A Modest Proposal About Chance[i]

Jenann Ismael

jtismael@u.arizona.edu

DRAFT!!! Comments most welcome

Before the 17th century, there was not much discussion, and little uniformity in conception, of natural laws. The rise of science in 17th century, Newton's mathematization of physics, and the provision of strict, deterministic laws that applied equally to the heavens and to the terrestrial realm had a profound impact in transforming the philosophical imagination. A philosophical conception of physical law built on the example of Newtonian Mechanics became quickly entrenched. Between the 17th and 20th centuries, there was a great deal of philosophical interest in probabilities, but probabilities were mostly regarded as having something to do with the management of opinion, not as having a fundamental role in science.[ii] Probabilities made their first appearance in an apparently ineliminable way in the laws of a fundamental theory with the advent of quantum mechanics. Quantum probabilities have come to be called 'chances' in the philosophical literature, and the interpretation of quantum probabilities has been one of the central problems in philosophy of science now for almost a century. There continue to be hold-outs that insist that there must be an underlying probability-free replacement for the quantum mechanics and Bohmians have had some success in formulating a deterministic alternative to quantum mechanics, but most physicists accept that the probabilistic character of the quantum mechanical laws is likely to be retained in any successor theory. While physics has adjusted itself comfortably to the existence of ineliminably probabilistic laws, philosophy has not managed arrive at a stable interpretation of quantum probability. The difficulty is that there are a number of constraints that an interpretation of chance must satisfy, constraints that appear to be partially definitive of the concept and it proves to be extraordinarily difficult to meet them simultaneously. Below I'll introduce the difficulty and suggest an unorthodox response that depends on challenging the virtually unquestioned dogma that chance is a physically fundamental form of probability.

Quantum mechanics

I'll remain within the context of standard, non-relativistic quantum mechanics. In this setting, chances provide the link between the fundamental level of physical description and the measurement results that mark the points of empirical contact between theory and world. Quantum mechanical states are represented by mathematical objects called wave functions, and Born's Rule is a rule that generates, from the state of a physical system, the chance that a measurement on the system would yield a given result. Any interaction in which an observation is made on a system, or information extracted from the system is considered a measurement. Reiterated application of Born's rule generates 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 e 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?" 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 can take 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 chance of the event has different values at different times, can flip around as much as you please up until the time at which the event is slated to occur, when it takes on a fixed value of 0 or 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. They are not things that we had to discover about the distribution of chances at our world. 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 identified and formulated in his Principal Principle (PP). [iii] PP says that if we know what the chance of an event is, and we have no crystal balls or magical sources of information from the future, we should adopt the chance of e as our credence. Now, if chances to guide beliefs we have to have ways of forming beliefs about chances. 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.

Before we can state Bernoulli law, we have to make a distinction. One of the most persistent sources of confusion in the discussion of chance is the failure to distinguish general from single case probability. General probabilities are also sometimes called indefinite probabilities and single-case probabilities called definite probabilities. General probabilities apply to classes of events. The basic form is conditional. The indefinite probability of B among A's is written pr(B/A).[iv] Single-case probabilities, by contrast, apply to particular events rather than classes, and the basic form is unconditional. Think of the difference between the general probability that an arbitrary roll of a pair of dice comes up double sixes, given that it is unbiased versus the single case probability that this particular roll of these particular die on this day comes up double sixes. The former is conditional, it doesn't change over time, there is an explicit reference class provided. The latter is unconditional, it has different values at different times. Single-case probabilities a direct connection to credence; general probabilities bear on credence only by way of their connection to single-case probabilities. The Reference Class Problem for probabilities is the problem of determining which general probability to use to determine the single case probability, since there are indefinitely many general probabilities that pertain to an event.[v]

Chance is a single-case, time-dependent probability. The chance distribution at t is given by a function that assigns real numbers to events. $Ch_t(e)$ is the chance that an event e has at time t. [vi] If e the event that a coin toss to be made at noon on Jan. 1, 2009 comes up heads, $Ch_{now}(e)$ is the present chance that the toss lands heads. In ordinary speech, we often suppress the temporal parameter, letting the conversational context decide the time at which the chance is assessed (if you ask me what chance I think you have of winning a bet predicated on the coin toss, we both understand that you are asking for the chance that pertains at the time of asking). So again, general probabilities apply to classes and don't generally have a time index. Single-case probabilities are condition, time-dependent, and pertain to particular events.

