looks* disorderly. More precisely, a sequence will be pseudorandom if it passes sufficiently good tests for randomness, in that 'no efficient algorithm can distinguish [pseudorandom] output from a truly random sequence' (Vadhan 2012: §1.1). So a pseudorandom sequence is not product random, because there exists some pattern underlying the sequence. But that pattern is not able to be efficiently exploited – so for all practical purposes, the sequence is as good as random. The interest of pseudorandomness lies in the question as to whether the use

² The same sorts of worries about the randomness of outcomes of very high and very low chance arise again here. It is very plausible to think that if your best estimate of whether an outcome will happen is no better than chance, then you are not able to predict that outcome. But nearly-extreme-chance outcomes certainly appear predictable; and indeed, it is often possible to know propositions that have high but not extreme chance (Hawthorne and Lasonen-Aarnio, 2009). So it can turn out (i) that the best credence to have in a high chance outcome is just equal to its chance, which makes that outcome unpredictable; but (ii) that for all practical purposes simply assuming that it will occur is the best strategy; and (iii) it may even be that, though unpredictable, one can know the outcome will occur. This is one more nice illustration of the sometimes uncomfortable fit between traditional and Bayesian epistemology.

of pseudorandom sequences might *replace* the use of random sequences in many applications, for example in cryptography. But since pseudorandom sequences are, by construction, neither produced by chance nor genuinely patternless, they are not random in the core sense that concerns us here.

2. Algorithmic Randomness

Laplace observed that, when tossing a coin,

if heads comes up a hundred times in a row, then this appears to us extraordinary, because the almost infinite number of combinations that can arise in a hundred throws are divided in regular sequences, or those in which we observe a rule that is easy to grasp, and in irregular sequences, that are incomparably more numerous. (Laplace, 1951: 16–7)

Modern theories of product randomness take up two themes from Laplace – first, that random sequences are not governed by a rule (whether or not ‘easy to grasp’); second, that random sequences are more numerous than non-random ones.

Following the literature on **algorithmic randomness**, we will restrict our attention to binary sequences of outcomes, like sequences of outcomes of tosses of a fair coin. The mathematical literature on random sequences is extensive; for further details, see Eagle 2013; Dasgupta 2011; Li and Vitanyi 2008; Nies 2009; and references therein.

As mentioned above, von Mises’ version of frequentism required, to avoid circularity, a characterisation of product randomness that did not involve an antecedently given probability function. The characterisation he offered applies only to infinite sequences, and it is convenient to begin by considering these. In fact, we will confine our attention to infinite binary sequences (sequences of 0s and 1s), the sort that might model an infinite series of coin tosses. The set of all infinite binary sequences we call the **Cantor space**.

2.1 Infinite Random Sequences: Von Mises’ Approach

Von Mises’ approach illustrates one of the Laplacean themes: the **unruliness** of random sequences. Considering his example of regular milestones spaced along a road, of which a probabilistic account would be inappropriate, he claims that

the essential difference between the sequence of the results obtained by casting dice and the regular sequence of large and small milestones consists in the possibility of devising a method of *selecting the elements*

so as to produce a fundamental change in the relative frequencies. (von Mises, 1957: 24)

The rule governing a non-random sequence can be exploited to construct a method to select a biased partial sequence from the original sequence: for example, the rule *select every element which occurs after nine small stones* can be used to select a partial sequence of the milestones in which the limit relative frequency of large stones is 1, even though we never use that attribute directly in selecting outcomes for the subsequence. However, in a random sequence, claims von Mises, we are unable to use such a method – the only way to obtain a biased partial sequence of outcomes is to choose outcomes based on their attributes directly.

Accordingly, he offers a characterisation of random sequences as those in which the **limiting frequencies are invariant in partial sequences:**

these limiting values must remain the same in all partial sequences which may be selected from the original one in an arbitrary way. Of course, only such partial sequences can be taken into consideration as can be extended indefinitely, in the same way as the original sequence itself. Examples of this kind are, for instance, the partial sequences formed by all odd members of the original sequence, or by all members for which the place number in the sequence is the square of an integer, or a prime number, or a number selected according to some other rule, whatever it may be. The only essential condition is that the question whether or not a certain member of the original sequence belongs to the selected partial sequence should be settled *independently of the result* of the corresponding observation, i.e., before anything is known about this result. We shall call a selection of this kind a *place selection*. The limiting values of the relative frequencies in a collective must be independent of all possible place selections. By place selection we mean the selection of a partial sequence in such a way that we decide whether an element should or should not be included without making use of the attribute of the element, i.e., the result of our game of chance. (von Mises, 1957: 24–5)

It is trivial to observe that – without further restrictions on what kind of place selections are admissible – there will be no random sequences, because no limit frequencies are invariant under arbitrary place selections. For if a place selection is

just any function f from natural numbers into $\{0,1\}$,³ then there will be a place selection that selects a biased subsequence, since

any increasing sequence of natural numbers $n_1 < n_2 < \dots$ defines a corresponding selection rule, ... given an arbitrary sequence of 0s and 1s ... there is among the selection rules the one which selects the 1s of the given sequence, so the limit frequency is changed (Martin-Löf, 1969a: 27)

Von Mises does implicitly place restrictions on appropriate place selections. The empirical basis for von Mises' claim that there exist sequences which are appropriate for probabilistic treatment lies in the **principle of the impossibility of a gambling system** (von Mises, 1957: 25): in some systems, there is no recipe for deciding when to bet on outcomes that ensures more successes than chance alone. (The 'gambler's fallacy' is an unusually simple gambling system, but its lack of success is entirely representative.) The 'recipe' *bet only on 1s* does exist, in some abstract mathematical sense, but it is not one that can be followed by anyone. So von Mises' implicit restriction is to those place selections that can be genuinely implemented by a prospective gambler, making use only of information in the gambler's possession (information about previous outcomes). Von Mises made his original proposal before Turing and others made the notion of an **algorithm** precise, but it was very natural to retrospectively read into his work a restriction to **computable** place selections:

