Determinism

December 1, 2014


1 Introduction

A system is deterministic just in case the state of the system at one time fixes the state of the system at all future times. A system is indeterministic just in case it is not deterministic. The question whether systems (or models or the world) are deterministic or indeterministic has concerned philosophers and scientists from the very beginning of philosophical and scientific thinking and still concerns them today. This article focuses on three recent discussions on determinism in the philosophy of science. First, determinism and predictability will be discussed (Section 2). Then, second, the paper turns to the topic of determinism, indeterminism, observational equivalence and randomness (Section 3). Finally, third, there will be a discussion about deterministic probabilities (Section 4). The paper will end with a conclusion (Section 5).

2 Determinism and Predictability

It has often been believed that determinism and predictability go together in the sense that deterministic systems are always predictable. Determinism is an ontological thesis. Predictability – that the future states of a system can be predicted – is an epistemological thesis. An illustration of mixing together determinism and predictability is the following famous quote by Laplace:

“We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes” (Laplace 1951: 4).

Here the first sentence is about the ontological thesis of determinism and the remainder of the quote concerns the epistemological notion of predictability.

However, a closer look reveals that determinism and predictability are very different notions. In particular, in recent decades chaos theory has highlighted that deterministic systems can be unpredictable in various different ways. Chaos theory is a field of study in mathematics that is part of dynamical systems theory and was developed in the second half of the twentieth century. As a mathematical theory it has applications in several disciplines including physics, meteorology, climate science, engineering, biology and economics.

Chaos theory studies the behavior of systems that are deterministic but at the same time show complicated behavior and are random and hence also unpredictable. Furthermore, chaos theorists
and philosophers have often claimed that chaotic systems are special in the sense that they are unpredictable in a way other deterministic systems are not (cf. Stone 1989; Smith 1998; Werndl 2009).

Let us now have a closer look in which sense deterministic systems can be unpredictable and what the unpredictability associated with chaotic dynamical systems amounts to. For illustration purposes it will be good to have a very simple example at hand. Consider a model of the evolution of the daily amount of precipitation over time where the possible amount of daily precipitation is in the range $[0mm,10mm]$ and the amount of precipitation $x_{t+1}$ at day $t+1$ is obtained from the amount of precipitation $x_t$ at day $t$ by the following equation:

$$x_{t+1} = \begin{cases} 2x_t & 0 \leq x_t \leq 5 \\ 2(10 - x_t) & 5 < x_t < 10 \end{cases}$$  \hspace{2cm} (1)

In chaos theory this map is called the tent map and it is shown in Figure 1. Note that, clearly, the dynamics of the tent map is deterministic.

![Figure 1](image)

Let me now turn to a first concept of unpredictability for deterministic systems, called asymptotic unpredictability. When measuring initial states such as the initial temperature, there will always be a certain inaccuracy. Thus a measurement corresponds to an extended bundle of initial conditions that represents the possible initial states compatible with the measurement. A system is said to be asymptotically unpredictable when any extended bundle of initial conditions, no matter how small, eventually over time spreads out more than a specific diameter representing the prediction accuracy of interest. Such a system is unpredictable in the sense that for any measurements of the initial states (regardless how fine), it will not be possible to predict the system with the desired prediction accuracy for all points of time in the future (cf. Werndl: 2009). The tent map map is asymptotically unpredictable. Figure 2 illustrates this by showing how a relatively small bundle of initial states spreads out over the entire range of possible values after just four time steps. Hence the evolution of the daily amount of precipitation is not predictable for all times.

When presented with notions such as asymptotic predictability, it becomes clear how systems that evolve according to deterministic laws can still be unpredictable: if the system is deterministic, it can still be that one cannot predict the state of system sufficiently far in the
future because very close initial states eventually lead to entirely different outcomes. In discussions about asymptotic unpredictability, sometimes the further claim has been made that this is the unpredictability unique to chaotic systems (e.g. Stone 1989: 127). However, it is easy to see that this cannot be true. For instance, a system where the possible states are in $(0,\infty)$ and the evolution is given by $x_{t+1} = cx_t$ for $c > 1$ shows asymptotical unpredictability, but this system is not chaotic because it is not random or complicated in any sense (cf. Smith 1998; Werndl 2009).

