Quantum Probability; chance

Jenann Ismael
Centre for Time, University of Sydney
jtismael@u.arizona.edu

draft; please do not quote or circulate

In contrast to classical physics, probabilities play an explicit and conspicuous role in quantum mechanics providing the link between the fundamental level of physical description and the measurement results that mark the points of empirical contact between theory and world. The nature of quantum probabilities (a.k.a., 'chances'), however, has been a source of contention since the theory appeared and despite a proliferation of accounts, the metaphysical nature of chance remains unsettled. The difficulty is that there are a number of constraints on an interpretation of chance, constraints that appear to be partially definitive of the concept, and it proves to be extraordinarily difficult to meet them simultaneously. I'm going to show how those constraints give rise to an interpretive dilemma that confounds the extant options, and then propose a way out that hinges on the use of a more basic form of probability for whose existence I have argued elsewhere.[1]

Quantum mechanics

Quantum mechanical states are represented by mathematical objects called wave functions, and there is a rule - Born's Rule - for generating, from the state of a physical system, the chance that a measurement on the system would yield a given result. Reiterated application of Born's rule generates what I'll call a chance profile for a system: i.e., a probability assignment to the event of observing a given result in any measurement that can be performed on it. For any time t and any measurement event e (where a measurement event is the event of observing a particular result on the conclusion of a measurement, we can ask "what is the chance of e at t?"[1],[2] If e is at t or in t's past, it will have probability 0 or 1. If it is in t's future, it will have any real value in the open interval between 0 and 1. An event that has a chance of 1 or 0 at p, it retains that value for all subsequent times. If we trace a path through time, keeping track of the chance of a particular event – e.g., the event that some coin flip comes up heads, or that a measurement carried out at a particular time and place shows a positive result – the event has different chances at different times, can flip around as much as you please up until the moment of the event itself, when it takes on a fixed value of 0 or 1. [1]

These quantitative facts – that past events all have value 1 or 0, that an event that has value 1 or 0 at one time retains that value at later times – are important because they are not contingencies that had to be *discovered* about chance. They're things that anyone who 'grasps the concept' of chance will be able to tell you, and they are our best clues to its nature. Another such clue to the nature of chance is provided by a link to belief that David Lewis expressed in his Principal

[1]\"Probability in Classical Physics; the fundamental measure". I'll be assuming its existence here, and showing how it provides a way out of to difficulties of interpreting chance.
[2]I'll use 'measurement interaction' to mean the carrying out of a measurement, and 'measurement event' to mean the observance of a particular result.

Principle. He himself regarded the Principal Principle as the sole constraint, believing that it told us "all we know about the concept of chance". Not many have followed him here, and there are reasons for thinking he can't be right. If chances to *guide* beliefs we have to have ways of forming beliefs about chances that aren't just ways of forming beliefs about our own beliefs. That means that we have to have some idea of what counts as evidence for statements about chance and that evidence will have to be connected to the truth conditions for those statements. What plays that role is a connection to frequency given by Bernoulli's Law.[3] Bernoulli's Law relates indefinite probabilities to frequencies. Applied to this instance, it says that the relative frequency of a's in a typical ensemble of systems in ψ approaches the indefinite probability of a/ψ as the size of the ensemble increases, but the possibility of divergence remains, no matter how large the ensemble,[4] we now have three constraints that any interpretation of chance must satisfy;

1. **Quantitative constraints;** the chance of an event after it occurs is always 1 or 0, an event that has value 0 or 1 at one times retains that value for all future times.
2. **Bernoulli's Law:** the relative frequency of a's in a typical ensemble of systems in ψ approaches the indefinite probability of a/y as the size of the ensemble increases, but the *possibility* of divergence in any finite ensemble remains, no matter how large the ensemble.
3. **PP;** one should set one's credence in a at t, equal to the chance of a at t, no matter what else one knows, provided one has no information from the future. [5],

Support for all of these constraints, which we can treat as partially, provisionally definitive of chance, is that someone who denied them would not fully understand the concept. [6] Someone who knew that George Bush won the election in 2004,for example, but denied that the present chance of his doing so is 1, or someone who knew what the