To a player who would beat the wheel at roulette a system is unusable which corresponds to a mathematical function known to exist but not given by explicit definition; and even the explicit definition is of no use unless it provides a means of calculating the particular values of the function. ... Thus a [place selection] should be represented mathematically, not as a function, or even as a definition of a function, but as an effective algorithm for the calculation of the values of a function. (Church, 1940: 133)

Where x is some infinite sequence, let ' $x \upharpoonright i$ ' denote the initial segment of x of length i . Church proposes that an implementable place selection is an **effectively computable function** $\varphi : \langle x \upharpoonright (i-1), i \rangle \mapsto \{0,1\}$ that takes the value 1 infinitely many times (so always selects an infinite partial sequence from an infinite sequence). Let $x[\varphi]$ be the infinite partial sequence consisting of any x_j such that $\varphi(x \upharpoonright (j-1), j) = 1$. According to Church's reconstruction of von Mises, a sequence is random iff for every effectively computable place selection φ , the limiting frequency of every outcome in $x[\varphi]$ equals

³ I.e., if $f(n) = 1$, then and only then select the n^{th} member of the original sequence to belong to the partial sequence.

the limiting frequency of that outcome in x . More informally: there exists no effective method for choosing a subsequence of a random sequence that is biased with respect to the original frequencies. The random sequences in this sense are immune to gambling systems: there are no **followable rules** that enable even fallible and unreliable prediction of patterns of outcomes in a random sequence, or any partial sequence derived from it, except strictly in accordance with the statistical predictions of its probabilistic laws.

The question of whether there are any random sequences in this revised sense still needs to be addressed. As there are many fewer effectively computable functions than functions, it is easier to satisfy the requirement of invariance of limit frequencies under all computable place selections. In fact, there do exist infinite random sequences in this refined von Mises/Church sense. Indeed **almost all** infinite binary sequences have the feature of being frequency-invariant under computable place selections. More precisely, the set of von Mises/Church random sequences forms a **measure one** subset of the Cantor space. This follows quickly from a result due to Wald (1938), that any **countable** set of arbitrary place selections (whether computable or not) defines a set of von Mises random sequences which is measure one, given that there are only countably many computable place selections. Pursuing the unruly behaviour of random sequences has now led us to the other aspect of random sequences Laplace noted: the von Mises/Church random sequences are much more numerous than the even partially rule-governed sequences.

2.2 Infinite Random Sequences: The Typicality Approach

What if we had started instead from Laplace's other claim: that irregular sequences are more numerous than regular sequences? Suppose we know that a fair coin is to be tossed repeatedly; what sort of outcome sequence should we expect? Well, we ought to expect that heads should occur about as often as tails; that strings of heads and tails of equal length should occur about as frequently as each other; and so on. These are features that are **typical** of sequences generated by this sort of chance setup. Even if we have very little confidence in any specific sequence of outcomes occurring, we should be confident that some typical sequence will occur. This confidence derives ultimately from the fact that typical sequences are much more common than atypical ones.

This last claim needs to be handled with some care. Given an uncountably infinite set, like the Cantor space, there are many ways of measuring the size of its subsets. We want the typical sequences to form a measure one set of sequences, but *under which measure?* This is particularly important when we consider biased binary processes. The typical outcome of a series of flips of a coin biased to heads will be a decidedly

atypical product of a series of fair coin flips. The standard way of approaching this issue is to let the probability function associated with the underlying binary process determine a measure over the Cantor set; thus typicality of a sequence is relative to a probability function (Martin-Löf, 1966; Gaifman and Snir, 1982).

Obviously, and in contrast to the von Mises/Church account, such approaches to randomness rely on antecedent probabilities, and thus are generally unsuited to the reductive project we're considering. But there is a special case, where we define a measure not from an underlying chance, but from symmetries in the Cantor space itself. The measure thus produced is a very natural one over the Cantor space: the **Lebesgue measure** λ . The key idea is that the set of all sequences which begin with a given initial subsequence x_n of length n has Lebesgue measure $1/2^n$. (So the set of sequences beginning with a 1 has measure $1/2$, the set of sequences beginning 010 has measure $1/8$, etc.) Every particular infinite sequence $x \in X$ has Lebesgue measure zero (it must have measure smaller than any set of sequences sharing an initial subsequence with it, and $\lim_{n \rightarrow \infty} \frac{1}{2^n} = 0$). So indeed does any countable set of infinite binary sequences. The set of von Mises/Church random sequences has Lebesgue measure one, because its complement – the set of effectively computable binary sequences – is countable, connected as it is with the countable set of effectively computable functions, and thus has Lebesgue measure zero. The von Mises/Church random sequences are typical with respect to the Lebesgue measure over the Cantor space. (Henceforth, I will simply say they are **typical**.)

A typical sequence has some **measure one property**: one that is shared by almost all the sequences. But there are lots of distinct measure one properties. For example, the property of having 0 and 1 occur equally frequently is a property of almost all sequences. It is also, intuitively, a property that a random sequence should have. Intuitively, a random sequence should have the stronger property of **Borel normality** (Borel, 1909): every string of 1s and 0s of length n occurs equally often in x . (So 010 occurs as often as 101 and 111, etc.) Borel normality is also a measure one property. Indeed, every measure one property that has been explicitly considered is one that, intuitively, random sequences have. Generalising from this, Ville (1939) proposed that a random sequence should be a member of *all* measure one sets of sequences – the random sequences are **typical in every respect**.

The typicality approach, in this crude form, cannot work. For each individual sequence has measure zero, so its **complement** with respect to the Cantor space has measure one and excludes it; so the intersection of *all* sets of measure one is the empty set. No sequence is typical in every respect.

This is reminiscent of the way in which consideration of arbitrary functions trivialised von Mises' original theory of place selections. Martin-Löf (1966) proposed an analogous solution: rather than considering arbitrary measure one sets, he suggests that the random sequences should satisfy all **effective** measure one properties. A set has Lebesgue measure zero if there is a sequence of sets which converge to it, such that the measures of the sets converge to zero. If that sequence of sets is **effectively computable** (so that there is a computable function which computes what each member of the sequence is), then the set has effective measure zero. A sequence is **Martin-Löf random** iff it is not a member of any effective measure zero set – i.e., has all effective measure one properties. The Martin-Löf random sequences are **typical in every effectively determinable respect**. (More precisely: they are not effectively determinable to be atypical.) All the properties we have considered so far – Borel normality, non-computability, and the property of symmetric oscillation considered at the end of this section – are effective measure one.

