Synthese
DOI 10.1007/s11229-010-9750-2

Deterministic probability: neither chance nor credence

Aidan Lyon

Received: 10 January 2010 / Accepted: 11 May 2010

Abstract Some have argued that chance and determinism are compatible in order to account for the objectivity of probabilities in theories that are compatible with determinism, like Classical Statistical Mechanics (CSM) and Evolutionary Theory (ET). Contrarily, some have argued that chance and determinism are incompatible, and so such probabilities are subjective. In this paper, I argue that both of these positions are unsatisfactory. I argue that the probabilities of theories like CSM and ET are not chances, but also that they are not subjective probabilities either. Rather, they are a third type of probability, which I call *counterfactual probability*. The main distinguishing feature of counterfactual-probability is the role it plays in conveying important counterfactual information in explanations. This distinguishes counterfactual probability from chance as a second concept of objective probability.

Keywords Chance · Credence · Determinism · Objective probability · Probability concepts

1 Introduction

Some scientific theories are true of some deterministic worlds but nevertheless posit what appear to be objective probabilities.[1] Classical Statistical Mechanics (CSM) is a paradigm example of such a theory. According to CSM, thermodynamic systems

[1] A deterministic world is a world whose entire history supervenes on the world's laws of nature with the complete state of the world at any time (see Earman 1986).

A. Lyon (✉)
Philosophy Program, Research School of Social Sciences,
Australian National University, Canberra, Australia
e-mail: aidan@coombs.anu.edu.au

are made up of large numbers of particles, whose behaviour is completely described by the deterministic theory of Hamiltonian mechanics. However, CSM entails that for any gas determined to freely expand (for example), there is some probability that it won't expand. But how can that be? How can an event be *determined* to occur and yet have some probability of not occurring?

Evolutionary Theory (ET) is another theory compatible with determinism that posits what appear to be objective probabilities. There has been some debate concerning whether ET is compatible with determinism, but there seems to be a growing consensus on this matter—in the affirmative.[2] The problem here is similar to the one before: the gene frequencies of a population can be determined to evolve one particular way, but ET assigns some probability to them *not* evolving that particular way.

In both theories (and in others), we have the same general issue: an event is determined to occur, but some probability is assigned to it not occurring. On the face of it, this seems like a strange and important problem. So let us give the problem a name: *the paradox of deterministic probabilities*.[3]

In the literature, there are two general strategies for resolving the paradox. The first strategy involves arguing that our concept of chance is in fact compatible with determinism, and so there isn't any problem here (beyond the general problem of analysing chance). Those who pursue this first strategy include Levi (1990), Loewer (2001), Hoefer (2007), and Frigg and Hoefer (2009).[4] The second strategy is to argue that, despite appearances, the probabilities in question are subjective probabilities, and so there is no problem because there is nothing problematic with a positive subjective probability assigned to an event determined not to occur. Those who pursue this second strategy include Rosenberg (1994), Graves et al. (1999), Schaffer (2007), and Frigg (2008).[5]

These two general strategies coincide with what appears to be an unspoken but often made assumption: that probabilities are only chances or credences. One often sees/hears the argument: the relevant probabilities can't be credences (for whatever reason), so they must be chances (e.g., Loewer 2001). One also sees/hears the opposite argument: the relevant probabilities can't be chances (for whatever reason), so they must be credences (e.g., Schaffer 2007). In contrast, the argument in this paper will be: the relevant probabilities can't be chances (for reasons given in Sect. 2), they can't be credences either (for reasons given in Sect. 3), so they must be something else. The upshot of the argument is that we need to identify a third concept of probability (to be explained in Sect. 4) that is distinct from chance and credence.

---

2 See, e.g., Weber (2001, 2005), Rosenberg (2001), Millstein (2003a,b), and Sober (1984, 2010) for discussion.

3 Loewer uses this name for the problem as it appears in CSM (Loewer 2001, p. 612). I find the problem just as vexing in ET and any other theory it appears in, so I prefer to use the name for the general problem.

4 Sober (2010) argues for the compatibility of macro objective probability and determinism. As I argue later, this is different from the compatibility of chance and determinism.

5 Strictly speaking, Rosenberg (1994) and Graves et al. (1999) address not the paradox of deterministic probabilities, but a very similar problem: how a theory can assign probability values of "1/2" and the like to events that are assigned quantum mechanical probabilities very close to 1 or 0.

This third concept of probability is what I call *counterfactual probability*, and is a second type of objective probability. The thesis is therefore that we have at least two distinct concepts of objective probability: chance and counterfactual probability, along with at least one concept of subjective probability: normatively constrained credence.[6] The primary goal of this paper is to show that we have this third concept of probability in science, and to explain how it is distinct from chance and credence. (This is not to rule out the possibility that we have even more concepts that need identifying.) It is not a goal of the paper to defend an analysis of counterfactual probability—just as my goal is not to defend an analysis of chance either. That said, I will sketch an analysis of counterfactual probability at the end of Sect. 4 as this will help elucidate the concept. The analysis I sketch is an interpretation of probability based on similarity relations over ensembles of possibilities developed by Bigelow (1976, 1977).

2 Not chances

Those who claim that there are no chances in deterministic worlds often do not provide a supporting argument (e.g., Lewis 1986, p. 120; Popper 1982, p. 105). In fact, there seems to be only two such arguments in the literature: the so-called Laplacean demon argument due to Laplace (1814), and an argument due to Schaffer (2007). The Laplacean demon argument has been sufficiently criticised elsewhere (see Sober 1984, 2010), so that leaves Schaffer's argument as the only tenable one in the literature. I defend (a version of) that argument here.

Schaffer identifies certain platitudes we apparently have about chance, and then argues that deterministic conceptions of chance fail to satisfy enough of those platitudes to count as genuine chance. He identifies six such platitudes in the form of six different principles about chance, one of which is the Principal Principle (PP). I'm going to focus solely on the PP here, for three reasons (other than that of brevity). First, the PP is the most famous of the principles and seems to be accepted by many authors. Second, some of the platitudes that the other principles capture are already captured by the PP.[7] And third, there is some debate about some of the details of the PP in the literature that is quite pertinent to the debate about deterministic chance, and so it will be worthwhile focusing on clearing that up.

Before we get to the argument from the PP, we should see what the PP is about. The PP is the following constraint on initial conditional credences, due to Lewis (1980):

$$Cr\,(A|Ch_{tw}(A) = x \wedge E) = x \tag{1}$$

[6] A note on terminology: under the umbrella of "subjective probability", I include so-called Objective Bayesians. This is because on that view, probability is degree of belief, even though that degree of belief is normatively constrained (to uniqueness or near uniqueness) by what evidence one has. On the other hand, the logical interpretation is an objective interpretation because (on the view) it represents a partial entailment relation between propositions (or sentences) that is meant to be independent of what evidence anyone has.

