No interpretation of probability*

Wolfgang Schwarz
Draft, 03 March 2016

I argue that none of the usual "interpretations of probability" provides an adequate interpretation of probabilistic theories in science. Assuming that the aim of such theories is to capture noisy relationships in the world, we do not have to give them classical truth-conditional content at all: their probabilities can remain uninterpreted. Indirectly, this account turns out to explain what is right about the frequency interpretation, the best-systems interpretation, and the epistemic interpretation.

Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means.

— Attributed to Bertrand Russell

1 Interpreting science

Probabilities play a central role in many theories all across science, from quantum physics and statistical mechanics to chemical kinetics, systems biology, and evolutionary theory. If we want to take seriously what science tells us about the world, we have to ask what the probability statements in these theories mean.

Consider Boltzmann-style statistical mechanics. Here the objects of study are isolated physical systems consisting of a large number N of particles. The possible micro-states of such systems (with fixed energy) correspond to a region Γ in a $6N$-dimensional state space, each point of which specifies the precise location and momentum of every particle. According to the "basic postulate" of statistical mechanics, the probability with which the system's state lies in a subregion S of Γ is equal to the ratio $\mu(S)/\mu(\Gamma)$, where μ is a measure of volume (the Liouville measure) associated with the space.

The postulate in some sense identifies the probability $P(S)$, that the system's state lies in region S, with the quantity $\mu(S)/\mu(\Gamma)$. But the identification is not a stipulative definition: '$P(S) = x$' is not just shorthand for '$\mu(S)/\mu(\Gamma) = x$'. That would turn

* Ancestors of this paper were presented to a metaphysical conspiracy meeting in Mainz, the probability reading group at the ANU, and the DFG research group on causation, laws, dispositions and explanations at the University of Cologne. I thank the audiences at these events for helpful feedback.

the probabilistic predictions of statistical mechanics into trivial analytic truths. But statistical mechanics is an empirical theory that explains real-life phenomena such as the melting of ice cubes and the diffusion of soy milk in coffee. This empirical success depends on the identification of $P(S)$ with $\mu(S)/\mu(\Gamma)$ rather than infinitely other probabilistic quantities that could be defined on the same state space. In fact, it is commonly held that in order to yield the right empirical predictions, the identification of $P(S)$ with $\mu(S)/\mu(\Gamma)$ is not quite right and that μ should be replaced by another measure μ_p that gives lower weight to systems moving from a high entropy past to a low entropy future (see e.g. [Albert 2000: ch.4]).

So the identification of $P(S)$ with $\mu(S)/\mu(\Gamma)$ or $\mu_p(S)/\mu_p(\Gamma)$ is not just a definition. It seems to have empirical consequences. What exactly are these consequences? What does statistical mechanics say about the world when it says that that $P(S) = \mu_p(S)/\mu_p(\Gamma)$?

Some hold that there is a fundamental probabilistic quantity of *chance*, built into the very fabric of physical reality. It is supposed to measure the propensity of physical systems to evolve in one way rather than another. Tritium atoms, for example, are thought to have a propensity for decaying into helium. This propensity is taken to be an intrinsic physical property, not unlike mass or charge. It is not reducible to what we know or believe, nor to statistical facts about relative frequencies: even though the half-life of tritium is 12 years, it is perfectly possible (albeit improbable) that all tritium atoms in fact decay after less than 12 years. Non-trivial chance, on this conception, requires indeterminism in the laws of physics. If the laws of physics entail that a given atom will decay in exactly 8.7 years, the decay would no longer be chancy.

There is an extensive debate in philosophy over whether the idea of primitive chance is ultimately coherent and whether the probabilities in quantum physics can be interpreted as primitive chances. We do not need to enter this debate, for it is quite clear – and widely agreed – that the probabilities of statistical mechanics are not primitive chances. My cup of coffee can hardly be said to have a propensity to be, right now, in one micro-state rather than another. Indeed, classical statistical mechanics assumes a deterministic micro-dynamics. It certainly does not presuppose that the fundamental laws are stochastic.[1]

One might suggest that although the probabilities of statistical mechanics do not pick out primitive quantum-physical propensities, they pick out another primitive probabilistic quantity – call it "statistical mechanical chance". But there is no good reason to believe

1 [Albert 2000: ch.7] suggests that the probabilities of statistical mechanics, and indeed of all higher-level scientific theories, might be reduced to those of GRW quantum physics. A friend of primitive chance might therefore hold that all probabilistic theories in science are in fact – more or less indirectly – concerned with the primitive chances of GRW. But even if we grant Albert's controversial reductionism about special science, GRW is a highly speculative theory. What if GRW is not the true theory of fundamental physics? Nobody thinks that this would undermine the basic principles of statistical mechanics or evolutionary theory. So there must be an interpretation of these theories on which they do not presuppose the truth of GRW.

in such a quantity. Among other things, the quantity would seem to be epiphenomenal. The future state of a physical system is completely determined (to the extent that it is) by its present micro-state and the fundamental dynamical laws. It isn't sensitive to the values of any other fundamental quantity.[2] Moreover, it is plausible that the laws of statistical mechanics supervene on the fundamental structure of the world, in the sense that a world with the very same distribution of micro-properties and the same micro-laws couldn't have different statistical mechanical probabilities. Again this suggests that these probabilities are not fundamental.

So even those who believe in primitive chance still need an interpretation of the probabilities in statistical mechanics and other high-level scientific models – an interpretation that is neutral on the existence of primitive chance.