Martin-Löf was led to this idea by considering **significance tests** in statistics. An experimental outcome prompts a hypothesis to be rejected at a significance level α if the outcome falls into a previously specified critical region. A sequence which falls into an effective measure zero set is one that would effectively yield a statistically significant result at arbitrarily high levels of significance. That is, it would prompt us to reject the null hypothesis that the sequence is random with arbitrarily high significance, and so must really be non-random. In this reformulation, Martin-Löf's proposal is that a sequence is random iff it passes all recursive statistical tests for randomness.

To show that there are Martin-Löf random sequences, Martin-Löf proves first that there is a **universal** significance test: there exists an effectively specifiable sequence of sets $U = U_1, \dots$, such that for any other significance test $G = G_1, \dots$, there exists a constant c such that for all i , $G_{i+c} \subseteq U_i$. If a sequence is rejected by some test at some significance level, it will also be rejected by U at some related significance level. Since the non-random sequences are those which, for any significance level, are rejected by some test, they will all be rejected by the universal test too. This universal test thus directly establishes that the set of Martin-Löf non-random sequences has effective measure zero; so Martin-Löf random sequences exist and collectively have effective measure one.

* * *

What is the relationship between von Mises' approach and the typicality approach? Let's begin by considering another way to think about random sequences. Each

random sequence x corresponds to a **random walk** on a line, in which the walker begins at some starting point, and moves one unit left at stage i if the i -th element of x is 0, and moves to the right otherwise. The behaviour of this trajectory can be used to illustrate properties of random sequences. For example, the typical random walk ends up back at the starting point, since it will have moved left exactly as often as right. This corresponds to an infinite binary sequence having a limit relative frequency of $1/2$ for 0 and 1.

Intuitively, a genuinely random walk should also cross the starting point often enough so that it spends as much time, in the long run, on the left of its starting point as the right.⁴ A walk that never spent any time to the left of the starting point has too much structure to be truly random. Happily for Martin-Löf's theory, this property of **symmetric oscillation** does have effective measure one (Dasgupta, 2011: §3.4).

But Ville showed there exist infinite binary sequences which are von Mises/Church random, and yet have biased initial segments: while the limit frequency of 1s in the sequence and all infinite subsequences is $1/2$, the limit is approached 'from above', and the frequency of 1s in every finite initial subsequence is $\geq 1/2$ (Ville, 1939; see also Lieb et al., 2006; van Lambalgen, 1987). Such sequences have the same limit as the genuinely random sequences, but are not random in *how* they approach that limit. So there is a property that, intuitively, random sequences have (and that Martin-Löf random sequences have, in agreement with intuition), that some von Mises/Church random sequences lack. So while von Mises/Church randomness is necessary for Martin-Löf randomness, it is not sufficient (see also van Lambalgen, 1987: §4).

2.3 Finite Random Sequences

A major difficulty facing Martin-Löf's proposal is that it cannot accommodate random finite sequences, or **strings**. Every string can be effectively produced by some Turing machine (even if not particularly *efficiently*), so no string can be even von Mises/Church random.

But while the contrast between sequences exhibiting effectively exploitable patterns and 'completely lawless' random sequences does not apply to finite sequences, a related contrast does apply: between those strings which can be efficiently described,

⁴ Not only should a random sequence have a limit relative frequency of 1s equal to $1/2$; it should approach this limit by having equally as many finite initial subsequences in which the relative frequency is below $1/2$ as those in which the relative frequency is above $1/2$. The idea that random processes should spend equal time in all equally sized outcome states (the states here are *being on the left of the origin* and *being on the right*) is related to ergodicity and its relatives, discussed in §3.2 below.

and those cannot. This leads us to the idea that finite random sequences, like their infinite cousins, are patternless, or disorderly, and cannot be predicted (or exploited) by a simple rule.

Kolmogorov was the first to connect this idea of disorder with **incompressibility** (Kolmogorov, 1963; Kolmogorov and Uspensky, 1988). The informal idea is that a pattern in a string can lead to a short description of its contents, while disorderly strings cannot be described more effectively than by stating their contents. So while the string consisting of 1000 alternating 1s and 0s can be simply specified (I just specified it in nine words), some string consisting of 1000 1s and 0s with no pattern is most effectively presented by just listing the 1000 digits comprising it. This can be made more precise by thinking about **compressibility**.

If f is a **computable encoding function** from strings to strings, we say that a string δ is an **f -description** of a string σ iff $f(\sigma) = \delta$. A string σ is **compressed** by f if there is an f -description δ of σ such that $|\delta| < |\sigma|$, where ' $|\varphi|$ ' denotes the length of the string φ . We may define the **f -complexity** of a string σ , $C_f(\sigma)$, as the length of the shortest string δ that f -describes σ . A string σ is **random relative to f** iff it is f -incompressible, that is, if the f -complexity of σ is roughly equal to $|\sigma|$.

This gives us randomness relative to a fixed algorithm f . It is in the spirit of the Laplacean ideas with which we began; since strings random relative to f are incompressible, they obey no rule – at least, no rule that f can exploit to describe them compactly. The random strings are also more numerous. Assuming that a useful encoding f will produce an f -description of a string σ that is at least $j \geq 1$ shorter than σ , very few strings usefully compress: a proportion of at most 2^{-j} strings of a given length. Even with the most pitiful amount of compression, $j = 1$, we see that at most half the strings of a given length can be compressed by any algorithm f ; and the compressible (non-random) strings are sparser the more compression we demand.

An encoding f is **at least as good** as g iff there is some constant k_g such that for any string σ , $C_f(\sigma) \leq C_g(\sigma) + k_g$, so that the f -complexity of any string is within k_g of the g -complexity (where k_g is independent of σ). Kolmogorov showed that there is an encoding algorithm f which is at least as good, in this sense, as any other algorithm (Kolmogorov, 1965; Chaitin, 1966; Martin-Löf, 1969b). Kolmogorov called such an algorithm **asymptotically optimal**, because, since k_g is fixed independently of σ , it becomes asymptotically negligible as $|\sigma|$ increases.

