Chance versus Randomness

First published Wed Aug 18, 2010; substantive revision Thu Feb 8, 2018

Randomness, as we ordinarily think of it, exists when some outcomes occur haphazardly, unpredictably, or by chance. These latter three notions are all distinct, but all have some kind of close connection to probability. Notoriously, there are many kinds of probability: subjective probabilities ('degrees of belief'), evidential probabilities, and objective chances, to name a few (Hájek 2012), and we might enquire into the connections between randomness and any of these species of probability. In this entry, we focus on the potential connections between randomness and chance, or physical probability. The ordinary way that the word 'random' gets used is more or less interchangeable with 'chancy', which suggests this Commonplace Thesis—a useful claim to target in our discussion:

(CT)   Something is random iff it happens by chance.

The Commonplace Thesis, and the close connection between randomness and chance it proposes, appears also to be endorsed in the scientific literature, as in this example from a popular textbook on evolution (which also throws in the notion of unpredictability for good measure):

> scientists use chance, or randomness, to mean that when physical causes can result in any of several outcomes, we cannot predict what the outcome will be in any particular case. (Futuyma 2005: 225)

Some philosophers are, no doubt, equally subject to this unthinking elision, but others connect chance and randomness deliberately. Suppes approvingly introduces

> the view that the universe is essentially probabilistic in character, or, to put it in more colloquial language, that the world is full of random happenings. (Suppes 1984: 27)

However a number of technical and philosophical advances in our understanding of both chance and randomness open up the possibility that the easy slide between chance and randomness in ordinary and scientific usage—a slide that would be vindicated by the truth of the Commonplace Thesis—is quite misleading. This entry will attempt to spell out these developments and clarify the differences between

chance and randomness, as well as the areas in which they overlap in application. It will also aim to clarify the relationship of chance and randomness to other important notions in the vicinity, particularly determinism and predictability (themselves often subject to confusion).

There will be philosophically significant consequences if the Commonplace Thesis is incorrect, and if ordinary usage is misleading. For example, it is intuitively plausible that if an event is truly random it cannot be explained (if it happens for a reason, it isn't truly random). It might seem then that the possibility of probabilistic explanation is undermined when the probabilities involved are genuine chances. Yet this pessimistic conclusion only follows under the assumption, derived from the Commonplace Thesis, that all chancy outcomes are random. Another interesting case is the role of random sampling in statistical inference. If randomness requires chance, then no statistical inferences on the basis of 'randomly' sampling a large population will be valid unless the experimental design involves genuine chance in the selection of subjects. But the rationale for random sampling may not require chance sampling—as long as our sample is representative, those statistical inferences may be reliable. But in that case, we'd be in a curious situation where random sampling wouldn't have much to do with randomness, and whatever justification for beliefs based on random sampling that randomness is currently thought to provide would need to be replaced by something else.

A final case of considerable philosophical interest is the *frequentist* approach to objective probability, which claims (roughly) that the chance of an outcome is its frequency in an appropriate series of outcomes (Hájek 2012 §3.4). To avoid classifying perfectly regular recurring outcomes as chancy, frequentists like von Mises (1957) proposed to require that the series of outcomes should be random, without pattern or order. Frequentism may fall with the Commonplace Thesis: if there can be chancy outcomes without randomness, both will fail.

The Commonplace Thesis is central to all three examples. As it is widely accepted that probabilistic explanation is legitimate, that random sampling doesn't need genuine chance (though it can help), and that frequentism is in serious trouble (Hájek 1997), there is already some some pressure on the Commonplace Thesis. But we must subject it to closer examination to clarify whether these arguments do succeed, and what exactly it means to say of some event or process that it is random or chancy. Though developing further consequences of this kind is not the primary aim of this entry, it is hoped that what is said here may help to untangle these and other vexed issues surrounding chance and randomness.

---

1. Chance

To get clear on the connections and differences between chance and randomness, it would be good first to have some idea of what chance and randomness amount to. Interestingly, philosophical attention has focussed far more on chance than randomness. This may well be a consequence of the Commonplace Thesis. Whatever its source, we can appeal to a substantial consensus in the philosophical literature as to what kind of thing chance must be.

Carnap (1945) distinguished between two conceptions of probability, arguing that both were scientifically important. His 'probability$_1$' corresponds to an epistemic notion, nowadays glossed either (following Carnap himself) as *evidential* probability, or as *credence* or degree of belief. This is contrasted with Carnap's 'probability$_2$', which is the concept of a non-epistemic objective kind of probability, better known as *chance*.

There are many philosophical accounts of what actually grounds chance, as part of the minor philosophical industry of producing 'interpretations'—really, reductive analyses or explications—of probability. In this vein we have at least the *frequency* theory of Reichenbach (1949) and von Mises (1957) (and Carnap's own explication of probability$_2$ was in terms of frequencies), the *propensity* theory of Popper (1959) and Giere (1973), as well as many more recent accounts, notably Lewis' (1994) 'Best System' account of chance (see also Loewer 2004). There is no agreement over which, if any, of these accounts are right; certainly both the accounts mentioned face difficulties in giving an adequate account of chance. The consensus mentioned earlier is not over what actually plays the role of chance, but rather on the constraints that determine what that role is.

There can be such a consensus because 'chance' is not a technical term, but is rather an ordinary concept deployed in fairly familiar situations (games of chance, complicated and unpredictable scenarios, large ensembles of similar events, etc.). There is widespread agreement amongst native speakers of English over when 'chance' applies to a particular case, and this agreement at least indicates that there is a considerable body of ordinary belief about chance. One needn't take the deliverances of folk intuition as sacrosanct to recognise that this ordinary belief provides the starting point for philosophical accounts of chance. It may turn out that *nothing* fits the role picked out by these ordinary beliefs and their philosophical precisifications, yet even in that case we'd be inclined to conclude that chance doesn't exist, not that our ordinary beliefs

about what chance must be are incorrect.

Below, some of the theoretical principles that philosophers have extracted from commonplace beliefs about chance will be outlined. (In so doing, we freely use probabilistic notation and concepts; see the entry on interpretations of probability, Hájek 2012: §1, for background probability theory required to understand these expressions.) Two such constraints have been widely accepted since the early days of the philosophy of probability. Firstly, it is required that the mathematics of chance should conform to some standard mathematical theory of probability such as Kolmogorov's 1933 axiomatisation of the probability calculus (or some recognisable variant thereof, like Popper's axiomatisation of conditional probability). Secondly, chance should be objective: mind-independent and not epistemic or evidential. But a number of other constraints have been articulated and defended in the literature. (Schaffer 2007: §4 contains a useful discussion of these and other constraints on the chance role.) While these principles have been termed 'things we know about chance', this shouldn't be taken to preclude our discovering that there is no such thing as chance—rather, the philosophical consensus is that if there is any such thing as chance, it will (more or less) fit these constraints.

- Chance should regulate (that is, it is involved with norms governing) rational belief in line with Lewis' Principal Principle (Lewis, 1980), or his New Principle (Lewis, 1994; Hall, 2004). Where C is a reasonable initial credence function, and E is the evidence, the Principal Principle (omitting some complications) is this:

  (PP) $C(p \mid \ulcorner \mathrm{Ch}(p) = x \urcorner \wedge E) = x$;

  This principle says that rational initial credence should treat chance as an expert, deferring to it with respect to opinions about the outcome p, by adopting the corresponding chances as your own conditional degree of belief. The New Principle—adopted to deal with some problematic interactions between Lewis' metaphysics and the PP—advocates deferring to chance in a somewhat different way. This New Principle NP suggests (more or less) that rational initial credence should treat chance as an expert with respect to the outcome p, by adopting the conditional chances, on the admissible evidence, as your conditional degrees of belief on that same evidence, as in this principle (this claim corresponds to equation 3.9 in Hall 2004; see his discussion at pp. 102–5 for some important qualifications, and the connection to the formulation of the NP in Lewis 1994):

  (Chance-Analyst) $C(p \mid \ulcorner \mathrm{Ch}(p \mid E) = x \urcorner \wedge E) = x$.

  Non-reductionist views about chance, which take chances to be independent fundamental features of reality, can follow PP. Reductionists, who take the values of chances to be fixed entirely by other features of reality (normally frequencies and symmetries, but chance is typically constrained by them in a quite indirect manner), are for technical reasons (to do with undermining, see supplement A.1) ordinarily forced to adopt NP, and Chance-Analyst, as norms on credence, though in many ordinary cases, NP and PP give very similar recommendations. In either case, both formal principles give content to the intuitively plausible idea that chances should guide reasonable credence.[1]

- Chance should connect with possibility. Leibniz claimed that probability was a kind of 'graded possibility', and more recent authors have largely agreed. In particular, it seems clear that if an outcome has some chance of occurring, then it is possible that the outcome occurs. This intuition has been made precise in the Basic Chance Principle (BCP) (See supplement A.2 for further details on this principle):

  > Suppose $x > 0$ and $Ch_{tw}(A) = x$. Then A is true in at least one of those worlds w' that matches w up to time t and for which $Ch_t(A) = x$. (Bigelow et al., 1993: 459)

But one needn't accept precisely this version of the BCP to endorse the general thesis that chance and possibility must be linked—for other versions of this kind of claim, see Mellor (2000); Eagle (2011) and the 'realization principle' of Schaffer (2007: 124).

- Chance should connect with actual frequencies, at least to the extent of permitting frequencies to be good evidence for the values of chances. This may be through some direct connection between chance and frequency, or indirectly through the influence of observed outcome frequencies on credences about chances via the Principal Principle (Lewis, 1980: 104–6). But chance should not be identified with frequency—since a fair coin can produce any sequence of outcomes, there is no possibility of identifying chance with observed frequency. (Though of course a fair coin is overwhelmingly likely to produce roughly even numbers of heads and tails when tossed often enough). Moreover, there can be a chance for a kind of outcome even when there are very few instances of the relevant process that leads to that outcome, resulting in the actual frequencies being misleading or trivial (for example, if there is only a single actual outcome: Hájek 1997).

- When considering the connection between frequency and chance, not just any frequency will do. What is wanted is the frequency in relevantly similar trials, with the same kind of experimental setup. The relevance of frequencies in such trial is derived from the assumption that in such similar trials, the same chances exist: intra-world duplicate trials should have the same chances. This is closely related to the 'stable trial principle' (Schaffer, 2003: 37ff). Chances attach to the outcomes of trials, but the physical grounds of the chance lie in the physical properties of the trial device or chance setup.

More details on all of these principles can be found in this supplementary document:

Supplement A. Some Basic Principles About Chance

1.1 'Single-Case' Chance and Chance Processes

Chance, it is commonly said, is 'single-case objective probability'. Philosophers haven't been very clear on what is meant by 'single-case', and the terminology is slightly misleading, as it falsely suggests that perhaps multiple cases have less claim to their chances. The most minimal version of the claim is that, at least sometimes, an outcome can have a chance to be a result of an instance of a given kind of process, or *trial*, even though no other trials of that process occur. This is what we will mean by 'single case' chance. (A stronger claim is that the chance of an outcome resulting from a given process is an intrinsic property of a single trial. The stronger claim is inconsistent with standard versions of the frequency theory, and indeed it can be difficult to see how chance and frequency might be connected if that stronger claim were true.) Some have claimed that single-case chance is no part of objective probability; for example, von Mises (1957: p. 11) remarks that the 'concept of probability … applies only to problems in which either the same event repeats itself again and again, or a great number of uniform elements are involved at the same time.'. However, this is a theoretical judgment on von Mises' part, based on difficulties he perceived in giving an account of single-case chance; it is not a judgement derived from internal constraints on the chance role. And how could it be? For

> like it or not, we have this concept [of single-case chance]. We think that a coin about to be tossed has a certain chance of falling heads, or that a radioactive atom has a certain chance of decaying within the year, quite regardless of what anyone may believe about it and quite regardless of whether there are any other similar coins or atoms. As philosophers we may well find the concept of objective chance troublesome, but that is no excuse to deny its existence, its legitimacy, or its indispensability. If we can't understand it, so much the worse for us. (Lewis, 1980: 90)

A number of the constraints discussed above require the legitimacy of single-case chance. The objection to frequentist accounts of chance that the frequency might misrepresent the chance if the number of actual outcomes is too low apparently requires that there be non-trivial chances even for events resulting from a type of trial that occurs only very few times, and perhaps even can only occur very few times (Hájek 2009: 227–8). The strong connections between possibility and chance mooted by the BCP and variants thereof also require that there are single-case chances. For the BCP requires that, for every event with some chance, it is possible that the event has that same chance and occurs. As noted in Supplement A.2, this renders the chances relatively independent of the occurrent frequencies, which in turn requires single-case chance. For some single outcomes—for example, the next toss of a coin that is biased ⅔ towards heads—only very few assignments of credence are reasonable; in that case, we should, if we are rational, have credence of ⅔ in the coin landing heads. The rationality of this unequal assignment cannot be explained by anything like symmetry or indifference. Its rationality can be explained by the PP only if the single-case chance of heads on the next toss is ⅔. Moreover, the existence of this constraint on rational credence should have an explanation. Therefore the PP, if it is to play this necessary explanatory role, requires single-case chance.

It is the stable trial principle that has the closest connection with single-case chance, however. For in requiring that duplicate trials should receive the same chances, it is natural to take the chance to be grounded in the properties of that trial, plus the laws of nature. It is quite conceivable that the same laws could obtain even if that kind of trial has only one instance, and the very same chances should be assigned in that situation. But then there are well-defined chances even though that type of event occurs only once.

The upshot of this discussion is that chance is a *process* notion, rather than being entirely determined by features of the outcome to which the surface grammar of chance ascriptions assigns the chance. For if there can be a single-case chance of $\frac{1}{2}$ for a coin to land heads on a toss even if there is only one actual toss, and it lands tails, then surely the chance cannot be fixed by properties of the outcome 'lands heads', as that outcome does not exist.[2] The chance must rather be grounded in features of the process that can produce the outcome: the coin-tossing trial, including the mass distribution of the coin and the details of how it is tossed, in this case, plus the background conditions and laws that govern the trial. Whether or not an event happens by chance is a feature of the process that produced it, not the event itself. The fact that a coin lands heads does not fix that the coin landed heads by chance, because if it was simply placed heads up, as opposed to tossed in a normal fashion, we have the same outcome not by chance. Sometimes features of the outcome event cannot be readily separated from the features of its causes that characterise the process by means of which it was produced. But the upshot of the present discussion is that even in those cases, whether an outcome happens by chance is fixed by the properties of the process leading up to it, the causal situation in which it occurs, and not simply by the fact that an event of a given character was the product of that process.[3]

1.2 Physics and Chance

Do chances exist? The best examples of probability functions that meet the principles about chance are those provided by our best physical theories. In particular, the probability functions that feature in radioactive decay and quantum mechanics have some claim to being chance functions. In orthodox approaches to quantum mechanics, some measurements of a system in a given state will not yield a result that represents a definite feature of that prior state (Albert 1992). So, for example, an x-spin measurement on a system in a determinate y-spin state will not yield a determinate result reflecting some prior state of x-spin, but rather has a 0.5 probability of resulting in x-spin $= +1$, and a 0.5 probability of resulting in x-spin $= -1$. That these measurement results cannot reflect any prior condition of the system is a consequence of various *no-hidden variables* theorems, the most famous of which is Bell's theorem (Bell 1964; see the entry on Bell's theorem, Shimony 2009). Bell's theorem shows that the probabilities predicted by quantum mechanics, and

experimentally confirmed, for spin measurements on a two-particle entangled but spatially separated system cannot be equal to the joint probabilities of two independent one-particle systems. The upshot is that the entangled system cannot be represented as the product of two independent localised systems with determinate prior x-spin states. Therefore, there can be no orthodox local account of these probabilities of measurement outcomes as reflecting our ignorance of a hidden quality found in half of the systems, so that the probabilities are in fact basic features of the quantum mechanical systems themselves.[4]

The standard way of understanding this is that something—the process of measurement, on the Copenhagen interpretation, or spontaneous collapse on the GRW theory—induces a non-deterministic state transition, called *collapse*, into a state in which the system really is in a determinate state with respect to a given quality (though it was not previously). These transition probabilities are dictated entirely by the state and the process of collapse, which allows these probabilities to meet the stable trial principle. The models of standard quantum mechanics explicitly permit two systems prepared in identical states to evolve via collapse into any state which has a non-zero prior probability in the original state, which permits these probabilities to meet the BCP. And the no-hidden variables theorems strongly suggest that there is no better information about the system to guide credence in future states than the chances, which makes these probabilities play the right role in the PP. These basic quantum probabilities governing state transitions seem to be strong candidates to be called chances.

The foregoing argument makes essential use of collapse. The existence of collapse as an alternative rule governing the evolution of the quantum state controversial, and it is a scandal of quantum mechanics that we have no satisfactory understanding of why collapse (or measurement) should give rise to basic probabilities. But that said, the existence of well-confirmed probabilistic theories which cannot be plausibly reduced to any non-probabilistic theory is some evidence that there are chances. (Though the Everettian ('many-worlds') program of generating quantum probabilities from subjective uncertainty, without basic chance in the present sense, has recently been gaining adherents—see Barrett 1999; Wallace 2007.) Indeed, it looks like the strongest kind of evidence that there are chances. For if our best physical theories did not feature probabilities, we should have little reason to postulate them, and little reason to take chances to exist. This will become important below (§5), when we discuss classical physics. The conventional view of classical physics, including statistical mechanics, is that it does not involve basic probability (because the state transition dynamics is deterministic), and is not accordingly a theory that posits chances (Loewer 2001).[5] Below, we will examine this view, as well as some of the recent challenges to this conventional view. But there is at least enough evidence from fundamental physics for the existence of chances for us to adopt it already at this point as a defensible assumption.

2. Randomness

As mentioned in the introduction, some philosophers deliberately use 'random' to mean 'chancy'. A random process, in their view, is one governed by chance in the sense of the previous section. This generates stipulative definitions like this one:

> 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)

This process conception of randomness is perfectly legitimate, if somewhat redundant. But it is not adequate for our purposes. It makes the Commonplace Thesis a triviality, and thereby neither interesting in itself nor apt to support the interesting conclusions some have drawn from it concerning explanation or experimental design. Moreover,

The invocation of a notion of process randomness is inadequate in another way, as it does not cover all cases of randomness. Take a clear case of process randomness, such as one thousand consecutive tosses of a fair coin. We would expect, with very high confidence, to toss at least one head. But as that outcome has some chance of not coming to pass, it counts as process random even when it does. This is at variance with what we would ordinarily say about such an outcome, which is not at all unexpected, haphazard, or unpredictable. We could search for some refinement of the notion of process randomness that would reserve the word 'random' for more irregular looking outcomes. But a better approach, and the one we pursue in this entry, is to distinguish between randomness of the process generating an outcome (which we stipulate to amount to its being a chance process), and randomness of the *product* of that random process. In the case just envisaged, we have a random process, while the outomce 'at least one head in 1000 tosses' is not a random product.

