**The Range Conception of Probability and the Input Problem**

John T. Roberts

draft of June 24, 2013

**Abstract:**

Abrams, Rosenthal, and Strevens have recently presented interpretations of the objective probabilities posited by some scientific theories that build on von Kries’s idea of identifying probabilities with ranges of values in a space of possible states. These interpretations face a problem, forcefully pointed out by Rosenthal, about how to determine ‘input probabilities.’ I argue here that Abrams’s and Strevens’s attempts to solve this problem do not succeed. I also argue that the problem can be solved by recognizing the possibility of laws of nature that manifest themselves not as universal regularities, but instead as constraints on relative frequencies.

**Keywords:** probability, range conception, laws of nature, method of arbitrary functions, frequency, macroperiodicity

Recently, a number of authors have articulated interpretations of the macro-level probabilities found in non-fundamental sciences (such as biology, statistical mechanics, and the theory of games of chance), consistent with the assumption of determinism, that are descended from von Kries’s (1886) view of probabilities as sizes of regions within a logical space. Abrams (2012) defends such an interpretation under the name ‘mechanistic probability’; Strevens (2011) defends one under the name ‘microconstant probability.’ Following Rosenthal (2012), I will call the general approach the ‘range conception’ of probability. There are important differences among the proposals of these authors, but for present purposes I will focus on what they have in common. As both Abrams and Strevens acknowledge, and as Rosenthal points out with great force and clarity, the range conception faces an important challenge that may be called ‘the problem of input probabilities’ (though Rosenthal does not use this term). Abrams and Strevens have each offered solutions to this problem, but I will argue here that neither solution is successful. However I will argue that there is a simple and plausible way of extending their proposals that does solve the problem. Supplemented with this solution, the range conception is a very attractive way of interpreting macro-level probabilities.

**1. The Range Conception of Probability**

Consider a repeatable kind of process that results in two or more different kinds of outcome of interest. (For example, flips of a certain type of coin, with the outcomes ‘Heads’ and ‘Tails’; or, the lives of insects bearing a certain phenotype, with the outcomes ‘fails to reproduce,’ ‘has one offspring,’ ‘has two offspring,’ etc.) Suppose that the process is completely deterministic, so that a complete specification of the initial conditions – that is, the conditions at the time of the onset of a token instance of the process – together with the prevailing laws determines the state at the end of the process. Suppose that the state at the end of the process is sufficient to determine which outcome of interest has occurred. Despite the determinism, it seems that it may be useful to assign probabilities to the outcomes. (At any rate, we seem to have no trouble doing so in the cases of coin-flips and reproductive success, even though for all we know they might be deterministic.) The range conception takes this appearance seriously, and offers a realist analysis of the probabilities in question, grounded in the dynamics of the system.

Here is the basic idea: Suppose we have a process-type we are interested in, and a set of possible outcomes of the process we are interested in. Imagine the initial-state space for the type of process in question – a space each point of which represents one complete set of initial conditions for a possible instance of the process-type. It is intuitively helpful to imagine this space two-dimensional – a portion of a plane, for example. Each point represents a maximally specific way the process could start out. If the process is deterministic, then each one of these points is destined to lead to a particular outcome. Assign a color to each possible outcome, and imagine each point colored with the color corresponding to the outcome it is destined to lead to. Depending on the dynamics of the system, it might turn out that the result is a great swirly marbled surface, with regions of all the colors mixed together well. Now imagine that you look at a small patch of this great plane – quite small, but not too small (we can worry about what precisely this means later) – and measure the fraction of its area that consists of each color. In Strevens’s terms, what we are measuring is the *strike-ratio* of each color within our patch.[[1]](#footnote-1) If the colors are sufficiently ‘well-mixed’ throughout the plane, then it might be that no matter where you take this appropriately-small patch, these fractions will take approximately the same values. If this condition is satisfied, then we may say the dynamics of the process is *microconstant*.[[2]](#footnote-2) Suppose that this is so; then there is some fraction that is approximately equal to the strike ratio for a given outcome in every appropriately-small patch of the plane; we identify this fraction as the probability of that outcome for this type of process.[[3]](#footnote-3) If the dynamics are not microconstant, then we deny that these probabilities are well-defined at all.

Why is this a good idea? Intuitively, we may think of the small-but-not-too-small patches as the regions that are as small as, or a bit smaller than, the smallest regions within which we can effectively locate the initial state of a given token process, either by measuring it or by manipulating it. When we know that a process of this type is about to start, then – or, when we cause such a process to start – the most precise information that we can have about it is that it is located within a patch of this size. So if we want to try to predict what the outcome will be, the best information we have to on is that the trajectory of this token process will be as one point in a space x% of which leads to outcome A, y% of which leads to outcome B, etc. This means there will be a certain sort of lack of predictability about any single case, but at the same, it is plausible that in the long run, we should expect the frequencies of these outcomes to be pretty close to x% for A, y% for B etc. In short, this is just the sort of behavior that we ordinarily associate with a process that has a probability of x% of yielding outcome A, y% of yielding outcome B, and so on.[[4]](#footnote-4) On the other hand, when the strike-ratio is not constant, we do not have this same lack of predictability. There will be regions of the plane where outcome is A is far more common than x%, and regions where outcome Y is far more common than y%; by finding out which of these regions a given token process’s initial state lies in, we could make a far better prediction than you would think possible in a process governed by probabilities of x% for A and y% for B.