Fixing on an asymptotically optimal function u , we define **complexity** simpliciter: $C(\sigma) = C_u(\sigma)$. Since u is optimal, it is at least as good as the identity function; it

follows that there exists a constant k such that $C(\sigma) \leq |\sigma| + k$. On the other hand, we also know that the proportion of strings of a given length for which $C(\sigma) \leq |\sigma| - k$ is at most 2^{-k} . As the length of the strings increases, k remains constant; so there is a length such that for all strings of at least that length, $C(\sigma) \approx |\sigma| \pm k$. Even if k is quite large, it is fixed, and for any fixed k , almost all strings are longer than k . So almost all strings are not noticeably compressible when compared to their initial length, and almost all strings have complexity of approximately their length. The *typical* finite string is incompressible. Implicitly following Laplace, Kolmogorov proposes that incompressible strings, being both unruly and typical, are the random ones: a string σ is **C-random** iff $C(\sigma) \approx |\sigma|$.

As we have placed no constraints on which descriptions are permissible descriptions, a carefully designed compression algorithm could encode more information than is in the content of the description itself: an efficient decoding

might begin its operation by scanning all of δ to determine its length, only then to read the contents of δ bit for bit. In this way, the information δ is really worth $|\delta| + \log |\delta|$ bits, so it is clear that we have been cheating in calling $|\delta|$ the complexity of σ . (van Lambalgen, 1987: 736, notation altered)

To exclude this possibility, we follow Chaitin and Levin and restrict the permissible descriptions to those which are **prefix-free** with respect to u (Chaitin, 1966; Levin, 1976). An algorithm f is prefix-free iff no two f -descriptions are such that one is an initial segment of the other. (Think telephone numbers:⁵ no well-formed phone number is such that appending further digits to it yields another well-formed phone number.) With a prefix-free encoding, a decompression algorithm can begin processing the description while reading it, and recognise the end of the description without having to be explicitly told that it has ended. Accordingly, such algorithms only make use of $|\delta|$ bits in decoding a string. We can define the **prefix-free f -complexity** of a string as the length of the shortest prefix-free f -description that generates the string, and then define the **prefix-free complexity** of a sequence $K(\sigma) = K_{u^*}(\sigma)$ for some fixed prefix-free asymptotically optimal u^* (which do exist: see Downey and Hirschfeldt, 2010: ch. 3).

The requirement that descriptions be prefix-free reduces the number of permissible descriptions, and thus in general increases the complexity of sequences: $C(\sigma) \leq K(\sigma)$ (the relation between C and K is discussed in Downey and Hirschfeldt, 2010: ch. 4).

⁵ Or sentences of properly bracketed formal logic (Shapiro 2013: §2, Theorem 5).

But then the prefix-free incompressible sequences are even more numerous than the C -random sequences, and so play the functional role of random sequences even better. Accordingly, call a string **Kolmogorov random** iff $K(\sigma) \gtrsim |\sigma|$.

* * *

This inefficiency in K is actually of benefit, for it allows us to extend the definition of Kolmogorov randomness to infinite sequences. We can define an **infinite sequence** x as **prefix-free Kolmogorov random** iff every finite initial subsequence $x \upharpoonright i$ is prefix-free Kolmogorov random. (This definition fails for the original measure of complexity C , as C is subject to the phenomenon of **complexity oscillation**, where any sufficiently long string has non- C -random initial segments: Li and Vitanyi, 2008: §2.5.)

So defined, the class of infinite Kolmogorov random sequences can be shown to be non-empty. In fact, a striking result due to Schnorr shows that it is a familiar class: **the set of infinite Kolmogorov random sequences is precisely the set of Martin-Löf random sequences** (Schnorr, 1971).

This result shows a happy convergence between our two discussions of randomness. It is some evidence that we have captured the intuitive notion of randomness in a precise formal notion. Some claim Schnorr's result is good evidence for a 'Martin-Löf-Chaitin thesis' (Delahaye, 1993) that identifies the intuitive notion of randomness (for infinite sequences at least) with the precise notion of **Kolmogorov/Martin-Löf (KML) randomness** – analogous to Church's thesis identifying the intuitive notion of effective procedure with Turing machine/recursive function.

In the case of Church's thesis, every precise account that is even close to intuitively plausible turns out to be equivalent (Turing machines, recursive functions, abacus machines, ...). This is not true of randomness, where there are many distinct extensionally inequivalent yet precise accounts of randomness that more or less agree with intuition: from those relatively closely associated with KML randomness (like Schnorr randomness and von Mises randomness), to other proposals of more distant pedigree, such as epistemic accounts (Eagle, 2005) or accounts based on indeterminism (Hellman, 1978). The dispute over whether there is a single precise notion of randomness that answers perfectly to our intuitions about random sequences can be largely sidestepped for present purposes (but see Porter 2012). KML randomness is a reasonable and representative exemplar of the algorithmic approach to randomness, and it is adopted here as a useful working definition of randomness for sequences. None of the difficulties we're about to discuss concerning the connection between randomness and chance turn on the details of which

particular sequences get counted as random – most arise from the mismatch between intuitions about chance, and features common to *all* accounts of product randomness.

3. Chance and Randomness

The question before us now is this: what role can product randomness play in the theory of chance? Does chance require the existence – or potential existence – of random sequences, as orthodox frequentism supposes? Does the existence of a random sequence entail the existence of chance? I address both questions below.

3.1 Chance without randomness

Random sequences are, if the foregoing is correct, the kinds of outputs that are typical of a certain sort of chance process: **independent, identically distributed** (i.i.d.) trials of a process with two outcomes (a so-called **Bernoulli process**). But many chance processes don't have these features, and their characteristic outputs are not product random.