The introduction of product randomness helps us make sense of some familiar uses of 'random' to characterise an entire collection of outcomes of a given repeated process. This is the sense in which a *random sample* is random: it is an unbiased representation of the population from which it is drawn—and that is a property of the entire sample, not each individual member. If a random sample is to do its job, it should be irregular and haphazard with respect to the population variables of interest. We should not be able to predict the membership of the sample to any degree of reliability by making use of some other feature of individuals in the population. (So we should not be able to guess at the likely membership of a random sample by using some feature like 'is over 180cm tall'.) A random sample is one that is representative in the sense of being *typical* of the underlying population from which it is drawn, which means in turn that—in the ideal case—it will exhibit no order or pattern that is not exemplified in that underlying population.

While many random samples will be drawn using a random process, they need not be. For example, if we are antecedently convinced that the final digit of someone's minute of birth is not correlated with their family income, we may draw a random sample of people's incomes by choosing those whose birth minute ends in '7', and that process of choice is not at all random. To be sure that our sample is random, we may wish to use random numbers to decide whether to include a given individual in the sample; to that end, large tables of random digits have been produced, displaying no order or pattern (RAND Corporation 1955). This other conception of randomness, as attaching primarily to collections of outcomes, has been termed *product randomness*.

Product randomness also plays an important role in scientific inference. Suppose we encounter a novel phenomenon, and attempt to give a theory of it. All we have to begin with is the data concerning what happened. If that data is highly regular and patterned, we may attempt to give a deterministic theory of the phenomenon. But if the data is irregular and disorderly—random—we may offer only a stochastic theory. As we cannot rely on knowing whether the phenomenon is chancy in advance of developing a theory of it, it is extremely important to be able to characterise whether the data is random or not directly, without detouring through prior knowledge of the process behind it. We might think that we could simply do this by examination of the data—surely the lack of pattern will be apparent to the observer? (We may assume that patternlessness is good evidence for randomness, even if not entailed by it.) Yet psychological research has repeatedly shown that humans are poor at discerning patterns, seeing them in completely random data, and (for precisely the same reason, in fact) failing to see them in non-random data (Gilovich *et al*., 1985; Kahneman and Tversky, 1972; Bar-Hillel and Wagenaar, 1991; Hahn and Warren, 2009). So the need for an objective account of randomness of a sequence of outcomes is necessary for reliable scientific inference.

It might seem initially that giving a rigorous characterisation of disorder and patternlessness is a hopeless task, made even more difficult by the fact that we need to characterise it without using the notion of chance. (Otherwise we make CT trivial.) Yet a series of mathematical developments in the theory of *algorithmic randomness*, culminating in the early 1970s, showed that a satisfactory characterisation of randomness of a sequence of outcomes was possible. This notion has shown its theoretical fruitfulness not only in the

foundations of statistics and scientific inference, but also in connection with the development of information theory and complexity theory. The task of this section is to introduce the mathematical approach to the definition of random sequences, just as we introduced the philosophical consensus on chance in the previous section. We will then be in a position to evaluate the Commonplace Thesis, when made precise using theoretically fruitful notions of chance and randomness.

The fascinating mathematics of algorithmic randomness are largely unknown to philosophers. For this reason, I will give a fairly detailed exposition in this entry. Various points of more technical interest have been relegated to this supplementary document:

> [Supplement B. Further Details Concerning Algorithmic Randomness](#)

Most proofs will be skipped, or relegated to this supplementary document:

> [Supplement C. Proofs of Selected Theorems](#)

Fuller discussions can be found in the cited references.

Throughout the focus will be on a simple binary process, which has only two types of outcome $O = \{0, 1\}$. (The theory of randomness for the outcome sequences of such a simple process can be extended to more complicated sets of outcomes, but there is much of interest even in the question *which binary sequences are product random?*) A *sequence of outcomes* is an ordered collection of events, finite or infinite, such that each event is of a type in O. So a sequence $x = x_1 x_2 \ldots x_k \ldots$, where each $x_i \in O$. The set of all *infinite* binary sequences of outcomes is known as the *Cantor space*. One familiar example of a process the outcomes of which form a Cantor space is an infinite sequence of independent flips of a fair coin, where 1 denotes heads and 0 tails. Notions from measure theory and computability theory are used in the discussion below; an elementary presentation of the mathematics needed can be found in supplement [B.2](#).

2.1 Product Randomness: Random Sequences are Most Likely

Perhaps counterintuitively, we begin with the case of infinite binary sequences. Which of these should count as random products of our binary process? Each individual infinite sequence, whether orderly or not, has measure zero under the standard (Lebesgue) measure over the Cantor space. We cannot determine whether an individual sequence is random from considering what fraction it constitutes of the set of all such sequences. But, intuitively, almost all such infinite sequences should be random and disorderly, and only few will be orderly (an observation first due to Ville 1939). A *typical* infinite sequence is one without pattern; only exceptional cases have order to them. If the actual process that generate the sequences are perfectly deterministic, it may be that a typical product of that process is not random. But we are rather concerned to characterise which of all the possible sequences produced by any process whatsoever are random, and it seems clear that most of the ways an infinite sequence might be produced, and hence most of the sequences so produced, will be random. This fits with intuitive considerations:

> We arrange in our thought, all possible events in various classes; and we regard as *extraordinary* those classes which include a very small number. In the game of heads and tails, 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 1826)

This fertile remark underscores both that random sequences should be *unruly*, and that they should be *common*. In the present framework: the set of *non*-random sequences should have measure zero, proportional

to the set of all such sequences—correspondingly, the set of random sequences should have measure one (Dasgupta, 2011: §3; Gaifman and Snir, 1982: 534; Williams, 2008: 407–11).

This helps, but not much. For there are many measure one subsets of the Cantor space, and we need some non-arbitrary way of selecting a privileged such subset. (The natural option, to take the intersection of all measure one subsets, fails, because the complement of the singleton of any specific sequence is measure one, so for each sequence there is a measure one set which excludes it; therefore the intersection of all measure one sets excludes every sequence, so is the empty set.) The usual response is to take the random sequences to be the intersection of all measure one subsets of the space which have 'nice' properties, and to give some principled delimitation of which properties are to count as 'nice' and why.

For example, if a sequence is genuinely random, we should expect that in the long run it would tend to have features we associate with the outputs of (independent, identically distributed trials of) a chancy process. The sequence should look as disorderly as if it were the expected product of genuine chance. This approach is known accordingly as the *typicality* approach to randomness. Typicality is normally defined with respect to a prior probability function, since what is a typical series of fair coin toss outcomes might not be a typical series of unfair coin toss outcomes (Eagle 2016: 447). In the present case, we use the Lebesgue measure as it is the natural measure definable from the symmetries of the outcome space of the binary process itself.

A typical sequence should satisfy all of the various 'properties of stochasticity' (Martin-Löf 1966: 604). What are these properties? They include the property of large numbers, the claim that the limit frequency of a digit in a random sequence should not be biased to any particular digit. The (strong) *law* of large numbers is the claim that, with probability 1, an infinite sequence of independent, identically distributed Bernoulli trials will have the property of large numbers. If we concentrate on the sequence of outcomes as independently given mathematical entities, rather than as the products of a large number of independent Bernoulli trials, we can follow Borel's (1909) characterisation of the strong law. Let $S_n(x)$ be the number of 1s occurring in the first n places of sequence x (this is just $\sum_{k=1}^{n} x_k$), and let B be the set of infinite sequences x such that the limit of $S_n(x)/n$ as n tends to infinity is $\frac{1}{2}$. Borel's theorem is that B has measure one; almost all infinite sequences are, in the limit, unbiased with respect to digit frequency.

Clearly, the property of large numbers is a necessary condition for randomness of a sequence. It is not sufficient, however. Consider the sequence 10101010…. This sequence is not biased. But it is clearly not random either, as it develops in a completely regular and predictable fashion. So we need to impose additional constraints. Each of these constraints will be another property of stochasticity we should expect of a random sequence, including all other such limit properties of 'unbiasedness'.

One such further property is *Borel normality*, also defined in that paper by Borel. A sequence is Borel normal iff each finite string of digits of equal length has equal frequency in the sequence.[6] Borel proved that a measure one set of sequences in the Cantor space are Borel normal. Borel normality is a useful condition to impose for random sequences, as it has the consequence that there will be no predictable pattern to the sequence: for any string σ appearing multiple times in a sequence, it will as commonly be followed by a 1 as by a 0. This lack of predictability based on the previous elements of the sequence is necessary for genuine randomness. But again Borel normality is not sufficient for randomness. The *Champernowne* sequence (Champernowne 1933) is the sequence of digits in the binary representations of each successive non-negative integer:

$$011011100101110111\ldots$$

This is Borel normal, but perfectly predictable, because there is a general law which states what the value of the sequence at each index will be—not because it can be predicted from prior elements of the sequence, but

because it can be predicted from the index.

We must impose another condition to rule out the Champernowne sequence. We could proceed, piecemeal, in response to various problem cases, to successively introduce further stochastic properties, each of which is a necessary condition for randomness, eventually hoping to give a characterisation of the random sequences by aggregating enough of them together. Given the complex structure of the Cantor space, the prospects for success of such a cumulative approach seem dim. A more promising bolder route is to offer one stochastic property that is by itself necessary and sufficient for randomness, the possession of which will entail the possession of the other properties we've mentioned (the property of large numbers, Borel normality, etc.).

2.1.1 Randomness and Gambling Systems—von Mises' Account

The first detailed and sophisticated attempt at a bolder approach to defining randomness for a sequence with a single stochastic property was by von Mises (von Mises, 1957; von Mises, 1941). Suppose you were presented with any subsequence $x_1, \ldots, x_{n-1}$ of (not necessarily consecutive members of) a sequence, and asked to predict the value of x_n. If the sequence were really random, then this information—the values of any previous members of the sequence, and the place of the desired outcome in the sequence—should be of no use to you in this task. To suppose otherwise is to suppose that there is an exploitable regularity in the random sequence; a gambler could, for example, bet reliably on their preferred outcome and be assured of a positive expected gain if they were in possession of this information. A gambling system selects points in a sequence of outcomes to bet on; a successful gambling system would be one where the points selected have a higher frequency of 'successes' than in the sequence as a whole, so that by employing the system one can expect to do better than chance. But the failure of gambling systems to make headway in games of chance suggests that genuinely random sequences of outcomes aren't so exploitable. Von Mises, observing the empirical non-existence of successful gambling systems, makes it a condition of randomness for infinite sequences that they could not be exploited by a gambling system (his '*Prinzip vom ausgeschlossenen Spielsystem*'). The idea is that without what effectively amounts to a crystal ball, there is no way of selecting a biased selection of members of a random sequence.

Shorn of the inessential presentational device of selecting an outcome to bet on based on past outcomes, von Mises contends that it is a property of stochasticity that a random sequence should not be such that information about any initial subsequence $x_1 x_2 \ldots x_{k-1}$ provides information about the contents of outcome x_k. He formally implements this idea by defining a *place selection* as '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 [value] of the element' (von Mises 1957: 25). He then defines a random sequence as one such that every infinite subsequence selected by an admissible place selection retains the same relative digit frequencies as in the original sequence (so one cannot select a biased subsequence, indicating that this is a genuine property of stochasticity). In our case this will mean that every admissibly selected subsequence will meet the property of large numbers with equal frequency of 1s and 0s.[7] One way to characterise the resulting set of von Mises-random (vM-random) sequences is that it is the largest set which contains only infinite sequences with the right limit frequency and is closed under all admissible place selections. If the limit frequency of a digit is 1, say in the sequence $111\ldots$, it is true that every admissible place selection determines a subsequence with the same limit frequency. Von Mises intends this result, for this is what a random sequence of outcomes of trials with probability 1 of obtaining the outcome 1 looks like. This sequence does not meet the property of large numbers, however. So we modify von Mises' own condition, defining the vM-random sequences as the largest set of infinite sequences which have the limit frequency $\frac{1}{2}$, and which is closed under all admissible place selections.

Von Mises' original proposals were deliberately imprecise about what kinds of procedures count as

admissible place selections. This imprecision did not concern him, as he was disposed to regard the 'right' set of place selections for any particular random collective as being fixed by context and not invariantly specifiable. But his explicit characterisation is subject to counterexamples. Since 'any increasing sequence of natural numbers $n_1 < n_2 < \ldots$ 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 1969b: 27). This clearly violates von Mises' intentions, as he presumably intended that the place selections should be constructively specified, yet the notion of vM-randomness remains tantalisingly vague without some more concrete specification.

Such a specification arrived in the work of Church (1940), drawing on the then newly clarified notion of an effective procedure. Church observed that

> 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 [gambling system] 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)

Church therefore imposes the condition that the admissible place selections should be, not arbitrary functions, but *effectively computable* functions of the preceding outcomes in a sequence. Formally, we consider (following Wald 1938) a place selection as a function f from an initial segment $x_1 x_2 \ldots x_{i-1}$ of a sequence σ into $\{0, 1\}$, such that the selected subsequence $\sigma' = \{x_i : f(x_1 \ldots x_{i-1}) = 1\}$, Church's proposal is that we admit only those place selections which are *computable* (total recursive) functions.[8] Church's proposal applies equally to the sequences von Mises is concerned with, namely those with arbitrary non-zero outcome frequencies for each type of outcome; to get the random sequences, we again make the restriction to those normal binary sequences where the limit relative frequency of each outcome is $\frac{1}{2}$ (these are sometimes called the *Church stochastic* sequences).

As Church points out, if we adopt the Church-Turing computable functions as the admissible place selections, it follows quickly that the set of admissible place selections is countably infinite. We may then show:

> **Theorem 1 (Doob-Wald).** The set of random sequences forms a measure one subset of the Cantor space.

> [Proof]

Thus von Mises' conception of randomness was made mathematically robust (Martin-Löf 1969b). We can see that various properties of stochasticity follow from this characterisation. For example, we can show:

> **Corollary 1.** Every von Mises-random sequence is Borel normal.

> [Proof]

These successes for the approach to randomness based on the impossibility of gambling systems were, however, undermined by a theorem of Ville (1939):

> **Theorem 2 (Ville).** For any countable set of place selections $\{\phi_n\}$ (including the identity), there exists an infinite binary sequence x such that:

>   i. for all m, $\lim_{m \to \infty} (\sum_{k=1}^{m} (\phi_n(x))_k)/m = \frac{1}{2}$; but

ii. for all m, $(\sum_{n=1}^{m} x_n)/m > \frac{1}{2}$.

That is, for any specifiable set of place selections, including the total recursive place selections proposed by Church as the invariantly appropriate set, there exist sequences which have the right limit relative frequencies to satisfy the strong law of large numbers (and indeed Borel normality), as do all their acceptable subsequences, but which are biased in all initial segments.[9]

Why should this be a problem for random sequences? The property of large numbers shows that almost all infinite binary sequences have limit digit frequency $\frac{1}{2}$, but says nothing about how quickly this convergence happens or about the statistical properties of the initial segments. There certainly exist sequences that converge to $\frac{1}{2}$ but where $S_n(x)/n > \frac{1}{2}$ for all n (the sequence has the right mean but converges 'from above'). In the 'random walk' model of our random sequences, where each element in the sequence is interpreted as a step left (if 0) or right (if 1) along the integer line, this sequence would consist of a walk that (in the limit) ends up back at the origin but always (or even eventually) stays to the right. Intuitively, such a sequence is not random.

Such sequences do indeed violate at least one property of stochasticity, as it turns out that in a measure one set of sequences, $S_n(x)/n$ will be above the mean infinitely many times, and below the mean infinitely many times. So stated, this is the *law of symmetric oscillation* (Dasgupta, 2011: 13).[10] Since the law of symmetric oscillations holds for a measure one set of sequences, it is a plausible property of randomness (it is naturally of a family with other properties of stochasticity). Ville's result shows that von Mises' definition in terms of place selections cannot characterise random sequences exactly because it includes sequences that violate this law (so don't correspond to a truly random random walk). Indeed, such sequences don't even correspond to von Mises' avowed aims. As Li and Vitányi say (2008: 54), 'if you bet 1 all the time against such a sequence of outcomes, your accumulated gain is always positive'. As such, Ville-style sequences seem to permit successful gambling, despite the fact that they do not permit a system to be formulated in terms of place selections.

2.1.2 Randomness and Effective Tests: Martin-Löf-Randomness

Von Mises and Church identified a class of sequences, those with limit frequencies invariant under recursive place selections, that satisfied a number of the measure one stochastic properties of sequences that are thought characteristic of randomness. But the class they identified was too inclusive. The next insight in this area was due to (Martin-Löf 1966), who realised that rather than looking for another *single* property of sequences that would entail that the sequence met all the further conditions on randomness, it was simpler to adopt as a definition that a sequence is random iff the sequence has all the measure one properties of randomness that can be specified. Here again recursion theory plays a role, for the idea of *a measure one property of randomness that can be specified* is equivalent to the requirement that there be an effective procedure for testing whether a sequence violates the property. This prompts the following very bold approach to the definition of random sequences:

> **Martin-Löf Randomness**:
> A random sequence is one which cannot be effectively determined to violate a measure one randomness property (Downey and Hirschfeldt 2010: §5.2; Dasgupta 2011: §6.1; Porter 2016: 461–2).

Recalling the definition of effective measure zero from supplement B.2, Martin-Löf suggests that a random sequence is any that does not belong to any effective measure zero set of sequences, and thus belongs to every effective measure one set of sequences. An effective measure zero set of sequences will contain

sequences which can be effectively determined to have a 'special hallmark' (for example, having a '1' at every seventh place, or never having the string '000000' occurring as a subsequence). It is part of von Mises' insight that no random sequence will possess any of these effectively determinable special hallmarks: such hallmarks would permit exploitation as part of a gambling system. But Martin-Löf notices that all of the commonly used measure one properties of stochasticity are effective measure one. Any sequence which violates the property of large numbers, or the law of symmetric oscillations, etc., will do so on increasingly long initial subsequences. So the violation of any such property will also be a special hallmark of a non-random sequence, an indicator that the sequence which possesses it is an unusual one. Since the unusual properties of non-stochasticity in question are effective measure zero, we can therefore say that the random sequences are those which are not special in any effectively determinable way. To formalise this, Martin-Löf appeals to the language of significance testing. His main result is sometimes put as the claim that random sequences as those which pass all recursive significance tests for sequences (Schnorr 1971: §1)—they are never atypical enough to prompt us to reject the hypothesis that they are random. See supplement B.1.1 for more detail on this point.

