No interpretation of probability*

Wolfgang Schwarz
Draft, 10 November 2013

1 Chance: believers and skeptics

Tritium atoms are unstable. About 50 percent decay into helium within 12 years of coming into existence. For all we know, there is no way to predict in advance when any particular atom will decay. All we can say is that the atom has a certain *probability* of decaying within a given period of time. But what does this mean?

Some hold that there is an objective type of probability, *chance*, built into the very fabric of the physical world. It is supposed to capture the tendency of physical systems to evolve in a particular way. On this view, the decay probabilities of tritium atoms are not a measure of our subjective uncertainty or ignorance. They are genuine physical quantities. Like mass and charge, they figure in causal explanations and physical laws.

Others are skeptical about primitive physical chance. For the most part, these skeptics do not dismiss all talk about chance as meaningless or false. Instead, a popular idea is that statements about chance express lack of information: to say that there's a 50 percent probability (or chance) of an atom decaying is to say that given the available information, it would be reasonable to give 50 percent credence to this outcome (where what counts as "available" depends on the conversational context). Thus chance is rational degree of belief conditional on (contextually determined) information about ordinary, non-chancy matters. Also popular among skeptics are deflationist "Humean" interpretations that reduce chance facts to facts about occurrent events in the world. Most simply, chances might be identified with relative frequencies. A more sophisticated proposal in the same spirit is the *best-systems analysis*, which identifies chance with the probability function that figures in whatever scientific theory best combines certain theoretical virtues.[1]

I am a skeptic. But I am also a scientific realist, and that creates a problem. The skeptical proposals just reviewed may be plausible as interpretations or regimentations

* Ancestors of this paper were presented to the "Metaphysische Verschörung" club in Mainz, the probability reading group at the ANU, and the research group on Causality xxx at the University of Cologne. I thank the audience on these occasions for helpful feedback.

1 As with all such divisions, the line between believers and skeptics is not perfectly clear. For example, Lewis [1980] and Schaffer [2003, 2007] should arguably be classified as believers although they endorse Humean accounts of chance. Paradigm believers are Popper [1982] and Gillies [2000]; paradigm skeptics are de Finetti [1974] and Jeffrey [1992]. For the present paper, it doesn't matter where exactly we draw the line.

of our ordinary conception of chance. But they are not at all plausible as interpretations of probabilistic theories in science. Let me explain.

2 Interpreting science

Probability is everywhere in contemporary science. Many of our most successful and powerful scientific theories are probabilistic – not just in quantum physics, but also in thermodynamics, chemical kinetics, systems biology, and evolutionary theory, among others. If we want to take seriously what science tells us about the world, we have to ask what these theories mean.

Consider Boltzmann-style statistical mechanics. Here the object of study are isolated physical systems consisting of a large number N of particles. The possible micro-states of such a system (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. The basic postulate of statistical mechanics now says that the probability with which the system's state lies in any (measurable) 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. In statistical mechanics, '$P(S) = x$' is not just shorthand for '$\mu(S)/\mu(\Gamma) = x$', which would turn the probabilistic predictions of statistical mechanics into trivial analytic truths. Statistical mechanics is an empirical theory. Its predictions explain real-life phenomena such as the melting of ice cubes and the diffusion of 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, $\mu(S)/\mu(\Gamma)$ is not quite right and that the standard measure μ should be replaced by another measure μ_p which gives zero 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 $P(S) = \mu_p(S)/\mu_p(\Gamma)$?

Believers in primitive chance might suggest that '$P(S) = \mu_p(S)/\mu_p(\Gamma)$' is a statement about primitive chance: it says that the relevant system has chance $\mu_p(S)/\mu_p(\Gamma)$ of being in state S. But this interpretation clashes with popular assumptions about primitive chance, which is why most believers in primitive chance actually reject it.

For example, I mentioned that chance is supposed to be a dynamic quantity, reflecting the tendencies of physical systems to evolve in one way or another. Probabilities in statistical mechanics lack this fundamental dynamical nature. My cup of coffee hardly has a tendency to be, right now, in one micro-state rather than another. Classical statistical mechanics also assumes a deterministic micro-dynamics, while non-trivial primitive chance is usually taken to be incompatible with determinism.