[7] For example, one such principle is called the Future Principle, which basically says that only the future is chancy. This fact about chance is already captured by the PP with the fact that historical propositions are always admissible (see 2). (Some deny that the Future Principle is a basic truth about chance—e.g., Frigg and Hoefer (2009).)

for any A and for any admissible E—where Cr is a reasonable initial credence function and $Ch_{tw}(A)$ is the chance of A, at time t, at world w. What it means, exactly, for a proposition to be admissible is a tricky issue, and Lewis has no definition of admissibility (as he tells us himself)—just a sketch. When sketching what admissibility amounts to, Lewis gives us some examples and a general job-description:

> Historical propositions are admissible; so are propositions about the dependence of chance on history [e.g., the laws of nature]. Combinations of the two, of course, are also admissible. More generally, we may assume that any Boolean combination of propositions admissible at a time also is admissible at that time. Admissibility consists in keeping out of a forbidden subject matter—how the chance processes turned out—and there is no way to break into a subject matter by making Boolean combinations of propositions that lie outside it. (Lewis 1980, p. 276)

We'll come back to the issue of how we should understand admissibility later as it is essential to understanding what platitude(s) about chance the PP is meant to capture. But for now, we have enough details to go through the argument from the PP against deterministic chance.

The argument is a *reductio*. First, suppose Ch_{tw} is the chance function of a deterministic world w, and consider a particular chance assignment, $Ch_{tw}(A) = x$, where $0 < x < 1$, and A is true. Let $\mathcal{H}_{tw}$ be the entire history of world w through to time t, and let $\mathcal{L}_w$ be the deterministic laws of w. From the PP we have:

$$Cr\,(A|Ch_{tw}(A) = x \wedge \mathcal{L}_w \wedge \mathcal{H}_{tw}) = x \tag{2}$$

(We can substitute $\mathcal{L}_w \wedge \mathcal{H}_{tw}$ in for E because $\mathcal{L}_w \wedge \mathcal{H}_{tw}$ is assumed to be admissible.) Since $0 < x < 1$, it follows that:

$$0 < Cr(A|Ch_{tw}(A) = x \wedge \mathcal{L}_w \wedge \mathcal{H}_{tw}) < 1 \tag{3}$$

But since w is deterministic, $\mathcal{L}_w \wedge \mathcal{H}_{tw} \vDash A$, and so from the probability axioms:

$$Cr(A|Ch_{tw}(A) = x \wedge \mathcal{L}_w \wedge \mathcal{H}_{tw}) = 1 \tag{4}$$

which contradicts (3). So it looks like deterministic chance cannot satisfy the PP.

(As I mentioned earlier, Schaffer gives similar arguments to show that deterministic conceptions of chance fail to satisfy other principles that he takes to be constitutive of chance. The overall strategy to the argument, then, is: indeterministic conceptions of chance can satisfy all of the principles in question, while no deterministic conception of chance can do this, so there are is no such thing as deterministic chance. I will be simply focusing on the PP here, and assuming that indeterministic conceptions of chance can satisfy it.)

Hoefer gives the following response to the above argument[8]:

[8] I have changed Hoefer's notation slightly, to fit in with mine.

> [… T]his derivation is spurious; there is a violation of the correct understanding of admissibility going on here. For if $\mathcal{L}_w \wedge \mathcal{H}_{tw}$ entails A, then it has a big (maximal) amount of information pertinent as to whether A, and not by containing information about A's objective chance! So $\mathcal{L}_w \wedge \mathcal{H}_{tw}$, so understood, must be held inadmissible, and the derivation of a contradiction fails. (Hoefer 2007, p. 559)

To maintain that $\mathcal{L}_w \wedge \mathcal{H}_{tw}$ is inadmissible is to deny at least one of the following—*contra* Lewis: (i) $\mathcal{H}_{tw}$ is admissible, (ii) $\mathcal{L}_w$ is admissible, and (iii) admissibility is closed under Boolean combinations.

There are some technical issues regarding whether the laws are admissible (see Lewis 1994) and whether admissibility is closed under Boolean combinations (see Lyon 2009).

Fortunately, for our purposes, we can bypass these issues. It turns out that there is a similar argument against the thesis that CSM probabilities are chances that only relies on the admissibility of historical propositions. This new argument will therefore allow us to focus on whether historical propositions are admissible.

For this new argument, assume everything as before, but now assume that $Ch_{tw}(A) = x$ is a probability statement of CSM and that A is entirely about the state of the world at time t. (Such probability statements exist in CSM and I will say more about this shortly.) From the PP we have:

$$Cr(A|Ch_{tw}(A) = x \wedge \mathcal{H}_{tw}) = x \tag{5}$$

(We can substitute $\mathcal{H}_{tw}$ in for E, because $\mathcal{H}_{tw}$ is assumed to be admissible.) Since $0 < x < 1$, it follows that:

$$0 < Cr(A|Ch_{tw}(A) = x \wedge \mathcal{H}_{tw}) < 1 \tag{6}$$

But since A is entirely about the state of the world at time t, it follows that $\mathcal{H}_{tw} \vDash A$, and so from the probability axioms:

$$Cr(A|Ch_{tw}(A) = x \wedge \mathcal{H}_{tw}) = 1 \tag{7}$$

which contradicts (6).

The crucial step in the above argument is the premise that CSM entails that for some A, $0 < Ch_{tw}(A) < 1$ and A is entirely about the state of the world at time t. This is true because CSM assigns probabilities at time t over ways a system could be at time t.

Consider the standard CSM explanation for why ice cubes in warm water melt. For any such ice cube, it is overwhelmingly likely—but not certain—that its micro-state is one that evolves deterministically into a micro-state that corresponds to the "melted" macro-state. Even though the ice cube has micro-state s_1 at time t, a positive probability is assigned at time t to the ice cube having a different micro-state, s_2, at time t.[9] (So an example of A in the above argument would be any disjunction of micro-states compatible with the macro-state.) This feature of CSM probabilities results from the fact that we start with an initial probability distribution over possible initial conditions and conditionalise it on the *macro*-history of the world as history unfolds (see Loewer 2001, p. 618). This leaves all the micro-states compatible with the macro-history as open possibilities—even though only one micro-state obtains.