[3] Indefinite probabilities are also sometimes called general probabilities. The basic form is relativized, and indeed the same as that of relative frequencies. The indefinite probability of B among A's is written pr(B/A); in logical terms, we can say that "pr" is a variable-binding operator, binding the "*x*" in "prob(*Bx*/*Ax*)".
[4] Some have been hesitant to treat indefinite probabilities as basic out of deference to physics, where it is supposed that chance plays a physically fundamental role apparently ungrounded in any general relativized probability. It will be clear that I regard that as misguided. I hold with Kyburg (1974), Pollock (2006), and Hajek (check) that the only approach that can explain the epistemological foundations of probabilistic reasoning constructs definite probability out of indefinite probabilities.
[5] David Lewis excluded the application of PP to cases in which an agent has magical sources of information from the future, without providing an explicit characterization of what falls in this category.
[6] We might add to the list something that characterizes the role of chance in guiding action. I ignore this since it won't play a role here.

chances were but didn't think that they should guide his expectations,[7] or someone who didn't think that frequencies under the right conditions provided evidence for statements about chance, wouldn't be regarded as understanding what chance is, or would be regarded as using the word in a different way. [8] 1 supplies an analytic connection to credence, 2 supplies necessary connections between chances and the events they are the chances *of*, 3 provides a necessary, but ineliminably approximate and probabilistic connection between chance and frequency. [9]

The interpretive dilemma

The constraints are easily satisfied individually, but conjointly they present a dilemma that destabilizes both the standard reductive and non-reductive accounts of chance. The problem is that they specify connections between facts about chance and facts about categorical events - the events that they are chances *of* - which seems, on the one hand, too loose to permit reduction and, on the other, too tight to let us treat them as distinct existences. Why too loose to permit reduction? The natural candidates for reduction are frequencies. If the reference class is chosen appropriately, a frequentist account can secure the role of chance in guiding belief, and allow facts about frequencies to provide evidence for statements about chance. As reductions, however, frequentist accounts run up against the intractable fact that the link between frequencies and probabilities is irreducibly probabilistic. Identifying chances with frequencies in actual reference classes fails because of the possibility of divergence between the probabilities and the actual frequencies. Identifying chances with frequencies in specified hypothetical classes ends up almost inevitably begging the question by using probabilistic notions in specifying ensembles within which frequencies necessarily reflect probabilities. probabilities necessarily *typically* represent frequencies, and frequencies in *typical* ensembles necessarily represent frequencies, but reduction requires elimination of the 'typical'. Without a necessary connection to frequencies that can be specified without using probabilistic notions, there is no reduction.[10]

Why is the link between probability and frequencies too tight to permit primitivism? There are necessary quantitative constraints on the relations between an event and the chance of its occurrence. Ask yourself whether they can vary independently of one another, focusing on times after the event. If these are distinct existences, there ought to be a possible world the event occurs and the chances have any distribution you care to describe. We ought to be able to at least conceive of a world in which a coin comes up heads and has a non-zero

[7] Even the gambler who knows the chances are against him, but still believes he will win, when he's thinking clearly will concede that chance should guide credence, even though it doesn't always do so.
[8] This is soft evidence, only because conceptual connections are themselves soft.
[9] It's important to understand the nature of these constraints; someone who violates the constraints is making a conceptual mistake, not a mistake of fact, and that means that the constraints should be derivable from a successful analysis of the concept.
[10] Some would disagree. See Hajek http://setis.library.usyd.edu.au/stanford/entries/probability-interpret/ for an excellent summary of the critical literature.

chance after the fact of having come up tails. But the concept doesn't (or doesn't seem to) make room for the possibility.[1] (To see this, think in four-dimensional terms, suspend any intuitive interpretation, look at how the chances relate to the events they are chances of, and notice that there are certain combinations of categorical fact (the occurrence of e) and values for chance (chance of e's occurrence) to which you can assign no content, combinations that don't make sense).

This is *the* fundamental dilemma faced by any account of the metaphysical status of chances. The peculiar ontologically intermediate status that frustrates both reduction to categorical facts and primitivism. From a formal point of view, chances look like ordinary physical quantities (they are represented by real-valued functions, they evolve in time, and so on), and so it would be convenient if we could treat chance as a primitive quantity, and say 'so much the worse for the Humean ban on necessary connections between distinct existences. Accounts that identify chances with propensities, explicitly denying that propensities are grounded in intrinsic categorical properties seem to be doing that. That reaction fails, however, to appreciate the regulative role that the Humean ban plays in these contexts. It's something more than a metaphysical principle that we can accept or reject. It doesn't serve as a premise. If we reject the Humean ban, we no longer have a way of recognizing distinct existences, Indeed, it's hard to say in that setting what we might mean by 'distinct existence'. When we engage in trying to sort out our ideas about the metaphysical structure of the universe - ask questions of the form 'what is A?', attempt reductions, and so on - part of what we're trying to do is provide a compact, non-redundant catalogue of the categorical facts, and we use necessary connections of a non-nomological nature as a test for redundancy. Once we've got a catalogue of the non-chancy facts, we raise the question about the metaphysical status of chancy facts. Have they already been included? Not by simply including the frequencies, since by Bernoulli's law, they don't supervene on the frequencies. But if we add them as primitive elements, it looks like we've got the necessary connections that are our test for redundancy in the categorical basis.[1] Those working for reductions are responding to the fact that e and the chance of e do not behave like distinct existences. Primitivists are responding to the fact that reduction seems blocked by ineliminably probabilistic character of the link between chances and frequency. But neither can respond appropriately to pressures from the other side. What to do?