Let me mention some more arguments against the idea that statistical mechanics deals with primitive chance. First, even if the laws of quantum physics are probabilistic, the quantum physical probability that my cup is *right now* in a given micro-state is plausibly either zero or one. Yet statistical mechanics assigns non-trivial probability to hypotheses about the present micro-state. It is hard to see how a single, basic physical quantity could take several values for the same token event.

In response one might postulate a whole family of primitive physical probabilities. But this leads to a second problem: most of these probabilities would have to be epiphenomenal. The reason is that 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 simply isn't sensitive to the values of any further fundamental quantities.[2] More directly, 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 aren't fundamental.

Third, and perhaps most simply, the reason why physicists identify the probability of state S with $\mu_p(S)/\mu_p(\Gamma)$ is patently not that this matches the value of some basic physical quantity. The choice of $\mu_p(S)/\mu_p(\Gamma)$ is justified in part by its formal simplicity and in part by the fact that it fits the empirical phenomena.

Of course, this last point is true for probabilistic theories in all parts of science, including quantum mechanics – which is one major reason for skepticism about primitive chance. Consider the main realist version of quantum mechanics that postulates probabilistic laws: the GRW theory of [Ghirardi et al. 1986] and [Pearle 1989]. GRW assumes that the deterministic Schrödinger dynamics for physical systems is occasionally punctured by random "jumps", where the probability of a jump within a given time span is a function of the system's mass. This function is one among many equally real and objective candidates; it is chosen on grounds of theoretical simplicity and fit with empirical data.

In any case, even if the primitivist interpretation were plausible for probabilities in GRW, we can safely conclude that it does not work for statistical mechanics and other "high-level" probabilistic theories. So even for realists about chance, our question is still largely open: what do probabilistic theories in science say about the world?[3]

2 The emphasis here is on *fundamental*. The future state of my coffee if certainly sensitive to its present temperature (say), 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.

3 David Albert [2000] argues that the probabilities of Statistical Mechanics, and indeed of all higher-level

Skeptical accounts such as the best-systems analysis promise a unified interpretation of probabilistic theories. On the best-systems account, chance is defined as the probability function that figures in 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.[4] This is easily applied not only to fundamental theories in physics, but also to theories about the statistical behaviour of macroscopic objects.

Unfortunately, the resulting interpretation of probabilistic theories once again looks rather implausible. The problem I have in mind is especially striking for fundamental theories. Let’s pretend for the sake of concreteness that GRW states that tritium atoms have a 50 percent probability of decaying within 12 years. What is the content of this statement? What is required for the statement to be true? If ‘probability’ is interpreted along the lines of the best-systems account, the statement is true at a world w iff *whichever physical theory best combines the virtues of simplicity, strength and fit with respect to the events in w assigns probability 0.5 to tritium atoms decaying within 12 years.* This is not the kind of proposition I would expect to find in the basic laws of physics.

Why not? One reason is the somewhat diffuse intuition that the basic laws of physics should specify quantitative relations between fundamental physical properties. Consider the Born rule in Copenhagen-style quantum physics, which states that measurements have a special effect on physical systems. It is widely agreed that this is not a credible law because *measurement* is an gerrymandered, anthropocentric, and not at all fundamental physical kind. The same is true for probability terms if they are interpreted along the lines of the best-systems account.

Indeed, many advocates of the best-systems account accept that the concept of a *best system* is anthropocentric, reflecting our contingent epistemic perspective and interests. In addition, the concept is certainly vague. There are many ways of spelling out the virtues of simplicity, strength and fit, and many ways of balancing them against each other. It is hard to believe that one of these ways is somehow objectively privileged. But then our probabilistic laws would also be vague: their precise content would depend on an arbitrary choice of how to rank theories.

With respect to the basic laws of physics, a further problem is that we expect those lows to be explanatory bedrock. Why do opposite charges attract? Perhaps there is no

scientific theories, can be reduced to those of GRW – assuming GRW is the true theory of fundamental physics. Thus a friend of primitive chances might hold that all probabilistic theories in science are (more or less directly) concerned with the primitive chances of quantum physics. But even if we grant Albert’s controversial reductionism about special science, GRW is a highly speculative theory. What if it turns out that the true theory of fundamental physics looks quite different? 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.