Bernoulli's law is a theorem that relates general probabilities to frequencies. It says that the relative frequency of A's in a typical ensemble of B's approaches Pr(A/B) as the size of the ensemble increases. This is good news; it gives us a necessary, probabilistic link between probabilities and frequencies. But there is a bad news addendum; there are other theorems that tell us that that link cannot be strengthened. The possibility of divergence remains, no matter how large the ensemble.[vii] This gives us the third constraint the interpretation of chance, the chance-frequency link. The chance-frequency link says that if S is in state ψ, there's a necessary, but approximate and probabilistic relation between the chance of observing e in an x-measurement on S and the relative frequency of e in x-measurements on systems in ψ.

We now have three clues to the nature of chance that we apply as constraints on interpretation;

1. **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.[viii]

2. **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.

3. **Chance-frequency link:** 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.

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 categorical facts.

The interpretive dilemma[ix]

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 non-probabilistic, non-modal facts about actual measurement results - 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? It's a fact about the logic of 'chance' that the overall pattern of actual categorical fact can be one that has a low, medium, or high chance of occurrence. Bernoulli's law explicitly allows for the possibility that the chances may diverge arbitrarily far from the frequencies, which is a way of saying that it is a fact about the logic of 'chance' that the very same distribution of actual events is logically compatible with multiple (indeed, in principle, unlimited) chance distribution. The link between the actual pattern of events and the chances is irreducibly and irremediably probabilistic.[x]

Why is the link between chance and categorical facts 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 chance after the fact of having come up tails. But the concept doesn't (or doesn't seem to) make room for the possibility. 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 presents a dilemma to any account of the nature of quantum probabilities. Chances have a peculiar ontologically intermediate status that seems to frustrate both reduction to categorical facts and primitivism. From a formal point of view, chances look so much like ordinary physical quantities (they are represented by real-valued functions, they are assigned to particular systems, and they evolve in time) that it would be nice 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.[xi] That reaction fails, however, to appreciate the regulative role that the Humean ban plays in these contexts. If we reject the Humean ban, we no longer have a way of recognizing distinct existences. When we engage in ontology - ask questions like what is chance? what are colors? what is goodness? - part of what we're trying to do is provide a compact, non-redundant catalogue of basic fact. In this context, necessary connections of a non-nomological nature function as a sign of redundancy. Here is a familiar pattern of argument: Tom asserts that A-properties are metaphysically distinct from B-properties. Alice responds that if A and B-properties were distinct existences, there ought to be possible worlds just like ours in which the B facts are just as they are, but the A ones are wholly different, and vice versa. To put it another way, the A-facts ought to vary independently of the B-facts across the space of possible worlds. If they fail to do so, Tom is wrong. The Humean ban functions in this pattern of argument a test for ontological redundancy, and without it, we no longer have any methodological foothold for addressing claims of metaphysical distinctness.

A rather large literature has built up around the interpretation of chance. There are a number of accounts on the books. A few papers of David Lewis in the early 1980's brought the topic under special scrutiny because of the difficulty he saw with incorporating chance into his metaphysical framework.[xii] But as things stand, by general consensus, there is no clear forerunner, and no uncontroversially satisfactory interpretation of chance. Each one of the existing interpretations suffers from well-known problems, each comes with a bullet to bite.[xiii] Those working for reductions are responding to the fact that e and the chance of e do not 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 categorical facts. But neither can respond appropriately to pressures from the other side.

The solution

There's a lesson that one learns in physics or math when one feels backed into a corner with a difficult problem, try the obvious solution. You'll be surprised how often it works. There's a way out of these difficulties that is so obvious it strikes one at first as trivial, and it takes a little unpacking to see what it accomplishes. The solution is to reject the unquestioned and unsupported dogma that chance is physically fundamental. 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 propose that w reject both primitivism and reduction to categorical facts and define chance\ as follows

Def; $Ch_t(e)=Pr_G$ (e/pre-t history)

Where $Ch_t(e)$ is the probability of e assessed at t, and Pr_G (A/B) is the general probability that a random pick from the B's will yield an A.