The solution

If the only options were to treat chance as a primitive quantity or reduce it to categorical facts, we would be stuck. But those aren't the only options. The conceptual character of the quantitative constraints provided clear symptoms of non-basicness, so reduction is needed, but it's not going to be reduction to categorical facts. I've argued elsewhere that any theory, deterministic or not, that can be used as a basis for prediction in the absence of Laplacian knowledge of the state of the universe comes equipped with a measure over state-space that represents the probability that a random pick from a specified population will yield a system in a state that falls within a given subspace; $pr(A/B)=_{def}$(the probability that a random pick from B-systems will generate a system in A). It's not hard, using this measure, to find a definition that gets the extension of chance right.

Def $Ch_t(e) =_{def} Pr(e/\text{pre-t history})$

Def guarantees (i) that historical events have a chance of 1 or 0, depending on whether they are included in history, (ii) that events that follow deterministically from historical events get probability 1, and (iii) that events that are incompatible with events that follow deterministically from historical events get probability 0. It will assign all other events values in the (closed) interval between 0 and 1, and the numbers it will assign will reflect the frequency of e-type events following similar histories, by Bernoulli's law.[11]

That takes care of the first two constraints on the interpretation of chance, but it leaves the third. The satisfaction of the first two constraints follows almost trivially from the definition, but why should a function that satisfies those also satisfy the third? Why, that is to say, should a function with the quantitative characteristics described in the first two constraints guide credence? Even if we grant that credence has to be a probability function, there are indefinitely many such functions derivable from the fundamental measure, any one of which can formally fulfill the role defined by PP. There is, for example, pance, where $pance_p(e)$= pr(e/the events that occur in Australasia before the 01/01/2001).[12] Or fance, where $fance_p(e) =_{def}$ Pr(e/contents of p's future light cone). What makes chance better suited than either of these as a guide to belief? Indeed, fance is the time-reversed counterpart of chance, and insofar as we are interested in predicting the future, the fances do better than the chances. For that matter, there's one function that seems better suited than any to guide belief, *viz.* Truth(e)=pr(1 if e occurs, 0 otherwise).

What is called for here is an explanation of why chance plays the role defined by PP in our epistemic lives. One explanation might point to something that distinguishes it intrinsically from other probability functions. A different type of explanation, the one that I think is correct, preserves the metaphysical parity of chances with other functions of the fundamental measure and explains the special importance *we* accord to chance by pointing to particularities of our epistemic situation. *Chance is probability tailored to guide belief for agents like us*, agents that have access to specific, potentially detailed information about the past from mementos and memories of past observations, but no independent source of information about the future.[13] Physical theories are specialized tools for squeezing information out of this accumulated storehouse of historical information: information about the unobserved past and parts of the universe that we don't have direct observational contact with, but mostly information about the future, i.e., information that can be used

[11] The global definition gets its meaning from application on the local scale, the chance of e on S is the probability a random pick from a typical ensemble of systems that shares all dynamically relevant features of this one will yield a system in which e.
[12] All events that occurred in Australasia before that date have a pance of 1, events incompatible with those have zero pance, all other events have pance in the interval $0<x<1$.
[13] For a precise recent attempt to characterize this epistemic asymmetry and its physical basis, see Albert, *Time and Chance*.

in decision as a basis for action.[14] They identify regularities that get expressed as laws and combined with the values of known variables and an objective measure to generate predictions.