Winsberg identifies this as the fundamental problem with understanding CSM probabilities objectively:

> The fundamental problem with understanding the probabilities in [CSM] to be objective is that we are meant to posit a probability distribution over a set of possible initial states, while we suppose, at the same time, that in fact only one of these initial states actually obtained. (Winsberg 2008, p. 2)

For ease of reference, I will call this property of the time index of the probability function being the same as the time of some events that are assigned (non-trivial) probability the *synchronicity* of CSM probabilities.[10]

Note that in the above argument from the PP against CSM chances, we made no mention of the laws of nature.[11] This means we made no assumption concerning the determinism/indeterminism of world w. So we have an argument for why CSM probabilities are not chances that has nothing to do with determinism. The issue of the compatibility of chance and determinism is therefore somewhat of a red-herring. The literature typically focuses on the determinism of CSM probabilities as a big problem for understanding CSM probabilities as chances. However, a more fundamental problem is their synchronicity. (Incidentally, the probabilities of quantum statistical mechanics also have this synchronicity, but note that the probabilities of quantum mechanics (standardly interpreted) do not.)

Some will respond to the above argument with an objection similar to Hoefer's objection to the original argument: there is a violation of the correct understanding of admissibility here; $\mathcal{H}_{tw}$ entails A, so it contains a maximal amount of information pertinent to A, and so must be inadmissible. We therefore need to address the issue of

[9] By "positive probability" I mean infinitesimal or measure—zero probability—depending on what the right way to think about such probabilities turns out to be. This is technical issue that isn't important here—see e.g., Hájek (2003) for further discussion.

[10] This is not to say that CSM does not assign *diachronic* probabilities (i.e., probabilities to events after t)—in fact it must, in order to make predictions. However, these past and future probabilities are derived from the present probabilities (all talk of past, present and future is relative to t). For example, the probability at time t that the entropy of a gas is high at time t' ($>t$) is the probability that the gas has a micro-state at time t that evolves into another micro-state at t' that corresponds to a macro-state of high entropy.

[11] In personal communication, Roman Frigg has objected that the chance-statement "$Ch_{tw}(A) = x$" smuggles in the laws of nature, because CSM probabilities are defined in terms of the phase-flow, which is defined in terms of Hamiltonian mechanics. While it is true that CSM probabilities *can* be expressed in terms of the phase-flow, it doesn't follow that a statement involving CSM probabilities has the laws of nature as part of its content. An analogy will help make this clearer. The probability of "heads" of a coin-flip can be expressed in terms of the mechanics that govern the coin-flip (e.g., see Diaconis 1998). However, the statement "the probability of "heads" is 1/2" doesn't entail anything about those mechanics—one can know that the probability of "heads" is 1/2, without knowing anything about the mechanics of the coin.

how we ought to understand admissibility (but we can do so without having to worry about the admissibility of laws and the Boolean closure of admissibility).

The notion of admissibility does most of the platitude-capturing work in the PP. One of the intuitions it seems we have about chance is that it "locks" credence in a very robust way. For example, when I find out that the chance of a coin-flip landing "heads" is 1/2, my credence in "heads" is 1/2 and it doesn't change from 1/2 when I acquire new evidence. This is one of the examples Lewis considers in his chance questionnaire at the beginning of Lewis (1980), where he first introduces the PP:

> A certain coin is scheduled to be tossed at noon today. You are sure that this chosen coin is fair: it has a 50% chance of falling heads and a 50% chance of falling tails. [… But] you have plenty of seemingly relevant evidence tending to lead you to expect that the coin will fall heads. This coin is known to have a displaced center of mass, it has been tossed 100 times before with 86 heads, and many duplicates of it have been tossed thousands of times with about 90% heads. Yet you remain quite sure, despite all this evidence, that the chance of heads this time is 50%. To what degree should you believe the proposition that the coin falls heads […]?
> Answer. […] 50% […]. *To the extent that uncertainty about outcomes is based on certainty about their chances, it is a stable, resilient sort of uncertainty—new evidence won't get rid of it.* (my emphasis) (Lewis 1980, pp. 264–265)

The admissibility of historical propositions is meant to capture this fundamental intuition about chance. Once we know the chance is 1/2, our credence is 1/2, and getting any other information about the world up to the current time won't affect this. As far as intuitions about chance go, they don't come much more straightforward than this. Nevertheless, no doubt, some will want to deny this—e.g., Strevens (2006) denies that historical propositions are admissible for related reasons. If that is the case, then it seems that the debate cannot proceed in a fruitful way since we have different intuitions about chance.

This deadlock becomes obvious when one begins to give an analysis of the probabilities in question. For example, Eagle objects to propensity analyses of chance that rule out deterministic chance, while Hájek objects to frequency analyses of chance that *allow* deterministic chance:

> Classical statistical mechanics proposes non-trivial probabilities, and yet is underlaid by a purely deterministic theory. To deny that these probabilities are "real" is simply to come into conflict with one of the starting points of any genuine inquiry into the nature of probability: that it should explain the empirical success of probabilistic theories like statistical mechanics. It is a heavy burden on the propensity theorist to explain why these 'pseudo-probabilities', given that they are the best fillers of the role available (as far as explanation and prediction go), should be denied the umbrella of probability. (Eagle 2004, pp. 386–387)

> *Determinism*, it would seem, is incompatible with *intermediate* (objective) probabilities: in a deterministic world, nothing is chancy, and so all objective chances

> are 0 or 1. But determinism is no obstacle to there being relative frequencies that lie between these values. (emphasis in original) (Hájek 1997, p. 81)

There is a sense in which this debate is merely terminological. Those who argue that chance and determinism are incompatible seem to intend to refer to a special kind of objective probability—the sort of probability one finds in quantum mechanics (standardly interpreted). Those who argue that chance and determinism are compatible intend to refer to any kind of objective probability, the sort of probability that science can discover—e.g., the probabilities one finds in CSM, ET, and many other theories.[12] Perhaps it is best, then, to drop the word "chance" altogether, and instead speak of the probabilities in different theories or applications. This way, we can address the conceptual issues without getting caught up in merely terminological matters. For example, we can ask: "Do we have the same concept of probability in quantum mechanics as we do in statistical mechanics?" without using the term "chance".

Having said that, though, for the rest of this paper I will stick with the usage of "chance" that Lewis, Schaffer and others prefer. This is because once we lay out all the platitudes we seem to have about chance, it appears that indeterministic conceptions of chance satisfy these platitudes better than deterministic ones can (see e.g., Schaffer 2007). The important point, though, is that this doesn't entail that the probabilities of CSM or ET are not "real" or objective. They are not chances, but they are still objective probabilities. They can constrain credence in a way similar to how chance constrains credence. For example, Loewer describes a version of the PP which he calls PP_{macro} which says (roughly) that one's credence in A should be the probability that CSM assigns to A provided one has no *macroscopically inadmissible* information. The difference between chance and other objective probabilities is captured by the above quote from Lewis: new evidence will never get rid of the uncertainty set by chances, but it will for other objective probabilities.