It is easy to see that Def satisfies the second and third of the constraints on the interpretation of chance. It follows immediately from the definition that the chance of any event assessed after its occurrence will always be 0 or 1. And since Pr_G is a probability function, it is going to satisfy the probability axioms, and the chance-frequency link will hold.

To establish the extensional adequacy of Def, it remains only to show that it satisfies PP. The non-trivial part of establishing that Def satisfies the constraints on interpretations is saying why the measure defined by Def should play the role characterized by the Principal Principle in guiding belief. After all, there are an indefinite number of functions of Pr_G, any of which have the right form. Consider, for example,

Pance(e)=pr_G(e/pre-2001 history of Australia)
$Fance_t(e)=pr_G$(e/post-t history)
Trance(e)=pr_G(e/all of history)

Out of all of these functions, what makes chance peculiarly suited for the role carved out by PP? The answer is that chance is the only one of these functions that trumps *all* and *only* historical information. For creatures like us, historical information is in principle available, whereas information from the future is ordinarily out of bounds. Chance guides belief because it is probability conditioned on all in principle available information to a situated agent. As such, it is the only function of Pr_G that trumps all and only such information. Belief guided by Pance(e) and Fance(e) would lead us to assign probabilities other than 1 to historical events that we know occurred, or non-zero probabilities to historical events we know did not occur. Belief guided by Trance(e) will not lead to those kinds of mistake, but there is a different problem.

Information about the trances of future events are in general beyond our ken, and hence unavailable to guide present belief.

Pr_G(A/B)

Def does not give us a reduction of facts about chance to non-probabilistic facts. It gives us an interpretation of chance in terms of a different, and purportedly more fundamental, form of probability Pr_G. The questions that need to be answered are: what is Pr_G(A/B)? What reason do we have to believe in it? What sort of progress has been made by reducing chance to Pr_G(A/B)? Taking these in turn: Pr_G(A/B) is the general, conditional probability that a random pick from some proscribed volume of phase space (B) will yield a system that – at that time, or some later time –in another such volume (A). [xiv] It effectively defines the notion of randomness according to a particular theory.

The case for recognizing Pr_G(A/B) needs to be made independently and in more detail, and has been given elsewhere. But the argument is that such a measure is tacitly and routinely invoked as a matter of practical necessity in classical contexts with deterministic laws because without such a distribution the laws are virtually impotent to generate expectations.[xv] Global deterministic dynamical laws determine what is *possible,* but they don't tell us how to divide opinion among the possibilities. In any realistic situation, there is an indefinite number of present states that the world can occupy and that will mean that there an indefinite number of future states into which it can evolve.[xvi] This is as true for a proscribed subsystem of the universe as for the universe as a whole because the state of any proscribed subsystem is subject to innumerable possibilities of unanticipated interference. Laws that tell us only which future states are possible, given our present knowledge will give an infinite undifferentiated set of possibilities with no distinction between significant and negligible possibilities or between relevant and irrelevant influences.

In practice, we deal with ignorance by randomizing over the values of unknown variables. But talk of randomness is not well-defined without a choice pf Pr_G. The choice of Pr_G(A/B) defines what is meant by randomness and guides the assignment of probabilities where there is ignorance.[xvii] To see this, let W be a collection of physically possible worlds (or physically possible histories for the world), let S be the set of worlds in which the system of interest exists and that satisfy all known properties (internal, relational, and external) of the system, and let W* be the complement of S in W. There are indefinitely many distributions over W*. When we talk about randomising over the values of unknown variables, we are talking about choosing a distribution over W*. It doesn't help to adopt an indifference principle, indifference principles can be applied consistently only over *equiprobable* alternatives. Nor does it help to say that we should give a flat distribution because there are different ways of partitioning any set of possibilities and a flat distribution over one partition will yield an uneven distribution over another.

There is a long history of attempts to justify the choice of Pr_G(A/B), and different views about whether and how the choice can be justified, but there is no way around the fact that some choice is needed.