Frequentism gives a central role to randomness. Probabilities only exist in those mass phenomena that can be idealised to a random sequence of outcomes, namely collectives: 'we shall not speak of probability until a collective has been defined' (von Mises 1957: 18). But this approach prioritises the overall sequence over the individual outcomes; and interesting chance phenomena that arise at the level of individual trials of the repeated processes –**single case chances** – are overlooked (Hájek, 2009; Jeffrey, 1992). The thrust of the literature surrounding frequentism has been that single-case chance must be part of any satisfactory account of physical probability. The upshot for randomness is that, unless the single case chances satisfy some significant constraints, the resulting outcome sequences need not be, and typically will not be, random.

Consider physically realistic chance processes in which the outcomes are not **independent**. One might consider the sequence of states of weather on successive days: the chance of a sunny day, given a sunny day yesterday, is higher than the unconditional chance of a sunny day. Sequences of non-independent outcomes are susceptible to gambling systems (e.g., bet on a sunny day tomorrow when it's been sunny today), and are thus non-random. Such a sequence cannot be idealised to a random collective, and so is not a suitable example of the mass phenomena that von Mises' takes as the proper subject matter of the theory of probability. If we are to understand the chances involved in weather patterns, they cannot be extracted from sequences of weather outcomes. One might reasonably ask, whence do these chances derive? For there is no necessity that, whenever there is a sequence of

dependent trials, there must also be some sequence of independent trials; but that is exactly what von Mises' approach demands.⁶

This illustrates a major problem facing any theory of chance which, like frequentism, makes the existence of random sequences partly constitutive of chance: namely, that many chance processes don't give rise to random sequences in the sense developed in §2. For another example, consider **biased** processes. These were not a problem for von Mises' original theory of randomness, because invariance of frequencies under place selections is independent of what the frequency is. But if we accept the conclusions of §2, that von Mises' theory is intuitively inadequate, and that a typicality approach is more successful, then we face the problem that such approaches need a probability measure with respect to which sequences are typical. Perhaps it is acceptable to appeal to the Lebesgue measure in a purely mathematical account of randomness, as it is an *a priori* symmetry of the space of sequences. But biased chances are not *a priori* features of the space of outcomes, but contingent and *a posteriori* features of a particular chance device. If randomness is partly constitutive of chance, *most* chance devices – any other than fair coins – don't usually produce random outcomes, because they usually produce outcomes that are atypical with respect to the Lebesgue measure. Given that, it is extremely hard to see how such chance devices end up involving chances at all, except by offering some reduction of unequal apparent chances to some combination of basic outcomes which are equally likely. Ultimately, then, such views will end up being a version of the **classical theory of probability**,⁷ because such views require that *a priori* symmetries in the outcome space underlie all random sequences, and thus all genuine chances.

Because frequentism gives randomness such a significant role in the foundations of chance, many cases with which frequentism has difficulty are also cases in which there is chance without randomness. Consider Hájek's **problem of ordering**: 'Limiting relative frequency depends on the order of trials, whereas probability does not' (Hájek 2009: 219; see also Hellman 1978). The same is true of randomness: whether a

⁶ Actual weather supervenes on some underlying mosaic of more localized outcomes, and perhaps in the actual case the frequentist *can* find some sequence which can be idealized to a random collective. But it is surely possible that there be a situation in which there is a pattern in the underlying mosaic that suggests probabilistic dependence between those outcomes. The frequentist, shackled with attempting to fit all probabilistic phenomena into the Procrustean bed of collectives, cannot handle such a case, while more liberal theories may accommodate it – even other Humean theories which ground probability ultimately in patterns of outcomes (Lewis 1994; Loewer 2004).

⁷ 'The theory of chance consists in reducing all the outcomes of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the outcome whose probability is sought.' (Laplace, 1951: 6–7). See also Zabell, this volume.

sequence is random depends on the order of the outcomes, and randomness is not necessarily preserved under permutation of outcome order. However, whether the outcomes are the product of a chancy or random process is invariant under permutation; so here we have cases in which there is process randomness without product randomness.

The preceding examples show a systematic disconnection between certain sorts of chance process and the existence of random outcome sequences. But, given the existence of single-case chances, cases of chance without randomness arise even in the most favourable circumstances for the idea that chance requires randomness. For any particular chance process could yield an **unrepresentative** outcome sequence, which doesn't reflect the underlying chances, and in which frequencies are not good evidence for the values of the chances. Those sequences will not be typical with respect to the underlying chances, and hence non-random. Such cases show don't show any systematic failure – they are, by construction, atypical of their underlying chances. But everyone except the frequentist accepts that atypical outcomes can happen.

3.2 Randomness without chance

Consider short sequences: they are not compressible, since in general a prefix-free encoding of a short sequence will be about as long as the sequence itself. Accordingly, all short sequences are (KML) random. But not all short sequences involve a random process. Since some short sequences are essentially short, consisting of unrepeatable or seldom repeatable outcomes, there are some random sequences that are essentially random despite not being chancy. So it cannot be a necessary condition on randomness that it require chance. Perhaps we might respond by denying that all short sequences are random, treating them as a **degenerate case** of incompressibility and restricting randomness to non-degenerately incompressible sequences. But then all short sequences turn out to be non-random; and if there is an unrepeatable chance event, there will be chance without randomness. If we consider the outcomes alone, either all short sequences are random or none of them are. But as some unrepeatable outcomes are chancy, and some are not, whichever way we opt to go with respect to randomness of the singleton sequences of such outcomes we will discover cases in which we have random processes without random products, or *vice versa*.

Other cases of randomness without chance arise from some of the more exotic possibilities of classical physics. One sort of possibility arises from **classical indeterminism**. It is well known that classical Newtonian particle mechanics is indeterministic: that specifying the positions and momenta of all particles at a time,

together with the laws, does not suffice to determine the future evolution of the system (Earman, 1986; Norton, 2008). But this classical indeterminism does not involve chance; nothing in the physics *requires* a probability distribution over these undetermined outcomes, rather than simply holding that such outcomes are nomologically possible. We could revise classical mechanics by adding a measure over outcomes (though which one?). But ordinary *unrevised* classical mechanics isn't incomplete just because it does not provide chances for our credences to reflect.