This means that we have the general concept of objective probability, under which many subconcepts fall. Chance is once such concept. How many others are there? I argue in Sect. 4 that we have at least one other such concept: counterfactual probability. To properly understand the applications of probability in our scientific theories, we may need to identify yet more concepts.[13] However, I argue that at least the probabilities of CSM and ET are best understood as counterfactual probabilities.

All of this, however, assumes that we should understand these probabilities objectively in the first place—i.e., that they are not subjective probabilities. This is a controversial assumption, so in the next section I will argue for it.

[12] Frigg and Hoefer (2009) argue for deterministic chance but admit that "deterministic chance" seems like an oxymoron (p. 1), and that chance and determinism seem to be incompatible (p. 2). However, they give no argument for why chance and determinism are compatible—despite our intuitions—that goes beyond the usual argument that the probabilities in question cannot be subjective, therefore they must be chances. They write: "The values of these probabilities are determined by how things are, not by what we believe about them. In other words, these probabilities are chances, not credences" (p. 2).

[13] For example, one referee of this article pointed out that the probabilities of coalescence theory ought be counted as objective probabilities, but may not be chances or counterfactual probabilities.

3 Not credences

Here is an argument for why CSM probabilities are objective: There are certain regularities in the world that are objective facts. CSM explains those regularities, making reference to probabilities. Those probabilities must therefore be objective.[14]

Here are some notable instances of this argument. First, Popper:

> Since the physical possibility of this event [the spontaneous reversion of a gas into a bottle] cannot be doubted, we explain the experimental fact that the process is irreversible by the extreme *improbability* of a spontaneous reversion into the bottle. And since the fact to be explained—the irreversibility of the process—is an objective experimental fact, the probabilities and improbabilities in question must be objective also. (emphasis in original) (Popper 1982, p. 107)

Second, Loewer (rhetorically):

> Consider, for example, the statistical mechanical explanation of why an ice cube placed in warm water melts in a short amount of time. The explanation roughly proceeds by citing that the initial macro-condition of the cup of water + ice (these are objects at different temperatures that are in contact) is one in which on the micro-canonical probability distribution it is overwhelmingly likely that in a short amount of time the cup + ice cube will be near equilibrium; i.e., the ice cube melted. If the probability appealed to in the explanation is merely a subjective degree of belief then how can it account for the melting of the ice cube? What could your ignorance of the initial state of the gas have to do with an explanation of its behaviour? (Loewer 2001, p. 611)

And finally, Albert (even more rhetorically):

> Can anybody seriously think that it is somehow *necessary*, that it is somehow *a priori*, that the particles that make up the material world must arrange themselves in accord with *what we know*, with what *we happen to have looked into*? Can anybody seriously think that our merely being *ignorant* of the exact micro-conditions of thermodynamic systems plays some part in *bringing it about*, in *making it the case*, that (say) *milk dissolves in coffee*? How could that *be*? (all emphasis in original) (Albert 2000, p. 64)

It will be convenient to break this argument up into premise and conclusion form:

P1 CSM gives probabilistic explanations of objective facts.

P2 The probabilities in probabilistic explanations of objective facts must be objective.

C Therefore, the probabilities in CSM probabilistic explanations are objective.

[14] See Sober (2010) for another, related argument, which uses a deterministic, but probabilistic model for a coin flip. That argument relies on there being non-trivial micro-probabilities in a deterministic world, which incompatibilists would deny.

P1–C is valid, and it seems to adequately capture the argument in the above passages. So we need to examine the premises.

Schaffer objects to **P1**:

> There just is no probabilistic explanation in the offing here. What explains the melting of the ice cube is the complex deterministic process that runs from the ice cube's entering the water to its melting, whose myriad details we can only guess. The only probability involved in the deterministic process of an ice cube melting is the measure of our ignorance of the real micro-explanation. (Schaffer 2007, p. 136)

And so does Frigg, in response to the above passage from Albert (2000):

> Of course the cooling down of drinks and the boiling of kettles has nothing to do with what anybody thinks or knows about them; but they have nothing to do with the probabilities attached to these events either. Drinks cool down and kettles boil because the universe's initial condition is such that under the dynamics of the system it evolves into a state in which this happens. All we need to explain why things happen is the initial condition and the dynamics. (Frigg 2008, p. 680)

Schaffer and Frigg seem to make two claims in response to the argument **P1–C**:

(i) CSM does not give probabilistic explanations of irreversible phenomena.
(ii) The initial conditions and the fundamental dynamics of the world are what really explain irreversible phenomena.

It seems they agree that *if* the explanations in question are probabilistic explanations, then the probabilities involved *are* objective. But they deny the antecedent of this conditional. In terms of our argument: **P2** is not disputed, but **P1** is.[15]

Schaffer goes on to explain why authors have been misled into thinking **P1–C** is a good argument. He draws a distinction between *probabilistic explanation* and *probability of explanation* (Schaffer 2007, p. 119). A probabilistic explanation is an explanation where probability plays an explanatory role. For example, an explanation of some phenomenon that involved a quantum mechanical probability as an explanatory factor is a probabilistic explanation.[16] In contrast, a probability of explanation is "merely an ignorance measure over various nonchancy explanatory paths" (ibid.). ("Nonchancy" because we are focusing on deterministic settings.) The real explanation for why, say, an ice cube melted is a complex, non-chancy micro-physical explanation. Because we are ignorant of the micro-physical details, we assign subjective probabilities to various epistemically possible micro-physical explanations. According to Schaffer, it is a mistake to think that this is a genuine probabilistic explanation; all we have here are probabilities of explanations—an ignorance measure over various possible "real" explanations.

Schaffer is not alone in drawing this distinction. Railton (1981), for example, draws the same distinction:

[15] And so also an epistemic interpretation of explanation (e.g., Hempel's Inductive-Statistical model) is not on the table here.

[16] Assuming a suitable indeterministic interpretation of quantum mechanics.

> At first blush, one might think that whenever statistics or probabilities are involved in explanatory practice one is dealing with a form of probabilistic explanation. However, this illusion is quickly shed once one recognizes the variety of ways in which statistics and probabilities figure in explanatory activities. Perhaps the commonest use of statistics and probabilities in connection with explanation is epistemic: they are used in the process of assembling and assessing evidence for causal and non-causal explanations alike. Somewhat less common, but still important, are [those cases in which] statistics and probabilities are used in providing explanatory information about causal and non-causal processes and their initial conditions. In some cases of the last sort we have genuine probabilistic explanation, specifically, in those cases where information is provided about a physically indeterministic process. (Railton 1981, p. 254)

Both Schaffer and Railton claim that only in the cases where we have indeterministic processes can we have genuine probabilistic explanations, and one cannot conclude that an explanation is probabilistic merely from the fact that it involves a probability.