4 See e.g. [Lewis 1994], [Loewer 2004], [Hoefer 2007], [Cohen and Callender 2009] for different versions of the account.

deeper scientific explanation. That's just how things are. By contrast, if the best system assigns probability x to A, this 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.

Moreover, consider theories that don't just make outright probability statements. Suppose a theory states what Lewis [1980] calls *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. On the present proposal, such a statement means that if H is the history of the world up to t, then the best theory assigns probability x to A at t'. But that can't be right. Assume H is the true history of the world up to t. By Modus Ponens, it should follow that the best theory assigns probability x to A at t'. But by assumption, the best theory does *not* assign outright probability to A at t'; it only specifies history-to-chance conditionals.[5]

Finally, the present proposal would collapse the important difference between the hypothesis that something is merely true and the hypothesis that it is nomologically necessary. Newton's second law, for example, says that $F = ma$, not that it is nomologically necessary that $F = ma$. But if we interpret $P(A) = x$ as claiming that the best system assigns probability x to A, 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).

One might respond that the best-systems analysis needn't be understood as spelling out the truth-conditions of probability statements. Perhaps it only "fixes the reference": it identifies chance by a certain role, without revealing the nature of the quantity that occupies the role. This is plausible for other theoretical terms. Consider inertial mass. Perhaps our concept of inertial mass can be analyzed in terms that we don't expect to find in the laws of physics, including terms for conscious experience (as suggested in [Chalmers 2012]). The analysis would identify inertial mass by its role in our experience of the world – roughly, as the property responsible for the fact that some things are harder to accelerate and slow down than others. This 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). Newton's second law can now be understood as a statement directly about that quantity.

But if we assume that probability terms similarly refer to a fundamental physical quantity, we are back to the previous proposal on which probabilistic theories are statements about a primitive probabilistic quantity. Alternatively, one might hold that

5 Does any serious scientific theory contain history-to-chance conditionals? Arguably not. Stochastic dynamical theories generally assign probabilities to future states relative to past (or present) states, but these "conditionals" are better understood as a special type of conditional probability. Versions of the present problem still arise whenever a theory contains suitably complex statements involving probabilities. I would feel uncomfortable if I had to declare from the armchair that this can never happen.

the referent are non-fundamantal quantities such as $\mu(S)/\mu(\Gamma)$, which would turn the relevant theories into empirically empty tautologies.

I conclude that the best-systems account fails to answer our question. It does not offer a plausible interpretation of probabilistic theories in science. It can't tell us what these theories say about the world.[6]

In section 1 I mentioned another skeptical proposal which holds that when a scientific theory assigns probability x to A, this means that it would be rational to have credence x in A. This runs into essentially the same problems as the best-systems interpretation. For example, the laws of physics would now involve *normative*, *psychological* terms, which is incredible. Moreover, such laws couldn't be explanatory bedrock: if it is rational to have credence x in A, and this is an epistemically contingent fact about the world (as the laws are supposed to be), then it must be due to some further fact about the world.

3 Science 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? What would a world have to be like for it to be true that a certain physical system has probability $\mu_p(S)/\mu_p(\Gamma)$ of being in state S?

I want to suggest that we should reject these questions. Probability statements in scientific theories are not meant to represent a special kind of fact. They are not meant to be true or false.

The idea is that we should broaden our conception of scientific theories. On the traditional realist conception, a good scientific theory registers important truths about the world – interesting and counterfactually robust patterns in the history of occurrent events. This is easy if the patterns are crisp: all Fs are Gs, whenever a system is in state S_1 it will later be in state S_2, whenever a phenotype has frequency x in generation 1 it has frequency y in generation 2. But what if the world is more complicated? For example, what if two quantities F and G are strongly and robustly correlated, but the

6 You may notice that I have focused on fundamental theories when arguing against the best-system interpretation and on non-fundamental theories when arguing against the primitive chance account. This suggests a mixed approach on which fundamental theories talk about primitive chance, and non-fundamental theories about best-system probabilities. But most of the problems for the best-system interpretation also arise for non-fundamental theories. For example, we could not allow for higher-level history-to-chance theories, we would collapse the distinction between p and *necessarily* p, we would have to accept that the laws of genetics (say) involve terms that are vague and anthropocentric. In addition, we would face at least some of the objections to the primitive chance interpretation, such as the problem of non-dynamical probabilities: as illustrated by Bohmian mechanics, there is no guarantee that the probabilities of fundamental physics are dynamic. In sum, the mixed approach does not seem to fare significantly better than the unified proposals. And of course it is not available to skeptics about primitive chance.