There is no shortage of proposals on the market. An old classic is the frequency interpretation that identifies probabilities with actual or counterfactual frequencies. But the frequency interpretation runs into a large number of well-known problems (see e.g. [Hájek 1997], [Hájek 2009]). Scientific practice clearly allows probabilities to deviate from actual frequencies, especially if the relevant conditions are rare. Turning to counterfactual frequencies only seems to makes things worse, since it is hard to give a coherent definition of counterfactual frequencies that preserves their status as a worthy object of empirical study.

A more sophisticated relative of the frequency interpretation is the best-systems interpretation, first developed by David Lewis ([Lewis 1986: 128], [Lewis 1994]). Here the probability of an outcome is defined as the probability assigned to the event by whatever empirical theory best combines the virtues of simplicity, strength, and fit, where the latter measures the extent to which the theory assigns high probability to actual events. Lewis originally restricted the analysis to dynamical probabilities in fundamental physics, but various authors have argued that it can be extended to statistical mechanics and other areas of science (see e.g. [Loewer 2004], [Loewer 2007], [Schrenk 2007], [Hoefer 2007], [Cohen and Callender 2009]).

Another popular proposal is the epistemic interpretation that treats probability statements as statements about rational belief. Roughly speaking, to say that an event has a 50 percent probability is here taken to mean that it would be rational to be 50 percent confident that the event will occur. The epistemic interpretation has been especially influential in statistical mechanics (see e.g. [Jaynes 1957], [Uffink 2011]).

I will argue that none of these proposals is plausible as an interpretation of probabilistic

2 The emphasis here is on *fundamental*. The future state of my coffee is certainly sensitive to its present temperature, although temperature does not figure in the fundamental laws. This is because temperature is realized by more fundamental physical properties. If the coffee's temperature had been different, then its micro-state would also have been different, which would have made a difference to the coffee's future.

theories in science. They all make the same mistake: they assume that probabilistic theories are in the business of making straightforward, apodictic claims about the world. I will suggest that this is not the right way to understand probabilistic theories. The probability claims in scientific theories are not meant to be true or false, and thus do not need an interpretation. The idea may sound radical and revisionary, but it turns out to be quite ecumenical. We will see that the frequency account, the best-systems account and the epistemic account all have a natural place in the resulting picture.

2 Troubles

Let me begin with a somewhat curious (and so far overlooked) problem for the best-systems account. The problem is easiest to see in the application for which the account was originally developed: stochastic dynamical theories in fundamental physics.

The leading version of quantum physics that postulates a stochastic dynamics is the GRW theory of [Ghirardi et al. 1986] and [Pearle 1989]. GRW assigns probabilities to certain transitions between states of physical systems, as represented by their wavefunction. The details don't matter for our purpose. For concreteness, let's imagine GRW directly postulates an 0.5 probability for tritium atoms to decay within 12 years. What does that mean?

The best-systems account defines physical probability (what Lewis calls "chance") indirectly via scientific theories. Imagine a list of all possible physical theories, understood as deductively closed systems of sentences in a suitable language. Given full information about the world, each such system can be evaluated for correctness, simplicity, strength, and other theoretical virtues. If a theory involves probabilities, we can also evaluate the extent to which the theory assigns high probability to actual events. Suppose some probabilistic theory T comes out best, on balance, by those criteria. The best-systems account now defines the true probability of an event as whatever probability this "best system" T assigns to the event.

Here is how Lewis puts it:

> Consider deductive systems that pertain not only to what happens in history, but also to what the chances are of various outcomes in various situations [...]. Require these systems to be true in what they say about history. We cannot yet require them to be true in what they say about chance, because we have yet to say what chance means; our systems are as yet not fully interpreted. [...] [S]ome systems will be simpler than others [...] some will be stronger than others: some will say either what will happen or what the chances will be when situations of a certain kind arise, whereas others will fall silent both about the outcomes and about the chances. And further, some will fit the actual course of history better than others. That is, the chance of that course

> of history will be higher according to some systems than according to others. [...] The virtues of simplicity, strength, and fit trade off. The best system is the system that gets the best balance of all three. [...] [T]he laws are those regularities that are theorems of the best system. But now some of the laws are probabilistic. So now we can analyse chance: the chances are what the probabilistic laws of the best system say they are. [Lewis 1994: 233f.]

On this approach, what is the truth-conditional content of a statement such as $P(\mathit{Decay}) = 0.5$? What does the statement say about the world? To be sure, it says that the probability of decay (i.e., of a tritium atom decaying within 12 years) is 0.5. Everyone agrees about that. What we want to know is what that means: does it mean that a fundamental physical measure of chance assigns value 0.5 to the decay event? Or does it mean that the relative frequency of tritium atoms decaying within 12 years is 0.5? According to the best-systems account, it means neither of these things. Rather, $P(\mathit{Decay}) = 0.5$ seems to mean the following:

(*) Whichever physical theory best combines the virtues of simplicity, strength, fit, etc. assigns probability 0.5 to tritium atoms decaying within 12 years.

The truth-conditional equivalence is straightforward. By the best-systems account, the objective chance of an event is defined as whatever probability the best system assigns to the event. So if the best system assigns probability 0.5 to an event – as (*) says – then the event's chance must be 0.5. Conversely, if the chance is 0.5, that must be the value the best system assigns to the event. So, if probability statements in physical theories are interpreted along the lines of the best-systems account, then our imagined law $P(\mathit{Decay}) = 0.5$ is analytically equivalent to (*). The problem is that (*) is not the kind of proposition I would expect to find in the basic laws of physics.[3]

3 I have assumed that theories directly specify probabilities to events. Lewis [1980] instead suggests that stochastic dynamical theories specify "history-to-chance conditionals", saying that *if* H is the history of the world up to time t, *then* event A at t' has probability x. Stochastic dynamical theories generally do assign probabilities to future states relative to past or present states, but arguably these "conditionals" are better understood as a kind of conditional probability. Nevertheless, suppose we follow Lewis. The hypothesis that a particular tritium atom has an 0.5 chance of decaying within the next 12 years then comes out as equivalent not to (*) but to something like (**):

(**) The true history of the world up to now is such that whichever physical theory best combines the virtues of simplicity, strength, fit, etc. specifies that if the history up to now is just like that then there is an 0.5 probability that the tritium atom will decay within the next 12 years.