Choosing Pr_G is equivalent to choosing a standard that identifies partitions of W* that divide it into equiprobable classes. In physical contexts, this standard gets embodied in the metrical structure of the phase space and such strong intuitive prejudices are operative in dividing possibilities into equiprobable classes, that the choice of $Pr_G(A/B)$ usually goes unmentioned and unnoticed. In classical mechanics, it didn't come to light until disputes in the foundations of statistical mechanics bought it under scrutiny. The result of those disputes is wide recognition now that a form of statistical probability is both compatible with determinism and needed to derive a dynamics for larger-than-point-sized volumes of phase space. We can picture these higher-level dynamics as migration patterns which tell us how a class of systems distributed in a given way across a finite volume of phase space typically redistributes itself over time.[xviii] To generalize the conclusion that the higher level migration patterns are what are needed to derive expectations not only in the specialized context of statistical mechanics but virtually any real-life calculation we need only observe that the scientist is always working with systems whose state can only be isolated in a finite volume of the universal phase space.[xix]

What lends $Pr_G(A/B)$ empirical content is the link to frequencies given by the Bernoulli Principle. The bad news addendum to the principle, however, blocks a direct reduction to frequencies. In my view Pr_G should be treated, alongside the physical laws, as a primitive postulate whose correctness is justified by the success of the theoretical package of which it is a part. The success of package as a whole confers whatever confirmation they possess upon its elements. And confirmation is judged not by compatibility with the laws, but by the ability of the theory to predict the evidence with high probability. Note, however that to say that Pr_G is a primitive theoretical postulate is to remain neutral on whether it can be reduced or eliminated, or analysed as a particular form of idealized subjective probability at a metascientific level. Just as it is possible for reasonable and well-educated philosophers to accept that laws are physically fundamental and disagree about the nature of laws, whether facts about laws can be reduced to or analysed in terms of some more fundamental set of facts, or whether we have good reason to believe that there are any laws, so it is possible for reasonable and well-educated philosophers to agree that a probability measure of the form $Pr_G(A/B)$ is physically fundamental and disagree about what probability is, whether facts about probability can be reduced to or analysed in terms of non-probabilistic facts, and whether we have good reason to believe in probability. Questions about which structures are physically fundamental can be separated out and treated independently. They are questions abut the components needed to have a working theoretical package.[xx]

But if Pr_G is being accepted as a primitive postulate, we need to say what progress has been made by reducing chance to $Pr_G(A/B)$. The definition of chance in terms of $Pr_G(A/B)$ reduces it to a form of statistical probability that also underwrites the probabilities of statistical mechanics, and explains the epistemic role of chance. Since $Pr_G(A/B)$ bears no necessary connections to categorical facts, and this resolves the dilemma generated by treating chance as fundamental. It is also a nice feature of Def that since Pr_G (A/B) is not time-dependent by defining chance in terms of Pr_G (A/B), Def preserves the temporal symmetry of the fundamental theoretical postulates. There's a very real worry that, because chances of past events always have values of 0 or 1, while chances of future events range in the (closed) interval between 0 and 1, admitting chance into physical theory is recognizing a fundamental form of temporal asymmetry; it is saying that the past is in some fundamental respect different from the future). Whether you think that is a virtue of the account will depend on other commitments.

It should be noted, however, that although Def reduces chance to a form of statistical probability, it is not the sort of reduction that Einstein envisaged. His hope was to reproduce the quantum mechanical probabilities from deterministic laws and a probabilistic postulate on the model of statistical mechanics. The proposal here accepts the ineilimably probabilistic character of the dynamical laws, but argues that determinism is no respite from probability. Even deterministic theories cannot make due in practical terms without a probabilistic postulate.

Why treat $Pr_G(A/B)$ as the basic object?