Werndl (2009) has argued that approximate probabilistic irrelevance is the kind of unpredictability that is unique to chaos. Unlike asymptotic unpredictability, approximate probabilistic irrelevance is a probabilistic concept of unpredictability. According to this concept, any measurement (i.e. knowledge of the initial states that the system may currently be in) is irrelevant for practical purposes for predicting outcomes sufficiently far in the future, i.e. makes it neither more nor less likely that the outcome is in any region of phase space of interest. This means that not only is it impossible to predict with certainty in which region the system will end up in the sufficiently distant future, but that also for practical purposes knowledge of the currently possible initial states neither lowers, nor heightens, the probability that the system will end up in a certain region of phase space in the sufficiently distant future.

As illustrated in Figure 2, the tent map is subject to this kind of unpredictability: The prior or default probability distribution of the tent map is the uniform distribution over the possible amounts of precipitation, i.e. the interval $[0mm,10mm]$. Suppose, for instance, that the knowledge of the amount of precipitation today is represented by the uniform distribution over $[0mm,10/8mm]$. Further, suppose one would like to know how likely it is that the amount of precipitation will be in some range, e.g. $[70/8mm,10mm]$ (i.e. in the set $B$) in the future. Then, if the prediction lead time is sufficiently long, e.g., four days, all one can say is that the probability of the amount of precipitation being in $[70/8mm,10mm]$ is $1/8$ as given by the prior or default probability distribution. That is, knowledge that the system is currently in $[0mm,10/8mm]$ will be entirely irrelevant for predicting whether the amount of precipitation will be in $[70/8mm,10mm]$ after four time steps. Approximate probabilistic irrelevance shows us another way in which deterministic systems can be unpredictable. Werndl (2009) has argued that this kind of unpredictability is also unique to chaotic systems, i.e., that only chaotic systems and no other deterministic systems show. Approximate unpredictability is a strong form of unpredictability and its discussion raises the question whether deterministic systems can be random and have properties similar to those of indeterministic processes, to which we now turn.

3 Determinism, Indeterminism and Randomness

Consider again our example of the model of the evolution of the daily amount of precipitation, where $[0mm, 10mm]$ is the range of possible amounts of precipitation, the dynamics is given by equation (1) and the probability of a certain outcome is measured by the uniform probability
Figure 3

measure over [0mm, 10mm]. Note that, as already emphasized, we can never measure states with infinite precision. Thus possible measurements correspond to a coarse-graining of the space of possible states. Figure 3 (top) shows a measurement with only two possible outcomes, while Figure 3 (bottom) shows a much finer measurement with eight possible outcomes. So when a deterministic system is observed, all we see is a sequence of observed outcomes. Suppose that our model of the evolution of the daily amount of precipitation is found to be in good agreement with the observations. Can we then be certain that the dynamics underlying the evolution of the daily amount of precipitation is deterministic? Or could there be observational equivalence between deterministic and indeterministic processes and could it be that the evolution of the daily amount of precipitation is governed by an indeterministic process?

To make progress on these questions, it needs to be made clear what an indeterministic model and process is and what observational equivalence amounts to. The focus here will be on stochastic models and processes. Consider the example where a coin is tossed every day, which corresponds to a two-valued Bernoulli model. A stochastic model such as a two-valued Bernoulli model consists of a set of possible outcomes (e.g., "head" or "tail") and \( Z_t \) denotes the outcome of the process at time \( t \) (e.g., whether the coin landed "head" or "tail" at day \( t \)). The probability distributions \( P(Z_t = e) \) give one the probability that the outcome of the process is \( e \) at time \( t \) (e.g., that the probability of tossing "head" today is 1/2); conditional probability distributions \( P(Z_t = e \mid Z_r = d) \) give one the probability that the outcome is in \( e \) at \( t \) given that it was \( d \) at \( r \) (e.g., that the probability of tossing "head" today is 1/2 given that I tossed "head" yesterday).

Now a deterministic model such as the model of the evolution of the daily amount of precipitation and a stochastic model such as the sequence of coin tosses are said to be observationally equivalent just in case the stochastic model and the deterministic model relative to the coarse-graining corresponding to the possible measurements give the same predictions (cf. Werndl 2009a). More specifically, the predictions obtained from the stochastic model are the probability distributions over the sequence of outcomes. Concerning the deterministic model, recall that a probability measure is defined over all possible states. Consequently, the predictions derived from the deterministic model relative to a certain coarse-graining (representing the possible measurements) are the probability distributions over the sequences of observations of the deterministic system. Hence what is meant by the phrase that the deterministic model and the stochastic model give the same predictions is that the possible observed values of the stochastic system and deterministic system are the same, and that the probability distributions over the sequences of observations of the deterministic model and the sequences of outcomes of the stochastic model are the same.