If we apply this interpretation to the probability statements in theories like GRW, which on Lewis's assumption have the form $H \to P(\mathit{Decay}) = 0.5$, they come out as equivalent to $H \to (**)$. The problems I am about to raise carry over, *mutatis mutandis*, to this interpretation. In addition, there is at least one further problem: if the best system is anything like GRW, it plausibly implies conditionals

Why not? One reason is that I expect the basic laws of physics to specify relations between fundamental physical quantities. That is what's wrong, for example, with the Copenhagen interpretation of quantum mechanics according to which the basic laws attribute a special role to measurements: *measurement* is a gerrymandered, anthropocentric, and not at all fundamental physical kind. The same is true for *probability* as interpreted by the best-systems account. The theoretical virtues that go into the definition of a best system are not part of fundamental physical reality. Indeed, proponents of the best-systems account often emphasize the anthropocentric character of the interpretation, the fact that it reflects our contingent epistemic perspective. Moreover, there are many ways of spelling out the virtues, and of balancing them against each other. It is hard to believe that one of these ways is somehow objectively privileged. On the best-systems interpretation, the precise content of the GRW laws would depend on arbitrary choices in the ranking of theories.

The problem here is not that on the best-systems account, what counts as a (probabilistic or non-probabilistic) physical law might depend on somewhat arbitrary and anthropocentric facts. That is true, and widely accepted by advocates of the best-systems account. The present problem is that the *content* of probabilistic laws now involves gerrymandered, anthropocentric notions.

A second point that worries me about fundamental laws like (*) is that I expect the basic laws of physics to be explanatory bedrock. Why do opposite charges attract? Perhaps there is no deeper scientific explanation. That's just how things are. By contrast, if the basic laws say that $P(\mathit{Decay}) = 0.5$ and what this means is that the best system assigns probability 0.5 to *Decay*, then that is clearly not a basic fact. It is explained by patterns of occurrent events in the history of the world together with the relevant standards for evaluating theories.

Again, the problem should not be conflated with a superficial similar but different problem: that the best-systems account makes the laws depend on occurrent facts, while many people intuit that the dependence goes the other way. We must distinguish the claim *that p is a law* from the simpler claim *that p*. On the best-systems account, *that it is a law that opposite charges attract* is made true by patterns in occurrent events. But the simpler claim *that opposite charges attract* is not a statement about laws; it may well be explanatory bedrock. The problem is that this can no longer be said for probabilistic claims such as $P(\mathit{Decay}) = 0.5$. If that claim is analysed as (*) then it clearly has an underlying explanation.

This brings me to a third worry: the best-systems account threatens to collapse the important difference just mentioned between the claim that something is merely true and

of the form $H \to P(\mathit{Decay}) = 0.5$ for various non-actual histories H such that the best system for H-worlds does not imply $H \to P(\mathit{Decay}) = 0.5$. The conditional $H \to (**)$ is then false, yet the laws of nature are supposed to be true!

the claim that it is nomologically necessary. Newton's second law, for example, says that $F = ma$, not that it is nomologically necessary that $F = ma$. However, if we interpret $P(Decay) = 0.5$ as (*) then the statement can't be true without also being part of the best system and hence a law (on the best-systems account of laws).

All these problems arise because the best-systems account of physical probability has implications for the *content* of probabilistic laws.[4] By comparison, the problems do not arise for the best-systems account of laws because physical theories generally do not involve the term 'law'; whatever we say about the meaning of 'law' therefore can't have implausible consequences about the content of physical laws. Probabilistic laws in science evidently do contain the term 'probability', or 'P'. Our starting point was the question what that term means: what does statistical mechanics say about the world when it says that $P(S) = \mu(S)/\mu(\Gamma)$? What does GRW say when it assigns such-and-such probability to transitions between physical states? The best-systems account suggests an answer: it suggests that the probability statements in GRW mean that whichever theory best combines such-and-such virtues assigns such-and-such probability to the relevant transitions. But that, I have argued, is problematic.

Analogous problems arise for the epistemic interpretation of physical probability. Suppose, as before, that $P(Decay) = 0.5$ is (an instance of) a fundamental physical law. On a simple-minded Bayesian interpretation, the law would state that some not-further-specified individual assigns subjective degree of belief 0.5 to the decay event. That is clearly absurd. The laws of nuclear decay are not merely statements about what some person happens to believe; they can be true even if no-one has the relevant degrees of belief. More sophisticated epistemic accounts interpret probability statements as statements about what it would be rational to believe. $P(Decay) = 0.5$ would mean that one should assign degree of belief 0.5 to *Decay*, or something along these lines. But that is still incredible. *Normative psychological* notions should not figure in the fundamental laws of physics! As above, the relevant propositions also do not seem to be explanatory bedrock. If it is rational to have degree of belief 0.5 in certain events, and this is an epistemically contingent fact about the world (as physical laws are supposed to be), then surely there must be a further explanation of why that degree of belief would be adequate.

One might respond that the epistemic account or the best-systems account should not be understood as spelling out the truth-conditions of probability statements in physical theories. Fair enough. Indeed, I will argue below that both accounts are essentially correct when it comes to interpreting discourse about physical probability in ordinary English.