This phenomenon may yield cases of product randomness without chance. By preparing a system repeatedly in a given state that does not determine its future state, and classifying the possible outcomes into two classes, '0' and '1', it is physically possible that we obtain a random binary sequence as the system evolves over time. But we then have randomness without a chance distribution over the outcomes. The possibility of randomness requires only two distinguishable possible outcomes and the ability to produce arbitrary sequences of such outcomes. Chance requires *in addition* that some measure be ascribable to each outcome. Chance is a sort of quantitative physical possibility; since a physical theory can be indeterministic without having a probability distribution over the different possible outcomes, we can produce a random sequence by an indeterministic but non-chancy process. These sorts of cases may induce us to query our initial regimentation in §1, which identified process randomness with chanciness.⁸ But splitting chance and process randomness further undermines any putative connection between product randomness and chance.

Another interesting possibility permitted by classical physics is **chaotic dynamics**. A (discrete) dynamical system can be characterised by four parameters: a set of basic states X , a σ -algebra Σ on X (intuitively, the possible outcomes in the system, corresponding to regions of the basic space X), some measure μ such that $\mu(X) = 1$, and a deterministic evolution map T from X onto X , which captures the lawlike evolution of states over one time step (Berkovitz and Frigg, 2006). Following Werndl (2009), a dynamical system is chaotic iff it is **mixing**: for all $A, B \in \Sigma$ (where ' T^n ' denotes applying T n times):

$$\lim_{n \rightarrow -\infty} \mu(T^n(B) \cap A) = \mu(B)\mu(A).$$

Suppose B is the region in which our system was found n time steps ago; a system is mixing if, in the limit as n increases, the system's having been in B becomes independent of its now being in A (Berkovitz and Frigg, 2006: 676). In the limit, such

⁸ That is, we may consider these sorts of case as involving process randomness because of the presence of indeterminism, even if they do not involve chance.

systems are maximally unpredictable (in the sense of Eagle, 2005: 775), because states of the distant past are increasingly credentially irrelevant for future states, assuming one sets prior credences in accordance with the μ -measure of the system (though that is consistent with the system being quite predictable in the short run). A system is **Bernoulli** when the states of the system are independent of all past (and future) states:

$$\mu(T^{-1}(B) \cap A) = \mu(A)\mu(B).$$

The trajectory through the state space of a Bernoulli system is just like a sequence of fair coin tosses: the past states of the system are independent of future states, and credence in them is equal to their μ -measure. Obviously all Bernoulli systems are mixing.

What's of interest for us is that chaotic systems do not require indeterminism: the evolution map T is a one-one function, and so the time evolution is deterministic. Given the additional but widely believed premise that chance requires indeterminism (Lewis, 1986; Schaffer, 2007: 120), this entails that chaos does not require chance or random processes. The motivation for saying that these systems are not process random is that the underlying classical non-probabilistic dynamics provides a physically complete account of the behaviour of any particular system – i.e., sufficient to predict, and causally explain, the behaviour of the system without need of additional supplementation.

But a Bernoulli system can generate a random sequence of outcomes, because the macroscopically-described trajectory of a chaotic system through the state space (its macroscopic history) provides resources too meagre to predict its future macroscopic behaviour. These sequences can be random: they are certainly at least Borel normal, since any complex pattern of successive states A will be an element of Σ , the limit frequency with which such patterns occur will be equal to $\mu(A)$, and the frequency with which one pattern follows any other will also be $\mu(A)$.

Sampled infrequently, even mixing systems can give rise to random sequences of outcomes. Some non-Bernoulli systems of great physical interest, such as the Lorenz model of atmospheric convection (Smith, 1998: §1.4), display *rapid* mixing (Luzzatto et al., 2005). Sampled at ordinary enough intervals, therefore, such systems can provide physical plausible examples of deterministic, and arguably non-chancy, random outcomes.

In such cases, admittedly, measures that look rather like chances are involved. The measure μ is formally like a probability measure, and moreover seems to be used to regulate credences about which macrostates the system will be in – reminiscent of the way chance regulates credence (Lewis, 1986). This has prompted some to *deny* the additional premise that chance requires indeterminism, and argue that measures over the state space in statistical mechanics, and other dynamical theories, should be understood as chances (Clark, 1987; Loewer, 2001; Meacham, 2005). There are also general arguments aiming to show that chance and determinism are compatible (Eagle, 2011; Glynn 2010). If these arguments are successful, these random sequences will after all be associated with a chance distribution. However, it is fair to say that not everyone is convinced that deterministic chance is possible – the received view remains that the measures involved in deterministic dynamical systems are *ignorance* measures over ensembles of macroscopically indiscriminable microstates, and so not objective chances after all. (And if there is deterministic chance, we open up the prospect of yet further examples of chance without randomness: cases where the deterministic laws coexist with chances, but in which sequences of outcomes generated in accordance with those laws is predictable and exploitable by a gambling system).

An example of randomness in a deterministic setting (though not that of classical physics) which does not make any appeal to a measure is given by Humphreys. Adapting the theorem of Ville's mentioned in §2, he proves that there exists 'a theory which is deterministic in the sense of Montague and which has as a model a sequence which is random in the sense of von Mises/Church' (Humphreys, 1978: appendix). This isn't quite KML randomness, but it is a suggestive result. The theorem exploits the fact that the standard Montague-Earman definition of determinism requires that the evolution map of a dynamical system be a function (Earman, 1986; Montague, 1974); but it does not require that it be computable. Here, given the explicit construction of the evolution function, it is difficult to accept that the limit frequency in the resulting sequence can play the chance role; so we have a case, albeit rather artificial, of randomness without chance.

3.3 Randomness as evidence for chance

We've seen plausible examples in which chancy processes do not produce random sequences – either by chance, in the case of Bernoulli processes, or because the family of chancy processes includes many whose typical outcomes do not show the hallmarks of random sequences.

We've also seen examples of randomness without chance. Some of these examples, such as the statistical mechanical cases, may fail to persuade, because we needed to

use objective probability measures that behave in many ways like chances; and these, it may be argued, play the chance role well enough to deserve the name. Even so, some examples of randomness without chance remain, and support the claim that chancy processes are unnecessary for random products.