In other words, an explanation involving probability is not automatically a probabilistic explanation—it could be a probability of explanation. And that is where those who endorse **P1**–**C** have gone wrong: they saw explanations that involved probabilities and concluded that they must be probabilistic explanations, when in fact they are only probabilities of explanations.

There are two serious problems with this response to the argument **P1**–**C**. The first problem is that once we get clear on what exactly it means for the probability distribution in a CSM explanation to be a "measure of our ignorance", it becomes clear that CSM probabilities cannot be probabilities of explanations. I will argue that CSM probabilities do not reflect our ignorance in a way that is appropriate for understanding them as probabilities of explanations. I therefore conclude that there must be probabilities in probabilistic explanations, and so there can be probabilistic explanations even for deterministic processes. The second problem is with the idea that the initial conditions plus the deterministic laws form "*the* real" explanation for why an ice cube melted. I will address the first problem in the remainder of this section and the second problem in the next section as it provides a natural spring-board to identifying some of the conceptual role counterfactual probability plays in explanations.

Both Schaffer and Frigg claim that CSM probabilities are measures of our ignorance. We have already seen that Schaffer makes this claim. As for Frigg:

> The universe has exactly one initial microcondition, and there is nothing chancy about this condition. How, then, can we understand a probability distribution over initial conditions? The only answer seems to be that this distribution reflects our ignorance about the system's initial microcondition; all we know is the system's initial macrostate, and so we put a probability distribution over the microconditions compatible with that macrostate that reflects our lack of knowledge. (Frigg 2008, p. 679)

But what, exactly, does it mean for a probability distribution to "reflect our ignorance"?

One possible answer is that it means that the probability distribution assigns probabilities to propositions that perfectly match our own personal degrees of belief in those propositions. This is implausible, though. It has been shown that our personal degrees of belief systematically fail to satisfy the standard probability axioms (Kahneman et al. 1982), and yet the probability distribution of CSM does satisfy them.

Another possible answer, then, is: it means that the probability distribution assigns probabilities to propositions that we *ought* to have as our personal degrees of belief. This is a normative account, in contrast to the descriptive one just entertained. The plausibility of this answer depends on how the normative claim is fleshed out. For instance, if I somehow happen to know the precise micro-physical details of an ice cube, then my credences should *not* be aligned with the probabilities of CSM. In fact, if the ought-claim is merely a norm of *rationality*, then, arguably, I don't even need to *know* the micro-physical details for the norm not to apply to me; I merely need to have certain *beliefs* about them. So, for this answer to have any plausibility, the normative claim must be relativised to a certain type of epistemic state. Schematically: if one believes Φ, and nothing stronger, then one's credences should be aligned with the probabilities of CSM.

The question then is: What is Φ? A seemingly plausible answer is that Φ is everything we *actually* know. We never actually know the micro-physical details of thermodynamic systems, we only ever get to know their macro-states. So, given this level of ignorance, it seems plausible that our credences should be equal to the CSM probabilities. This proposal seems to be closer to what Schaffer and Frigg are getting at.

However, consider the standard CSM explanation for why an ice cube initially frozen at time t is melted at some later time, t'. At t, it was highly likely that the ice cube was in a micro-state that would evolve into another micro-state at t' that corresponds to the "melted" macro-state; but there is also some probability that this is not the case. Put another way, there is a non-trivial probability assigned to those possible micro-states that evolve into micro-states that correspond to the ice cube not being melted. *But we know that the ice cube melted.* We therefore know that the ice cube was *not* in a micro-state that would evolve into a micro-state that corresponds to the "frozen" macro-state. We are not as ignorant as the CSM probability distribution allegedly makes us out to be.[17]

The point is that while we never know what the micro-state of a thermodynamic system is, we know enough about its macro-states to rule out certain micro-states that the CSM distribution assigns non-trivial probability to. In personal communication, Carl Hoefer has raised the concern that this problem could easily be resolved by simply

[17] A similar, but distinct point has been recognised in the literature on probability in evolutionary theory by Roberta Millstein. She writes:

> [T]his "ignorance" interpretation overlooks the fact that we are aware of more causal factors than are included in the transition probability equation; for example, we know things about the predator and the color of the butterflies. Thus, we choose to ignore these causal factors, rather than being ignorant of them. (Millstein 2003b, p. 1321)

This is slightly different to my point though. The difference is that when we use the probabilities of CSM to make *predictions*, the probability distribution does seem to appropriately reflect the extent of our ignorance. However, even when we make predictions with the transition probability equations that Millstein writes of, we often know more causal factors than those which the models represent.

conditionalising the distribution at t—call it ρ—on the macro-history at t', to create a new distribution, ρ'. The idea is that ρ' assigns probability 0 to all of the problematic micro-states, and therefore adequately reflects our ignorance concerning the system. The claim that ρ' reflects our ignorance in such a situation is much more plausible than the claim that ρ reflects our ignorance in the situation. However, in the standard CSM explanation for why an ice cube melts, it is ρ that appears in the explanation—*not* ρ'. And that is the crux of the matter: while we may use ρ' in order to *predict* facts about the world, we use ρ to *explain* facts about the world, *even though* ρ does not reflect our ignorance. This point has also been made in the literature concerning the probabilities of ET. Sober (1984) has argued that we have two uses of probabilities in ET: for predictions and explanations, and that the probabilities we use for explanations in ET will sometimes differ from the ones we use to make predictions. Sometimes the probability distribution that we use in an explanation in ET does not account for (i.e., has not been conditionalised on) everything we know, or believe. The same goes for CSM.

Also in personal communication, Jonathan Schaffer has suggested that Φ is not all that we know, but only what we know at time t—i.e., the macro-state of the system at time t. This certainly seems plausible, for the purpose of *prediction*—i.e., assigning credences at time t to ways the system might behave. However, this line of thought would undermine Schaffer's original proposal regarding *explanation*: that CSM probabilities are probabilities of explanations, and not probabilities in probabilistic explanations. This is because according to this version of the account, CSM probabilities are no longer of a kind with other canonical examples of probabilities of explanations.