4 By 'content', I mean intensional, truth-conditional content. One can individuate content more finely so that analytic equivalence does not imply sameness of content. In that sense, '$P(Decay) = 0.5$' and (*) plausibly differ in content. The problems I have raised all concern the coarse-grained, truth-conditional content of probabilistic laws.

However, the question I have raised is precisely a question about the truth-conditions of probability statements in physical (and non-physical) theories. I want to know what these theories say about the world. Changing the topic is not an answer.

So much the worse, you might say, for reductionist accounts of physical probability. Believers in primitive chance have an attractive alternative: $P(Decay) = 0.5$ simply means that there is an 0.5 chance of tritium atoms decaying within 12 years, and that can't be analyzed in terms of anything else. However, this interpretation, too, faces serious problems – for example, when it comes to explaining the link between physical probability and rational degree of belief (see [Lewis 1994]). Moreover, recall that the primitive chance account is at best applicable to a very narrow range of scientific theories: stochastic dynamical theories in fundamental physics. It says nothing about the probabilities in Bohmian mechanics, statistical mechanics, chemical kinetics or systems biology.

What about a mixed approach then: GRW talks about primitive chance, while the other theories talk about best-systems probabilities or epistemic probabilities? I agree that we should not take for granted that a unified interpretation can be given for probability statements in all areas of science. However, most of the problems I just raised for the best-systems interpretation and the epistemic interpretation are not specific to GRW; similar problems arise for, say, probabilistic models in genetics. In addition, the mixed approach would face all the problems of postulating primitive chances.

What else could we try? One might suggest that the best-systems account or the epistemic account should not be understood as *analyses* of physical probability, but rather as “fixing the reference”: they identify physical probability by a certain role, without revealing the nature of the quantity that occupies that role. This kind of story is familiar and plausible for other theoretical terms. Perhaps our concept of inertial mass can be analyzed in terms that we don't expect to find in the laws of fundamental physics. The analysis might identify inertial mass by its role in our experience of the world – roughly, as the property responsible for the fact that we find some things harder to accelerate than others. The role is realized by a fundamental physical quantity (as it turns out, by the very same quantity that also plays the role associated with the distinct concept of gravitational mass). The content of Newton's second law is arguably a proposition directly about that quantity. Unfortunately, we can't assume that probability terms similarly refer to fundamental physical quantities. That is at most plausible for the probabilities in GRW, and even here it is highly controversial. We could alternatively take the referent to be a non-fundamantal quantity such as $\mu(S)/\mu(\Gamma)$, but that would turn the probability statements in the relevant theories into empirically empty tautologies.

I do not claim to have refuted any – let alone all – candidate interpretation of probabilistic theories. Most advocates of the best-systems account, in particular, have come to accept that the account has counterintuitive consequences, so they might accept the problems I have raised as further bullets that have to be bitten. Nonetheless, I hope

I have said enough to motivate trying something new.

3 Theories without truth

I began with a question: what do probabilistic theories in science say about the world? What would a world have to be like for it to be true that tritium atoms have a 50 percent probability of decaying within 12 years? I want to suggest that we should reject the question. Probability statements in scientific theories do not express a special kind of fact. They are not meant to be true or false.

The idea is that we broaden our conception of scientific theories. On the traditional realist conception, scientific theories aim to register important truths about the world – interesting and robust patterns in the observable phenomena and in what lies behind these phenomena. The task is straightforward if the relevant patterns are crisp: all Fs are Gs, whenever a system is in state S_1 it will later be in S_2, whenever a phenotype has frequency x in one generation it has frequency y in the next generation. But what if the world is more complex? What if two quantities F and G are strongly and robustly correlated, but the value of G on any given occasion is not completely determined by the value of F, nor is there a simple formula for how G is determined by F together with other salient features of the situation? We could simply refrain from saying anything about the connection between the quantities. But then we might fail to capture an important fact about the world. What is a scientist supposed to do if she notes (or suspects) an interesting, robust, but *noisy* relationship between two quantities? How can she express such a relationship in a scientific theory?

This is where probability enters the picture. Let's allow our scientist to specify a probabilistic relationship between F and G, perhaps by adding a noise term to an algebraic equation. The result is a probabilistic model or theory. The point of the model is to capture the noisy, stochastic relationship between F and G. It is not to capture a crisp relationship between F, G, and third quantity P. This is why we could not find a sensible answer when we asked what that quantity might be: primitive propensity, best-systems probability, rational credence, or what have you. All these answers misunderstand the point of probabilistic models.

When a scientist puts forward a probabilistic model, she commits herself to the assumption that the model fares well, on balance, in terms of simplicity, strength, fit and other relevant virtues. But this is not the *content* of her model. Her model doesn't say of itself that it maximizes theoretical virtues, or that it captures noisy relationships in the world. In order to serve its purpose, it is enough that the model contains a probability function. The function does not need an interpretation.

Consider a toy example. Our object of study is a series of events with two kinds of outcome, call them "heads" and "tails". (If you want, imagine that this is all there is in the universe. At any rate, nothing else false in scope of our inquiry.) There are a million outcomes in total, arriving in seemingly random order with heads having a stable relative frequency of around 0.8. How could we model this noisy pattern? We could simply list all individual outcomes in the order in which they arrive. But such a list would be unwieldy, and it wouldn't reveal any patterns in the data; among other things, it would give us little insight into how the series might have continued beyond a million outcomes. For many purposes, it might therefore be useful to put forward a probabilistic model. Specifically, we could put forward a model that assigns probability 0.8 to heads on each toss, independent of the other outcomes. The model's probability for heads and tails then closely matches their relative frequency, but the probabilities are not meant to stand for relative frequencies. Indeed, by treating the events as independent the model assigns positive probability to many sequences of heads and tails (such as 1000 tails in a row) that never occur in the series at all. Nor are the probabilities meant to stand for fundamental propensities. The events in question may or may not be generated by an underlying deterministic mechanism; the usefulness of our model doesn't depend on that. Nor are the probabilities meant to stand for rational credence, or anything else. The point of our probabilistic model is, as I said, to capture a noisy pattern in the world.