Together, the examples in §§3.1–3.2 make the case that the distinctions drawn in §1 between product and process randomness mark genuine metaphysical distinctions, and neither is plausibly reducible to the other.

Nothing in my discussion undermines the sensible position that randomness of some sequences of outcomes is **good evidence** for the involvement of a chancy process. The appearance of a series of outcomes typical of a sequence of independent and identically distributed trials is generally good evidence that there was such a sequence of trials. Conversely, seeing a sequence of outcomes governed by a recognisable law will disconfirm any probabilistic hypothesis – not conclusively, but substantially. Since the typical outcomes of repeated equiprobable i.i.d. trials are random sequences, we ought sensibly *expect* such processes to produce a random sequence of outcomes. Accordingly, Hellman says:

The link, then, between mathematical and physical randomness is epistemic and only that. Observations of mathematically non-random sequences can be used to decide when further explanation in terms of as yet undiscovered causal factors is wanting. But, in no sense is any notion of mathematical randomness serving as an explication for ‘ultimate physical randomness’, whatever that might be. (Hellman, 1978: 86)

Taking ‘mathematical’ and ‘physical’ to mean product and process randomness respectively, this conclusion seems inescapable.

Predictably, this epistemic connection between randomness and chance resembles that between frequencies and chance. Relative frequencies are good evidence for the chances; known chances should lead us to expect certain frequencies; and extreme frequencies might prompt us to look for non-chance hypotheses. But though frequency evidence is good evidence for chance hypotheses – and though successful accounts of chance must explain *why* it’s such good evidence – frequentism is implausible as a reductive account of chance, and any proposed reduction of chance to randomness is equally implausible.

By definition, the Kolmogorov complexity function K cannot be a recursive function. Hence there is no effective positive conclusive test for randomness even of finite sequences. Even if our evidence is in fact a random sequence of outcomes, the

proposition that the sequence is random will not generally be in our evidence, as (for all we know) there may be a short description of the sequence that we are unaware of. That an evidence sequence is random is no less a theoretical hypotheses about it than the hypothesis that it was produced by a chance process. In that light, it seems sensible to neglect randomness, and focus directly on how likely a chance process is to have been involved in the production of a given sequence, conditional on the contents of that evidence sequence.⁹

* * *

That said, more sophisticated descendants of frequentism, such as Lewis' **best systems analysis** of chance (Lewis 1994; Schwartz, this volume), might enable us to make more precise claims about the bearing of randomness on chance. One application of randomness in a Humean metaphysics of chance might be in analysing Lewis' notion of **fit**: namely, the proposal that a probability function fits some outcomes just when those outcomes are random with respect to that function (§2.2; see also Elga 2004). Another application might be in understanding when to invoke chances, given an austere fundamental ontology of only the 'spatiotemporal arrangement of local qualities throughout all of history, past and present and future' (Lewis, 1994: 474). One attractive suggestion might be that we ought to adopt a probabilistic description as best balancing strength and simplicity when that underlying arrangement of local qualities is complex, in the precise sense of being Kolmogorov incompressible (§2.3). For complex sequences of outcomes preclude there being simple non-probabilistic descriptions that are strong enough to carry useful content about the outcome sequence. Yet even if randomness can play these important roles in a Humean metaphysics of chance, that falls short of the kind of reduction originally envisaged by von Mises.

Acknowledgements

For helpful comments on earlier versions of this chapter, I'd like to thank Al Hájek, an audience at a session of the Paris-Diderot Philosophy of Math Seminar on algorithmic randomness, particularly Chris Porter and Sean Walsh, students in my Oxford graduate seminar on philosophy of probability, Wo Schwartz and others in a session of the ANU philosophy of

⁹ We are generally in a better position to know that a process is chancy when we draw that conclusion on the basis of consideration of its physical characteristics, than when we draw that conclusion by performing a double abduction, first inferring that a sequence of outcomes is in fact random from the fact that it seems random, and then inferring that the process behind that random sequence is chancy.

The irony of concluding a chapter that spends so much time talking about randomness by recommending that it be ignored does not escape me.

probability reading group, and colleagues from Adelaide and Flinders for comments on a written draft.

References

- Albert, D.Z. (2000) *Time and Chance*. Cambridge, MA: Harvard University Press.
- Berkovitz, J. and R. Frigg (2006) 'The Ergodic Hierarchy, Randomness and Hamiltonian Chaos', *Studies in History and Philosophy of Modern Physics* 37: 661–91.
- Borel, E. (1909) 'Les Probabilités Dénombrables Et Leurs Applications Arithmétiques', *Rendiconti Del Circolo Matematico Di Palermo* 27: 247–71.
- Chaitin, G. (1966) 'On the Length of Programs for Computing Finite Binary Sequences', *Journal of the Association for Computing Machinery* 13: 547–69.
- Church, A. (1940) 'On the Concept of a Random Sequence', *Bulletin of the American Mathematical Society* 46: 130–5.
- Clark, P. (1987) 'Determinism and Probability in Physics', *Proceedings of the Aristotelian Society, Supplementary Volume* 61: 185–210.
- Dasgupta, A. (2011) 'Mathematical Foundations of Randomness', in P. Bandyopadhyay and M. Forster (eds.), *Handbook of the Philosophy of Science. Volume 7: Philosophy of Statistics*, Amsterdam: North-Holland, 641–710.
- Delahaye, J.-P. (1993) 'Randomness, Unpredictability and Absence of Order: the Identification by the Theory of Recursivity of the Mathematical Notion of Random Sequence', in Dubucs, J.-P. (ed.), *Philosophy of Probability*. Dordrecht: Kluwer, 145–67.
- Downey, R. and D.R. Hirschfeldt (2010) *Algorithmic Randomness and Complexity*. New York: Springer.
- Eagle, A. (2005) 'Randomness Is Unpredictability', *The British Journal for the Philosophy of Science* 56(4): 749–90.
- Eagle, A. (2011) 'Deterministic Chance', *Nous* 45: 269–99.
- Eagle, A. (2014) 'Chance Versus Randomness', Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), <http://plato.stanford.edu/archives/spr2014/entries/chance-randomness/>.
- Earman, J. (1986) *A Primer on Determinism*. Dordrecht: D. Reidel.
- Elga, A. (2004) 'Infinitesimal Chances and the Laws of Nature', *Australasian Journal of Philosophy* 82: 67–76.
- Gaifman, H. and M. Snir (1982) 'Probabilities Over Rich Languages, Testing and Randomness', *The Journal of Symbolic Logic* 47: 495–548.
- Glynn, L. (2010) 'Deterministic Chance', *British Journal for the Philosophy of Science* 61: 51–80.