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 would 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 this 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 point of the resulting probabilistic model is to capture a noisy, stochastic pattern in the world. It is not to capture a crisp pattern involving F, G and third quantity P. This is why we could not find a sensible answer when we asked what this quantity might be: primitive chance, best-system probability, rational credence, or what have you. All these answers misunderstood the point of probabilistic models.

Of course, when a scientist puts forward a probabilistic model, she commits herself to the claim that it fares well, on balance, in terms of simplicity, informativeness 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 doesn't need an interpretation.

From this perspective, we can see that there is something deeply right about the best-system account. The account assumes that given full information about the world, one can compare and rank possible scientific theories without assigning an interpretation to their probability terms. Only the non-probabilistic parts of a theory can therefore be regarded as true or false. The probabilistic parts can only be evaluated by criteria such as simplicity, strength and fit, where, roughly, *simplicity* measures syntactic complexity, *strength* the number of relevant actual events in its scope, and *fit* the extent to which the theory assigns high probability to actual events. The best-systems account now goes on to define *chance* as the probability function that figures in whatever theory wins the competition. This step is obviously crucial if we seek an "interpretation of chance". But that is not our topic. Interesting scientific theories do not contain the term 'chance', and in any case I have argued that we should not try to find an interpretation for the probability terms in scientific models.

Admittedly, this perspective on scientific theories may be unfamiliar and therefore somewhat counter-intuitive. We are used to thinking that respectable scientific theories simply represent the world as being a certain way, for example (as I said above) by stating connections between fundamental quantities. On the present view, this is not quite true for probabilistic theories. If a theory "states" a probabilistic connection between fundamental quantities, it doesn't really state anything, in the sense that it doesn't 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, these beliefs are always apodictic: the expert beliefs 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. These partial beliefs are not outright beliefs with a special probabilistic content. Believing something to a given degree is not to have a full belief about a physical quantity, or about one's own state of mind. Accordingly, a system of partial beliefs is in the first place not true or false, but more or less *close* to the truth (more or less *accurate*): a good expert assigns high degree of belief to true propositions and low degree of belief to false ones. The same is true for probabilistic theories in science. A good theory scores well in terms of fit, which means that it assigns high probability to true propositions and low probability to false ones.

Before I move on, let me briefly address the stock objection to every proposal that some apparent statement lacks truth-conditions: if, say, '$P(A)=x$' doesn't have truth-conditional content, how do we interpret complex statements such as 'either $P(A)=x$ or $P(A)=y$', or (returning to an earlier example) 'if H, then $P(A)=x$'?

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 theories. Such theories can usually expressed in a language that contains some terms of natural language as well as special-purpose technical vocabulary. If the theory is probabilistic, my proposal entails that the probability terms should count as special-purpose technical vocabulary, and that they should not be construed as referring expressions. Now we can still ask about cases where a theory contains disjunctive or conditional statements involving probability. The question is how such a theory should be evaluated for simplicity, strength, fit and other theoretical virtues. That is the only "interpretation" probabilistic theories have, on my proposal. So we need to ask, for example, how to measure a theory's probabilistic fit if it specifies that $P(A)$ equals either x or y. This might be an interesting question to ponder, but it is not terribly pressing question, since real theories generally don't contain statements like these (see note 5 above). Note also that if there is a problem here, then it is equally a problem for the standard best-systems account, which also assumes that one can evaluate theories for simplicity, strength and fit without first assigning a meaning to the probability terms.

I am not going to spell out precise standards for measuring a theory's virtues. I doubt that there are such precise and universal standards. If one theory is a little simpler, another a little more informative, and a third a little better integrated into other parts of science, we don't need to assume that there is a fact of the matter about which of them is objectively best. What actual scientists value in their models varies from discipline to discipline, from school to school, and even from person to person. I see no reason to regard this as a flaw.