To a first approximation, we can spell out what that means by following the best-systems analysis – but without assigning an interpretation to the probabilities. Imagine all possible ways of assigning probabilities to the members of our series. These are our "theories". Some of them are simpler than others. A theory that assigns probability 1 to every actual outcome and thus effectively lists the entire series is not very simple; a theory that treats the outcomes as independent is (other things equal) simpler than a theory that doesn't. And so on. We can also compare our theories in terms of strength. A theory that assigns probabilities only to individual outcomes is (other things equal) weaker than a theory that also assigns probabilities to sequences of outcomes. And we can compare our theories in terms of probabilistic fit. For theories that assign a probability to the entire sequence, we can follow Lewis and use that as a measure of fit. Finally, then, what it means for a theory to *capture the patterns* in our sequence is that the theory fares comparatively well, on balance, in terms of simplicity, strength, and fit.

If the aim of scientific theories is to capture possibly noisy patterns in the world, we don't need to interpret the theories' probabilities. In fact, interpreting the probabilities would get in the way of this aim, since it is unclear how a theory which states a crisp relationship between three quantities F, G, and P is supposed to capture a non-crisp pattern in the relationship between F and G alone.

Admittedly, the present view of scientific theories may be unfamiliar and therefore somewhat counterintuitive. We are used to thinking that respectable scientific theories explicitly represent the world as being a certain way, for example (as I said above) by stating relations between fundamental quantities. On the present account, this is not quite true for probabilistic statements in scientific theories. If a theory "states" a probabilistic connection between fundamental quantities, it doesn't really state anything, insofar as it does not make an apodictic, outright claim about the world.

As an analogy, it may help to imagine scientific theories as agents ("experts") with certain beliefs about the world. On the traditional conception of theories, the expert only has binary beliefs: she believes that all Fs are Gs, that whenever quantity A has value x, then B has value y, and so on. Now we also allow partial beliefs. The expert can be more or less confident that something is G given that it is F, or that B has value y if A has value x. The expert can be 80 percent confident that the first outcome in a series is heads. Such partial beliefs are not outright beliefs with a special probabilistic content. To believe something to a given degree is not to have a full belief about a physical quantity, or about one's own state of mind. As a consequence, a system of partial beliefs is in the first place not true or false, but more or less *close* to the truth. A good expert generally assigns high degree of belief to true propositions and low degree of belief to false ones. A range of "accuracy measures" have been proposed to render this kind of distance to the truth precise (see e.g. [Joyce 1998]). Such measures can be applied not only to probability functions that represent degrees of belief but also to uninterpreted probability functions in scientific theories, where they offer a natural approach to measuring fit. Like a good expert, a good theory should generally assign high probability to true propositions and low probability to false ones.[5]

One might think that an ideal expert assigns degree of belief 1 to every truth and degree of belief 0 to every falsehood. Accordingly, an ideal theory would have no need to involve probabilities. But a complete theory of all truths is not only beyond our reach, it is also not what we seek in scientific theories. Science is looking for *patterns* in the total history of the world, for simple yet powerful principles that allow predicting a wide variety of facts. If these patterns are suitably noisy, even an ideal theory will be probabilistic.

So that's my proposal. The probabilities in scientific theories do not have an interpretation. As a consequence, probabilistic theories cannot be true or false, except in their non-probabilistic parts. They can still be more or less simple, more or less unified, and more or less close to the truth, as measured by the difference between the (uninterpreted) probabilities and the actual events in the world. And that is all we need. The point of probabilistic models in science is to provide a simple and informative systematization of noisy patterns in the world. To serve that purpose, the models do not need to be assessable for truth and falsity.

5 Of course, an agent's subjective probability function is not "uninterpreted": it represents the agent's degrees of belief. The point of the analogy is that partial beliefs, like the probabilities in scientific theories, do not simply represent the world as being one way or another.

4 Capturing patterns

Let me say a little more on how we may understand the goal of "capturing noisy patterns". Above I explicated this notion by following the best-systems account: I suggested that a probabilistic theory captures a noisy pattern in certain events just in case it scores best, on balance, in terms of simplicity, strength, and fit among all possible theories of the relevant events. It is crucial for my proposal that a theory can do that without having truth-conditional content.

Intuitively, one might think that a theory's *strength* is to be measured by how many possibilities it rules out. This would plausibly require that the theory has truth-conditional content. But while the suggested notion of strength may be useful for certain applications (assuming one can find a sensible way of counting possibilities), there are independent reasons why it is not adequate in the context of either the best-systems account or my own proposal. In particular, we here need a measure of strength that is relative to a history of events. For example, consider a simple statement of the form *all Fs are Gs*. In worlds where there are many Fs, all of which are Gs, this may be a good candidate for an axiom in the best system. Not so in worlds where there are no Fs at all. The statement is equally true, and equally simple, in both kinds of worlds, but it provides much more valuable information in the first. All else equal, the relative strength of *all Fs are Gs* in a given world should therefore be greater the more Fs and the fewer Gs there are in the world. The same criterion could be used for statements of the form *The probability of an F being G is x*.