The chances "solve for history". They make explicit the predictive content of a measure and set of laws relative to a hypothetical history.[15] They tell us what to expect of a beam of electrons confronting a barrier in a two slit experiment, for example, what to expect of a ball rolled down an inclined plane, and in general what to expect in the future of a system whose past has been specified. The chances are – to borrow a useful phrase from Edwin Hutchins – 'partially prepared solutions to frequently encountered problems'. They're only partially prepared because the models of hypothetical systems for which we can actually calculate chances don't approach the complexity of systems we encounter in the messy business of real life. We never have the time or knowledge we need to calculate the chances for real systems, so we keep a stock of simplified hypothetical models on hand and use those to estimate the chances for real systems, by approximation and mxing. Our theory will tell us which variables are significant. Among those, we specify the values of known variables and derive expectations by a mixing procedure that randomizes over unknown ones, effectively treating the system of interest as a random pick from systems compatible with the values of known variables and letting the probability that a random pick will yield a system with value in the range a for a determine the weight, α, that we assign in forming the mixture: $pr(a)=\Sigma\alpha pr_i(a)$. Out in the wild, it's not PP that guides belief, but this mixture, which is a generalization of PP that says not 'adjust credence to chance' but 'adjust credence to your best estimate of the chances'.[16] This sort of generalization is all part of the familiar, if messy, business of applying hypothetical models to real world situations.

So what makes chance among all the probability functions that there are peculiarly suited to guide belief? Chances provide an explicit characterization of the information that history contains about the future, and for agents that have access in principle to historical information but no independent access to information about the future, they are unique in being both ascertainable and trumping all other ascertainable information. To let credence be guided by $pance_p(e)$= pr(e/the events that occur in Australasia before the 01/01/2001), for example, would be a bad policy because it would be ignoring information in our possession about what has happened since 2001. Even a stupid bookie could make money by getting us to bet on horses that lost races between 2001 and today. The problem with letting credence be guided by $fance_p(e) =_{def}$ Pr(e/contents of p's future light cone) or Truth(e)=pr(1 if e occurs, 0 otherwise) is different. It's not that it ignores information we have, but presupposes information we don't. We can't calculate the fances or the truths that without knowing what the future holds, and so we can't calculate the fances or truths

[14] This is not to say that this is *all* they are.
[15] In the special case of Markovian laws, the state of a system at any time screens off preceding states, so history is irrelevant given full specification of present state.
[16] I'm suppressing theoretical uncertainty; speaking here just of the prediction derived from a theory, utilizing its laws and employing its measure over state-space. But see my "Raid! The Big, Bad Bug dissolved", *Nous,* forthcoming.

without knowing the very events we want to use them to predict.

Novelty

If the novelty introduced by indeterministic laws isn't in the appearance of a new form of probability, or a new role for probabilities in inductive inference, wherein does it lie? It consists in the elimination of a degenerate case that is present in a deterministic setting. I've argued that metrical assumptions play an ineliminable role whenever we're working with volumes of phase space from which there emerge multiple physically possible trajectories. In deterministic theories, the number of such trajectories goes to 1 as the size of the volume goes to 0. So eliminating historical ignorance reduces all prediction to the degenerate case in which there is only one possibility.[17] In indeterministic theories, there are multiple trajectories through every point, so there is no degenerate case, and even if we eliminate historical ignorance, a measure is needed to set credence.

There is a way of putting this that makes contact with the technical work on hidden variables in quantum mechanics. In deterministic theories even when we are working with imprecise or incomplete information, there is a finer-grained description that locates the system in an underlying space with a single trajectory through every point. In the indeterministic case, there is no such underlying space, no fine-grained description with a single trajectory through every point that will let us reconstruct the probabilities at the higher level from a distribution over variables whose values are (in principle, jointly) measureable. We have to formulate the laws using a space that has multiple trajectories through almost every point, and that means that if we are to end up with an objective measure over possibilities, probabilities have to appear explicitly the statement of the laws.[18]

Can't we reinstate determinism trivially by 'discerning' variables so hidden that they can't be measured, using them to parameterize phase space with an intrinsic measure that generates the Born measure over quantum mechanical state descriptions? This is a natural suggestion, and the one that Einstein explored until his death. Under such an arrangement, quantum mechanical state descriptions would correspond to finite volumes of the underlying space in the same way that thermodynamic state-descriptions correspond to finite volumes of the phase space of statistical mechanics. The problems with the suggestion are well-known; there is a vast, and technically formidable literature in foundational discussions exploring the possibilities for hidden variable interpretations. There are unsettled questions, but we do know is that any such interpretation will exhibit non-locality or contextuality of some form at the fundamental level, and its very difficult both in physical terms and conceptually to work out the implications.[19]

All of this raises the question of why indeterministic laws have *seemed* to be introducing something wholly new into physics. Hard to say, but here are some reasons;

---

[17] Keeping in mind that it still plays role in the wider class of problems where we don't have full historical knowledge.

[18] I'm suppressing the real question of whether Born's Rule is properly called a dynamical law.