For example, consider why the dinosaurs are extinct: an asteroid probably hit the Earth 65 million years ago and killed them all. The "probably" involved in this explanation is clearly a probability of explanation. Given what we know, we can't rule out with certainty other possible explanations—e.g., that a large increase in volcanic activity 65 million years ago killed the dinosaurs. But note that given what we know, we can rule out the possibility that *nothing* killed the dinosaurs (i.e., that they are alive and well today), and no probability is assigned to this possibility. In contrast, we know that the ice cube doesn't remain frozen, but some probability *is* assigned to this possibility. Or consider another example: Someone was sick, they were given antibiotics, and their health improved. Why did their health improve? Answer: The antibiotics probably did their work. It could have been something else—e.g., that the bacteria were resistant to the antibiotics, the person was naturally immune to the infection and that is why they got better. Here the "probably" in "The antibiotics probably did their work" is a probability of explanation—an ignorance measure over various non-chancy explanatory paths.[18] And again, in contrast to the ice cube example, no probability is assigned to the person remaining ill because we know they got better.[19]

18 Thanks to John Matthewson for this example.

19 This is not to deny that there may be probabilistic explanations in the ballpark. For example, the probability of the patient getting better, given that they were given antibiotics may figure in a probabilistic explanation for why the patient got better.

The reference to probability in a probability of explanation reflects the extent of our ignorance concerning what happened—we use this probability to make the "best guess" at what happened. However, on this latest proposal concerning what Φ is, the probabilities of CSM are not *our* credences; they are not even what our credences *should* be, given our knowledge about the macro-states of the system. Rather, they are what our credences should be *at an earlier time*.[20] This is radically different to any other canonical example of a probability of explanation.

In fact, CSM probabilities play a role in explanations more analogous to the role chances play in canonical examples of probabilistic explanations. Consider the quantum mechanical probability that figures in the explanation for why half of a collection uranium-235 atoms have decayed after 704 million years (roughly the half-life of uranium-235). This probability clearly does not represent our current credences (i.e., after half of the atoms have decayed): according to the distribution, it is possible that *all* of the atoms decay. But we know that they don't all decay. This is completely analogous to how probability figures in the explanation for why an ice cube melted.

So while Schaffer is correct to draw a distinction between probabilistic explanations and probabilities of explanations, he is incorrect in placing CSM probabilities in the latter category. The claim that CSM probabilities are probabilities of explanations would be plausible if they could somehow be understood as our "ignorance measures", but they are clearly not *our* "ignorance measures"—we are not as ignorant as they would make us out to be. Since we have only distinguished between probabilistic explanations and probability of explanations, the only option left is to understand CSM probabilities as probabilities that figure in probabilistic explanations. And given that they play a role similar to the one that chances play in other canonical examples of probabilistic explanations, it seems appropriate not to worry about a third possibility.

4 Explanatory ecumenism and counterfactual probability

So far, I have argued that there can be probabilities in theories like CSM and ET that are not chances (Sect. 2), and not credences either (Sect. 3). These probabilities are therefore of a third type. While such a third concept of probability is typically missed by the literature, I am not alone in recognising that it exists. For example, Batterman (1992) clearly recognises that we need to identify a third concept of probability (and for similar reasons):

> On the old classical view, probabilities are due entirely to our ignorance of the system's true exact state. On the quantum theory, probabilities are irreducible, where this is understood in terms of propensities and the nonexistence of hidden variables. The simple examples discussed in the early sections, as well as the later discussion of explanation in [C]SM, indicate that probabilities may arise in some third way, *via instabilities and symmetries in the dynamics of the theory*. (emphasis in original) (Batterman 1992, p. 347)

[20] Given, presumably, what we knew or *would* have known about the macro-states at that time.

Having argued that we need to identify such a third concept of probability, I now want to say something positive about the concept, by outlining some of the conceptual role this third concept of probability plays. What I think is the distinguishing feature of this third concept is its role in conveying a certain type of counterfactual information in explanations. A brief detour through the philosophy of explanation will help elucidate this role.

As we saw before (Sect. 3), Schaffer claims that *the real* explanation of any given irreversible phenomenon is a micro-physical explanation. One reason to think this claim is true is if reductionism about explanation is true. However, there are good reasons not to be such a reductionist. For example, Putnam (1975) famously argues that the micro-physical explanation for why a square peg didn't fit through a round hole in a board misses something very important about why the peg didn't fit. The micro-physical explanation doesn't capture the fact that the exact micro-physical details do not matter: it is the squareness of the peg and the roundness of the hole that matter. Putnam goes further and argues that the micro-explanation is not even an explanation, or is a terrible one.

Jackson and Pettit (1992) and Sterelny (1996) give similar arguments for why there can be multiple explanations of some phenomena, but they stop short of Putnam's extra claim, adopting a more ecumenical approach to explanation.[21] For them, any given phenomenon can have multiple, equally good, explanations for why it occurred. In the case of Putnam's example, the micro and macro explanations are not competing explanations. Rather, they are quite complementary to each other. They provide two types of important information about the peg-board system. The micro-physical explanation tells us in full detail what stopped the peg from going through the hole; it tells which particles and which forces are actually responsible. The macro-physical explanation—the one in terms of the squareness of the peg and the roundness of the hole—tells us that it didn't matter which particles actually stopped the peg passing through the hole; no matter how the peg was rotated, there would be some set of particles and forces stopping it from passing through the hole.

Jackson and Pettit (1992) call these two types of information *modally contrastive information* and *modally comparative information*, respectively. The micro-physical explanation gives us modally contrastive information: it contrasts our world from other worlds, it identifies our world by specifying the actual micro-physical details of the peg-board system. The macro-physical explanation gives us modally comparative information: it compares our world with other worlds, it unites our world with other worlds by generalising the micro-physical details of the peg-board system. Or, more accurately, the micro-explanation gives *more* contrastive information than the macro-explanation does, and the macro-explanation gives *more* comparative information than the micro-explanation does. Both types of information can be important to an explanation.

If our goal is to convey as much modally contrastive information as possible when explaining some phenomenon, then in the ideal case, we give a micro-physical explanation. In a deterministic world, such an explanation will not involve (non-trivial)

[21] See also Sober (1999) for another pluralist approach to such explanations.

probabilities, for if it did, there would always be some other explanation, with more information about what the actual world is like. However, in an indeterministic world—like our quantum world—such explanations will typically involve (non-trivial) probabilities for there is no more information available to contrast the actual world with other worlds.[22] These facts allow us to identify an important aspect of the conceptual role chance plays in explanations: chances are those probabilities in explanations that maximise modally contrastive information.

However, often enough, our goal is not to maximise modally contrastive information. Often our goal is to convey some level of modally comparative information. When this is the case, we end up giving a macro-physical explanation. Typical examples of this in the literature are cases where the modally comparative information comes in an all-or-nothing form. For instance, in Putnam's example, the modally comparative information is the proposition that the peg is square and the hole is round. This tells us that the micro-physical details did not matter *at all*. However, there are cases where modally comparative information comes in degree form. These are cases where the micro-physical details matter or do not matter to some *degree*. Jackson and Pettit (1992) happen to give such an example: a flask full of water cracked because the water was boiling. When discussing this example, they write:

> … in being made aware of the boiling-water explanation, we learn something new: we learn that in *more or less all possible worlds* where the relevant causal process is characterized by involving boiling water, the process will lead to the flask cracking. (my emphasis) (Jackson and Pettit 1992, p. 15)

The CSM explanation for why an ice cube melted (for example) does something very similar: it tells us that the micro-physical details of the ice cube matter *very little*. This "modal robustness" is an important, contingent physical fact about thermodynamic systems. To not recognise this fact, and to focus only on the micro-physical details of the world, is to miss an important fact about the world.