Note that the best-systems account, too, assumes that one can evaluate theories for simplicity, strength, and fit without assigning an interpretation to the probabilities (see the above quote from [Lewis 1994: 234]). Lewis suggests to measure simplicity by syntactic complexity, strength by the variety of circumstances and outcomes for which a theory specifies probabilities, and fit as the probability a theory assigns to the entire history of the world. Lewis does not defend these criteria. Our observation that strength should be world-relative indicates that his criterion for strength is inadequate. His criterion for fit also runs into well-known problems in cases where theories assign either no probability or probability zero to the entire history of the world. [Elga 2004] suggests an alternative characterization of fit in terms of typicality; in [Schwarz 2014] I suggest a measure of fit that aggregates the differences between actual frequencies and theory-expected frequencies.

In general, it is fair to say that nobody has yet put forward fully satisfactory and precise criteria for simplicity, strength, and fit, and for how these are meant to trade off against each other. For the present proposal, this is less of an embarrassment than it is for the best-systems account itself. On the account I have put forward, the standards of simplicity, strength, and fit are only used to clarify the scientific aim of capturing patterns. This aim does not have to be absolutely precise and objective. We can allow that what scientists value in their theories is to some extent imprecise and varies from discipline to discipline, from school to school, or even from person to person.

I will not go further into how one might spell out the relevant notions of simplicity, strength, and fit. I do, however, want to highlight a further kind of goal that is often ignored in the literature on best systems.

Real scientific theories typically aim for more than a compact statistical summary of certain events. They try to shed light not only on how the events are distributed, but also on why they are distributed the way they are. Accordingly, the probabilities in scientific theories are generally motivated by underlying explanatory assumptions, often about how the relevant events come about. The binomial probabilities in the Wright-Fisher model of neutral evolution, for example, are not based on inductive generalizations from observed frequencies. Rather, they are motivated and explained by internal assumptions of the model.

A popular and powerful tool for motivating probabilities is the "method of arbitrary functions" (see e.g. [von Plato 1983]). Paradigm applications of the method are gambling devices such as roulette wheels or dice. These devices are built in such a way that any reasonably smooth frequency distribution over initial conditions is mapped by the dynamics of the system to approximately the same distribution over outcomes. The characteristic patterns in the observed outcomes can therefore be explained by the absence of very unusual patterns in the input conditions. Several authors have recently suggested that considerations along similar lines can also justify the probabilities in statistical mechanics and other scientific theories (see e.g. [Strevens 2003], [Myrvold 2012]).

What's important for our present topic is not so much how this or that probabilistic model can be justified, but the more general fact that we expect the probabilities in a model to have some such underlying justification. Among other things, this explains why we tend to hold fixed the adequacy of our models under counterfactual suppositions: on the supposition that a fair coin were tossed a million times, we expect the relative frequency of heads to be approximately 1/2.

The method of arbitrary functions, the ergodic theorem, and other popular ways of justifying probabilistic models explain why a model can be expected to have good probabilistic fit, but they do not provide an interpretation of the model's probabilities. Consequently, these explanations are often supplemented with an epistemic or frequentist interpretation of probability – leading to the usual problems for these interpretations. On the present approach, no supplementation is called for.

5 Theories, predictions, beliefs

At first glance, my proposal seems to create a host of problems. If probability statements don't have truth-conditional content, how can they be believed, disbelieved or conjectured? How can they be confirmed or disconfirmed by observation? How do we interpret complex sentences that embed statements about probability?

In response, I should first stress that my proposal does not concern the interpretation of probability statements in ordinary language. My topic is the interpretation of scientific models or theories. Arguably, such models are best understood not as linguistic constructions at all. If they are expressed in language, the relevant language will generally involve both terms from natural language and special-purpose technical vocabulary. On my proposal, probability terms should be treated as technical terms, and they should not be given an interpretation. I will say a little more on the interpretation of 'probability' in ordinary English below, but that is not the focus of my proposal.

So the problem with complex sentences only arises for complex sentences within a given scientific theory. That is, what if instead of assigning an outright probability to an event A, a theory merely states that the probability of A is *either x or y*? Or what if a theory says that *if H, then $P(A)=x$* (see note 3 above)?

On the present account probabilistic theories do not have a classical truth-conditional interpretation. They only need to be evaluated for simplicity, strength, probabilistic fit and other theoretical virtues. So we need to ask, for example, how to measure a theory's fit with respect to actual events in the world if the theory merely specifies that the probability of A is either x or y. This might be an interesting question to ponder, but it is not a terribly urgent question, since real theories rarely take that form.[6]

The issue of confirmation and belief is more serious. Fortunately, there is a simple and natural answer. Suppose a scientist proposes or endorses a probabilistic theory T. On the account I suggested, she thereby commits herself to the hypothesis that T provides a good systematization of the relevant patterns in the world. So the scientist commits herself to the truth not of T itself, but of a derivative proposition $\Box T$: that T fares well in terms of simplicity, strength, fit and other theoretical virtues. Unlike T, $\Box T$ is an ordinary (albeit vague) proposition. It can be true or false. It can be believed, disbelieved, conjectured, and denied. It can be confirmed and disconfirmed by empirical observations.

So what appear to be propositional attitudes towards a probabilistic theory T are really attitudes towards an associated proposition $\Box T$ – roughly, the proposition that T provides the best systematization of the relevant patterns in the world.

6 To the extent that there is a problem here, it is equally a problem for the best-systems account, which also assumes that one can evaluate theories for probabilistic fit without yet assigning a meaning to the probability terms.