There are a host of results showing that deterministic models are often observationally equivalent to stochastic models (Werndl 2009, 2011, 2013a). An example for this is our deterministic model of the evolution of the daily precipitation: relative to the observational accuracy shown in Figure 3 (top), it is observationally equivalent to a two-state Bernoulli model such as our example of the sequence of coin tosses (outcome B1 corresponds to "head" and outcome “B2” to “tail”).
Relative to the observational accuracy shown in Figure 3 (bottom), the deterministic model of the tent map is observationally equivalent to a Markov model. For Markov models the next outcome only depends on the previous outcome and no other outcomes (and Markov models are among the most widely used stochastic models in science).

In the case where a deterministic model relative to a certain measurement accuracy and a stochastic model are observationally equivalent, the question arises: which model is preferable? There would be underdetermination if the data equally supported the deterministic and the stochastic model. Suppes (1993) and Suppes and de Barros (1996) argue there is underdetermination in these cases.

Werndl (2013a,b) points out that one needs to distinguish between the currently possible observations (given the current technology etc), and the the observations which are possible in principle (assuming that there are no limits, in principle, on observational accuracy). She argues that relative to observations which are possible in principle it will always be clear whether the deterministic or stochastic model is preferable (cf. also Wüthrich 2009). However, she argues that matters are less clear relative to the currently possible observations. Here underdetermination could arise, but, in her view, underdetermination can still be avoided for the most commonly discussed examples of a choice between deterministic models of Newtonian theory and stochastic models. Her argument makes use of the idea of indirect evidence, which is best introduced with an example. The theory of natural selection is only about processes happening out there in nature. So data about artificial breeding cannot be derived from the theory of natural selection. Still, with evolutionary theory as a bridge, data about artificial breeding can (and are often taken to) provide indirect evidence for the theory of natural selection.

Laudan and Leplin (1991) emphasize that indirect evidence can blocks the conclusion of underdetermination. For instance, suppose that there is a hypothesis H that does not follow from evolutionary theory but that (together with auxiliary hypotheses) gives rise to the same predictions as the theory of natural selection. Then the theory of natural selection is preferable relative to evidence and there is no underdetermination between H and the theory of natural selection because only the theory of natural selection is additionally supported by indirect evidence from artificial breeding. Similarly, Werndl (2013a) argues, the deterministic models from Newtonian mechanics are supported by indirect evidence from similar Newtonian models but the stochastic models are not, and thus the former are preferable. This argument can also be illustrated with our example of the evolution of the daily precipitation: suppose that the tent map were derivable from a general well-confirmed theory of the climate but the stochastic model is not derivable from any more general theory. Then the deterministic model would receive indirect evidence from other similar models of the climate theory and would hence be preferable to the stochastic model.

Stochastic processes such as Bernoulli processes and Markov processes are random. The results of observational equivalence show that deterministic models can be observationally equivalent to Bernoulli models or Markov models. Hence these results show that deterministic models can show randomness properties similar or equal to those stochastic processes. Indeed, a major task of the mathematical field of ergodic theory was to investigate to what extent randomness properties of stochastic models can also be found in deterministic models, and the investigations showed that many randomness properties carry over to deterministic models (Ornstein and Weiss 1992). To provide a concrete example: recall the discussion of approximate probabilistic irrelevance as a kind of unpredictability in the previous section. Eagle (2005: 775) defines randomness as a strong form of unpredictability: an event is random just in case the probability of the event conditional on evidence equals the prior probability of the event. This idea is at the
heart of approximate probabilistic irrelevance. Consequently, this concept of unpredictability can be regarded as a certain kind of randomness, and this randomness can also be found in deterministic systems such as the tent map. Let me finally turn to the third main topic of this article: the question of deterministic probabilities.

4 Probability and Determinism

Philosophers have often questioned whether ontic probabilities (i.e. probabilities that are real features of the world) can exist given deterministic laws. The method of arbitrary functions, which has been developed and advocated, amongst others, by Hopf, Poincare, Reichenbach and von Kries, promises to show that determinism and probabilities are compatible (it is important to note that this method is only meant to apply to certain cases and not to all situations where there are ontic probabilities).