Note that the restriction to effective properties of sequences is crucial here. If we allowed, for example, the property *being identical to my favourite random sequence x*, that would define a test which the sequence x would fail, even though it is random. But it follows from our observations about von Mises randomness (which is still a necessary condition on randomness) that no effectively computable sequence is random (if it were, there would be a place selection definable from the algorithm that selected all the 1s in the sequence). So there is no effective test that checks whether a given sequence is identical to some random sequence.

The central result of Martin-Löf (1966) is the following:

> **Theorem 3 (Universal Tests and the Existence of Random Sequences).** There is a universal test for ML-randomness; moreover, only a measure zero set of infinite binary sequences fails this test. So almost all such sequences are ML-random.

[Proof]

The universal test does define an effective measure one property, but (unlike normality, or having no biased admissible subsequences), it is far from a naturally graspable property. Nevertheless, Martin-Löf's result does establish that there are random sequences that satisfy all the properties of stochasticity, and that in fact almost all binary sequences are random in that sense. Returning to Ville's theorem 2, it can be shown that all ML-random sequences satisfy the law of symmetric oscillations (van Lambalgen 1987a: §3.3). Hence the kind of construction Ville uses yields vM-random sequences which are not ML-random. All the ML-random sequences have the right limit relative frequencies, since they satisfy the effective measure one property of large numbers. So Martin-Löf random sequences satisfy all the intuitive properties we would expect of a sequence produced by tosses of a fair coin, but are characterised entirely by reference to the effectively specifiable measure one sets of infinite sequences. We have therefore characterised random sequences entirely in terms of the explicit features of the product, and not of the process that may or may not lie behind the production of these sequences.

There are other accounts that develop and extend Martin-Löf's account of randomness, in the same kind of framework, such as that of Schnorr (1971); for some further details, see supplement B.1.2.

2.2 Product Randomness: Random Sequences are Most Disorderly

For infinite binary sequences, the Martin-Löf definition in terms of effective tests is a robust and mathematically attractive notion. However, it seems to have the major flaw that it applies only to infinite binary sequences. (Since finiteness of a sequence is effectively positively decidable, and the set of all finite

sequences is measure zero, every finite sequence violates an effective measure one randomness property.) Yet ordinarily we are happy to characterise even quite small finite sequences of outcomes as random. As mentioned above (§2), there is room for doubt at our ability to do so correctly, as we seem to be prone to mischaracterise sequences we are presented with, and perform poorly when asked to produce our own random sequences. However, there is nothing in this literature to suggest that we are fundamentally mistaken in applying the notion of randomness to finite sequences at all. So one might think this shows that the Martin-Löf approach is too restrictive.

Yet there is something in the idea of ML-randomness that we might apply profitably to the case of finite sequences. Since being generated by an effective procedure is a measure zero property of infinite sequences, given that there are only countably many effective procedures, it follows immediately that no ML-random sequence can be effectively produced. This fits well with the intuitive idea that random sequences don't have the kind of regular patterns that any finite algorithm, no matter how complex, must exploit in order to produce an infinite sequence. This contrast between random sequences which lack patterns that enable them to be algorithmic generated, and non-random sequences which do exhibit such patterns, does not apply straightforwardly to the finite case, because clearly there is an effective procedure which enables us to produce any particular finite sequence of outcomes—simply to list those outcomes in the specification of the algorithm. But a related contrast does exist—between those algorithms which are simply crude lists of outcomes, and those which produce outcomes which involve patterns and regularities in the outcome sequence. This leads us to the idea that finite random sequences, like their infinite cousins, are not able to be generated by an algorithm which exploits patterns in the outcome sequence. The outcomes in random sequences are thus patternless, or *disorderly*, in a way that is intuitively characteristic of random sequences.

2.2.1 Kolmogorov Complexity and Randomness

Disorderly sequences, in the above sense, are highly *incompressible*. The best effective description we can give of such a sequence—one that would enable someone else, or a computer, to reliably reproduce it—would be to simply list the sequence itself. This feature allows us to characterise the random sequences as those which cannot be produced by a compact algorithm (compact with respect to the length of the target sequence, that is). Given that algorithms can be specified by a list of Turing machine instructions, we have some basic idea on how to characterise the length of an algorithm. We can then say that a random sequence is one such that the shortest algorithm which produces it is approximately (to be explained below) the same length as the sequence itself—no greater compression in the algorithm can be attained. This proposal, suggested by the work of Kolmogorov, Chaitin and Solomonov (KCS), characterises randomness as the algorithmic or informational *complexity* of a sequence. Comprehensive surveys of complexity and the complexity-based approach to randomness are Li and Vitányi 2008 and Downey and Hirschfeldt 2010: Part I. (See also Chaitin 1975, Dasgupta 2011: §7, Downey et al. 2006: §§1–3, Earman 1986: 141–7, Kolmogorov 1963, Kolmogorov and Uspensky 1988, Smith (1998: ch. 9) and van Lambalgen 1995.)

If f is effectively computable—a recursive function—let us say that δ is an f-description of a finite string σ iff f yields σ on input δ. 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. If there is no such δ, let the f-complexity of σ be infinite. f is thus a *decompression* algorithm, taking the compressed description δ back to the original string σ. There are obviously many different kinds of decompression algorithm. One boring case is the identity function (the empty program), which takes each string to itself. The existence of this function shows that there are decompression algorithms f which have a finite f-complexity for any finite string. Any useful decompression algorithm will, however, yield an output string significantly longer than the input description, for at least some input descriptions.

One example is this algorithm: on input of a binary string δ of length $4n$, the algorithm breaks the input down

into n blocks of 4, which it turns into an output sequence σ, as follows. Given a block $b_1, \ldots, b_4$, it produces a block of the symbol contained in b_1, the length of which is governed by the binary numeral $b_2 b_3 b_4$. So the block 1101 produces a string of five 1s. The output sequence is obtained by concatenating the output of successive blocks in order. Every string σ can be represented by this algorithm, since the string σ' which involves replacing every 1 in σ by 1001, and every 0 by 0001, will yield σ when given as input to this algorithm. So this algorithm has finite complexity for any string. But this algorithm can do considerably better; if the original string, for example, is a string of sixteen 1s, it can be obtained by input of this description: 11111111, which is half the length. Indeed, as reflection on this algorithm shows, this algorithm can reconstruct an original string from a shorter description, for many strings, particularly if they contain reasonably long substrings of consecutive 1s or 0s.

However, there is a limit on how well any algorithm can compress a string. If $|\sigma|$ is the length of σ, say that a string σ is *compressed* by f if there is an f-description δ such that $\sigma = f(\delta)$ and $|\delta| < |\sigma|$. If a useful decompression algorithm is such that for some fixed k, $|f(\delta)| \leq |\delta| + k$, so that f-descriptions are at least k shorter than the sequence to be compressed, then it follows that very few strings usefully compress. For there are 2^l strings σ such that $|\sigma| = l$; so there are at most 2^{l-k} f-descriptions; since f is a function, there are at most 2^{l-k} compressible strings. As a proportion of all strings of length l, then, there are at most $2^{l-k}/2^l = \frac{1}{2}^k$ compressible strings . This means that as the amount of compression required increases, the number of sequences so compressible decreases exponentially. Even in the most pitiful amount of compression, $k = 1$, we see that at most half the strings of a given length can be compressed by any algorithm f.

So our interest must be in those decompression functions which do best overall. We might hope to say: f is better than g iff for all σ, $C_f(\sigma) \leq C_g(\sigma)$. Unfortunately, no function is best in this sense, since for any given string σ with f-complexity $|\sigma| - k$, we can design a function g as follows: on input 1, output σ; on input $n\delta$ (for any n), output $f(\delta)$. (Generalising, we can add arbitrarily long prefixes of length m onto the inputs to g and have better-than$-f$ compression for 2^m sequences.) But we can define a notion of *complexity near-superiority* of f to g iff there is some constant k such that for any string, $C_f(\sigma) \leq C_g(\sigma) + k$. f is least as good as g, subject to some constant which is independent of the functions in question. If f and g are both complexity near-superior to each other, for the same k, we say they are *complexity equivalent*.

Kolmogorov (1965) showed that there is an optimal decompression algorithm:

> **Theorem 4 (Kolmogorov).** There exists a decompression algorithm which is near-superior to any other program. Moreover, any such optimal algorithm is complexity equivalent to any other optimal algorithm (see also Chaitin 1966 and Martin-Löf 1969a).

> [Proof]

Such a universal function Kolmogorov called *asymptotically optimal* (for as $|\sigma|$ increases, the constant k becomes asymptotically negligible).

Choose some such asymptotically optimal function u, and define the complexity (simpliciter) $C(\sigma) = C_u(\sigma)$. Since u is optimal, it is near-superior to the identity function; it follows that there exists a k such that $C(\sigma) \leq |\sigma| + k$. On the other hand, we also know that the number of strings for which $C(\sigma) \leq |\sigma| - k$ is at most $\frac{1}{2}^k$. We know therefore that all except $1 - 2^k$ sequences of length n have a complexity within k of n. As n increases, for fixed large k, therefore, we see that almost all sequences have complexity of approximately their length. All this can be used to make precise the definition of randomness sketched above.

> **Kolmogorov random**:

We say that a sequence σ is *Kolmogorov-random* iff $C(\sigma) \approx |\sigma|$.

It follows from what we have just said that there exist random sequences for any chosen length n, and that as n increases with respect to k, random sequences come to be the overwhelming majority of sequences of that length.

Theorem 5. A random sequence of a given length cannot be effectively produced.

[Proof]

An immediate corollary is that the complexity function C is not a recursive function. If it were, for any n, we could effectively compute $C(\sigma)$ for any σ of length n. By simply listing all such sequences, we could halt after finding the first σ for which $C(\sigma) \geq n$. But then we could effectively produce a random sequence, contrary to theorem 5.

The notion of Kolmogorov randomness fits well with the intuitions about the disorderliness of random sequences we derived from the Martin-Löf account. It also fits well with other intuitions about randomness — random sequences don't have a short description, so there is no sense in which they are produced in accordance with a plan. As such, Kolmogorov randomness also supports von Mises' intuitions about randomness being linked to the impossibility of gambling systems, as there will be no way of effectively producing a given random sequence of outcomes using a set of initially given data any smaller than the sequence itself. There is no way of predicting a genuinely random sequence in advance because no random sequence can be effectively produced, yet every predictable sequence of outcomes can (intuitively) be generated by specifying the way in which future outcomes can be predicted on the basis of prior outcomes. Moreover, because for increasing k the number of strings of length n which are random increases, and because for increasing n we can choose larger and larger k, there is some sense in which the great majority of sequences are random; this matches well the desiderata in the infinite case that almost all sequences should be random. Finally, it can be shown that the Kolmogorov randomness of a sequence is equivalent to that sequence passing a battery of statistical tests, in the Martin-Löf sense — indeed, that the Kolmogorov random sequences are just those that pass a certain universal test of non-randomness (Martin-Löf 1969a: §2).[11]

2.2.2 Prefix-Free Kolmogorov Complexity

The plain Kolmogorov complexity measure is intuitively appealing. Yet the bewildering variety of permitted f-descriptions includes many unmanageable encodings. In particular, for a given decompression algorithm f, there are f-descriptions γ and δ such that $\delta = \gamma^\frown \tau$, for some string τ. This is an inefficient encoding, because if γ can occur both as a code itself, and as the initial part of another code, then an algorithm cannot decode its input string 'on the fly' as soon as it detects a comprehensible input, but must wait until it has scanned and processed the entire input before beginning to decode it. An efficient coding, such that no acceptable input is an initial substring of another acceptable input, is called *prefix-free* (because no member is a prefix of any other member). A good example of an encoding like this is the encoding of telephone numbers: the telephone exchange can, on input of a string of digits that it recognises, immediately connect you; once an acceptable code from a prefix-free set has been input, no other acceptable code can follow it.

Prefix-free encodings are useful for a number of practical purposes, and they turn out to be useful in defining randomness also. (As we will see in §2.3, they are of special importance in avoiding a problem in the definition of infinite Kolmogorov random sequences.) The change is the natural one: we appeal, not to the plain complexity of a sequence in defining its randomness, but the so-called *prefix-free* complexity (Downey and Hirschfeldt 2010: §2.5ff; Dasgupta 2011: §8).

To fix ideas, it is useful to have an example of a prefix-free encoding in mind. Suppose we have a string $\sigma = x_1 \ldots x_k$ of length k. This is the initial part of the string $\sigma 1$, so if any string was an acceptable input, we would not have a prefix-free encoding. But if the code contained information about the *length* of the string encoded, we would know that the length of σ, k, is less than the length of $\sigma 1$. We can make this idea precise as follows (using a code similar to that used to a different end in the [proof](#) of Theorem [4](#)). Let the code of σ be the string $1^{[|\sigma|]} 0 \sigma$—that is, the code of a string consists of a representation of the length of the string, followed by a 0, followed by the string. This is clearly a prefix-free encoding.[12] This coding is not particularly efficient, but more compact prefix-free encodings do exist.

The notion of prefix-free complexity is defined in exactly the same way as plain complexity, with the additional restriction that the allowable f-descriptions of a string, given a decompression function f, must form a prefix-free set. With an appropriate choice of coding, we can get a set of f-descriptions which is monotonic with increasing length, i.e., if $|\gamma| < |\delta|$ then $|f(\gamma)| < |f(\delta)|$. Our definition goes much as before: The prefix free Kolmogorov complexity, given a decompression function f with a prefix-free domain, of a string σ, denoted $K_f(\sigma)$, is the length of the shortest f-description of σ (and infinite otherwise). Since $1^{[|\sigma|]} 0 \sigma$ is a finite prefix-free code for σ, we know there are at least some prefix-free decompression algorithms with finite K_f for every string. As before, we can show there exist better decompression algorithms than this one, and indeed, that there exists a universal prefix-free decompression algorithm u, such that for every other algorithm m there is a k such that $K_u(\sigma) \leq K_m(\sigma) + k$, for all σ (Downey and Hirschfeldt 2010: §2.5). We define $K(\sigma) = K_u(\sigma)$.

Since the set of prefix-free codes is a subset of the set of all possible codes, we should expect generally that $C(\sigma) \leq K(\sigma)$. On the other hand, we can construct a universal prefix-free algorithm u as follows. A universal Turing machine u' takes as input the Gödel number of the Turing machine we wish to simulate, and the input we wish to give to that machine. Let us concatenate these two inputs into a longer input string that is itself uniquely readable; and we then encode that longer string into our prefix-free encoding. The encoding is effectively computable, clearly, so we can chain a decoding machine together with our universal machine u'; on input an acceptable prefix-free string, the decoder will decompose it into the input string, we then decompose the input string into the Gödel number and the input, and run the machine u' on that pair of inputs. Depending on the particular encoding we choose, we may establish various bounds on K; one obvious bound we have already established is that $C(\sigma) \leq K(\sigma) < C(\sigma) + 2|\sigma|$. By using a more efficient prefix-free coding of a u'-description, we can establish better bounds.[13] (Some more results on the connection between K and C are in Downey and Hirschfeldt 2010: §§3.1–3.2.)

With prefix-free complexity in hand, we may define:

> **Prefix-free Kolmogorov Randomness**:
> A string σ is Prefix-free Kolmogorov random iff $K(\sigma) \geq |\sigma|$ (modulo an additive constant).

Again, there do exist prefix-free random sequences, since we know that there are plain random sequences, and given the greater length of a prefix-free encoding, we know that the prefix-free code of ordinary random sequence will be generally longer than an arbitrary code of it, and thus random too. Indeed, there will be more prefix-free random sequences because strings compress less effectively under K than C. Yet K and C behave similarly enough that the success of plain Kolmogorov complexity at capturing our intuitions about randomness carry over to prefix-free Kolmogorov randomness, and the label 'Kolmogorov random' has come to be used generally to refer to prefix-free Kolmogorov random sequences.

2.3 Schnorr's theorem: Kolmogorov and ML-randomness coincide

Both plain and prefix-free Kolmogorov randomness provide satisfactory accounts of the randomness of finite sequences. One difficulty arises, as I suggested earlier, when we attempt to extend plain Kolmogorov randomness to the case of infinite sequences in the most obvious way, that is, by defining an infinite sequence as Kolmogorov random iff all finite initial segments are Kolmogorov random. It would then turn out that no infinite sequence is random. Why? Because of the following theorem, which shows that there is no sequence such that all of its initial segments are random:

> **Theorem 6 (Martin-Löf 1966)**. For any sufficiently long string, there will always be some fairly compressible initial segments. (See also Li and Vitányi 2008: §2.5.1 and Downey and Hirschfeldt 2010: §2.1.)

> [Proof]

This dip in complexity of an initial subsequence will occur infinitely often in even a random infinite sequence, a phenomenon known as *complexity oscillation* (Li and Vitányi 2008: §2.5.1). This phenomenon means that 'it is difficult to express a universal sequential test precisely in terms of C-complexity' (Li and Vitányi 2008: 151), and the best that can be precisely done is to find upper and lower bounds expressible in terms of ordinary Kolmogorov complexity between which the set of ML-random sequences falls (Li and Vitányi 2008: § 2.5.3).

However, the phenomenon of complexity oscillation does not pose as significant a problem for prefix-free Kolmogorov complexity. Complexity oscillation does arise, but in fact the inefficiency of prefix-free encodings is a benefit here: 'K exceeds C by so much that the complexity of the prefix does not drop below the length of the prefix itself (for random infinite ω)' (Li and Vitányi 2008: 221). That is, while the complexity of some initial segments dips down, it always remains greater than the length of the prefix. So it can be that, uniformly, when x is an infinite sequence, for any of its initial subsequences $\sigma, K(\sigma) \geq |\sigma|$. This suggests that we can extend prefix-free Kolmogorov complexity to the infinite case in the straightforward way: *an infinite sequence x* is prefix-free Kolmogorov random iff every finite initial subsequence is prefix-free Kolmogorov random.

With this definition in hand we obtain a very striking result. The class of infinite prefix-free Kolmogorov random sequences is certainly non-empty. Indeed: it is just the class of ML-random sequences!

> **Theorem 7 (Schnorr)**. A sequence is ML-random iff it is prefix-free Kolmogorov random.