The "relevant patterns" are not just patterns in the phenomena. To be sure, a scientist might only half-heartedly and instrumentally "accept" a theory, confident that it captures interesting patterns in past and future observations, but agnostic about whether the entities it postulates are real and whether they display the patterns suggested by the theory. In contrast, to really endorse (say) GRW quantum mechanics, you have to believe (roughly) that the true state of an isolated physical system is accurately and completely characterized by its wavefunction, that the state mostly evolves in accordance with the Schrödinger equation, but that this evolution is occasionally punctured by collapse events whose frequency and outcome displays statistical regularities to which the probabilities in GRW are a good approximation. This (roughly) is the content of $\Box$GRW. It goes far beyond the hypothesis that GRW is a useful tool for predicting measurement outcomes.[7]

In general, $\Box T$ is closely related to propositions about disorder and relative frequency. Return to the toy example from section 3. Let T be the so far unnamed theory that assigns probability 0.8 to heads on each toss. T itself can't be true or false, but $\Box T$ can. What does $\Box T$ entail about the sequence of outcomes? It obviously depends on the precise meaning of the box. For concreteness, let's assume that $\Box T$ states that T is the best systematization of the sequence as measured by Lewis's [1994] criteria of simplicity, strength, and fit – setting aside the worries raised in the previous section. $\Box T$ then entails that about 80 percent of the tosses actually come up heads. For suppose the actual frequency is only 70 percent. Then T provides a significantly worse systematization of the sequence than a rival theory T' that assigns probability 0.7 to heads on each toss. Specifically, T' has much greater fit – the probability of 70 percent heads is approximately 8.7×10^{-4} according to T', but 8.4×10^{-12237} according to T – while it fares equally well in terms of simplicity and strength. $\Box T$ also entails that the sequence of outcomes does not have any conspicuous patterns. For example, it can't be 200000 heads followed by 800000 tails, or 200000 repetitions of HHHHT; in either case, it would be easy to specify the exact sequence, so a good systematization of the outcomes would not resort to probabilities at all. Finally, $\Box T$ plausibly entails that right after a heads outcome, the relative frequency of another heads is not too far from 80 percent; otherwise a theory that doesn't treat successive tosses as independent would have greater fit without too much a cost in simplicity.

So there is a tight connection between probabilistic theories and claims about relative frequency and disorder. If a scientist accepts our coin model T, she will expect an irregular sequence with about 80 percent heads and 20 percent tails. If the sequence turns out to be more regular or the frequencies different, the scientist will have to revise her attitudes towards T. It is therefore understandable that many science textbooks

7 So there is still an important contrast between scientific realism and anti-realism. It's just that what is at issue is strictly speaking not the truth (or truth in certain respects) of our best theories T, but the truth (etc.) of the associated propositions $\Box T$.

endorse some form of the frequency interpretation on which probability claims simply are claims about relative frequency.[8]

We can also see what is right about the epistemic interpretation. On the supposition that a theory T provides a good systematization of the relevant patterns in the world (i.e., on the supposition that $\Box T$ is true), a rational agent should generally align her credence with the theory's probabilities. To illustrate, suppose you know that the best systematization of our coin toss sequence is the theory T that treats the tosses as independent with a fixed probability 0.8 of heads. As we saw, this entails that the sequence is irregular with about 80 percent heads and 20 percent tails. Now consider, say, toss number 512. How confident should you be that this particular toss results in heads? In the absence of further relevant information, surely your credence should be about 0.8. Moreover, your credence should be fairly insensitive to information about other outcomes. For example, conditional on the assumption that toss number 511 lands tails, your degree of belief in heads on toss number 512 should still be about 0.8. (See [Schwarz 2014] for more details and generalizations of these observations.)

How could you have come to know $\Box T$, without having surveyed the entire sequence? The short answer is: by induction. Perhaps you have witnessed the first 10000 tosses, and found an irregular pattern of heads and tails with about 80 percent heads. All else equal, you would then be justified to assume that the same noisy regularities obtain in the unobserved parts of the sequence.[9] Remember also that real scientific theories typically aim for more than a mere summary of actual events. If we know the dynamics of roulette wheels, the method of arbitrary functions explains why it is reasonable to believe that a given probabilistic model captures the actual pattern of outcomes even without any direct information about those outcomes.

6 Conclusion

I have argued that none of the currently popular interpretations of probability yields an adequate understanding of probabilistic theories in science. These interpretations all assume that probability claims in science are claims about a particular probabilistic quantity, but it is hard to see what that quantity could be. I have suggested that we

8 The connection between scientific probability and relative frequency is sometimes presented as a consequence of the laws of large numbers. But it is not. The connection is a substantive fact that any account of scientific probability should explain. The laws of large numbers are theorems of the probability calculus that hold for every probability function whatsoever, no matter how out of touch it is with the frequencies in the world.

9 As Hume pointed out, there is no logical guarantee that the patterns will continue, and thus no logical refutation of an inductive skeptic who claims that they won't. In the literature on physical probability, it is sometimes assumed that in order to explain the connection between probability and rational belief, one must also refute the inductive skeptic (see e.g. [Salmon 1967], [Strevens 1999]). Unsurprisingly, the verdict then is that the task is impossible.

should stop looking for a candidate. The point of probabilistic theories is not to capture facts about some probabilistic quantity out in the world, but rather to capture noisy relationships between ordinary, non-probabilistic quantities.

On the resulting picture, probabilistic theories cannot be true or false, except in their non-probabilistic parts, but they can still be evaluated for simplicity, strength, fit and other theoretical virtues. To capture a noisy pattern in the world means to score (comparatively) high in terms of these virtues. If a theory captures a noisy pattern, we could say that it represents or predicts that pattern. In that sense, probabilistic theories do have representational content, even though their probability functions do not have an interpretation. The theory represents a pattern without stating that the pattern exists.

To fully accept a theory is to regard it as a good systematization of the relevant facts. Under normal conditions, this implies expecting a close fit between the theory's probabilities and actual (as well as counterfactual) frequencies. It also implies adopting the theory's probabilities as one's own degrees of belief.