I've said that what we need in practice to derive expectations outside the context of pen and paper problems in which all of the information is artificially stipulated, are migration patterns over phase space that deliver the probability that a system that starts out in some finite volume ends up in another volume some finite time later. These are not derivable in any context - deterministic or otherwise - without a probability measure put in somewhere. But one might ask: why treat $Pr_G(A/B)$ as the basic object? Why not some other form of probability? Why not, for example, a probability distribution over initial conditions? The first reason is that the idea of a probability distribution over initial conditions is problematic. Initial conditions occur only once, and a probability distribution over initial conditions is *prima facie* empirically empty. $Pr_G(A/B)$, by contrast, gets empirical content from a probabilistic link to frequencies in finite actual ensembles.[xxi] The second reason for treating $Pr_G(A/B)$ as basic is that it cannot be generated from a distribution over initial conditions in indeterministic contexts. When dynamical evolution multiplies the number of possibilities, an initial distribution will not determine a final one. So it is a special feature of deterministic laws that they can generate a distribution at later times from a distribution over initial conditions. This is a good reason for talking $Pr_G(A/B)$ as the more general and basic form.[xxii]

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 probabilistic 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. 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.

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.[xxiii]

All of this raises the question of why the interpretation of chance has operated under the dogma that chance is physically fundamental. It's not just that probabilistic assumptions work behind the scenes in deterministic contexts. Even where probabilities are quite explicitly being used, there is confusion about their status. Because probabilities are only needed in the deterministic case when there is historical ignorance, it is easy to slip into thinking that the probabilities represent facts about our epistemic states. That's not right, what they represent is statistical facts about migration patterns of typical ensembles in the general population, and we rely on these general facts when our specific knowledge gives out. The most important reason one suspects, however, is a holdover from a conception of physical law built on the Newtonian example, a conception of physical law that conceives dynamical evolution as a species of causal production, so that dynamical laws tell us how one state of the universe produces the next in a manner that is temporally asymmetric and involves ideas of causal sufficiency. In a relativistic setting, this conception of laws doesn't fit. Laws are more naturally thought of as global constraints with no intrinsic temporal direction and no special connection with time. The distinction between lawlike and contingent features of the world is central to any physical theory, but it's an accidental and quite special feature of the deterministic case that contingencies can be isolated on a space-like hyperplane, so that once we fix the state of any such hyperplane, the laws determine everything else. Fix the state on any hyperplane – past, present, or future – and everything else follows. From a relativistic perspective, there is no reason to expect the distinction between lawlike and contingent features of the manifold to line up so nicely with the division into space and time. In an indeterministic setting, contingency is distributed throughout space and time and from a relativistic point of view seems a much more natural arrangement. It seems that it should be regarded as that a surprising accident – something that needed to be explained - if the world *did* turn out to be deterministic.

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 is present but remains largely below the surface in a deterministic setting. accepting this prompts a methodological reorientation, shifting attention from chance to Pr_G. This shift of attention has implications for how we think of physical probability. First, unlike chance, Pr_G 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. Second, we no longer have to think of there being two distinct types of physical probability; the microcanonical probability measure invoked by statistical mechanics and chance. We have a single measure from which both can be derived.[xxiv] Finally, rejecting the dogma that chance is physically fundamental resolves the interpretive dilemma that chance suffers by focusing interpretive attention not on objective, single case, dynamical probability, but on a form of general statistical probability. Statistical probability provides a much more intuitive entry point into the circle of probabilistic concepts.

Applications

I have spoken in non- relativistically for intuitive ease, but I want to pause to say how to translate into a relativistic setting and mention a couple of applications. To translate into a relativistic setting, we identify systems with world lines, define chances at points, and substitute 'contents of past light cone' for 'history'. So now we have

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

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. 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. Let's see how this works. Suppose I'm an observer in an Everettian Universe about to carry out a spin measurement. For simplicity, we'll assume that there is actual splitting into separate spatiotemporally disconnected branches of the universe, and that there is no evolution between the time of the measurement and observation of a result. There are two possible observations for a post-measurement situated agent, each with some non-zero probability, and the observation of a particular result by such an agent will turn out to be a chancy matter (not determined by the laws, conditionalized on her back light cone).

Another nice application of specialized interest is that $Pr_G(A/B)$ gives us a natural intepretation of the notion of typicality that plays a central role in the Goldstein, Zanghi, and Durr (GZD) version of Bohmian Mechanics.[xxv] GZD explicitly disavow a construal of typicality in terms of probability on the grounds that it is empirically empty to talk about the probability of initial conditions. As Craig Callendar writes

"An instinctive negative reaction to assigning probabilities to initial conditions in a deterministic universe is natural and probably even healthy. DGZ are understandably reluctant to dub the universe probable, for it invites quasi-theological pictures of supernatural beings picking the universe out of a big urn."[xxvi]

$Pr_G(A/B)$ generates a probability distribution over initial conditions, but it gets its empirical content from the constraints it places on observed statistics in smaller than universe-sized samples.