> [Proof]

Schnorr's theorem is evidence that we really have captured the intuitive notion of randomness. Different intuitive starting points have generated the same set of random sequences. This has been taken to be evidence that ML-randomness or equivalently (prefix-free) Kolmogorov randomness is really the intuitive notion of randomness, in much the same way as the coincidence of Turing machines, Post machines, and recursive functions was taken to be evidence for *Church's Thesis*, the claim that any one of these notions captures the intuitive notion of effective computability. Accordingly, Delahaye (1993) has proposed the *Martin-Löf-Chaitin Thesis*, that either of these definitions captures the intuitive notion of randomness. If this thesis is true, this undermines at least some sceptical contentions about randomness, such as the claim of Howson and Urbach (1993: 324) that 'it seems highly doubtful that there is anything like a unique notion of randomness there to be explicated'.

There are some reasons to be suspicious of the Martin-Löf-Chaitin Thesis, despite the mathematically elegant convergence between these two mathematical notions. For one, there is quite a bit of intuitive support for accounts of randomness which do not make it primarily a property of sequences, and those other accounts are

no less able to be made mathematically rigorous (see especially the 'epistemic' theories of randomness discussed in §6.2, as well as theories of randomness as indeterminism discussed in §7.2). The existence of other intuitive notions makes the case of randomness rather unlike the supposedly analogous case of Church's Thesis, where no robust alternative characterisation of effective computability is available.

Even if we accept that randomness, like disorder, is at root a product notion, there are a number of candidates in the vicinity of the set identified by Schnorr's thesis that might also deserve to be called the set of random sequences. Most obviously, there is Schnorr's own conception of randomness (§2.1.2; supplement B.1.2). Schnorr (1971) suggests that, for technical and conceptual reasons, Schnorr randomness is to be preferred to Martin-Löf randomness as an account of the intuitive notion. While results that parallel the convergence of ML-randomness and Kolmogorov randomness have been given (Downey and Griffiths 2004), the relevant compressibility notion of randomness for Schnorr randomness was not known until quite recently, and is certainly less intuitively clear than Kolmogorov randomness. Moreover, since the set of ML-random sequences is a strict subset of the set of Schnorr random sequences, any problematic members of the former are equally problematic members of the latter; and of course there will be Schnorr random sequences which fail some Martin-Löf statistical test, which might lead some to reject the viability of Schnorr's notion from the start.

Schnorr's result showing the convergence between prefix-free Kolmogorov complexity and Martin-Löf randomness is very suggestive. As has become clear, the existence of other notions of randomness—incluing Schnorr randomness, as well as a number of other proposals (Li and Vitányi 2008: §2.5; Porter 2016: 464–6) —shows that we should be somewhat cautious in yielding to its suggestion.

This is especially true in light of a recent argument by Porter (2016: 469–70). He considers a certain schematic characterisation of computable functions, something like this: for every computable function f with property P_i, f is differentiable at x iff x corresponds to a random$_i$ sequence. It turns out that for each sense of randomness (ML randomness, Schnorr randomness, computable randomness, etc.), there is some corresponding property of computable functions. Most importantly, none of these properties look overwhelmingly more natural or canonical than the others. For example, computable functions of bounded variation are differentiable at the Martin-Löf random points, while nondecreasing computable functions are differentiable at the computably random points. The difficulty for the Martin-Löf-Chaitin Thesis is this: these results give us the notion of a typical sequence, with respect to the differentiability of various kinds of function. Unfortunately, these notions of typical sequence diverge from one another. Unlike Church's thesis, where all the notions of effective computability line up, here we have a case where various notions of a typical sequence do not line up with each other (though there is significant overlap). Porter concludes that 'no single definition of randomness can do the work of capturing every mathematically significant collection of typical points' (Porter 2016: 471).

That conclusion may well be justified. But we can largely sidestep the dispute over whether there is a single precise notion of randomness that answers perfectly to our intuitive conception of random sequence. Kolmogorov-Martin-Löf randomness is a reasonable and representative exemplar of the algorithmic approach to randomness, and it overlaps almost everywhere with any other plausible definition of randomness. It is adopted here as a useful working account of randomness for sequences. None of the difficulties and problems I raise below for the connection between random sequences and chance turns in any material way on the details of which particular set of sequences gets counted as random (most are to do with the mismatch between the process notion of chance and any algorithmic conception of randomness, with differences amongst the latter being relatively unimportant). So while the observations below are intended to generalise to Schnorr randomness and other proposed definitions of random sequences, I will explicitly treat only KML randomness in what follows.

3. The Commonplace Thesis Refined

The notions of chance and randomness discussed and clarified in the previous two sections are those that have proved scientifically and philosophically most fruitful. Whatever misfit there may be between ordinary language uses of these terms and these scientific precisifications, is made up for by the usefulness of these concepts. This is particularly so with the notion of randomness, chance being in philosopher's mouths much closer to what we ordinarily take ourselves to know about chance. On these conceptions, randomness is fundamentally a product notion, applying in the first instance to sequences of outcomes, while chance is a process notion, applying in the single case to the process or chance setup which produces a token outcome. Of course the terminology in common usage is somewhat slippery; it's not clear, for example, whether to count random sampling as a product notion, because of the connection with randomness, or as a process notion, because sampling is a process. The orthodox view of the process, in fact, is that it should be governed by a random sequence; we enumerate the population, and sample an individual n just in case the n-th outcome in a pre-chosen random sequence is 1. (Of course the sample thus selected may not be random in some intuitive sense; nevertheless, it will not be biased because of any defect in the selection procedure, but rather only due to bad luck.)

With these precise notions in mind, we can return to the Commonplace Thesis CT connecting chance and randomness. Two readings make themselves available, depending on whether we take single outcomes or sequences of outcomes to be primary:

> **CTa**:
> A sequence of outcomes happens by chance iff that sequence is random.

> **CTb**:
> An outcome happens by chance iff there is a random sequence of outcomes including it.

Given the standard probability calculus, any sequence of outcomes is itself an outcome (in the domain of the chance function defined over a σ-algebra of outcomes, as in standard mathematical probability); so we may without loss of generality consider only (CTb). But a problem presents itself, if we consider chancy outcomes in possible situations in which only very few events ever occur. It may be that the events which do occur, by chance, are all of the same type, in which case the sequence of outcomes won't be random. This problem is analogous to the 'problem of the single case' for frequency views of chance ([Hájek 2012](#): [§3.3](#)), because randomness is, like frequency, a property of an outcome sequence. The problem arises because the outcomes may be too few or too orderly to properly represent the randomness of the entire sequence of which they are part (all infinite random sequences have at least some non-random initial subsequences). The most common solution in the case of frequentism was to opt for a hypothetical outcome sequence—a sequence of outcomes produced under the same conditions with a stable limit frequency (von Mises 1957: 14–5). Likewise, we may refine the commonplace thesis as follows:

> **RCT**:
> An outcome happens by chance iff, were the trial which generated that outcome repeated often enough under the same conditions, we would obtain a random sequence including the outcome (or of which the outcome is a subsequence).

Here the idea is that chancy outcomes would, if repeated often enough, produce an appropriately homogenous sequence of outcomes which is random. If the trial is actually repeated often enough, this sequence should be the actual sequence of outcomes; the whole point of Kolmogorov randomness was to permit finite sequences to be random.

RCT is intuitively appealing, even once we distinguish process and product randomness in the way sketched above. It receives significant support from the fact that fair coins, tossed often enough, do in our experience invariably give rise to random sequences, and that the existence of a random sequence of outcomes is compelling evidence for chance. The truth of RCT explains this useful constraint on the epistemology of chance, since if we saw an actual finite random sequence, we could infer that the outcomes constituting that sequence happened by chance. However, in the next two sections, we will see that there are apparent counterexamples even to RCT, posing grave difficulties for the Commonplace Thesis. In §4 we will see a number of cases where there are apparently chancy outcomes without randomness, while in §5 we will see cases of apparent randomness where there is no chance involved.

A fundamental problem with RCT seems to emerge when we consider the fate of hypothetical frequentism as a theory of chance. For there seems to be no fact of the matter, in a case of genuine chance, as to what sequence of outcomes would result: as Jeffrey (1977: 193) puts it, 'there is no telling whether the coin would have landed head up on a toss that never takes place. That's what probability is all about.' Many philosophers (e.g., Hájek 2009) have followed Jeffrey in scepticism about the existence and tractability of the hypothetical sequences apparently required by the right hand side of RCT. However, there is some reason to think that RCT will fare better than hypothetical frequentism in this respect. In particular, RCT does not propose to *analyse* chance in terms of these hypothetical sequences, so we can rely on the law of large numbers to guide our expectation that chance processes produce certain outcome frequencies, in the limit, with probability 1; this at least may provide some reason for thinking that the outcome sequences will behave as needed for RCT to turn out true. Even so, one might suspect that many difficulties for hypothetical frequentism will recur for RCT. However, these difficulties stem from general issues with merely possible evidential manifestations of chance processes, and have nothing specifically to do with randomness. The objections canvassed below are, by contrast, specifically concerned with the interaction between chance and randomness. So these more general potential worries will be set aside, though they should not be forgotten—they may even be, in the end, the most significant problem for RCT (if the right kind of merely possible sequence doesn't exist, we must retreat to CTa or CTb and their problems).

4. Chance Without Randomness

4.1 Unrepresentative Outcome Sequences

It is possible for a fair coin—i.e., such that the chances of heads and tails are equal—to be tossed infinitely many times, and to land heads on every toss. An infinite sequence of heads has, on the standard probability calculus, zero chance of occurring. (Indeed, it has zero chance even on most non-standard views of probability: Williamson 2007.) Nevertheless if such a sequence of outcomes did occur, it would have happened by chance—assuming, plausibly, that if each individual outcome happens by chance, the complex event composed by all of them also happens by chance. But in that case we would have an outcome that happened by chance and yet the obvious suitable sequence of outcomes is not KML-random. This kind of example exploits the fact that while random sequences are a measure one set of possible outcome sequences of any process, chancy or otherwise, measure one does not mean *every*.

This counterexample can be resisted. For while it is possible that a fair coin lands only heads when tossed infinitely many times, it may not be that this all heads outcome sequence is a suitable sequence. For if we consider the counterfactual involved in RCT—what would happen, if a fair coin were tossed infinitely many times—we would say: it would land heads about half the time. That is (on the standard, though not uncontroversial, Lewis-Stalnaker semantics for counterfactuals: Lewis 1973), though the all-heads outcome sequence is possible, it does not occur at any of the nearest possibilities in which a fair coin is tossed infinitely many times.

If we adopt a non-reductionist account of chance, this line of resistance is quite implausible. For there is nothing inconsistent on such views about a situation where the statistical properties of the occurrent sequence of outcomes and the chance diverge arbitrarily far, and it seems that such possibilities are just as close in relevant respects as those where the occurrent outcome statistics reflect the chances. In particular, as the all-heads sequence has some chance of coming to pass, there is (by BCP) a physical possibility sharing history and laws with our world in which all-heads occurs. This looks like a legitimately close possibility to our own.

Prospects look rather better on a reductionist view of chance (Supplement A.3). On such a view, we can say that worlds where an infinite sequence of heads does occur at some close possibility will look very different from ours; they differ in law or history to ours. In such worlds, the chance of heads is much closer to 1 (reflecting the fact that if a coin were tossed infinitely many times, it might well land heads on each toss)— the coin is not after all fair. The response, then, is that in any situation where the reductionist chance of heads really is 0.5, suitable outcome sequences in that situation or its nearest neighbours *are* in fact *all* unbiased with respect to outcome frequencies. That is to say, they at least satisfy the property of large numbers; and arguably they can be expected to meet other randomness properties also. So, on this view, there is no counterexample to RCT from the mere possibility of these kinds of extreme outcome sequences. This response depends on the success of the reductionist similarity metrics for chancy counterfactuals developed by Lewis (1979a) and Williams (2008); the latter construction, in particular, invokes a close connection between similarity and randomness. (Lewis' original construction is criticised in Hawthorne 2005.)

However, we needn't use such an extreme example to make the point. For the same phenomenon exists with almost any unrepresentative outcome sequence. A fair coin, tossed 1000 times, has a positive chance of landing heads more than 700 times. But any outcome sequence of 1000 tosses which contains more than 700 heads will be compressible (long runs of heads are common enough to be exploited by an efficient coding algorithm, and 1000 outcomes is long enough to swamp the constants involved in defining the universal prefix-free Kolmogorov complexity). So any such outcome sequence will not be random, even though it quite easily could come about by chance. The only way to resist this counterexample is to refuse to acknowledge that such a sequence of outcomes can be an appropriate sequence in RCT. This is implausible, for such sequences can be actual, and can be sufficiently long to avoid the analogue of the problem of the single case, certainly long enough for the Kolmogorov definition of randomness to apply. The only reason to reject such sequences as suitable is to save RCT, but that is clearly question begging in this context. In this case of an unrepresentative finite sequence, even reductionism about chance needn't help, because it might be that other considerations suffice to fix the chance, and so we can have a genuine fair chance but a biased and non-random outcome sequence.

4.2 The Reference Class Problem

Any given token event is an instance of many different types of trial:

> It is obvious that every individual thing or event has an indefinite number of properties or attributes observable in it, and might therefore be considered as belonging to an indefinite number of different classes of things… (Venn 1876: 194)

Suppose Lizzie tosses a coin on Tuesday; this particular coin toss may be considered as a coin toss; a coin toss on a Tuesday; a coin toss by Lizzie; an event caused by Lizzie; etc. Each of these ways of typing the outcome give rise to different outcome sequences, some of which may be random, while others are not. Each of these outcome sequences is unified by a homogenous kind of trial; as such, they may all be suitable sequences to play a role in RCT.

This is not a problem if chance too is relative to a type of trial, for we may simply make the dependence on

choice of reference class explicit in both sides of RCT. If chances were relative frequencies, it would be easy enough to see why chance is relative to a type of trial. But chances aren't frequencies, and single-case chance is almost universally taken to be not only well-defined for a specific event, but unique for that event. We naturally speak of *the* chance that this coin will lands heads on its next toss, with the chance taken to be a property of the possible outcome directly, and not mediated by some particular description of that outcome as an instance of this or that kind of trial. Moreover, for chance to play its role in the Principal Principle (§1 and Supplement A.1), there must be a unique chance for a given event that is to guide our rational credence, as we have only one credence in a particular proposition stating the occurrence of that event. (Indeed, the inability of frequentists to single out a unique reference class, the frequency in which is the chance, was taken to be a decisive objection to frequentism.) On the standard understanding of chance, then, there is a mismatch between the left and right sides of the RCT. And this gives rise to a counterexample to RCT, if we take an event with a unique non-trivial single-case chance, but such that at least one way of classifying the trial which produced it is such that the sequence of outcomes of all trials of that kind is not random. The trivial case might be this: a coin is tossed and lands heads. That event happened by chance, yet it is of the type 'coin toss which lands heads', and the sequence consisting of all outcomes of that type is not random.

The natural response—and the response most frequentists offered, with the possible exception of von Mises[14]—was to narrow the available reference classes. (As noted in Supplement A.3, many frequentists were explicit that chances were frequencies in repetitions of natural kinds of processes.) Salmon (1977) appeals to objectively homogenous reference classes (those which cannot be partitioned by any relevant property into subclasses which differ in attribute frequency to the original reference class). Salmon's proposal, in effect, is that homogeneous reference classes are random sequences, the evident circularity of which will hardly constitute a reply to the present objection. Reichenbach (1949: 374) proposed to 'proceed by considering the narrowest class for which reliable statistics can be compiled', which isn't circular, but which fails to respond to the objection since it provides no guarantee that there will be only one such class. There may well be multiple classes which are equally 'narrow' and for which reliable statistics can be collected (Gillies, 2000: 816). In the present context, this will amount to a number of sequences long enough to make for reliable judgements of their randomness or lack thereof.

This objection requires the chance of an event to be insensitive to reference class. Recently, Hájek (2007) has argued that *no* adequate conception of probability is immune to a reference class problem, so that this requirement cannot be met. (For a related view of relativised chance, though motivated by quite different considerations, see Glynn 2010.) However, as Hájek notes, this conclusion makes it difficult to see how chance could guide credence, and it remains an open question whether a relativised theory of chance that meets the platitudes concerning chance can be developed.

4.3 Bias

The two previous problems notwithstanding, many have found the most compelling cases of chance without randomness to be situations in which there is a biased chance process. A sequence of unfair coin tosses will have an unbalanced number of heads and tails, and such a sequence cannot be random. But such a sequence, and any particular outcome in that sequence, happens by chance.

That such sequences aren't random can be seen by using both Martin-Löf- and Kolmogorov-style considerations. In the latter case, as we have already seen, a biased sequence will be more compressible than an unbiased sequence, if the sequence is long enough, because an efficient coding will exploit the fact that biased sequences will typically have longer subsequences of consecutive digits, and hence will not be random. In the former case, a biased sequence will violate at least one measure one property, on the standard Lebesgue measure on infinite binary sequences—in particular, a measure one subset of the Cantor space will be Borel normal (§2.1), but no biased sequence is Borel normal. So on the standard account of randomness,

no sequences of outcomes of a biased chance process are random, but of course these outcomes happened by chance.

One response to this problem is to try and come up with a characterisation of randomness which will permit the outcomes of biased chances to be random. It is notable that von Mises' initial characterisation of randomness was expressly constructed with this in mind—for him, a random sequence is one for which there is no admissible subsequence having a frequency differing from the frequency in the original sequence. This account is able to handle any value for the frequency, not only the case where each of two outcomes are equifrequent. Given that the Martin-Löf approach is a generalisation of von Mises', it is not surprising that it too can be adapted to permit biased sequences to be random. Consider a binary process with outcome probabilities $(p, 1 - p)$. The law of large numbers in a general form tells us that a measure one set of sequences of independent trials of such a process will have limit frequencies of outcomes equal to $(p, 1 - p)$. This measure is not the standard Lebesgue measure, but rather a measure defined by the chance function in question. We can similarly re-interpret the other effective statistical tests of randomness. Drawing as we did above (§2.1.2) on the language of statistical testing, we can characterise the random sequences as those which are not significant with respect to the hypothesis that the outcome chances are $(p, 1 - p)$—those which, as it were, conform to our prior expectations based on the underlying chances.