4 Theories, predictions, beliefs

I have argued that the point of probabilistic models in science is to provide a simple and informative systematization of noisy patterns in the world, and that to serve this purpose, such models don't need to be assessable for truth and falsity. At first glance, this may seem to create a host of problems. If probabilistic theories don't have truth-conditional content, how can they be believed, disbelieved or conjectured? How can they be confirmed or disconfirmed by observation?

Fortunately, there is a simple and natural answer to these worries. In fact, we will see that in contrast to many alternatives, the present account makes straightforward sense of how probabilistic hypotheses are tested in science. We will also see why both the best-systems interpretation and the epistemic interpretation come out as basically right for statements about chance in ordinary discourse.

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. T needn't be the absolutely *best* systematization, since the terms of comparison are vague, but at least it shouldn't be clearly worse than any alternative. So the scientist commits herself to the truth not of T itself, but of a derivative proposition $\Box T$ that T fares comparatively well in terms of simplicity, strength, fit and whatever other theoretical virtues are salient in the context. Unlike T, $\Box T$ is an ordinary (albeit vague) proposition. It can be true or false. It can be believed, disbelieved, conjectured or denied. It can be confirmed or disconfirmed by empirical observations.

So the simple answer to the above worries is that what appear to be propositional attitudes towards a probabilistic theory T are really attitudes towards an associated proposition $\Box T$. While $\Box T$ is not plausible as the content of T, it is quite plausible as something you're committed to if you endorse T.

It is important to be clear about the content of $\Box T$. First of all, $\Box T$ is not just the hypothesis that T provides a good systematization of our evidence, or of all evidence we could possibly gather. 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 isolated physical systems is accurately and completely characterized by their wavefunctions, and that these states evolve in accordance with the Schrödinger equation except for occasional and irregular jumps whose frequency and outcome displays statistical regularities to which the probabilities in GRW are a good approximation. This 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 randomness and relative frequency. Consider a toy example. Suppose a coin is tossed a million times, and let T be a theory that assigns probability 0.8 to heads on each toss, independent of the other outcomes. T itself can't be true or false, but $\Box T$ can. What does $\Box T$ entail about the sequence of outcomes, assuming, for concreteness, that we use Lewis's [1994] criteria of simplicity, strength and flesh out the meaning of the box. $\Box T$ 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 the alternative theory T' that assigns probability 0.7 to heads on each toss. Specifically, T' has much better fit (it assigns much higher probability to the actual outcomes[8]), 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. Plausibly, $\Box T$ also entails that right after a heads outcome, the relative frequency of another heads is again 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.

This is how probabilistic models are tied to randomness (irregularity) and relative frequencies. If a scientist endorses 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 attitude towards T.

Given these observations, we can see why many scientists sympathize with a frequency interpretation on which probability statements are analyzed as statements about relative frequencies in irregular sequences of outcomes. We know that this can't be quite right. Scientific practice clearly allows probabilities to deviate to some extent from actual frequencies, especially if the relevant conditions are rare. Turning to counterfactual frequencies only 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

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 $\Box T$.

8 The probability of 70% heads in a million tosses is approximately 8.7×10^{-4} according to T', but 8.4×10^{-12237} according to T.

study.[9]

So frequentism doesn't work. But this is not the end of the story. To echo a well-known passage from [Armstrong 1968: 68], whatever theory of scientific probability is true, it has a debt to pay, and a peace to be made, with frequentism. Any theory of scientific probability must explain the connection between probability, randomness and frequencies. The proposal I have offered can pay this debt.[10]

We can also vindicate another popular idea in science: that the probabilities in scientific theories are epistemic, representing appropriate degrees of belief given imperfect information. Return to the coin toss example. Suppose you know that the best systematization of the outcomes is our theory T which 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.

Generalizing, it is not hard to show that on the assumption that a given theory T provides the best systematization of relevant patterns in the world (i.e., given $\Box T$), a rational agent without unusual information should align her credence with the theory's probabilities (see [Schwarz forthcoming]).

Of course a further question is how one could 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. You would then be justified, all else equal, to assume that the same noisy regularities obtain in the unobserved parts of the sequence.[11]

A similar story might be told for Statistical Mechanics. Our experience of the world reveals statistical patterns in the distribution of microstates among macroscopic systems: the proportion of reasonably isolated systems whose state lies in a region S of their state

9 See also [Hájek 1997] and [Hájek 2009] for further arguments against the frequentist interpretation. However, note that Hájek is mostly interested in whether frequentism is a plausible interpretation of his pre-theoretic, intuitive concept of *chance*. We shouldn't take it for granted that this concept plays a role in our best scientific theories.