The method of arbitrary functions is best introduced with an example (cf. Strevens, 2011). Consider a simple wheel of fortune that is painted in an equal numbers of very small equal-sized white, light grey and dark grey sections. The wheel is given a certain initial velocity, and when it comes to rest, a fixed pointer tells one the outcome (white or light grey or dark grey). We immediately tend to think that the probability of the outcome "white", "light grey" and "dark grey" is 1/3, despite the fact that the dynamics of the wheel is deterministic (or, if quantum effects crop up, the dynamics is at least approximately deterministic, and all that will be said carries over to this case).

A more detailed analysis of the wheel of fortune can substantiate this judgement. First of all, we have to look at the dynamics of the wheel, i.e. how initial velocities give rise to certain outcomes. As shown in Figure 4, what is distinctive is what Strevens (2003) calls a microconstant dynamics, i.e. given small ranges of initial velocities, the proportion of initial velocities leading to the outcomes “white“, "light grey“ and "dark grey“ is 1/3, respectively. Second, we have to look at how the wheel of fortune is prepared in a certain initial velocity. We can model the preparation of the system by a probability distribution p over the initial velocities. Usually our knowledge about this initial probability distribution is very limited. Furthermore, different ways of spinning a wheel by different persons etc. can be expected to correspond to different initial probability distributions. However, all this does not matter if the plausible assumption holds that all the possible probability densities p that we might employ do not fluctuate drastically on a very small region (Strevens, 2003, calls probability densities with this property “macro-periodic“). Given a microconstant dynamics and a macro-periodic probability density, the probabilities for the outcomes "white“, "light grey“ and “dark grey“ will all be approximately 1/3. This is illustrated by Figure 5, which shows two very different initial probabilities that both lead to probability 1/3 for the outcomes "white“, "light grey“ and “dark grey“.
In conclusion: even though the wheel of fortune is governed by deterministic equations, there are still ontic probabilities and these are explained by a microconstant dynamics and a class of possible probability distributions that are macro-periodic. Evidently, the method of arbitrary functions is of particular relevance when there is a class of possible initial densities. The prime example to which the method was applied by Hopf, Poincare, van Kries and Reichenbach etc. are games of chance. Next to this, it has also been suggested that it can make sense of deterministic probabilities in statistical mechanics, ecology and the social sciences (Abrams, 2012; Strevens, 2003; Werndl, 2013).

From a philosophical point of view a crucial question is how to interpret the probabilities of the method of arbitrary functions. And how to interpret these probabilities will depend on how one interprets the initial probability distribution. Thus let us ask: how should the initial probability distributions be interpreted? For lack of space, we can only discuss here three proposals (another proposals can be found in Rosenthal 2010, 2012).

Abrams' (2012) answer to this question is based on the frequencies obtained by actual inputs. More specifically, he argues that the correct input probability distribution is the micro-constant probability measure that minimises differences between probabilities and the frequencies obtained by the actual inputs. Yet it is unclear why there is a need for such a minimisation procedure (and it is also not clear why the measure has to be strictly microconstant). Furthermore, as Abrams' account is based on actual frequencies, it inherits many problems of finite frequentism (Rosenthal, 2010, 2012; Myrvold, 2011, 2014). For instance, there are cases where the first 50 tosses of a coin are highly unusual in the sense that they suggest that the coin is biased while it is not (which would become clear if further coin tosses were made). Abrams' account cannot make sense of such judgements since it is based on actual frequencies.