To approach the topic of randomness of biased sequences through Kolmogorov complexity, suppose we are given—somehow—any computable probability measure λ over the set of infinite binary sequences (that is, the probability of a given sequence can be approximated arbitrarily well by a recursive function). A sequence σ is λ-incompressible iff for each n, the Kolmogorov complexity of the length n initial subsequence of σ (denoted σ_n) is greater than or equal to $-\log_2(\lambda(\sigma_n))$. Where λ is the Lebesgue measure (supplement B.2), it follows that

$$-\log_2(\lambda(\sigma_n)) = -\log_2(1/2^n) = n,$$

so we get back the original definition of Kolmogorov complexity in that special case. With this generalised definition of Kolmogorov randomness in hand, it turns out that we can show a generalisation of Schnorr's theorem (§2.3): a sequence σ is λ-incompressible iff σ is ML-random with respect to λ.[15] In the framework of supplement B.1.1, a sequence is ML-random in this generalised sense iff the measure, under our arbitrary computable measure λ, of the nth sequential significance test is no greater than $1/2^n$. (There are some potential pitfalls, suggesting that perhaps the generalisation to an arbitrary computable measure is an overgeneralisation: see supplement B.1.3 for some details.)

While the above approach, with the modifications suggestion in the supplement taken on board, does permit biased random sequences, it comes at a cost. While the Lebesgue measure is a natural one that is definable on the Cantor space of sequences directly, the generalisation of ML-randomness requires an independent computable probability measure on the space of sequences to be provided. While this may be done in cases where we antecedently know the chances, it is of no use in circumstances where the existence of chance is to be inferred from the existence of a random sequence of outcomes, in line with RCT—for every sequence there is some chance measure according to which it is random, which threatens to trivialise the inference from randomness to chance. As Earman (1986: 145) also emphasises, this approach to randomness seems to require essentially that the chanciness of the process producing the random sequence is conceptually prior to the sequence being random. This kind of process approach to randomness has some intuitive support, and we will return to it below (§6.2), but it risks turning RCT into an uninformative triviality. By contrast, the Lebesgue measure has the advantage of being *intrinsically* definable from the symmetries of the Cantor space, a feature other computable measures lack.

The main difficulty with the suggested generalisation to biased sequences lies in the simple fact that biased

sequences, while they might reflect the probabilities in the process which produces them, simply don't seem to be random in the sense of disorderly and incompressible. The generalisation above shows that we can define a notion of disorderliness that is relative to the probabilities underlying the sequence, but that is not intrinsic to the sequence itself independently of whatever measure we are considering. As Earman puts it (in slightly misleading terminology):

> [T]here is a concept of randomness and a separable concept of disorder. The concept of disorder is an intrinsic notion; it takes the sequence at face value, caring nothing for genesis, and asks whether the sequence lacks pattern.… By contrast, the concept of randomness is concerned with genesis; it does not take the sequence at face value but asks whether the sequence mirrors the probabilities of the process of which it is a product. There is a connection between this concept of randomness and the concept of disorder, but it is not a tight one. (Earman 1986: 145)

As we might put it: Kolmogorov randomness is conceptually linked to disorderliness, and while we can gerry-rig a notion of 'biased disorder', that doesn't really answer to what we already know about the incompressibility of disorderly sequences. While we might well regard a sequence featuring even numbers of heads and tails, but produced by successive tosses of an *unfair* coin, to be biased in some sense with respect to the underlying measure of process behind it, it is still plausible that this *unrepresentativeness* of the sequence isn't conceptually connected with disorder in any interesting sense. It is very intuitive, as this remark from Dasgupta suggests, to take that biasedness—the increased orderliness of the sequence—to contrast with randomness:

> if for a sequence x this limiting frequency exists but is not equal to 1/2, then, in view of our underlying fair coin model, x would clearly be biased, not random. … Thus, it is natural to view this 'stochastic law of unbiasedness' as a 'stochastic law of randomness'. (Dasgupta 2011: §3.2)

As the bias in a chance process approaches extremal values, it is very natural to reject the idea that the observed outcomes are random. (As an example, we may consider human behaviour—while people aren't perfectly predictable, and apparently our behaviour doesn't obey non-probabilistic psychological laws, nevertheless to say that people act randomly is incorrect.) Moreover, there is a relatively measure-independent notion of disorder or incompressibility of sequences, such that biased sequences really are less disorderly. We can define a measure-dependent notion of disorder for biased sequences only by ignoring the availability of better compression techniques that really do compress biased sequences more than unbiased ones. To generalise the notion of randomness, as proposed above, permits highly non-random sequences to be called random as long as they reflect the chances of highly biased processes. So there is at least some intuitive pull towards the idea that if randomness does bifurcate as Earman suggests, the best deserver of the name is Kolmogorov randomness in its original sense. But this will be a sense that contrasts with the natural generalisation of ML-randomness to deal with arbitrary computable probability measures, and similarly contrasts with the original sense of randomness that von Mises must have been invoking in his earliest discussions of randomness in the foundations of probability.

In light of the above discussion, while there has been progress on defining randomness for biased sequences in a general and theoretically robust way, there remain difficulties in using that notion in defence of any non-trivial version of RCT, and difficulties in general with the idea that biased sequences can be genuinely disorderly. But the generalisation invoked here does give some succour to von Mises, for a robust notion of randomness for biased sequences is a key ingredient of his form of frequentism.

4.4 Dependence: Randomness is Indifferent to History

A further counterexample to RCT, related to the immediately previous one, is that randomness is indifferent

to history, while chance is not. Chance is history-dependent. The simplest way in which chance is history-dependent is when the conditions that may produce a certain event change over time:

> Suppose you enter a labyrinth at 11:00 a.m., planning to choose your turn whenever you come to a branch point by tossing a coin. When you enter at 11:00, you may have a 42% chance of reaching the center by noon. But in the first half hour you may stray into a region from which it is hard to reach the center, so that by 11:30 your chance of reaching the center by noon has fallen to 26%. But then you turn lucky; by 11:45 you are not far from the center and your chance of reaching it by noon is 78%. At 11:49 you reach the center; then and forevermore your chance of reaching it by noon is 100%. (Lewis, 1980: 91)

But there are more complicated types of history-dependence. In Lewis' example, the property which changes to alter the chances is how close the agent is to the centre. But there are cases where the property which changes is a previous outcome of the very same process. Indeed, any process in which successive outcomes of repeated trials are not *probabilistically independent* will have this feature.

One example of chance without randomness involves an unbiased urn where balls are drawn *without replacement*. Each draw (with the exception of the last) is an event which happens by chance, but the sequence of outcomes will not be random (because the first half of the sequence will carry significant information about the composition of the second half, which may aid compressibility). But a more compelling example is found in stochastic processes in which the chances of future outcomes depend on past outcomes. One well-known class of such process are Markov chains, which produce a discrete sequence of outcomes with the property that the value of an outcome is dependent on the value of the immediately prior outcome (but that immediately prior outcome screens off the rest of the history). A binary Markov chain might be the weather (Gates and Tong 1976): if the two possible outcomes are 'sunny' and 'rainy', it is plausible to suppose that whether tomorrow is rainy depends on whether today was rainy (a rainy day is more likely to be preceded by another rainy day); but knowing that today was rainy arguably makes yesterday's weather irrelevant.

If a Markov chain is the correct model of a process, then even when the individual trial outcomes happen by chance, we should expect the entire sequence of repeated trials to be non-random. In the weather case just discussed, we should expect a sunny day to be followed by a sunny day, and a rainy day by a rainy one. In our notation 11 and 00 should be more frequent than 10 or 01. But the condition of Borel normality, which all random sequences obey, entails that every finite sequence of outcomes of equal length should have equal frequency in the sequence. So no Borel normal sequence, and hence no random sequence, can model the sequence of outcomes of a Markov chain, even though each outcome happens by chance.

4.5 Pseudorandom Sequences

At least some non-random sequences satisfy many of the measure one properties required of random sequences. For example, the Champernowne sequence, consisting of all the binary numerals for every non-negative integer listed consecutively (i.e., 011011100101110111…), is Borel normal. This sequence isn't random, as initial subsequences of reasonable length are highly compressible. But it looks like it satisfies at least some desiderata for random sequences. This sequence is an attempt at producing a *pseudorandom* sequence—one that passes at least some statistical tests for randomness, yet can be easily produced. (The main impetus behind the development of pseudorandom number generators has been the need to efficiently produce numbers which are random for all practical purposes, for use in cryptography or statistical sampling.) Much better examples exist than the Champernowne sequence, which meet more stringent randomness properties.[16] One simple technique for generating pseudorandom sequences is a *symbol shift* algorithm (Smith 1998: 53). Given an initial 'seed' numeral $s_1, s_2, \ldots, s_n$, the algorithm simply spits out the

digits in order. Obviously this is useless if the seed is known, or can in some way be expected to be correlated with the events to which one is applying these pseudorandom numbers. But in practical applications, the seed is often chosen in a way that we do expect it to carry no information about the application (in simple computer pseudorandom number generators, the seed may be derived in some way from the time at which the seed is called for). With a finite seed, this sequence will obviously repeat after a some period. A symbol shift is the simplest possible function from seed to outcome sequence; better algorithms use a more complicated but still efficiently computable function of the seed to generate outcome sequences with a longer period, much longer than the length of the seed (e.g., the 'Mersenne twister' of Matsumoto and Nishimura 1998 has a period of $2^{19937} - 1$).

If the seed is not fixed, but is chosen by chance, we can have chance without randomness. For example, suppose the computer has a clock representing the external time; the time at which the algorithm is started may be used as a seed. But if it is a matter of chance when the algorithm is started, as it well may be in many cases, then the particular sequence that is produced by the efficient pseudorandom sequence generator algorithm will be have come about by chance, but not be random (since there is a program which runs the same algorithm on an explicitly given seed; since the seed is finite, there will be such a program; and since the algorithm is efficient, the length before the sequence produced repeats will be longer than the code of the program plus the length of the seed, making the produced sequence compressible). Whether the seed is produced by chance or explicitly represented in the algorithm, the sequence of outcomes will be the same—one more way in which it seems that the chanciness of a sequence can vary while whether or not it is random remains constant. (Symbol shift dynamics also permit counterexamples to the other direction of RCT—see §5.2.)

Much the same point could have been made, of course, with reference to any algorithm which may be fed an input chosen by chance, and so may produce an outcome by chance, but where the output is highly compressible. (One way in which pseudorandom sequence generators are nice in this respect is that they are designed to produce highly compressible sequences, though non-obviously highly compressible ones). The other interesting thing about those algorithms which produce pseudorandom sequences is that they provide another kind of counterexample to the epistemic connection between chance and randomness. For our justification in thinking that a given sequence is random will be based on its passing only finitely many tests; we can be justified in believing pseudorandom sequence to be random (in some respectable sense of justification, as long as justification is weaker than truth), and justified in making the inference to chance via RCT. But then we might think that this poses a problem for RCT to play the right role epistemically, even if it were true. Suppose one sees a genuinely random sequence and forms the justified belief that it is random. The existence of pseudorandom sequences entails that things might seem justificatorily exactly as they are and yet the sequence not be random. But such a scenario, arguably, defeats my knowing that the sequence is random, and thus defeats my knowing the sequence to have been produced by chance (and presumably undermines the goodness of the inference from randomness to chance).

5. Randomness Without Chance

The counterexamples to RCT offered in §§4.3–4.4 suggest strongly that the appeal of RCT depends on our curious tendency to take independent identically distributed trials, like the Bernoulli process of fair coin tossing, to be paradigms of chance processes. Yet when we widen our gaze to encompass a fuller range of chance processes, the appeal of the right-to-left direction of RCT is quite diminished. It is now time to examine potential counterexamples to the other direction of RCT. There are a number of plausible cases where a random sequence potentially exists without chance. Many of these cases involve interesting features of classical physics, which is apparently not chancy, and yet which gives rise to a range of apparently random phenomena. Unfortunately some engagement with the details of the physics is unavoidable in the following.

One obvious potential counterexample involves coin tossing. Some have maintained that coin tossing is a deterministic process, and as such entirely without chances, and yet which produces outcome sequences we have been taking as paradigm of random sequences. This will be set aside until §7, where the claim that determinism precludes chance will be examined.

5.1 Short Sequences

For many short sequences, even the most efficient prefix-free code will be no shorter than the original sequence (as prefix-free codes contain information about the length of the sequence as well as its content, if the sequence is very short the most efficient code may well be the sequence itself prefixed with its length, which will be longer than the sequence). So all short sequences will be Kolmogorov random. This might seem counterintuitive, but if randomness indicates a lack of pattern or repetition, then sequences which are too short to display pattern or repetition must be random. Of course it will not usually be useful to say that such sequences are random, mostly because in very short sequences we are unlikely to talk of the sequence at all, as opposed to talking directly about its constituent outcomes.

Because of the modal aspect of RCT, for most processes there will possibly be a long enough sequence of outcomes to overcome any 'accidental' randomness due to actual brevity of the outcome sequence. But for events which are unrepeatable or seldom repeatable, even the merely possible suitable reference classes will be small. And such unrepeatable events do exist—consider the Big Bang which began our universe, or your birth (*your* birth, not the birth of qualitatively indistinguishable counterpart), or the death of Ned Kelly. These events are all part of outcome sequences that are necessarily short, and hence these events are part of Kolmogorov random sequences. But it is implausible to say that all of these events happened by chance; no probabilistic theory need be invoked to predict or explain any of them. In the case of Kelly's death, for example, while it may have been a matter of chance when he happened to die, it was not chance that he died, as it is (physically) necessary that he did. So there are random sequences—those which are essentially short—in which each outcome did not happen by chance.

The natural response is to reject the idea that short sequences are apt to be random. The right hand side of RCT makes room for this, for we may simply insist that unrepeatable events cannot be repeated often enough to give rise to an adequate sequence (whether or not the inadequate sequence they do in fact give rise to is random). The problem here is that we can now have chance without randomness, if there is a single-case unrepeatable chance event. Difficulties in fact seem unavoidable. If we consider the outcomes alone, either all short sequences are random or none of them are; there is no way to differentiate on the basis of any product-based notion between different short sequences. But as some single unrepeatable events are chancy, and some are not, whichever way we opt to go with respect to randomness of the singleton sequences of such events we will discover counterexamples to one direction or another of RCT.

5.2 Chaotic Dynamics

The simple symbol shift dynamics in §4.5 had a finite seed, which allowed for chance without randomness. Yet there seem to be physical situations in which a symbol shift dynamics is an accurate way of representing the physical processes at work. One simple example might be a stretch and fold dynamics, of the kind common in chaos theory (Smith 1998: §4.2). The classic example is the *baker's transformation* (Earman 1986: 167–8; Ekeland 1988: 50–9). We take a system the state of which at any one time is characterised by a point (p, q) in the real unit square. We specify the evolution of this system over time as follows, letting ϕ be the function governing the discrete evolution of the system over time (i.e., $s_{t+1} = \phi(s_t)$):

$$\phi(p, q) = \begin{cases} (2p, q/2), & \text{if } 0 \le p \le \frac{1}{2} \\ (2p - 1, (1 + q)/2), & \text{if } \frac{1}{2} \le p \le 1 \end{cases}$$

This corresponds to transforming the unit square to a rectangle twice as wide and half the height, chopping off the right half, and stacking it back on top to fill the unit square again. (That this transformation reminded mathematicians of baking says something about their unfamiliarity with the kitchen—similar transformations, where the right half is 'folded' back on top, are more realistic.) If we represent the coordinates p and q in binary, the transformation is this:

$$\phi(0.p_1 p_2 \ldots, 0.q_1 q_2 \ldots) = (0.p_2 p_3 \ldots, 0.p_1 q_1 q_2 \ldots).$$

So this is a slight variant on a simple symbol shift, as the p-coordinate is a symbol shift to the right, while the q coordinate is, in effect, a symbol shift to the left.[17]

One important feature of this dynamics is that it is measure preserving, so that if X is a subset of the unit square, $\mu(X) = \mu(\phi(X))$. (This is easily seen, as the symbol shift dynamics on the basic sets of infinite binary sequences is measure-preserving, and each coordinate can be represented as an infinite binary sequence.) Define the set $L = \{(p, q) : 0 \le p < \frac{1}{2}\}$. We see that $(p, q) \in L$ iff $p_1 = 0$. Since p can be represented by an infinite binary sequence, and a measure one set of finite binary sequences is Borel normal, we see that almost all states of this system are such that, over time, $\mu(\phi(s) \in L \mid s \in L) = \mu(\phi(s) \in L)$ —that is, whether the system is in L at t is probabilistically independent of its past history. Furthermore, $\mu(L) = \mu(\overline{L})$. The behaviour of this system over time, with respect to the partition $\{L, \overline{L}\}$, is therefore a Bernoulli process, exactly like a sequence of fair coin tosses—a series of independent and identically distributed repetitions of a chance process. If the RCT is true, then a system which behaves exactly like a chance process should have as its product a random sequence. So the sequence of outcomes for the baker's transformation (under this partition) is a random sequence.

But unlike a sequence of fair coin tosses, supposing them to involve genuine chance, the baker's transformation is entirely *deterministic*. Given a particular point (p, q) as an initial condition, the future evolution of states of the system at each moment is time t is determined to be $\phi^t(p, q)$. So while the product produced is random, as random as a genuine chance process, these outcomes do not happen by chance; given the prior state of the system, the future evolution is not at all a matter of chance. So we have a random sequence without chance outputs. (Indeed, given the symbol shift dynamics, the evolution of the system over time in $\{L, \overline{L}\}$ simply recapitulates successive digits of the starting point.) To be perfectly precise, the trial in this case is sampling the system at a given time point, and seeing which cell of the coarse grained partition it is in at each time. This is a sequence of arbitrarily repeated trials which produces a random sequence; yet none of these outcomes happens by chance.[18] To put the point slightly differently: while the sequence of outcomes is random, there is a perfectly adequate theory of the system in question in which probability plays no role. And if probability plays no role, it is very difficult to see how chance could play a role, since there is no probability function which serves as norm for credences, governs possibility, or is non-trivial and shared between intrinsic duplicate trials. In short, no probability function that has the features required of chance plays a role in the dynamics of this system, and that seems strong reason for thinking there is no chance in this system.