[19] See http://setis.library.usyd.edu.au/stanford/entries/kochen-specker/ for a discussion of hidden variable results.

- because probabilities are only needed in the deterministic case when there is historical ignorance, we get to thinking that the probabilities *represent* facts about our epistemic states,
- because of the way we formulate theories: we give laws that apply to the degenerate case, exploiting the metrical structure of phase space in mixing, so that probabilities remain hidden in deterministic contexts, and
- because there is a long philosophical tradition of thinking of laws in terms that don't carry over naturally to an indeterministic setting, terms that seem to preclude anything but a strict and necessary connection, and an important adjustment in metaphysical ideas is needed to accommodate the indeterministic case.[20]

Methodological reorientation

I've been arguing that we've been treating chance as a surprising and anomalous protrusion of probability into an otherwise probability-free environment, where in fact it's a special case of something more fundamental and perfectly pervasive, the tip of a probabilistic iceberg that remains largely below the surface in classical physics. If I am right, this should prompt a methodological reorientation, shifting attention from chance to the fundamental measure. This shift of attention has implications for how we think of physical probability. First, unlike chance, the fundamental measure is not inherently dynamical. If we can talk about the probability of evolving into from one subspace of phase space into another, we can equally talk about the degree of overlap between two subspaces: the probability of being in B, given that one is in A.[21] Second, we no longer have to think of there being two distinct types of physical probability; the canonical measure invoked by statistical mechanics and chance. We have a *single* measure from which both can be derived. Finally – and perhaps most significantly - the fundamental measure has close intuitive links to the concept of information that are canonized in the basic definitions information theory. The basic unit of information (the bit), for example, is defined as a choice between *equally probable* alternatives. This link suggests a promising intuitive entry into the circle of probabilistic concepts, opening up potentially fruitful avenues of exploration. So long as we conceive of chance as an exotic quantum mechanical phenomenon, I believe we'll get nowhere with it.

Applications

---

[20] We could say that the chance e on S is a measure of the propensity of a system with S's past to give rise to e-containing futures, but this requires caution. It is one thing to hold that the chance of e on S is a measure of the strength of the propensity of systems with S's history up to t show e when measured; it is another to identify it with the intrinsic ground of the propensity. The difference is crucial if we want to secure constraints on their relations to the events they are chances of. The intrinsic grounds of propensities bear no necessary connection to the events to which they give rise.

[21] The probabilities of statistical mechanics combine both. We start with the probability of being in a certain microstate, given that one is in a given macrostate, and use that, in conjunction with microdynamical laws to derive transition probabilities for macrostates.

Before closing, I want to pause to say how to translate into a relativistic setting and mention an application having to do with the interpretation of probabilities in an Everett Universe. I have spoken in non-relativistically for intuitive ease, but all that is needed to translate is to identify systems with world lines, define chances at points, and substitute 'contents of past light cone' for 'history'. So now we have

Def* $Ch_p(e) =_{def} Pr(e/\text{the contents of p's past light cone})$

Traditional Laplacian definitions characterize chancy events as those that are undetermined by the dynamical laws from the preceding state of the universe, leaving no room for chancy events in a universe like Everett's governed by global deterministic laws. One of the virtues of **Def***is that it separates the existence of chance from determinism, tying it to the light cone structure which gives us chances in an Everett Universe, the determinism of the dynamical laws notwithstanding. [22]

If we replace the Laplacian definition with:

An event e that occurs at p is **chancy** just in case its occurrence cannot be predicted with certainty by application of the dynamical laws to the contents of p's past light cone.

We preserve a connection between chance and predictability but break the connection with determinism because the connection between predictability and determinism is lost in the context of a light cone structure that imposes greater restrictions on the availability of information. The result is that if we apply the definition in an Everett universe, we get chancy events despite the determinism of the global dynamical laws.

Another issue that has been difficult for decoherence-based Everett style interpretations of quantum mechanics is justifying the Born Measure. It is a special case of a generalized measure over state-space present also in classical contexts, whose only justification is that it makes explicit the predictive content of the projectible patterns in the manifold when combined with historical information.

Conclusion

Together, this paper and "Classical Probability: the fundamental Measure" provide a unified account of probability in classical and quantum physics that recognizes a species of probability in the form of a measure over phase space, treating chance as a special case: $Ch_p(e) =_{def} Pr(e/\text{the contents of p's past light cone}$. The definition is the only account I know that gets the pattern of relationships to categorical events, frequencies, and credence right.

[22] There are (at least) two problems about probability in an Everett Universe; one is making room for probability, one is the quantitative problem. The quantitative problem requires separate treatment.