Strevens (2011) claims that for nearly all long series of trials of the system the initial conditions constitute a set that is macroperiodically distributed. He emphasizes that the initial distributions should just provide a summary of the actual occurrences of initial states and should have nothing to do with probabilities. He regards it as crucial that the initial distributions are not interpreted as probabilities because what is needed is an interpretation of the probabilities of the method of arbitrary functions that arises from non-probabilistic facts. Note that Strevens has to appeal to the condition that nearly all long series of trials produce macroperiodically distributed sets to avoid obvious objections (such as that we need an account why in an experimental situation that differs only in unimportant details, the same probabilities will arise).
Strevens' proposal is original and worthwhile, but there are also some problems. One problem can be illustrated with the example of the initial velocities of the wheel of fortune. When the wheel is spun repeatedly in the same context, one will find that the frequency distribution of the initial velocities approximates a certain density distribution. Because of this, scientists postulate that there is a probability distribution that describes the probability of preparing the wheel in a certain initial velocity. This probability distribution is useful from a predictive perspective in the sense that it usually gives correct predictions about the frequencies of initial velocities produced in the future in the same context. Yet since for Strevens there is nothing more than the actual occurrences of the initial velocities, he cannot make sense of such a predictive power. Another problem looms in the 'nearly-all' condition. Strevens makes this condition more precise by claiming that the actual distributions are macro-periodic in nearly all relevantly close possible worlds, where "nearly-all" is measured in terms of the Lebesgue measure, which measures "ways of altering the world". Yet it remains unclear why one can formally assign a measure to "ways of altering the world" and, even if one can do this, why the Lebesgue measure is the correct measure to use (cf. Rosenthal, 2010, 2012).

In my opinion, a more promising possibility is to interpret initial distributions as probability distributions that are physical quantities characterizing the particular situation at hand (as has been suggested by Szabó, 2007; see also Sober, 2010). In more detail: the concept of ontic probabilities can be reduced to ordinary physical quantities (as a consequence, the precise meaning of the probabilities will depend on the context of application). Hence, the probabilities of the method of arbitrary functions are simply physical quantities that characterize particular physical situations at hand.

It has been argued above that the method of arbitrary functions can provide an explanation of how probabilities arise out of determinism. Given this, let me now address the prominent worry that deterministic probabilities lead to a violation of the Principal Principle – a principle that establishes a connection between chances and credences (e.g. Schaffer, 2007). According to the Principal Principle, the credence of a rational agent in the occurrence of an event E should equal the chance of E as long as the agent has no inadmissible knowledge about the truth of E. More precisely: for all events E, all P and all K

$$cr_t(E | P\&K) = p$$

(2)

where 'cr_t' stands for the agent's credence at time t, P is the proposition stating that the chance that E occurs is p and K is an arbitrary admissible proposition. Now the crucial question is how an 'admissible proposition' is defined. Lewis (1986) famously suggested that laws of nature as well as historical information about the exact state of a system up to time t always count as admissible. Hence, given deterministic laws, the credences in equation (2) would be either 0 and 1 and the existence of nontrivial probabilities would lead to a violation of the Principal Principle. Because a violation of the Principal Principle is regarded as unacceptable, a common conclusion drawn from is that there are no deterministic probabilities.

This conclusion is too quick, and there are alternative and better ways to characterize an admissible proposition. In particular, Frigg and Hoefer (2014: 4) propose the following definition: a proposition K is admissible with respect to event E and chance setup S iff "K contains only the sort of information whose impact on reasonable credence about E, if any, comes entirely by way of impact on credence about the chances of those outcomes" (see also Glynn, 2010). Given this alternative definition, as desired, the Principal Principle (2) comes out as true.
5 Conclusion

The questions of determinism and indeterminism have concerned philosophers and scientists from the very beginning of philosophical and scientific thinking. This article illustrates that this topic is still very much a relevant one: as philosophy and science change and progress, there is always more that can be discovered and learnt about determinism and indeterminism. This article has focused on three recent discussions on determinism in the philosophy of science. First, determinism and predictability was discussed. It was emphasized that determinism and prediction are very different notions, and it was shown how the recent mathematical field of chaos theory has shed light on the various ways in which deterministic systems can be unpredictable. The second topic of the paper was determinism indeterminism, observational equivalence and randomness. Here results were presented that show that deterministic and indeterministic models can be observationally equivalent and the question was discussed how to choose between deterministic and indeterministic models. Further, it was argued that certain randomness properties of indeterministic systems carry over to deterministic systems. Finally, the third topic the article focused on was deterministic probabilities. Here it was argued that the method of arbitrary function is promising for understanding how deterministic probabilities can arise and that deterministic probabilities are not in violation of the Principal Principle.

Note on contributors

Charlotte Werndl is Professor for Logic and Philosophy of Science at the University of Salzburg and Visiting Professor at the Department of Philosophy, Logic and Scientific Method at the London School of Economics. She specialises in philosophy of science, philosophy of climate science, philosophy of physics, philosophy of statistics and philosophy of mathematics.

References