Railton (1981) has also pointed out this fact about CSM in connection with explanation:

> This illuminates a modal feature of the causal processes involved and therefore a modal feature of the relevant ideal explanatory texts: this sort of causal process is such that its macroscopic outcomes are remarkably insensitive (in the limit) to wide variations in initial microstates. The stability of an outcome of a causal process in spite of significant variation in initial conditions can be informative about an ideal causal explanatory text in the same way that it is informative to learn, regarding a given causal explanation of the First World War, that a world war would have come about (according to this explanation) even if no bomb had exploded in Sarajevo. This sort of robustness or resilience of a process is important to grasp in coming to know explanations based upon it. (Railton 1981, p. 251)

[22] They will not *necessarily* involve probabilities as there can be non-probabilistic indeterministic worlds (see Norton 2006).

However, Railton ranks explanations according to their quality (or what he calls their "explanitoriness") (Railton 1981, p. 240). At one end of the ranking, we have the best explanations (the most explanatory), which maximise what Railton calls *explanatory information*. At the other end are the worst explanations, those that convey no explanatory information. By "explanatory information", Railton appears to intend to refer to the same concept that we have been using "modally contrastive information" to refer to. So, on this picture, CSM explanations are not as good as explanations involving initial conditions and dynamical laws. However, I prefer to distinguish between two types of explanatory information: modally contrastive information, and modally comparative information. On this picture, CSM explanations are explanatory in a way that is different from the way that the explanations involving the initial conditions and dynamical laws are explanatory, and so the former are neither better nor worse than the latter.

Probability plays a similar conceptual role in the explanations of ET, for ET is also a theory that conveys modally comparative information in its explanations. As Sterelny and Kitcher write:

> In principle, we could relate the biography of each organism in the population, explaining in full detail how it developed, reproduced, and survived, just as we could track the motion of each molecule of a sample of gas. But evolutionary theory, like statistical mechanics, has no use for such a fine grain of description: the aim is to make clear the central tendencies in the history of evolving populations [...]. (Sterelny and Kitcher 1988, p. 345)

Both theories are in the business of abstracting away from particular details of their systems of interest and capturing and explaining the general behaviour of those systems.[23] In the case of ET, it may have been that a population evolved one particular way because of some detailed sequence of organism and environment interactions, but that sequence may have been largely irrelevant to the final outcome. That is, the population would have still evolved the way it did, had the sequence been slightly different, because of (say) fitness differences in the population.

So sometimes we use probabilities to express modally comparative information about a system, which is a certain kind of counterfactual information about the system. Such probabilities are objective since they express objective facts about the system in question, and they play a role in explanations that is distinct from the role chances play in explanations. For lack of a better term, I call these probabilities *counterfactual probabilities*. Counterfactual probability is not probability in some counterfactual situation; rather it is a measure of how *robust* a proposition is under a class of counterfactual situations.[24] (Sterelny (1996) calls explanations that convey modally comparative information *robust process explanations*. In this terminology, counterfactual probability is a measure of *how* robust a robust process is.) A rough analogy with logical probability may make it clearer how these probabilities are counterfactual. Logical probabilities (if they exist) generalise logical entailment: "$P(A|B) = x$" generalises

23 Sober (1984, pp. 118–134) also makes this point.

24 One referee for this article suggested calling chances "maximally fine-grained" and counterfactual probabilities "more coarse-grained", which is a useful way of thinking about the difference.

"$B \vdash A$". In a similar way, counterfactual probabilities generalise counterfactuals: "$P(A|B) = x$" generalises "$B \Box\!\!\rightarrow A$".

Now that chance and counterfactual probability are distinguished, the next question is how to analyse them.[25] Answering this question in full detail requires the space of another paper. However, I would like to suggest one approach to answering this question, as I think it may help make the distinction between chance and counterfactual probability clearer. To help fix our ideas, let us assume that a propensity interpretation (e.g., Popper 1959) or a best-system analysis (e.g., Lewis 1994) is correct for chance. How, then, should we analyse counterfactual probability? The main distinguishing role that counterfactual probability plays is that it conveys important counterfactual information in explanations. So an analysis that is closely connected to the analysis of counterfactuals seems like a promising first start. Fortunately, an analysis along these lines has already been developed by Bigelow (1976, 1977). Roughly speaking, Bigelow takes the similarity relation that appears in popular analyses of counterfactuals and uses it to induce a similarity metric over an ensemble of possibilities. This metric is then used to measure the size of propositions in that ensemble, and then this measure is used to build a probability function. Roughly speaking, on this account, the probability of a proposition is proportional to the size of the proposition (measured according to the similarity relation, which is taken as primitive), which is precisely the sort of thing we are after. I think this is a very promising approach, and it would be worth exploring in further detail in another paper.

5 Conclusion

In this paper I have not developed an analysis of counterfactual probability, just as I have not developed an analysis of chance. My goal here has been primarily to show that we need to identify and distinguish these two concepts of objective probability in our scientific theories. How we should analyse these two concepts—that is, what theory we should give as to what in the world makes chance statements and counterfactual probability statements true—is another matter. It may be that propensity facts make chance statements true, and frequency facts make counterfactual probability statements true. Or it may be that frequency facts make both sorts of statements true. My suspicion is that a best-system analysis of chance (Lewis 1994) and a similarity-relation analysis of counterfactual probability (Bigelow 1976, 1977) will be the best approach to take.

Identifying more than one concept of objective probability has two significant advantages over not doing so. First, we can avoid the unacceptable conclusion that the probabilities of some our scientific theories are subjective just because there are

[25] The two concepts differ in other respects as well—not just by the role they play in probabilistic explanations. For example, counterfactual probabilities won't satisfy the PP and yet chances do. Counterfactual probabilities will satisfy some of the platitudes identified by Schaffer (2007), but not all of them (e.g., the counterfactual probabilities of ET will not satisfy the Lawful Magnitude Principle if there are no laws of biology). I lack the space here for a full discussion of how counterfactual probabilities satisfy or fail to satisfy Schaffer's platitudes. However, all I need for the purposes of this paper is that chance and counterfactual probability differ in at least *some* respects.

no deterministic chances. And second, we can avoid the unacceptable conclusion that chance and determinism are compatible just because the probabilities of our scientific theories are not subjective. In other words, by distinguishing two types of objective probabilities, we can avoid the paradox of deterministic probabilities (see Sect. 1).