A couple of objections are worth addressing before closing. The first objection is that **Def** tells us that chance the chance of e at p is the probability obtained from Pr_G by conditionalizing on historical information accessible from p. But nobody ever in fact possesses full historical information, so chances are epistemically inaccessible in practice, if not in principle. How, then can they guide belief in the manner dictated by PP? The response is that PP correctly captures the conceptual connection between chance and belief, but what we apply in practice is a generalization of PP that accounts for historically based ignorance by forming a mixture of chance functions obtained by conditionalizing on (appropriately weighted) epistemically possible histories. Belief, that is to say, is guided by the chances if they are known, and our best estimate of the chances if they are not.[xxvii]

There is an interesting addendum that arises from a counter-response due t0 Alan Hajek. Hajek observed that the response threatens to undermine the answer that was given earlier to the question of why Chance(e) was peculiarly suited to play the role carved out by PP in guiding belief. Consider in particular, the trances which are obtained from Pr_G by conditionalizing on all of history: Trance(e)=pr_G(e/all of history). The reason for denying that these guide belief in the manner dictated by PP was that if we don't have crystal balls, we don't in general know what the trances are. But once we allow that in practice we don't know what the chances are, either, that opens the door for arguing that Trances do as good a job guiding belief as the chances.[xxviii] Trance(e) agrees with Chance(e) on past events and it does at least as well as the chances on future events. It is certainly true that if we know the Trances we ought to adopt them as our credences, no matter what other information we have, since the Trance(e) assigns probability 1 to all and only truths. If the only objection to adopting the trances as credence was that we don't in general know what the trances are (indeed we are precisely as ignorant of the trances are of the future), we have opened the door to saying that we accommodate ignorance of the Trances by forming a mixture of Trances obtained by conditionalizing on epistemically possible futures.

This is correct. What is interesting, however, is that provided we use the same weights over epistemically possible futures in forming our best estimate of the trances and chances, our best estimate of the future Trances will be quantitatively indistinguishable from our best estimate of the chances. To say that our best estimate of the chances will be indistinguishable from our best estimate of the trances, however, is in its turn to say that in absence of information from the future, our best estimate of the chances is also our best estimate of the truth, since Trance(e) assigns probability 1 to all and only truths. And that is precisely the result that we wanted. The chances and the trances come apart only in the presence of crystal balls. Recall that PP said that provided that we have no crystal balls or magical sources of information from the future, we should adopt the chances as our credences. And this just reinforces our explanation of the epistemic role of chance. We see now that we can derive a slightly more precise version of PP from the analytic and unqualified claim that truth should guide belief by adding a substantive, general claim about our epistemic position.

■If you know the truth, no matter what other information you have, adopt it as credence.
■The only specific information a situated human agent has about the world is information about past events.
■If you know the chances, provided you have no magical information from the future, adopt them as credence.

Truth reduces to chance when we form a mixture over epistemically possible futures.

The second objection charges that we've reduced one form of probability to another, but haven't made any progress explicit reduction of probabilistic facts to non-probabilistic ones. To this I reply as before that the goal wasn't to reduce probabilistic facts to non-probabilistic ones, but to find a stable interpretation of chance. Physics can and, in my view should, leave Pr_G unreduced, treating it on a par with other physical modalities. To do so leaves entirely open whether Pr_G can be reduced to non-probabilistic facts, treated at metascientific level as a form of idealized subjective probability, or accepted as an objective and fundamental feature of the world. It's a separate question, but one to which a proper understanding of which forms of probability are physically fundamental is a prelude.

---

[i] I am grateful to audiences at the Arizona Ontology Conference, the Perimeter Institute for Theoretical Physics in Canada and Australian National University, and especially Alan Hajek and Guido Bacciagaluppi for insightful comments.