- Hawthorne, J. and M. Lasonen-Aarnio (2009) 'Knowledge and Objective Chance', in Greenough, P. and Pritchard, D. (eds), *Williamson on Knowledge*. Oxford: Oxford University Press, 92–108.
- Hájek, A. (2009) 'Fifteen Arguments Against Hypothetical Frequentism', *Erkenntnis* 70: 211–35.
- Hellman, G. (1978) 'Randomness and Reality', in Asquith, P.D. and Hacking, I. (eds), *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 2, University of Chicago Press, 79–97.
- Humphreys, P.W. (1978) 'Is "Physical Randomness" Just Indeterminism in Disguise?', in Asquith, P.D. and Hacking, I. (eds), *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 2, University of Chicago Press, 98–113.
- Jeffrey, R.C. (1992) 'Mises Redux', in *Probability and the Art of Judgement*. Cambridge: Cambridge University Press, 192–202.
- Kolmogorov, A.N. (1963) 'On Tables of Random Numbers', *Sankhyā* 25: 369–76.
- Kolmogorov, A.N. (1965) 'Three Approaches to the Definition of the Concept "Quantity of Information"', *Problemy Peredachi Informatsii* 1: 3–11.
- Kolmogorov, A.N. and V.A. Uspensky (1988) 'Algorithms and Randomness', *SIAM Theory of Probability and Applications* 32: 389–412.
- van Lambalgen, M. (1987) 'Von Mises' Definition of Random Sequences Revisited', *Journal of Symbolic Logic* 52: 725–55.
- Laplace, P.-S. (1951) *Philosophical Essay on Probabilities*. New York: Dover.
- Levin, L.A. (1976) 'Various Measures of Complexity for Finite Objects', *Soviet Mathematics Doklady* 17: 522–6.
- Lewis, D. (1979) 'Counterfactual Dependence and Time's Arrow', *Nous* 13(4): 455–76.
- Lewis, D. (1986) 'A Subjectivist's Guide to Objective Chance', in *Philosophical Papers*, Volume 2. Oxford University Press, 83–132.
- Lewis, D. (1994) 'Humean Supervenience Debugged', *Mind* 103: 473–90.
- Li, M. and P.M.B. Vitanyi (2008) *An Introduction to Kolmogorov Complexity and Its Applications* (3rd ed.). Berlin and New York: Springer Verlag.
- Lieb, E.H., D. Osherson, and S. Weinstein. (2006) 'Elementary Proof of a Theorem of Jean Ville', eprint arXiv:cs/0607054 [cs.CC].
- Loewer, B. (2001) 'Determinism and Chance', *Studies in History and Philosophy of Modern Physics* 32: 609–20.
- Loewer, B. (2004) 'David Lewis's Humean Theory of Objective Chance', *Philosophy of Science* 71: 1115–25.
- Luzzatto, S., I. Melbourne, and F. Paccaut. (2005) 'The Lorenz Attractor Is Mixing', *Communications in Mathematical Physics* 260: 393–401.

- Martin-Löf, P. (1966) 'The Definition of Random Sequences', *Information and Control* 9(6): 602–19.
- Martin-Löf, P. (1969a) 'The Literature on Von Mises' Kollektivs Revisited', *Theoria* 35: 12–37.
- Martin-Löf, P. (1969b) 'Algorithms and Randomness', *Revue De l'Institut International De Statistique*, 37: 265–72.
- Meacham, C.J.G. (2005) 'Three Proposals Regarding a Theory of Chance', *Philosophical Perspectives* 19: 281–307.
- Mises, von, R. (1957) *Probability, Statistics and Truth*. New York: Dover.
- Montague, R. (1974) 'Deterministic Theories', in Thomason, R.H. (ed.), *Formal Philosophy*. New Haven: Yale University Press, 303–59.
- Nies, A. (2009) *Computability and Randomness*. Oxford Logic Guides 51. Oxford: Oxford University Press.
- Norton, J. (2008) 'The Dome: an Unexpectedly Simple Failure of Determinism', *Philosophy of Science* 75: 786–98.
- Porter, C. (2012) *Mathematical and Philosophical Perspectives on Algorithmic Randomness*. PhD thesis, University of Notre Dame: Notre Dame, Indiana.
- Rathmanner, S. and M. Hutter (2011) 'A Philosophical Treatise of Universal Induction', *Entropy* 13: 1076–136.
- Schaffer, J. (2007) 'Deterministic Chance?', *The British Journal for the Philosophy of Science* 58: 113–40.
- Schnorr, C.P. (1971) 'A Unified Approach to the Definition of Random Sequences', *Theory of Computing Systems* 5: 246–58.
- Shapiro, S. (Winter 2013) 'Classical Logic', in Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*, plato.stanford.edu/archives/win2013/entries/logic-classical/.
- Smith, P. (1998) *Explaining Chaos*. Cambridge: Cambridge University Press.
- Solomonoff, R. (1964) 'A Formal Theory of Inductive Inference, Parts I and II', *Information and Control* 7: 1–22, 224–54.
- Vadhan, S. P. (2012) 'Pseudorandomness', *Foundations and Trends® in Theoretical Computer Science* 7: 1–336.
- Ville, J. (1939) *Étude Critique de la Notion Collectif*, Paris: Gauthier-Villars.
- Wald, A. (1938) 'Die Widerspruchfreiheit Des Kollektivbegriffs in Der Wahrscheinlichkeitsrechnung', *Ergebnisse Eines Mathematischen Kolloquiums* 8: 38–72.
- Werndl, C. (2009) 'What Are the New Implications of Chaos for Unpredictability?', *The British Journal for the Philosophy of Science* 60: 195–220.