I have argued that what distinguishes two particular types of objective probabilities are the roles they play in probabilistic explanations. Chances are those probabilities in explanations that maximise modally contrastive information. Counterfactual probabilities are those probabilities in explanations that give some level of modally comparative information in degree form.[26] To account for probability of explanations (see Sect. 3), we need the concept of subjective probability. Therefore, to account for explanations in science that involve probabilities generally, we need to identify at least three concepts of probability: subjective probability, chance, and counterfactual probability.

Acknowledgements I'd like to especially thank Alan Hajek and Jonathan Schaffer for incredibly helpful and extensive comments on earlier versions of this paper. I'd also like to thank Marshall Abrams, Jeremy Butterfield, Fabrizio Cariani, David Chalmers, Mark Colyvan, Anthony Eagle, Kenny Easwaran, Christopher Eliot, Branden Fitelson, Roman Frigg, Carl Hoefer, Jack Justus, John Matthewson, Daniel Nolan, Mike Titelbaum, and audiences at the LSE Sigma Club, 2009 Foundations of Uncertainty Conference in Prague, 2009 Meeting of the Australasian Association for Philosophy, University of Sydney Current Projects Seminar, and the Australian National University PhilSoc Seminar.

References

Albert, D. (2000). *Time and chance*. Cambridge, MA: Harvard University Press.

Batterman, R. (1992). Explanatory instability. *Noûs, 26*, 325–348.

Bigelow, J. (1976). Possible worlds foundations for probability. *Journal of Philosophical Logic, 5*(3), 299–320.

Bigelow, J. (1977). Semantics of probability. *Synthese, 36*(4), 459–472.

Diaconis, P. (1998). A place for philosophy? The rise of modeling in statistical science. *Quarterly of Applied Mathematics, 56*(4), 797–805.

Eagle, A. (2004). Twenty-one arguments against propensity analyses of probability. *Erkenntnis, 60*, 371–416.

Earman, J. (1986). *A primer on determinism*. Dordrecht: Reidel.

Frigg, R. (2008). Chance in Boltzmannian statistical mechanics. *Philosophy of Science, 75*(5), 670–681.

Frigg R. & Hoefer C. (2009) Determinism and chance from a Humean perspective . In D. Dieks, W. Gonzalez, S. Hartmann, M. Weber, F. Stadler, & T. Uebel (Eds.), *The present situation in the philosophy of science*. Berlin: Springer.

Graves, L., Horan, B., & Rosenberg, A. (1999). Is indeterminism the source of the statistical character of evolutionary theory?. *Philosophy of Science, 66*, 140–157.

Hájek, A. (1997). 'Mises Redux'-redux: Fifteen arguments against finite frequentism. *Erkenntnis, 45*, 209–227.

Hájek, A. (2003). What conditional probability could not be. *Synthese, 137*(3), 273–323.

Hoefer, C. (2007). The third way on objective probability: A sceptic's guide to objective chance. *Mind, 116*(463), 549–596.

Jackson, F., & Pettit, P. (1992). In defense of explanatory ecumenism. *Economics and Philosophy, 8*(1), 1–21.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

[26] This is not to say that these roles are *all* that there is to chance and counterfactual probability. For example, both are also guides to credences.

Laplace, P. (1814). *A philosophical essay on probabilities*. New York: Dover (1951)
Levi, I. (1990). Chance. *Philosophical Topics, 18*(2), 117–149.
Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. II). Berkeley: University of California Press.
Lewis, D. (1986). A subjectivist's guide to objective chance. In *Philosophical papers* (Vol. II, pp. 83–132). Oxford: Oxford University Press.
Lewis, D. (1994). Humean supervenience debugged. *Mind, 103*, 473–490.
Loewer, B (2001). Determinism and chance. *Studies in History and Philosophy of Modern Physics, 32*, 609–620.
Lyon, A. (2009). *Three concepts of probability*. Ph.D. thesis, Australian National University.
Millstein, R. (2003a). How not to argue for the indeterminism of evolution: A look at two recent attempts to settle the issue. In A. Hüttemann (Ed.), *Determinism in physics and biology* (pp. 91–107). Paderborn: Mentis.
Millstein, R. L. (2003b). Interpretations of probability in evolutionary theory. *Philosophy of Science, 70*, 1317–1328.
Norton, J. (2006). The dome: An unexpectedly simple failure of determinism. *Philsci-Archive*.
Popper, K. R. (1959). The propensity interpretation of probability. *The British Journal for the Philosophy of Science, 10*(37), 25–42.
Popper, K. (1982). *Quantum theory and the schism in physics*. Totowa, NJ: Rowman and Littlefield.
Putnam, H. (1975). Philosophy and our mental life. In *Readings in philosophy of psychology*. Cambridge, MA: Harvard University Press.
Railton, P. (1981). Probability, explanation, and information. *Synthese, 48*(2), 233–256.
Rosenberg, A. (1994). *Instrumental biology or the disunity of science*. Chicago: University of Chicago Press.
Rosenberg, A. (2001). Discussion note: Indeterminism, probability, and randomness in evolutionary theory. *Philosophy of Science, 68*(4), 536–544.
Schaffer, J. (2007). Deterministic chance?. *The British Journal for the Philosophy of Science, 58*(2), 113–140.
Sober, E. (1984). *The nature of selection*. Chicago: University of Chicago Press.
Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science, 66*(4), 542–564.
Sober, E. (2010). Evolutionary theory and the reality of macro probabilities. In E. Eells & J. Fetzer (Eds.), *Probability in science*. La Salle, IL: Open Court.
Sterelny, K. (1996). Explanatory pluralism in evolutionary biology. *Biology and Philosophy, 11*, 193–214.
Sterelny, K., & Kitcher, P. (1988). The return of the gene. *The Journal of Philosophy, 85*(7), 339–361.
Strevens, M. (2006). Probability and chance. *Encyclopedia of philosophy* (2nd ed.). Detroit: Macmillan Reference USA.
Weber, M. (2001). Determinism, realism, and probability in evolutionary theory. *Philosophy of Science, 68*(3), S213–S224.
Weber, M. (2005). Darwinism as a theory for finite beings. In V. Hösle & C. Illies (Eds.), *Darwinism and philosophy* (pp. 275–297). Notre Dame: University of Notre Dame Press.
Winsberg, E. (2006). Probability and chance. *Studies in History and Philosophy of Modern Physics*. doi:10.1016/j.shpsb.2008.05.005.