The baker's transformation provides a simple model of deterministic *macro-randomness*—a system which has a μ-measure-preserving temporal evolution, and produces a sequence of coarse grained states that have the Bernoulli property. It is a question of considerable interest whether there are physically more realistic systems which exhibit the same features. We may conceive of an n-particle classical (Newtonian) system as

having, at each moment, a state that is characterised by a single point in a $6n$-dimensional state space (each point a $6n$-tuple characterising the position and momentum of each particle). The evolution of the system over time is characterised by its *Hamiltonian*, a representation of the energetic and other properties of the system. The evolution under the Hamiltonian is μ-measure-preserving (by Liouville's theorem), so it might be hoped that at least some systems could be shown to be Bernoulli too. Yet for closed systems, in which energy is conserved over time, this is not generally possible. Indeed, for closed systems it is not generally possible to satisfy even a very weak randomness property, *ergodicity*. A system is ergodic just in case, in the limit, with probability one, the amount of time a system spends in a given state is equal to the (standard) measure of state space that corresponds to that state (Earman 1986: 159–61; Sklar 1993: 60–3; Albert 2000: 59–60). While a Bernoulli system is ergodic, the converse entailment does not hold; if the system moves only slowly from state to state, it may be ergodic while the state at one time is strongly dependent on past history (Sklar 1993: 166–7). While ergodicity has been shown to hold of at least one physically interesting system (Yakov Sinai has shown that the motion of hard spheres in a box is ergodic, a result of great significance for the statistical mechanics of ideal gases), a great many physically interesting systems cannot be ergodic. This is the upshot of the so-called KAM theorem, which says that for almost all closed systems in which there are interactions between the constituent particles, there will be stable subregions of the state space—regions of positive measure such that if a system is started in such a region, it will always stay in such a region (Sklar 1993: 169–94). Such systems obviously cannot be ergodic.

The upshot of this for our discussion may be stated: 'there are no physically realistic classical systems which exhibit even ergodicity, and so no physically realistic classical system can exhibit randomness. The baker's transformation is a mathematical curiosity, but is not a genuine case of randomness without chance, since systems like it are not physically possible.' This response is premature. There are physically interesting systems to which the KAM theorem does not apply. Open or dissipative systems, those which are not confined to a state space region of constant energy, are one much studied class, because such systems are paradigms of chaotic systems. The hallmarks of a chaotic dissipative system are two (Smith 1998: §1.5):

1. There exists at least one set of states in state space A (an *attractor*) such that when the system is started initially in its neighbourhood $N(A)$, the trajectory of the system will in the limit end up in A ($\lim_{t \to \infty} \phi^t N(A) \subseteq A$); and
2. The system displays sensitive dependence on initial conditions: that is, in some set of state space points that are all within some arbitrary distance δ of each other, there are at least two points whose subsequent trajectories diverge by at least ε after some time t.[19]

There are physically realistic classical systems which exhibit both of these features, of which the best known is perhaps Lorenz's model of atmospheric convection (Smith 1998: §1.4; Earman 1986: 165). The combination of these two features permits very interesting behaviour—while the existence of an attractor means that over time the states of the system will converge to the region of the attractor, the sensitive dependence on initial conditions means that close states, at any time, will end up arbitrarily far apart. For this to happen the attractor must have a very complex shape (it will be a region of small measure, but most of the state space will be in the neighbourhood of the attractor). More importantly for our purposes, a system with these characteristics, supposing that the divergence under evolution of close states happens quickly enough, will yield behaviour close to Bernoulli—it will yield rapid *mixing* (Luzzatto *et al*. 2005).[20] Roughly, a system is mixing iff the presence of the system in coarse state at one time is probabilistically independent of its presence in another coarse state at another time, *provided there is a sufficient amount of time between the two times*. This is weaker than Bernoulli (since the states of a Bernoulli system are probabilistically independent if there is any time between them), but still strong enough to plausibly yield a random sequence of outcomes from a coarse grained partition of the state space, sampled infrequently enough. So we do seem to have physically realistic systems that yield random behaviour without chance. (See also the systems discussed in Frigg 2004.)

Indeed, the behaviour of a chaotic system will be intuitively random in other ways too. The sensitive dependence on initial conditions means that, no matter how accurate our finite discrimination of the initial state of a given chaotic system is, there will exist states indiscriminable from the initial state (and so consistent with our knowledge of the initial state), but which would diverge arbitrarily far from the actual evolution of the system. No matter, then, how well we know the initial condition (as long as we do not have infinite powers of discrimination), there is another state the system could be in for all we know that will evolve to a discriminably different future condition. Since this divergence happens relatively quickly, the system is unable to be predicted. (Anecdotally, at least, Lorenz' model of the weather seems borne out by our inability to reliably predict future weather even a few days from now.) Insofar as randomness and lack of reliable prediction go hand in hand, we have another reason for thinking there is product randomness here (§6.2).

Just as before, the classical physical theory underlying the dynamics of these chaotic systems is one in which probability does not feature. So we are able to give an adequate characterisation of the physical situation without appeal to any probabilities fit to play the chance role. Given well-enough behaved boundary conditions, this system is also deterministic (though see §5.3), and that may also be thought to preclude a role for non-trivial chances. So again we have randomness in the performance, though none of the outcomes happened by chance.

Two avenues of resistance to the line of thought in this section suggest themselves. The first is to maintain that, despite the truth of determinism for the baker's transformation and classical physics (modulo worries to be addressed in the following section), there are still, or at there least may be, non-trivial chances in these theories. The proposal is that the possibility remains of *deterministic chance*, so that from the fact of determinism alone it does not follow that we have a counterexample to RCT. The outcomes may be determined to happen by the prior conditions, but (so the suggestion goes) they may still happen by chance. This radical proposal is discussed below, in §7. It should be noted, however, that even if it is true that deterministic chance is possible, that observation is far from establishing that the physical theories under discussion here are ones in which deterministic chance features. That some deterministic theories may have chances is no argument that all will, and particularly in very simple cases like the baker's transformation, there doesn't seem much point in an invocation of deterministic chance: chances will be trivial or redundant if classical physics is true. The second avenue of resistance is to claim that there really is chance here—it is chance of the *initial conditions* which is recapitulated in the random sequence of outcomes. While the Lebesgue measure over the set of initial conditions in our models is formally like a probability function, to assume it yields genuine chances is a fairly controversial thesis (for a contrary opinion, see Clark 1987). Other initial conditions could have obtained; still it seems wrong to think that (somehow) there was a chance process which terminated in the selection of the initial conditions which happened in fact to obtain at our world. Rather, that the initial conditions are what they are seems to be a brute fact. If there are to be chances, then, they cannot be dynamical chances, the kind that is familiar from physics, and from our discussion in §1.2. Some recent arguments in favour of the possibility of chancy initial conditions are discussed in the following supplementary document:

> Supplement D. Chance and Initial Conditions

But whether or not the idea of chancy initial conditions can be made to work, the fact remains that at most one outcome in the random sequence—the first one—happens by chance. The subsequent states do not, yet RCT is committed to there being (dynamical, transition) chances in those state transitions.

5.3 Classical Indeterminism

Despite the neat picture of classical determinism drawn in the previous section, it is well known that classical

physics is not in fact deterministic. These cases of indeterminism do not undermine the applications of classical mechanics in the previous section. But classical indeterminism may provide problems of its own for RCT. Useful further material on the topic of this section can be found in the entry on causal determinism (Hoefer 2010: §4.1).

For present purposes, indeterminism occurs when the state of the system at one time does not uniquely fix the state the system will be in at some future time (see §7). To show indeterminism in the classical case, it suffices to give a state of some system at a given time and to specify two future states that are incompatible with each other and yet both states are consistent with Newton's laws of motion and the initial state.

To help us in this task, it is useful to note one fact about Newtonian mechanics: the laws are *time-reversal invariant* (Albert 2000: ch. 1). That is, for every lawful trajectory through the state space, there is another lawful trajectory which results from the first by mapping every instantaneous state in the first trajectory to its image state in which the particles have the same positions but the signs on the components of momentum are reversed, and running the trajectory in reverse order. These image states are those where the particles are in the same positions but moving in exactly the opposite direction. So for every lawful process, that process running backwards is also lawful. (If these trajectories are so lawful, why don't we see them?—that is the deep question of thermodynamic asymmetry, discussed briefly in supplement D.)

Two examples serve to illustrate the possibility of classical indeterminism. A very elegant recent example is provided by 'Norton's dome' (Norton 2003; 2008). A point mass is at rest at the apex of a dome (of certain shape) at t^*. One obvious solution to Newton's equations of motion is that it continues to do so for all moments $t > t^*$. But, Norton points out, there is another solution: that while at $t = t^*$ the mass is at rest, at every moment $t > t^*$, the mass is moving in some direction. But this means the mass spontaneously moves off in some arbitrary direction at an arbitrary time.[21] Determinism is clearly violated: for some given time t^*, there is a state at t' where the particle remains at the apex of the dome; and there are many incompatible states where at t' the particle is somewhere else on the surface of the dome. Nothing in the conditions at t^* fixes which of these many future states will be the one to eventuate. An easy way to understand the dome example is to consider its time reversal: a ball is given some initial velocity along the surface of the dome toward the apex. Too little, and it falls short; too great, and it overshoots. Just right, however, and the ball comes to rest precisely at the apex, and remains at rest. The time reversal of this system is the original dome example.

A more exotic example involves 'space invaders' (Earman 1986: ch. III). These are particles that are at no spatial location at time t, and therefore form no part of the state at time t, but which travel in 'from' spatial infinity and by time t' have a location. We can see the example more clearly if we invoke time-reversal invariance. Consider two point particles, a and b, at rest and in the vicinity of one another at t^*. From t^* onwards, force is applied to a in such a way that the velocity of a away from b increases without bound. This is possible, because there is no upper bound on velocity in classical physics. Indeed, let the velocity of a increase fast enough that by some finite time t', a has no finite velocity, and is thus 'at' spatial infinity. The time-reversal of this system has the particle a having no location at t', but having a location and continuously decreasing velocity at every moment $t > t'$, until it comes to rest at t^*. This system violates determinism in the sense given above. The state at t' consists of a single particle b at rest. That state could be followed at t^* by the space-invaded state just described; or it could be followed at t^* by the incompatible state of b simply being at rest some more. Nothing in the laws rules out either of these transitions. Of course, the model isn't particularly physically realistic—where does the force acting on a come from? But physically more realistic systems exhibiting the same general structure have been given; Earman mentions one due to Mather and McGehee (1975) involving four point particles moving in such a way that the forces they exert on each other in collisions end up unboundedly far from one another in a finite time (see also Saari and Xia 1995).

While classical mechanics is thus indeterministic, it is importantly not chancy. There is no reason to think that we either need to or can assign a probability distribution over the possible future states in our indeterministic systems. Norton says this about his dome:

> One might think that … we can assign probabilities to the various possible outcomes. Nothing in the Newtonian physics requires us to assign the probabilities, but we might choose to try to add them for our own conceptual comfort. It can be done as far as the direction of the spontaneous motion is concerned. The symmetry of the surface about the apex makes it quite natural for us to add a probability distribution that assigns equal probability to all directions. The complication is that there is no comparable way for us to assign probabilities for the time of the spontaneous excitation that respect the physical symmetries of solutions. Those solutions treat all candidate excitation times T equally. A probability distribution that tries to make each candidate time equally likely cannot be proper—that is, it cannot assign unit probability to the union of all disjoint outcomes.[22] Or one that is proper can only be defined by inventing extra physical properties, not given by the physical description of the dome and mass, Newton's laws and the laws of gravitation, and grafting them unnaturally onto the physical system. (Norton 2003: 9–10)

The point Norton makes about the impossibility of a uniform distribution over a countable set of time intervals holds also for the time at which we might expect the space invader to occur in the second type of case. Thus, it seems, we have indeterminism without chance.

We can use these constructions to come up with counterexamples to RCT. Let a dome system be prepared, and left for a period of 5 seconds. If the ball remains at rest on the apex, call the outcome '0'. If the ball moves, and so is at the end of 5 seconds somewhere on the dome other than the apex, call the outcome '1'. Both outcomes are physically possible final states of the system after 5 seconds. If the system is repeatedly prepared in this state, it is physically possible to get a random sequence of outcomes of these repeated trials. Certainly the outcomes cannot be produced by some finite algorithm, since the indeterministic dynamics permits as physically possible every sequence of outcomes, including those that differ at some place from every algorithmically compressible sequence. In the infinite future case, it is physically possible for the system to produce every infinite binary sequence, but at most countably many of these are non-random. So it is physically possible for these setups to produce a random sequence of outcomes in the KML sense. But we then have randomness without a chance distribution over the outcomes. Randomness only requires two distinguishable possible outcomes and the possibility of production of arbitrary sequences of such outcomes. Chance requires two distinguishable outcomes, each of which has some chance. These cases show that chance and possibility come apart—there are cases where there are two possible outcomes of a process, neither of which has any chance at all (not even a chance of zero).

5.4 Drawing Without Replacement

The last objection draws on a remark made in §4.4. Consider a large finite urn full of black and white balls. The number of balls is sufficiently large that the sequence of outcomes of draws is long enough to be random. So suppose we make a random selection from this urn, drawing balls without replacement until the urn is empty. The resulting sequence of outcomes is, or at least could be, random—it had better be, since this sequence meets all the conditions to be a simple random sample from the population. (We could attach a number to each member of an experimental population, and give the n-th member a pill of an active substance (respectively, a placebo) if the n-th draw is black (white).) But in this process, the outcomes become less and less chancy as the number of balls diminishes. Since there are only finitely many balls, there will a draw such that the chance of it being black is either 1 or 0, and so whatever outcome eventuates did not occur by chance. But then we have a random sequence that includes an outcome (drawing a black ball, say) which did not happen by chance, contrary to RCT.

One response is to say that this last outcome did happen by chance, since at the start of the drawing there was a positive chance that a white ball would be drawn last, and a positive chance that a black ball would be. This response neglects the time-dependence of chances. If there were n balls initially, and we let 'Lastie' name the ball that was in actual fact drawn last, then we can say: the chance that Lastie is drawn last was $1/n$ initially, and after the m-th draw it was $1/(n-m)$, until it reached 1 and remained there. At that last stage it was no longer a matter of chance whether Lastie would be drawn last; it is the only ball left in the urn. RCT maintains that a given outcome happens by chance iff it is part of a random sequence. At every time, the event of Lastie being drawn last is part of a random sequence. But then there is at least one time at which Lastie being drawn last is part of random sequence but, at that time, it did not happen by chance. (Alternatively, of course, it could be maintained that even events with trivial chances happen by chance. But this would permit again the problem of biased sequences; a string of all heads tossed by a two-headed coin could then be random, which it is not.)

6. Saving the Thesis: Alternative Conceptions of Chance and Randomness

The discussion in §§4–5 leaves RCT in a doubtful position. But it may be that the problems for RCT are due more to some defect in the theories of chance and randomness sketched in §§1–2. As noted earlier, there are alternative conceptions of chance and randomness that have some appeal and might perhaps save RCT. They won't have much to say about the modal problems for merely possible random sequences mentioned when RCT was first introduced. But perhaps the other objections can be avoided. The problems for RCT arise fundamentally because of the split between product randomness and process chance. Closing this gap promises to aid RCT. This following two subsections will consider product conceptions of chance and process based conceptions of randomness.

6.1 Product Chance

The frequency theory is a product conception of chance. An outcome-type has a chance, according to von Mises (1957), just in case it is part of a random sequence of outcomes in which that outcome type has a robust limit relative frequency. So chance can't be separated from randomness; it in fact requires randomness. Moreover, since having a limit relative frequency of $\frac{1}{2}$ is a measure one property of infinite binary sequences, all random sequences define collectives. (Those infinite binary sequences which do not converge on a limit frequency are non-random.) Yet the problems with frequentism as a theory of chance are well known (Hájek, 1997; 2009; Jeffrey, 1977)—we have come across some of them above—and to save RCT at the price of accepting frequentism has not attracted many.

But the prospects are more promising for Humean views of chance (Lewis 1994; Loewer 2004), like the reductionist views discussed in supplements A.1, A.3, and D. These are a species of product conception of chance, for a possible world will not feature chances unless the best (simplest, best fitting, and most informative) description of what events occur in that world involves a probability function. Two worlds couldn't differ in their chances without differing also somewhere in the events which occur. Chance thus supervenes on the actual outcomes in a given world, but not necessarily in a direct way—in the service of simplicity, some deviation from the actual frequency values may yield a description that is better overall. There is considerable debate over whether a Humean best systems account can explain all of the things we know about chance. Lewis thought that chance was the 'big bad bug' for his broadly Humean world view (though he thought that the NP, discussed in supplement A.1, debugged Humean Supervenience: Lewis 1994), and there has been considerable debate over whether or not the PP or the BCP can be accounted for by the Humean, as the references in the supplement A attest. Moreover, there are problems about whether the

notion of a probability distribution 'fitting' a set of outcomes makes sense (Elga 2004). But suppose a best systems account can be made to work.

The role of simplicity in Humean views is important for randomness. For if a world contains a genuinely random sequence of outcomes, there will be no short description of that sequence. Those descriptions which do not attempt to describe what happens in all its particularity, but instead involve a probability distribution that makes the sequence a typical one with respect to that probability function, will be less informative but much shorter, and still fit just as well. So it seems that if a world contains a random sequence of outcomes, the best theory will be one which involves probabilities, and in virtue of being part of the best theory, those probabilities will be chances. That makes the right to left direction of RCT plausible. The other direction would hold if this route through simplicity was the only way in which chances could enter into the best system. Could there be a world with chances in which there was no random sequence? Here things are rather murkier for the Humean. For the hallmark of the best systems approach to chance is that, by involving simplicity, it avoids some of the problems for pure frequentism. In particular, an unrepresentative outcome sequence (a short one, or highly biased one) need not force the chances to take the value of the actual frequencies. Suppose a world contains two intrinsic duplicate coins, one of which is tossed many times, landing heads about half the time; the other is tossed few times and lands tails every time. The second coin's outcome sequence has a heads frequency of zero. It is a strength of the best systems account that we may still say the chance of heads was $\frac{1}{2}$, because the coin is embedded in a world where another, similar, coin did have the appropriate outcome frequencies, and it is overall a simpler theory to say that both of these coins obeyed the same probabilistic laws than that the second one was an oasis of determinism. But this case provides a problem for RCT—it looks like the second coin toss is not part of any random sequence of outcomes (since a few all tails tosses is not random), but it has a chance.

We should not let even the partial success of best systems analysis in preserving RCT sway us. The Humean account of chance is perfectly compatible with the existence of bias and with non-independent sequences of trials; RCT is not. The problem just mentioned arises even in the best circumstances for RCT, where there is at least one actual unbiased fair coin sequence. The existence of such a problem could have been predicted. The best systems account deviates from pure frequentism precisely in trying to accommodate the single-case process intuitions that, as we saw in §1, characterise chance. Every success in this direction brings this broadly product conception of chance closer to a process conception, and will therefore be a potential opportunity for counterexamples to RCT to emerge.