10 Some authors seem to think the demand can be met by simply pointing at the laws of large numbers. But these laws are valid for every probability function, no matter how out of touch it is with actual frequencies and no matter how regular or irregular the actual outcomes.

11 As Hume pointed out long ago, 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 scientific 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 this can't be done.

space is generally close to the ratio $\mu_p(S)/\mu_p(\Gamma)$. We also have indirect reason to believe in this kind of regularity: given what we know about the microphysical laws and the age and beginnings of the universe, it is reasonable to expect that those statistical regularities should arise. (This may be part of what makes them *robust* regularities: regularities that would have persisted if things had been a little different.) So we are justified to believe that statistical mechanics captures interesting statistical patterns in the world, and therefore to align our degrees of belief with the theory's probabilities. Here, too, the epistemic interpretation is not far from the truth. Any reason to endorse statistical mechanics is at the same time a reason to regard its probabilities as good prescriptions for belief.[12]

These connections between theoretical probability and rational belief also explain why theoretical probabilities are called 'probabilities'. In ordinary language, probability statements generally seem to be statements about rational degree of belief. We have seen that to endorse a probabilistic theory T means to endorse the theory's probabilities as rational degrees of belief – i.e., as *probabilities* in the ordinary sense of the term. In this way, the epistemic interpretation of probabilistic theories is almost right. The same is true for the best-systems interpretation, since to endorse T also means to accept that whatever theory provides the best systematization of the relevant patterns assigns high probability to A.[13]

5 Probability and actuality

Humeanism in metaphysics is the rejection of primitive modality. Humeans believe that truths about what *must* or *could* or *would probably* be the case are ultimately grounded in truths concerning what actually *is*. The suggestion I have defended does not presuppose a Humean metaphysics. In principle, it is compatible with the idea that what is systematized by probabilistic theories are regularities not just in the actual world but in a larger sphere of nomically possible world (although this might weaken some of the advantages I discussed in the previous section). It is certainly compatible with the existence of fundamental powers and other non-Humean whatnots. But it does support

12 I am glossing over a large number of subtleties here concerning the interpretation of statistical mechanics; see e.g. [Winsberg 2008], [Frigg 2008] and [Callender 2011] for some of these issues. It may also be worth noting that scientific theories often assign probabilities to "outcomes" A under "conditions" C, where it is much easier to find out whether C obtains than whether A obtains, which fits the idea that objective probability is the credence one ought to have given easily available information.

13 On some occasions, statements about probability in the context of developing or applying probabilistic theories might also be statements about a quantity like $\mu_p(S)/\mu_p(\Gamma)$ with which probability is identified in the relevant theory.

the Humean cause insofar as it entails that the probabilities in science do not stand for any such whatnots.

It may also support Humeanism in another respect, by helping to answer what is often regarded as the most serious problem for Humeanism. The problem is this. Imagine a possible world that consists of nothing but a million atomic "coin toss" events, with 80 percent heads and 20 percent tails. On Humean accounts of chance, the chance of heads on each toss must be determined by the actual sequence of outcomes – in this type of case, the chance is presumably identical to the relative frequency. Hence the assumption that the chance of heads on each toss is 0.8 *entails* that about 80 percent of the tosses result in heads. Intuitively however, the assumption about chance entails nothing at all about actual outcomes. Given that the chance of heads is 0.8, it could still happen that the coins land all tails, or all heads, or 50 percent heads and 50 percent tails.

To some extent, sophisticated Humean theories such as the best-systems account can explain where this intuition comes from. It is arguably driven by a simple chance-credence principle stating that rational credence ought to match the known chances. The reasoning might be spelled out as follows: if the tosses are independent and the chance of heads is 0.8, then every sequence of heads and tails has non-zero chance; so every sequence *might come about*; all tails, for example, has a miniscule but positive chance of $0.2^{1000000}$, and deserves a corresponding amount of credence.