[ii]Statistical mechanics did bring probabilities into physics, but not at the fundamental level and not in a way that challenged the strict necessity of the fundamental laws.

[iii] Lewis regarded the Principal Principle as the sole constraint, believing that it told us "all we know about the concept of chance". There is controversy about the correct form of the principle. The differences won't matter for our purposes. But see Hall (1994), Thau (1994), Roberts (2001), Vranas (2002) for proposed revisions, including Lewis (1994). I defend the original version of PP in Ismael (2007). The form adopted here is the one originally presented by Lewis and defended to the end of his life as capturing the epistemic role of chance, though he eventually adopted a closely related principle (Hall's NP) as a revisionary proposal to make it compatible with Humeans Supervenience (see note ix).

[iv]In logical terms, we can say that "Pr" is a variable-binding operator, binding the "x" in "Pr (Bx/Ax)". I've suppressed the variables in the text.

[v]In this case, there is the general probability (a roll of the dice comes up double sixes/ the dice are fair), or that (a roll shows double sixes/it occurs on a Monday), or that (a roll shows double sixes/Sam says it will show double sixes). One very striking example of the lack of clarity about the distinction between general and single case probabilities is that the term 'Born Probability' is used, without disambiguation to refer both to the general conditional probabilities of the form pr(a/Ψ) derived from Born's Rule and to the single-case unconditional probability of a for a particular system in a state Ψ. Ask a physicist which one he means, and you're more likely than not to get a blank stare. It makes no difference when one is calculating because the general, conditional probability of (a/Ψ) the single-case unconditional probability of some particular a in the future of a system in Ψ are quantitatively undistinguishable. But it makes a difference to interpretation. The tow have a different logical form, they bear different relations to categorical facts, they treat past events differently, and so on

[vi] Note that the t, here, is the time at which the chance is *assessed*, rather than the time at which the event is slated to occur.

[vii] Thanks to Alan Hajek.

[viii] 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. 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 chances were but didn't think that they should guide his expectations, 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.

[ix] The interpretive dilemma distils a rather large body of literature to bare bones. For historical reasons, a good part of the philosophical discussion of chance is preoccupied with the compatibility of chance with Humean Supervenience (HS). HS is an independent metaphysical commitment, and it is not itself a constraint on the interpretation of chance. Lewis himself, the original defender of HS, held that the viability of HS hinged on the existence of an account of chance compatible with it, rather than the other way around.

[x] This is going to raise questions about the Best Systems Analysis of theories (BSA) which purports to provide truth conditions for statements about probability and laws and all of the apparently modal implications of a physical theory in terms of patterns in the manifold of categorical fact. It turns out, on this account, that to say that a statement L is a law is to say that the best overall systematisation of actual fact treats L as a law. And to say that it the chance that a certain photon passes a filter is .99 is to say that the best overall systematisation of actual fact entails that the chance that the photon passes the filter is .99. The

BSA is, in my view, best viewed as a metascientific view. If the BSA is going to be successful, it has to be able to recognize that even though the acceptance conditions of a theory can be given in terms that advert only to the pattern of actual events, statements about law and probability have modal implications that can't be cashed out or reduced to claims about actual events. Getting the internal logic of a theory right depends on recognizing these modal implications.

[xi] Accounts that identify chances with propensities, explicitly denying that propensities are grounded in intrinsic categorical properties seem to be doing that.

[xii]See Lewis (1994)

[xiii]There are many excellent critical discussions and surveys of the standard influential interpretations of chance. See Pollock (1990), Strevens (1999), and Callendar (ms). Hajek provides an especially nice general survey of interpretations of probability. One of the difficulties in assessing these views is that some of them (propensity interpretations) are clearly intended as interpretations of single-case probabilities but others (e.g., frequency views) are better understood as interpretations of general probabilities. This confusion runs deep in the literature, and there's so little explicit clarity on the matter I gave up on attempts to gothrough and sort out intentions.

[xiv] I remain agnostic on whether the backward probability $Pr_G(A/B)$ where B occurs after A is essential. In an indeterministic setting, the backward probabilities may not be given. In a deterministic setting, they follow from the forward probabilities. See Bacciagaluppi.