6.2 Process Randomness: Epistemic Theories

As mentioned at the beginning of §2, sometimes 'random' is used in a process sense. There have been some philosophical approaches to randomness which attempt to take this seriously, but which do not take it to be merely equivalent to 'chancy' and thus trivialise RCT. The most popular such approach is to connect randomness with *indeterminism*, and to defend RCT by arguing directly that indeterminism yields both chance and randomness. Prospects for that approach will be discussed in §7.

The next most discussed view of process randomness is an epistemic one. Random processes on this view are those whose outcomes we cannot know in advance; that is, random processes are *unpredictable* (Eagle 2005; Kyburg 1974: ch. 9).[23] Here is one clear expression of the view:

> At the most basic level, we say an event is random if there is no way to predict its occurrence with certainty. Likewise, a random process is one for which we are not able to predict what happens next. (Frigg 2004: 430)

For this kind of view to yield the right results, we cannot count as 'being able to predict a process' if we

merely guess rightly about its outcomes. For that reason, prediction must involve some notion of *reasonableness*; it must be rational for the agent to make the predictions they do. For example, Eagle (2005: 770) insists that it is the physical theory accepted by the predicting agent that makes certain posterior credences reasonable; simply guessing will not be reasonable, even if it's correct.

This definition overlaps considerably with those definitions of randomness canvassed in §2. In particular, if a process is predictable, that will make available a winning betting strategy on the sequence of outcomes of that process, which cannot therefore be a KML-random sequence of outcomes. So unpredictability of the generating process is a necessary condition on KML-randomness of any concrete outcome sequence.

But unpredictability is not sufficient, for it may be that we cannot know everything true of the future outcomes of the process, and those truths might preclude a KML-random sequence. One way to see this draws on our discussion of chaotic dynamics. Let's say that a system exhibits *apparent* dependence on initial conditions if states indiscriminable to us may end up arbitrarily far apart after some relatively short period of time. (Another definition, if knowledge obeys a margin for error principle, would be: a system exhibits apparent dependence on initial conditions if, for all we know, it is sensitively dependent on initial conditions.) Apparent dependence on initial conditions can obtain even if sensitive dependence on initial conditions does not. For there may well be a size v such that, under some partition of the set of states into regions of measure smaller than v, any two points starting in the same cell of the partition evolve to future states which are close to one another—as long as v is small with respect to our abilities to discriminate. A system which is apparently but not sensitively dependent on initial conditions will be unpredictable to us, but there will exist an algorithm that, given some finite specification of the initial conditions, generates the future evolution of the system perfectly. (The key is that the finite specification must involve more precise data than we could know to characterise the system.) This sequence, if long enough, is not KML-random, but it is unpredictable.[24]

Because of the considerable overlap between the unpredictably generated sequences and the KML-random sequences (Frigg 2004: 431), many roles the latter can play will be played by the former too. Eagle (2005) further argues that the unpredictably generated sequences are a better fit to the theoretical role of randomness, and claims on that basis that *randomness is unpredictability*. One nice thing about this thesis is that, being a process notion, it directly connects with chance processes. We can thus directly evaluate the original thesis connecting chance and randomness, CT, in the form:

> **(CTU)** A process is chancy iff it is not rationally predictable.

The left-to-right direction of CTU looks relatively secure when we attend just to independent and identically distributed trials. But when the trials are not independent, like the examples in §4.4, future outcomes can happen by chance, even if knowing the past states of the system does put one in a position to better predict the future outcomes. The right-to-left direction of CTU is in even worse straits. For if KML-randomness of a sequence entails its unpredictable generation, then every counterexample to RCT which involves KML-randomness without chance will also involve unpredictability without chance, and also constitute a counterexample to CTU. There is no succour to be found for defenders of RCT in this conception of randomness.

7. Chance, Randomness, and Determinism

One last hope for the thesis that chancy outcomes are random comes from the connection of both notions with indeterminism. Consider this argument:

> **P1**:

An outcome happens by chance iff it is produced by an indeterministic process.

P2:
A possible sequence of outcomes is random iff a repeated indeterministic process could produce all of the outcomes.

RCT:
Therefore, an outcome happens by chance iff there is a possible random sequence of outcomes, produced by the same process, of which it is a member.

This argument is valid. If the premises are true, we have a direct argument for RCT. We have no direct response to the objections raised in §§4–5, but somehow, if this argument succeeds, those objections must miss the point. The premises have some initial plausibility (though P2 is dubious: surely a properly indeterministic process could produce any sequence of outcomes whatsoever, including many non-random sequences? We discuss this further in §7.2). The thesis that indeterminism is necessary and sufficient for chance has long been a popular claim. And randomness and indeterminism also seem to have a close connection. But to evaluate them, we will need to be more precise than we have been about determinism.

Earman-Montague Determinism:
A scientific theory is *deterministic* iff any two sequences of states in models of that theory which share some state at some time share every state at every time. A theory is *indeterministic* iff it is not deterministic; equivalently, if two systems can be in the same state at one time and evolve into distinct states. A system is (in)deterministic iff the theory which completely and correctly describes it (of which it is a model) is (in)deterministic. (Earman 1986; Montague 1974)

Determinism so stated is a supervenience thesis: as Schaffer (2007: 115) puts it, 'A world w is deterministic iff: for all times t in w, the total occurrent history of w supervenes on the occurrent state of w at t together with the laws of w.'

With this in mind, we now evaluate the premises of this argument. There seem to be significant doubts about both directions of both premises. The upshot is that indeterminism offers little comfort for the defender of RCT.

7.1 Chance and Determinism

It is very nearly a piece of philosophical orthodoxy that non-trivial objective chances require indeterminism. The view is seldom defended; even those who trouble to state the view explicitly (Lewis, 1980: 120) don't go to the further effort of justifying it, perhaps because it seems obvious. After all, under determinism, someone who knew the past and the laws would be in a position to know with certainty every future outcome. So our use of probability under determinism must be purely subjective, a side-effect of our (perhaps unavoidable) ignorance of the past or the laws. If this orthodoxy is correct, at least the left-to-right direction of P1 would be true.

However, there has recently been a considerable amount of work by philosophers defending the thesis that chance and determinism are consistent. Many of the topics dealt with above feature in their arguments. Loewer (2001) draws on the best systems analysis of chance to argue that, in worlds like ours (with entropy-increasing processes, and apparently fair coin tosses, etc.), the best description involves some probabilistic component which deserves to be called chance. Clark (1987) draws on how we use the phase space measure μ to govern our expectations about the behaviour of Bernoulli and mixing systems in classical statistical mechanics, and argues (in effect) that this is an objective probability function despite the deterministic underlying physics. Various other proposals for deterministic chance have been developed (Eagle 2011;

Glynn, 2010; Hoefer, 2007; Ismael, 2009; Sober, 2010). The general technique is to argue that there are probability distributions over outcomes that can play the chance role, even in impeccably deterministic theories. Many of these philosophers are sympathetic to reductionism about chance, which permits the compatibility of chance and determinism. For the fact that the entire history of a world supervenes on any moment of its history, as determinism states, apparently entails nothing one way or another about whether the best description of that entire history involves chances (this is related to the 'no-connection' argument of Schaffer 2007: 115). If there is such a thing as deterministic chance, however, P1 is false.

Anti-reductionists about chance have generally found these arguments less persuasive (Popper 1992; Black 1998; Weiner and Belnap 2006). In particular, it is in many ways hard to reconcile the BCP (and RP) with deterministic chance. Doing so will require that there is a physically possible world (sharing the same laws) which matches ours in occurrent facts up until t but diverges thereafter; but if determinism is true, such a divergence is not possible. If that world matches ours at any time, it matches it at all times. So, it seems, 'only an incompatibilist function can fit the RP' (Schaffer 2007: 130). In any case, the debate over whether deterministic 'chance', and reductionism about chance more generally is ongoing (see further the discussion in supplement A); the status of the left-to-right direction of P1 is at best unsettled.

The same cannot be said for the right-to-left direction. For the discussion in §5.3 showed that there are indeterministic theories without chances, those where indeterminism is secured by the existence of alternative future possibilities, but where those possibilities collectively do not permit or require a probability distribution over them. A more controversial class of cases of indeterminism without chances comes from those who reject *universalism* about chance: the thesis 'that chance of truth applies to any proposition whatever' (Lewis, 1980: 91). If universalism is false, there may be indeterministic situations where the alternative outcomes are those to which chance does not apply. Von Mises rejected universalism, because he thought that chance applied properly only to 'mass phenomena'; in an indeterministic world where the indeterministic process occurs only once, for example, the theory of probability does not apply. Hoefer (2007) also holds a view something like this, rejecting chances for those processes which don't have stable enough outcome patterns.[25]

7.2 Randomness and Determinism

We could give an argument for P2 if it could be shown that, from the above definition of determinism, we could conclude (i) that only random sequences could occur under indeterminism, and (ii) that random sequences could only occur under indeterminism. However, both parts of this claim are problematic.

> **Theorem 8 (Humphreys 1978).** There is a theory which is deterministic in the sense of Montague which has as a model a system which produces sequences which are random in the sense of von Mises/Church.

The proof of this theorem relies on the fact that a theory can be non-trivially deterministic without being computable. There is an arithmetically definable function which governs the evolution of the system over time (in Humphrey's construction, the instantaneous state of the system is an arithmetically definable function of the time, which ensures that any two models agreeing at one time will agree at all times, guaranteeing determinism). But the function is not effectively computable, so no algorithm can produce the sequence of states that this system goes through.[26] The physical significance of such uncomputable functions is not clear (though see Pour-El and Richards 1983), but the possibility of a deterministic physics which features such equations of motion is enough to undermine a close connection between randomness and indeterminism. This shows that claim (ii) above is false.

Moreover, since we've already seen (§4.1) that it is possible for a chancy and indeterministic process to

produce a non-random sequence of outcomes, and such a sequence would not be random, we also have a counterexample to claim (i). Claim (i) could be saved if we made a more robust claim about what could happen under indeterminism. There is a sense in which, while it is possible that a fair coin could land heads an infinite number of times, it *would not*. That is, the counterfactual 'If I tossed the coin an infinite number of times, it wouldn't land all heads' is apparently true. There is some question whether the counterfactual really is true; Lewis (1979a) and Williams (2008) argue that it is, while Hawthorne (2005) argues that it is not. But if it is, the way lies open to defend a modified counterfactual version of P2:

> **P2′**:
> A possible sequence of outcomes is random iff it is not the case that, were an indeterministic process to be repeated infinitely, it would not produce that sequence of outcomes.

But this is highly controversial; and the problem for claim (ii) would still stand.

If we are to accept this argument, then, we shall have to take P2 as an independent truth about randomness. Analyses of randomness as indeterminism, which take P2 to be analytic, have been given: to their detriment, if the foregoing observations are correct. Hellman (1978: 83) suggests that randomness is 'roughly interchangeable with "indeterministic"', while Ekeland (1988: 49) says 'the main feature of randomness is some degree of independence from the initial conditions. … Better still, if one performs the same experiment twice with the same initial conditions, one may get two different outcomes'.

However, it has been argued that this view of randomness as indeterminism makes it difficult to understand many of the uses of randomness in science (Eagle 2005: §3.2). This view entails that random sampling, and random outcomes in chaotic dynamics, and random mating in population genetics, etc., are not in fact random, despite the plausibility of their being so. It does not apparently require fundamental indeterminism to have a randomised trial, and our confidence in the deliverances of such trials does not depend on our confidence that the trial design involved radioactive decay or some other fundamentally indeterministic process. Indeed, if Bohmians or Everettians are right (an open epistemic possibility), and quantum mechanics is deterministic, this view entails that nothing is actually random, not even the most intuitively compelling cases. This kind of view attributes to scientists a kind of error theory about many of their uses of the term 'random', but as yet the philosophical evidence adduced to convict scientists of this pervasive error is not compelling.

One reason for the continuing attractiveness of the thesis that randomness is indeterminism may be the fact that, until quite recently, there has been a tendency for philosophers and other to confuse unpredictability and indeterminism. Laplace's original conception of determinism was an epistemic one:

> [A]n intelligence which could comprehend all the forces by which nature is animated and the respective situation of all the [things which] compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies in the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as well as the past, would be present to its eyes. (Laplace 1826: p. 4)

This kind of account still resonates with us, despite the fact that with the Montague-Earman definition we now have a non-epistemic characterisation of determinism. Since random sequences will, almost always, be unpredictable, it is clear why we might then taken them to be indeterministic. But once we keep clear the distinction between predictability and determinism, we should be able to avoid this confusion (Bishop, 2003; Schurz, 1995; Werndl, 2009).

8. Conclusion

From what we have seen, the commonplace thesis cannot be sustained. It would in many ways have been nice if chance of a process and randomness of its product had gone hand in hand—the epistemology of chance would be much aided if it invariably showed itself in random outputs, and we could have had a tight constraint on what outcomes to expect of a repeated chance process, to say nothing of the further interesting consequences the thesis may have for random sampling or probabilistic explanation, mentioned in the introduction. But the counterexamples to the thesis in §§4–5 show that it is false, even in its most plausible form. Various attempts to salvage the thesis, by appeal to non-standard accounts of chance or randomness, fail to give us a version of the thesis of much interest or bearing on the issues we had hoped it would illuminate. A final attempt to argue directly for the thesis from the connections between chance, randomness, and determinism also failed, though it does shed light on all three notions. It is safest, therefore, to conclude that chance and randomness, while they overlap in many cases, are separate concepts.

This is not to say that there is no link between KML-randomness and physical chance. The observation of a random sequence of outcomes is a defeasible incentive to inquire into the physical basis of the outcome sequence, and it provides at least a prima facie reason to think that a process is chancy (though recall §4.5). Moreover, if we knew that a process is chancy, we should expect (eventually, with high and increasing probability) a random sequence of outcomes. Conversely, a sequence of outcomes that appears predictable, compressible, and rule-governed will be strong disconfirming evidence for any hypothesis to the effect that the sequence was produced by chance alone. Hellman concludes

> 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 randomness' to be product randomness, and 'physical randomness' to mean process randomness (chanciness), this conclusion seems unavoidable.

A parallel with the relationship between frequencies and chances is tempting and unsurprising. Relative frequencies are good but not infallible indicators of the chances, and the existence of outcome frequencies strictly between 0 and 1 is evidence that chance processes are involved in the production of those outcomes. But frequentism is implausible as a reductive account of chance. And it is no more plausible to think that chance is present iff random sequences of outcomes are.[27] An evidential and epistemic connection between chance and randomness falls well short of the conceptual connection proposed by the Commonplace Thesis with which we started.

Bibliography

Agafonov, V. N., 1968, 'Normal sequences and finite automata', *Soviet Mathematics Doklady*, 9: 324–25.
Albert, David Z., 1992, *Quantum Mechanics and Experience*, Cambridge, MA: Harvard University Press.
——, 2000, *Time and Chance*, Cambridge, MA: Harvard University Press.
Arntzenius, Frank and Ned Hall, 2003, 'On What We Know About Chance', *British Journal for the Philosophy of Science*, 54: 171–9.
Bar-Hillel, Maya and Willem A. Wagenaar, 1991, 'The Perception of Randomness', *Advances in Applied Mathematics*, 12: 428–54.
Barrett, Jeffrey A., 1999, *The Quantum Mechanics of Minds and Worlds*, Oxford: Oxford University Press.
Beall, JC and Greg Restall, 2006, *Logical Pluralism*, Oxford: Oxford University Press.
Bell, John S., 1964, 'On the Einstein-Podolsky-Rosen paradox', *Physics*, 1: 195–200.

Berkovitz, Joseph, Roman Frigg, and Fred Kronz, 2006, 'The Ergodic Hierarchy, Randomness and Chaos', *Studies in History and Philosophy of Modern Physics* 37: 661–91.

Bigelow, John, John Collins and Robert Pargetter, 1993, 'The Big Bad Bug: What are the Humean's Chances?', *British Journal for the Philosophy of Science*, 44: 443–62.

Bishop, Robert C., 2003, 'On Separating Predictability and Determinism', *Erkenntnis*, 58: 169–88.

Black, Robert, 1998, 'Chance, Credence, and the Principal Principle', *British Journal for the Philosophy of Science*, 49(3): 371–85.

Boolos, George S., John P. Burgess, and Richard C. Jeffrey, 2003, *Computability and Logic*, Cambridge: Cambridge University Press, 4th edition.

Borel, Émile, 1909, 'Les Probabilités Dénombrables et Leurs Applications Arithmétiques', *Rendiconti del Circolo Matematico di Palermo*, 27: 247–71.

Butler, Joseph, 1736, *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*, London: Knapton.

Callender, Craig, 2009, 'Thermodynamic Asymmetry in Time', in *The Stanford Encyclopedia of Philosophy* (Spring 2009 edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/spr2009/entries/time-thermo/>.

Carnap, Rudolf, 1945, 'The Two Concepts of Probability', *Philosophy and Phenomenological Research*, 5: 513–32.

Chaitin, Gregory, 1966, 'On the Length of Programs for Computing Finite Binary Sequences', *Journal of the Association for Computing Machinery*, 13: 547–69.

——, 1975, 'Randomness and Mathematical Proof', *Scientific American*, 232: 47–52.

——, 1987, 'Incompleteness Theorems for Random Reals', *Advances in Applied Mathematics*, 8: 119–46.

Champernowne, D. G., 1933, 'The Construction of Decimals Normal in the Scale of Ten', *Journal of the London Mathematical Society*, 8: 254–60.

Church, Alonzo, 1940, 'On the Concept of a Random Sequence', *Bulletin of the American Mathematical Society*, 46: 130–135.

Clark, Peter, 1987, 'Determinism and Probability in Physics', *Proceedings of the Aristotelian Society, Supplementary Volume*, 61: 185–210.

Dasgupta, Abhijit, 2011, 'Mathematical Foundations of Randomness', in Prasanta Bandyopadhyay and Malcolm Forster (eds.), *Philosophy of Statistics* (Handbook of the Philosophy of Science: Volume 7), Amsterdam: Elsevier, pp. 641–710.

Davies, Martin and Lloyd Humberstone, 1980, 'Two Notions of Necessity', *Philosophical Studies*, 38: 1–30.

de Finetti, Bruno, 1964, 'Foresight: Its Logical Laws, Its Subjective Sources', In *Studies in Subjective Probability*, Henry E. Kyburg, Jr. and Howard E. Smokler (eds.), New York: Wiley, pp. 93–158.

Delahaye, Jean-Paul, 1993, 'Randomness, Unpredictability and Absence of Order', In *Philosophy of Probability*, Jacques-Paul Dubucs (ed.), Dordrecht: Kluwer, pp. 145–67.

Doob, J. L., 1936, 'Note on Probability', *Annals of Mathematics*, 37: 363–7.