On the best-systems account, the simple chance-credence principle is (demonstrably) valid in ordinary scientific cases: if you know that an outcome has chance x, and you don't have unusual further information, your credence in the outcome should also be x. So it is not surprising that this connection should have become part of our very concept of chance. However, when we consider very large events or very small worlds, the usual alignment between best-system probabilities and rational belief breaks down. If you know that the best systematization of a sequence of coin tosses assigns 80 percent probability to heads, your credence in heads on toss number 512 should be 0.8, but your credence in heads on *every* toss should be 0, not $0.8^{1000000}$. (More sophisticated formulations of the chance-credence alignment, such as the "New Principle" of [Lewis 1994] and [Hall 1994], may hold universally, but they no longer licence the inference from 'positive chance' to 'might happen'.)

The best-systems account thus has a plausible error theory for the anti-Humean intuitions. The present account can go even further.

First of all, the anti-Humean may still insist that her concept of chance is tied to the simple chance-credence principle. But why should we make room for such a quantity in serious physics or metaphysics? Given that no such quantity is required – or even useful – for a realist interpretation of science, wouldn't it be presumptuous for philosophers to insist that physics has overlooked some fundamental quantity because its list of fundamentals includes nothing that plays the intuitive chance role?

Secondly, there is a sense in which the present account actually vindicates the anti-Humean intuitions. Again consider a theory T that assigns probability 0.8 to heads. On the account I defended, T entails nothing about actual outcomes. It does not entail that around 80 percent of the tosses come up heads. Worlds in which the relative frequency of heads is 0.5 or even 0 are perfectly sensible models of T, just as worlds where the frequency is 0.8. On the metaphor of theories as experts, by the lights of T, all of these are possible ways things, although of them are more credible than others. That the frequency of heads is around 0.8 is entailed not by T itself, but by the assumption $\Box T$ that T provides the best systematization of the actual outcomes. And *that* entailment is hardly controversial.

6 Conclusion

None of the currently popular accounts of objective chance yields a plausible interpretation of probabilistic theories in science. I have argued that such theories should not be understood as hypotheses about a special subject matter, objective probability. Their point is rather to provide an efficient systematization of noisy relationships in the world.

On this picture, probabilistic theories can't be true or false; they can only be evaluated for simplicity, strength, fit and other theoretical virtues. To endorse a theory is not to regard it as true. It is to regard it as a good systematization of the relevant facts, and consequently to adopt its probabilities as one's own degrees of belief.

My proposal does not entail that objective chance, as conceived by believers, does not exist. But it entails that current science gives no reason to believe in any such thing. Lewis [1980: 83] was wrong when he claimed that "the practice and the analysis of science" requiress a concept of objective chance.

References

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

David M. Armstrong [1968]: *A Materialist Theory of the Mind*. London: Routledge

Craig Callender [2011]: "The past histories of molecules". In C. Beisbart and S. Hartmann (Eds.) *Probabilities in Physics,* Oxford: Oxford University Press, 83–113

David J. Chalmers [2012]: *Constructing the World*. Oxford: Oxford University Press

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

Bruno De Finetti [1974]: “Bayesianism: its unifying role for both the foundations and applications of statistics”. *International Statistical Review*, 42: 117–130

Roman Frigg [2008]: “Chance in Boltzmannian Statistical Mechanics”. *Philosophy of Science*, 75: 670–681

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

Donald Gillies [2000]: *Philosophical Theories of Probability*. London: Routledge

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

Ned Hall [1994]: “Correcting the Guide to Objective Chance”. *Mind*, 103: 505–517

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

Richard Jeffrey [1992]: *Probability and the Art of Judgment*. Cambridge: Cambridge University Press

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. Reprinted in Lewis’s *Philosophical Papers*, Vol. 2, 1986.

— [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

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

Karl Popper [1982]: *Quantum theory and the schism in physics*. Totowa, NJ: Rowman and Littlefield

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

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

— [2007]: “Deterministic Chance?” *British Journal for the Philosophy of Science*, 58: 113–140

Wolfgang Schwarz [forthcoming]: “Proving the Principal Principle”. In A. Wilson (ed.), *Asymmetries of Chance and Time*

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

Eric Winsberg [2008]: “Laws and chances in statistical mechanics”. *Studies in History and Philosophy of Modern Physics*, 39: 872–888