So the best-systems interpretation, the frequency interpretation, and the epistemic interpretation are not entirely off the mark. They all misrepresent the content of probabilistic theories, but they capture important aspects of what a rational agent must believe who accepts a probabilistic theory. Probabilistic laws do not say of themselves that they have various theoretical virtues, but accepting the laws plausibly involves believing that they do.

What about probability statements in ordinary language? Officially my proposal is neutral on this question. I have some sympathy for the view (defended e.g. in [Maher 2010]) that most ordinary statements about probability are normative epistemic statements. That is, by saying that there is a 90 percent probability of rain, I would typically express or recommend a corresponding degree of belief. Since accepting a theory goes hand in hand with taking the corresponding degrees of beliefs to be rationally adequate, we can see how the ordinary sense of 'probability' relates to the probabilities in scientific theories, and thus why the latter are called 'probabilities'.[10]

Finally, what about the probabilities of statistics, as they figure for example in parameter estimation from noisy data? Again, the proposal I have made does not speak directly to this question. In principle, it is compatible with both Bayesian and frequentist accounts. However, it might offer a new perspective on the interpretation of "likelihoods": the probability of data given some hypothesis. These likelihoods are often (on frequentist accounts, always) derived from general models of the experimental setup. My proposal straightforwardly applies to those models. It suggests that such model-based likelihoods are neither degrees of beliefs nor frequencies. Just as frequentists insist, they track objective features of the world. But they are also closely related to rational degrees of

10 Evidently, these remarks are not even a sketch of an actual semantics/pragmatics for probability statements in natural language; the details are hard (see e.g. [Lassiter 2011], [Rothschild 2012]).

belief, for the reasons I have reviewed. Specifically, on the assumption $\Box M$ that some statistical model M captures the noisy relationship between hypotheses and experimental data, it is generally rational to set one's credence in data E given hypothesis H to equal the probability that M assigns to E given H. Bayesian reasoning therefore goes through essentially as before.

References

David Albert [2000]: *Time and Chance.* Cambridge (Mass.): Harvard University Press

Jonathan Cohen and Craig Callender [2009]: "A Better Best System Account of Lawhood". *Philosophical Studies*, 145(1): 1–34

Adam Elga [2004]: "Infinitesimal Chances and the Laws of Nature". In Frank Jackson and Graham Priest (Eds.) *Lewisian Themes: The Philosophy of David K. Lewis,* Oxford: Oxford University Press, 68–77

G. Ghirardi, A. Rimini and T. Weber [1986]: "Unified dynamics for micro and macro systems". *Physical Review D*, 34: 470–491

Alan Hájek [1997]: ""Mises redux" – redux: Fifteen arguments against finite frequentism". *Erkenntnis*, 45(2-3): 209–227

— [2009]: "Fifteen Arguments Against Hypothetical Frequentism". *Erkenntnis*, 70(2): 211–235

Carl Hoefer [2007]: "The Third Way on Objective Probability: A Skeptic's Guide to Objective Chance". *Mind*, 116: 549–596

Edwin T Jaynes [1957]: "Information Theory and Statistical Mechanics". *Physical Review*, 106(4): 620–630

James Joyce [1998]: "A Nonpragmantic Vindication of Probabilism". *Philosophy of Science*, 65: 575–603

Daniel Lassiter [2011]: "Gradable epistemic modals, probability, and scale structure". In *Proceedings of SALT,* vol 20. 197–215

David Lewis [1980]: "A Subjectivist's Guide to Objective Chance". In Richard Jeffrey (Ed.), *Studies in Inductive Logic and Probability* Vol. 2, University of California Press, Berkeley.

— [1986]: *Philosophical Papers II*. New York, Oxford: Oxford University Press

— [1994]: “Humean Supervenience Debugged”. *Mind*, 103: 473–490

Barry Loewer [2004]: “David Lewis's Humean Theory of Objective Chance”. *Philosophy of Science*, 71: 1115–1125

— [2007]: “Counterfactuals and the Second Law”. In H. Price and R. Corry (Eds.) *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited,* New York: Oxford University Press, 293–326

Patrick Maher [2010]: “Explication of Inductive Probability”. *Journal of Philosophical Logic*, 29: 593–616

Wayne C. Myrvold [2012]: “Probabilities in Statistical Mechanics: What are they?” URL http://philsci-archive.pitt.edu/9236/

P. Pearle [1989]: “Combining stochastic dynamical state-vector reduction with spontaneous localization”. *Physical Review*, 39A: 2277–2289

Daniel Rothschild [2012]: “Expressing Credences”. In *Proceedings of the Aristotelian Society,* vol 112. 99–114

Wesley Salmon [1967]: *The Foundation of Scientific Inference.* Pittsburgh: University of Pittsburgh Press

Markus A. Schrenk [2007]: “A Lewisian Theory for Special Science Laws”. In Helen Bohse and Sven Walter (Eds.) *Philosophie: Grundlagen und Anwendungen. Ausgewählte Beiträge aus den Sektionen der GAP 6,* Paderborn: Mentis

Wolfgang Schwarz [2014]: “Proving the Principal Principle”. In A. Wilson (Ed.) *Chance and Temporal Asymmetry,* Oxford: Oxford University Press, 81–99

Michael Strevens [1999]: “Objective Probability as a Guide to the World”. *Philosophical Studies*, 95: 243–275

— [2003]: *Bigger than Chaos.* Cambridge (Mass.): Harvard University Press

Jos Uffink [2011]: “Subjective probability and statistical physics”. In C. Beibart and S. Hartmann (Eds.) *Probabilities in physics,* Oxford: Oxford University Press, 25–50

Jan von Plato [1983]: “The method of arbitrary functions”. *British Journal for the Philosophy of Science*, 34: 37–47