[xv] This is a qualified conclusion. One can imagine laws that retain their predictive power even in the absence of precise knowledge, but the Newtonian dynamical laws do not. So the claim is that global determinism *by itself* places very weak constraints on the evolution of local subsystems of the universe. For examples of laws that do retain predictive power even in the absence of precise knowledge, consider laws that entail that any system takes the shortest path through phase space at a fixed speed to a final state S, no matter what it's starting state and in a manner that is entirely immune to influence. To multiply examples, one just has to limit interaction and influence.

[xvi]The reference is to: "Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective positions of the beings which compose it, if moreover this intelligence were vast enough to submit these data to analysis, it would embrace in the same formula both the movements of the largest bodies in the universe and those of the lightest atom; to it nothing would be uncertain, and the future as the past would be present to its eyes." (Laplace (1812-1820)).

[xvii] A little more precisely, let S be a system of interest, let C be a summary of everything that is known about S, and let **A**=$\{A_1, A_s, A_3, \ldots\}$ be an unknown variable with values $\{A_1, A_s, A_3, \ldots\}$. Randomising over **A**-values, holding C fixed, involves assigning each value of A the probability that a random pick from systems that satisfy C will yield that value. But that is just to say that it involves assigning the probability distribution over $\{A_1, A_s, A_3, \ldots)$ generated by $Pr_G(A_n/C)$.

[xviii]To derive these higher-level migration patterns for any distribution, it suffices to say how a flat distribution evolves.

[xix]Although it is still counter to orthodoxy, the idea that objective probability may be not just a quantum mechanical phenomenon is gaining some currency. Loewer is the strongest advocate. Loewer uses the term 'chance' to refer to any form of objective physical probability. I have reserved it to refer specifically to single-case quantum probability.

[xx] The recent debate between Maudlin and Loewer about the nature of laws is instructive in this regard. See Maudlin (2007a and 2007b) Loewer (2001 and 1997). Both agree that laws are physically fundamental structures. As do, for example, David Armstrong (1983) and Bas van Fraassen (1990). But they hold a wide range of views about the nature of laws. Maudlin denies that there is any physics-independent platform from which to raise questions about ontology, and regards physically fundamental structures as ontological bedrock. Loewer follows Lewis in thinking that laws (and everything else that exists) reduce to patterns in the Humean manifold of events. Armstrong holds that laws are relations between universals. One might just as well accept that they are primitive physical postulates and hold that they are convenient fictions, nodes in an uninterpreted calculus for deriving predictions about observable structures, or the real structures that underlie and explain the world's observable properties.

[xxi] See Durr, D., Goldstein, S., Tumulka, R., and Zanghì, N. (1992 and forthcoming), and Callendar (webarchive). Also, Price.

[xxii]And it is a special feature of bideterministic laws that they can generate a distribution at one time from a distribution at any other. The Newtonian laws are bideterministic

[xxiii]See Held (webarticle) for an excellent discussion of hidden variable results.

[xxiv] Classical statistical mechanics consists of three postulates: 1) the fundamental dynamical laws, 2) a uniform probability distribution the micro-canonical distribution- over the possible phase points at the origin of the universe, and 3) a statement characterizing the origin of the universe as a special low-entropy condition.(this is using Albert's formulation (Albert, 2000)) $Pr_G(A/B)$ is equivalent to the microcanonical probability distribution. And if we put a low entropy initial state in for B, it will give us the thermodynamic laws. It's not a nice feature of statistical mechanics that we don't get the right statistics unless we conditionalize on a low entropy starting state. But – for better or worse, and for reasons that this is not the place to enter into – that's the way the theory is organized.

[xxv] Durr, D., Goldstein, S., Tumulka, R., and Zanghì, N. (1992 and forthcoming).

[xxvi] Callendar (webarchive), p. 23.

[xxvii]See Ismael (2007) for fuller discussion.

[xxviii]The arguments that Pance(e)=pr_G(e/pre-2001 history of Australia) and $Fance_t$(e)=pr_G(e/post-t history) don't guide belief remain in place. Pance would have us assigning crazy credences, and Fance(e) would have us assigning credences other than 1 (or 0) to events we know have occurred (or have failed to occur).