Downey, Rod and E. Griffiths, 2004, 'Schnorr randomness', *Journal of Symbolic Logic*, 69: 533–54.

Downey, Rod and Denis R. Hirschfeldt, 2010, *Algorithmic Randomness and Complexity*, Berlin: Springer.

——, ——, André Nies and Sebastiaan A. Terwijn, 2006, 'Calibrating Randomness', *Bulletin of Symbolic Logic*, 12: 411–91.

Eagle, Antony, 2004, 'Twenty-One Arguments Against Propensity Analyses of Probability', *Erkenntnis*, 60: 371–416.

——, 2005, 'Randomness is Unpredictability', *British Journal for the Philosophy of Science*, 56: 749–90.

——, 2011, 'Deterministic Chance', *Noûs*, 45: 269–99.

——, 2016, 'Probability and Randomness', In *The Oxford Handbook of Probability and Philosophy*, Alan Hájek and Christopher Hitchcock (eds.), Oxford: Oxford University Press, pp. 440–59.

Earman, John, 1986, *A Primer on Determinism*, Dordrecht: D. Reidel.

Eells, Ellery, 1983, 'Objective Probability Theory Theory', *Synthese*, 57: 387–442.

Ekeland, Ivar, 1988, *Mathematics and the Unexpected*, Chicago: University of Chicago Press.

Elga, Adam, 2004, 'Infinitesimal Chances and the Laws of Nature', *Australasian Journal of Philosophy*, 82: 67–76.

Feller, William, 1945, 'The Fundamental Limit Theorems in Probability', *Bulletin of the American Mathematical Society*, 51: 800–32.

——, 1950, *An Introduction to Probability Theory and Its Applications*. New York: Wiley.

Frigg, Roman, 2004, 'In What Sense is the Kolmogorov-Sinai Entropy a Measure for Chaotic Behaviour? —Bridging the Gap Between Dynamical Systems Theory and Communication Theory', *British Journal for the Philosophy of Science*, 55: 411–34.

Futuyma, Douglas J., 2005, *Evolution*. Cumberland, MA: Sinauer.

Gaifman, Haim, 1988, 'A Theory of Higher Order Probabilities', in *Causation, Chance and Credence*, volume 1, Brian Skyrms and William Harper (eds.), Dordrecht: Kluwer, pp. 191–219.

——, and Marc Snir, 1982, 'Probabilities Over Rich Languages, Testing and Randomness', *Journal of Symbolic Logic*, 47: 495–548.

Gates, P. and H. Tong, 1976, 'On Markov Chain Modeling to Some Weather Data', *Journal of Applied Meteorology*, 15: 1145–51.

Giere, Ronald N., 1973, 'Objective Single-Case Probabilities and the Foundations of Statistics', 4, Amsterdam: North-Holland, 467–83.

Gillies, Donald, 2000, 'Varieties of Propensity', *British Journal for the Philosophy of Science*, 51: 807–35.

Gilovich, Thomas, Robert Vallone, and Amos Tversky, 1985, 'The Hot Hand in Basketball: On the misperception of random sequences', *Cognitive Psychology*, 17: 295–314.

Glynn, Luke, 2010, 'Deterministic Chance', *British Journal for the Philosophy of Science*, 61: 51–80.

Gowers, Timothy, (ed.), 2008, *The Princeton Companion to Mathematics*, Princeton: Princeton University Press.

Hacking, Ian, 1965, *The Logic of Statistical Inference*, Cambridge: Cambridge University Press.

Hahn, Ulrike and Paul A Warren, 2009, 'Perceptions of Randomness: Why Three Heads Are Better Than Four', *Psychological Review*, 116: 454–61.

Hájek, Alan, 1997, '"Mises Redux"-Redux: Fifteen Arguments Against Finite Frequentism', *Erkenntnis*, 45: 209–227.

——, 2007, 'The Reference Class Problem is Your Problem Too', *Synthese*, 156: 563–85.

——, 2009, 'Fifteen Arguments Against Hypothetical Frequentism', *Erkenntnis*, 70: 211–235.

——, 2012, 'Interpretations of Probability', in *The Stanford Encyclopedia of Philosophy* (Winter 2012 edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/win2012/entries/probability-interpret/>.

Hall, Ned, 1994, 'Correcting the Guide to Objective Chance', *Mind*, 103: 505–17.

——, 2004, 'Two Mistakes About Credence and Chance', in *Lewisian Themes*, Frank Jackson and Graham Priest (eds.), Oxford: Oxford University Press, pp. 94–112.

Hawthorne, John, 2005, 'Chance and Counterfactuals', *Philosophy and Phenomenological Research*, 70: 396–405.

—— and Maria Lasonen-Aarnio, 2009, 'Knowledge and Objective Chance', in *Williamson on Knowledge*, Patrick Greenough and Duncan Pritchard (eds.), Oxford: Oxford University Press, pp. 92–108.

Hellman, Geoffrey, 1978, 'Randomness and Reality', in *PSA 1978*, volume 2, Peter D. Asquith and Ian Hacking (eds.), Chicago: University of Chicago Press, pp. 79–97.

Hoefer, Carl, 2007, 'The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance', *Mind*, 116: 549–96.

——, 2010, 'Causal Determinism', *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/spr2010/entries/determinism-causal/>.

Howson, Colin and Peter Urbach, 1993, *Scientific Reasoning: the Bayesian Approach*, Chicago: Open Court, 2nd edition.

Humphreys, Paul W., 1978, 'Is "Physical Randomness" Just Indeterminism in Disguise?', in *PSA 1978*, volume 2, Peter D. Asquith and Ian Hacking (eds.), Chicago: University of Chicago Press, pp. 98–113.

–––, 1985, 'Why Propensities Cannot be Probabilities', *Philosophical Review*, 94: 557–70.

Ismael, Jenann, 1996, 'What Chances Could Not Be', *British Journal for the Philosophy of Science*, 47: 79–91.

–––, 2008, 'Raid! Dissolving the Big, Bad Bug', *Noûs*, 42: 292–307.

–––, 2009, 'Probability in Deterministic Physics', *Journal of Philosophy*, 106: 89–109.

Jeffrey, Richard C., 1977, 'Mises Redux', in *Basic Problems in Methodology and Linguistics*, Robert E. Butts and Jaakko Hintikka (eds.), Dordrecht: D. Reidel, pp. 213–222.

Joyce, James M., 1998, 'A Nonpragmatic Vindication of Probabilism', *Philosophy of Science*, 65: 575–603.

–––, 2007, 'Epistemic Deference: The Case of Chance', *Proceedings of the Aristotelian Society*, 107: 187–206.

Kahneman, Daniel and Amos Tversky, 1972, 'Subjective Probability: A judgment of representativeness', *Cognitive Psychology*, 3: 430–54.

Kautz, Steven M., 1991, *Degrees of random sets*, Ph.D. thesis, Cornell University, [Available online].

Kolmogorov, A. N., 1929, Über das Gesetz des iterierten Logarithmus', *Math. Ann.*, 101: 126–35.

–––, 1933, *Grundbegriffe der Wahrscheinlichkeitrechnung*, Ergebnisse Der Mathematik; translated as *Foundations of Probability*, Chelsea Publishing Company, 1950.

–––, 1963, 'On Tables of Random Numbers', *Sankhya*, 25: 369–376.

–––, 1965, 'Three Approaches to the Definition of the Concept "Quantity of Information"'. *Problemy Peredachi Informatsii*, 1: 3–11.

–––, and V. A. Uspensky, 1988, 'Algorithms and Randomness'. *SIAM Theory of Probability and Applications*, 32: 389–412.

Kratzer, Angelika, 1977, 'What 'must' and 'can' Must and Can Mean'. *Linguistics and Philosophy*, 1: 337–55.

Kyburg, Jr., Henry E., 1974, *The Logical Foundations of Statistical Inference*, Dordrecht: D. Reidel.

–––, 1978, 'Subjective probability: Criticisms, reflections, and problems', *Journal of Philosophical Logic*, 7: 157–80.

Laplace, Pierre-Simon, 1826, *Philosophical Essay on Probabilities*, New York: Dover, 1951

Levi, Isaac, 1980, *The Enterprise of Knowledge*, Cambridge, MA: MIT Press.

Levin, L. A. and A. K. Zvonkin, 1970, 'The Complexity of Finite Objects and the Basing of the Concepts of Information and Randomness on the Theory of Algorithms', *Uspekhi Matematicheskikh Nauk*, 25 (6 (156)): 85–127.

Lewis, David, 1973, *Counterfactuals*, Oxford: Blackwell's.

–––, 1979a, 'Counterfactual Dependence and Time's Arrow', in his *Philosophical Papers*, volume 2, Oxford: Oxford University Press, 1986, pp. 32–66.

–––, 1979b, 'Scorekeeping in a Language Game', *Journal of Philosophical Logic*, 8: 339–59.

–––, 1980, 'A Subjectivist's Guide to Objective Chance', in his *Philosophical Papers*, volume 2, Oxford: Oxford University Press, 1986, pp. 83–132.

–––, 1994, 'Humean Supervenience Debugged', *Mind*, 103: 473–90.

Li, Ming and Paul M. B. Vitányi, 2008, *An Introduction to Kolmogorov Complexity and Its Applications*, Berlin and New York: Springer Verlag, 3rd edition.

Loewer, Barry, 2001, 'Determinism and Chance', *Studies in History and Philosophy of Modern Physics*, 32: 609–20.

–––, 2004, 'David Lewis' Humean Theory of Objective Chance', *Philosophy of Science*, 71: 1115–25.

–––, 2007, 'Counterfactuals and the Second Law', in *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Huw Price and Richard Corry (eds.), Oxford: Oxford University Press, pp. 293–326.

Luzzatto, Stefano, Ian Melbourne, and Frederic Paccaut, 2005, 'The Lorenz Attractor is Mixing', *Communications in Mathematical Physics*, 260: 393–401, [Preprint available online].

Martin-Löf, Per, 1966, 'The Definition of a Random Sequence', *Information and Control*, 9: 602–619.

–––, 1969a, 'Algorithms and Randomness', *Review of the International Statistical Institute*, 37: 265–72.

––––, 1969b, 'The Literature on von Mises' Kollektivs Revisited', *Theoria*, 35: 12–37.

Mather, J. and R. McGehee, 1975, 'Solutions of the collinear four-body problem which become unbounded in finite time', in *Dynamical Systems, Theory and Applications*, (Series: Lecture Notes in Physics, Volume 38), J. Moser (ed.), Berlin: Springer Verlag, pp. 573–97.

Matsumoto, Makoto and Takuji Nishimura, 1998, 'Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator', *ACM Transactions on Modeling and Computer Simulation*, 8(1): 3–30.

Mellor, D. H., 2000, 'Possibility, Chance and Necessity', *Australasian Journal of Philosophy*, 78: 16–27.

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

Montague, Richard, 1974, 'Deterministic Theories', in *Formal Philosophy*, Richmond H. Thomason (ed.), New Haven: Yale University Press.

Nies, André, 2009, *Computability and Randomness*, Oxford: Oxford University Press.

Norton, John D., 2003, 'Causation as Folk Science', *Philosophers' Imprint*, 3(4), URL = [Available online].

––––, 2008, 'The Dome: An Unexpectedly Simple Failure of Determinism', *Philosophy of Science*, 75: 786–98.

Pathak, Noopur, 2009, 'A computational aspect of the Lebesgue differentiation theorem', *Journal of Logic and Analysis*, 1 (9): 1–15 [Available online].

Popper, Karl, 1959, 'A Propensity Interpretation of Probability', *British Journal for the Philosophy of Science*, 10: 25–42.

––––, 1992, *Quantum Theory and the Schism in Physics*, New York: Routledge.

Porter, Christopher P., 2016, 'On Analogues of the Church-Turing Thesis in Algorithmic Randomness', *The Review of Symbolic Logic*, 9 (3): 456–79.

Pour-El, B. M. and I. Richards, 1983, 'Non-computability in Analysis and Physics', *Advances in Mathematics*, 48: 44–74.

RAND Corporation, 1955, *A Million Random Digits with 100,000 Normal Deviates*, New York: Free Press.

Reichenbach, Hans, 1949, *The Theory of Probability*, Berkeley: University of California Press.

Saari, Donald G. and Zhihong (Jeff) Xia, 1995, 'Off to Infinity in Finite Time', *Notices of the American Mathematical Society*, 42: 538–46.

Salmon, Wesley C., 1977, 'Objectively Homogeneous Reference Classes', *Synthese*, 36: 399–414.

Schaffer, Jonathan, 2003, 'Principled Chances', *British Journal for the Philosophy of Science*, 54: 27–41.

––––, 2007, 'Deterministic Chance?' *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.

––––, 1977, 'A survey of the theory of random sequences', in R. E. Butts and J. Hintikka (eds.), *Basic Problems in Methodology and Linguistics*, Dordrecht: D. Reidel, pp. 193–210.

Schurz, Gerhard, 1995, 'Kinds of Unpredictability in Deterministic Systems', in *Law and Prediction in the Light of Chaos Research*, P. Weingartner and G. Schurz (eds.), Berlin: Springer, pp. 123–41.

Schwarz, Wolfgang, 2016, 'Best System Approaches to Chance', In *The Oxford Handbook of Probability and Philosophy*, Alan Hájek and Christopher Hitchcock (eds.), Oxford: Oxford University Press, pp. 423–39.

Shimony, Abner, 2009, 'Bells' Theorem', in *The Stanford Encyclopedia of Philosophy* (Summer 2009 edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/sum2009/entries/bell-theorem/>

Sinai, Yakov G., 1992, *Probability Theory*, Berlin and Heidelberg: Springer-Verlag. Translated by D. Haughton.

Sklar, Lawrence, 1993, *Physics and Chance*, Cambridge: Cambridge University Press.

Skyrms, Brian, 1980, *Causal Necessity*, New Haven: Yale University Press.

Smith, Peter, 1998, *Explaining Chaos*. Cambridge: Cambridge University Press.

Sober, Elliott, 2010, 'Evolutionary Theory and the Reality of Macro Probabilities'. in *Probability in Science*, Ellery Eells and James H. Fetzer (eds.), Dordrecht: Springer, 133–61, [Preprint available online]

Stalnaker, Robert, 1978, 'Assertion'. in P. Cole (ed), *Syntax and Semantics 9*, New York: New York

Academic Press, 315–32.

Strevens, Michael, 1999, 'Objective Probability as a Guide to the World', *Philosophical Studies*, 95: 243–75.

Suppes, Patrick, 1984, *Probabilistic Metaphysics*, Oxford: Blackwell.

Talbott, William, 2008, 'Bayesian Epistemology', in *The Stanford Encyclopedia of Philosophy* (Fall 2008 edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/fall2008/entries/epistemology-bayesian/>.

Thau, Michael, 1994, 'Undermining and Admissibility', *Mind*, 103: 491–503.

van Lambalgen, Michiel, 1987a, *Random Sequences*, Ph.D. thesis, University of Amsterdam, [Available online].

–––, 1987b, 'Von Mises' Definition of Random Sequences Revisited', *Journal of Symbolic Logic*, 52: 725–55.

–––, 1995, 'Randomness and Infinity', Tech. Rep. ML-1995-01, ILLC, University of Amsterdam, URL = [Available online (in compressed Postscript)].

Venn, John, 1876, *The Logic of Chance*, London: Macmillan and Co., 2nd edition.

Ville, J., 1939, *Étude Critique de la Notion Collectif*, Paris: Gauthier-Villars.

von Mises, Richard, 1941, 'On the Foundations of Probability and Statistics', *Annals of Mathematical Statistics*, 12: 191–205.

––– 1957, *Probability, Statistics and Truth*. New York: Dover.

Wald, Abraham, 1938, 'Der Widerspruchsfreiheit des Kollektivbegriffes', *Actualités Scientifique et Industrielles*, 735: 79–99.

Wallace, David, 2007, 'Quantum Probability from Subjective Likelihood: Improving on Deutsch's Proof of the Probability Rule'. *Studies in History and Philosophy of Modern Physics*, 38: 311–32.

Wang, Yongge, 1996, *Randomness and Complexity*, Ph.D. Thesis, University of Heidelberg, Available online].

Weatherson, Brian, 2010 'David Lewis', *The Stanford Encyclopedia of Philosophy* (Summer 2010 Edition), Edward N. Zalta (ed.), URL =<https://plato.stanford.edu/archives/sum2010/entries/david-lewis/>.

Weiner, Matthew and Nuel Belnap, 2006, 'How Causal Probabilities Might Fit Into Our Objectively Indeterministic World', *Synthese* 149(1): 1–36.

Werndl, Charlotte, 2009, 'What Are the New Implications of Chaos for Unpredictability?' *British Journal for the Philosophy of Science*, 60: 195–220.

Williams, J. R. G., 2008, 'Chances, Counterfactuals, and Similarity', *Philosophy and Phenomenological Research*, 77: 385–420.

Williamson, Timothy (2006), 'Indicative Versus Subjunctive Conditionals, Congruential Versus Non-Hyperintensional Contexts', *Philosophical Issues*, 16: 310–33.

–––, 2007, 'How Probable is an Infinite Sequence of Heads?' *Analysis*, 67: 173–80.

Academic Tools

- How to cite this entry.
- Preview the PDF version of this entry at the Friends of the SEP Society.
- Look up this entry topic at the Internet Philosophy Ontology Project (InPhO).
- Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

- Brattka, Vasco, Joseph S. Miller and André Nies, 2011, 'Randomness and Differentiability' arXiv:1104.4465 [Available online]

- Kautz, S., 1991, *Degrees of Random Sets*, Ph.D. Thesis, Cornell University.
- Lieb, E. H., Daniel Osherson, and Scott Weinstein, 2006, 'Elementary proof of a theorem of Jean Ville' arXiv:cs/0607054v1 [Available online].
- Nies, André, 2011, 'Randomness and computable analysis: Results and open questions', talk at Computability, Complexity and Randomness, Cape Town, [Slides Available Online].
- Norton's Dome
- Reaber, Grant, 2010, 'The Deference-Based Conception of Rational Belief Updating'.
- Reimann, Jan, and Theodore A. Slaman, 2008, 'Measures and Their Random Reals.'
- Rute, Jason, 2011, 'Randomness and the Lebesgue Differentiation Theorem', talk at Southern Wisconsin Logic Colloquium, [Slides available online].

Related Entries

Bell's Theorem | chaos | computability and complexity | determinism: causal | epistemology: Bayesian | Lewis, David | probability, interpretations of | statistical physics: philosophy of statistical mechanics | time: thermodynamic asymmetry in

Acknowledgments