What is more, in intuitively simple cases involving gambling devices, this approach seems to lead to the right answers for what seem like the right reasons. For example, consider the familiar example of the wheel of fortune discussed by Poincaré.[[5]](#footnote-5) A wheel is divided into a large number of small compartments of equal width, alternating red and black; a ball is spun round the wheel by a human croupier with enough force to around the wheel many times until it finally comes to rest in front of some compartment. Suppose that the final position of the ball is determined entirely by the speed initially imparted to it. Then the device is fully deterministic, but nevertheless it seems correct to ascribe probabilities of ½ to each outcome. Human croupiers (and similar devices) cannot control their spin speeds with complete precision; the best they can do is locate these spins within ranges that contain roughly equally many speeds that would lead to Red outcomes as would lead to Black outcomes. The range conception yields the same probabilities for what is essentially the same reason.

What I have sketched in this section is the basic idea behind the range conception of probability; it undergoes a lot of refinement before it evolves into either Abrams’s or Strevens’s proposal. In the next section we will look at one aspect of that refinement.

**2. The Problem of the Input Probabilities**

One obvious problem, as Rosenthal (2012), Abrams (2012) and Strevens (2011) have pointed out, is that it seems to take for granted what we might call an ‘input probability distribution.’ Suppose we know that within any suitably small region of the space of possible initial states of some process-type, the points that would lead to outcome A take up approximately fraction r of the room; why should that justify us in thinking that the probability of outcome A is r? Well, if the underlying micro-dynamics are deterministic, then the probability of outcome A should be equal to the probability that the process starts out in one of the initial states that inevitably lead to A. And those initial states make up a fraction r of the whole space of possible initial states. So it appears that we need to assume that the probability of the process beginning in one region of the state space is proportional to the size of that region. But what justifies us in assuming that probability is distributed over possible initial states in this way?

There are really two problems here; which one is most relevant depends on how you want to use the range conception. If you want to use it to give a metaphysics of objective probability – by which I just mean an interpretation of the objective probability statements found in science – then the problem is that you have taken for granted that there is such a thing as the distribution of probabilities over initial states (i.e., the input probability distribution), and you have not given any interpretation of *that* probability distribution. Rosenthal expresses this when he writes about the worry that this approach to probability

merely describe[s] certain interesting cases of the transformation of probability distributions, but for this very reason cannot possibly be used for an analysis of probability. Given an initial-state space, and given a probability distribution on it, the outcome probabilities are fixed, but interpretational tasks of any sort are simply pushed back from the latter to the former probabilities… [T]he situation is still ‘probabilities in – probabilities out.’ (Rosenthal 2012, p. 230)

On the other hand, perhaps your ambition is not to give an interpretation of probability – maybe all you want to do is describe ‘certain interesting cases of the transformation of probability distributions’; maybe you just want to do what Strevens called ‘physics of probability’ by contrast with ‘metaphysics of probability.’[[6]](#footnote-6) Still, there is a problem. The macro-level probabilities that we get out of the range conception depend not only on there being well-defined input probabilities, but also on certain assumptions about them. In particular, as we saw above, it appears that we must assume that the probability that the process will begin within any given region of the state space is proportional to the size of that region. This is an appealingly simple and general assumption about the form the input probability distribution should take. But what justifies us in assuming that it is true? After all, there are many different possible ways of distributing probability over space; the distribution that assigns probabilities proportional to sizes of regions is only one of them.

 But perhaps we do not really need to assume that the probability of any region of the initial-state space is proportional to the size of that region; perhaps a weaker assumption will do. Consider again the wheel of fortune. Its initial-state space is just a portion of the real line, containing the points corresponding to the range of possible initial spin-speeds the croupier could impart to the wheel. As before, imagine the points colored with the color corresponding to the result, Red or Black, that they inevitably lead to; so the line consists of very small, alternating red and black intervals. Suppose that the input probability density is approximately constant over every interval that is small enough that it contains at most two of the adjacent alternating red-and-black-producing intervals. Then for each such adjacent pair, the probability of the outcome Red, given that the spin speed starts out somewhere within the range represented by that pair, will be approximately equal to the strike-ratio for Red. Since the same is true for every pair, and these pairs exhaust the entire initial-state space, it follows that the probability of Red is equal to the strike-ratio for Red (and similarly for Black). Therefore, to justify the identification of probabilities with strike-ratios, it is not necessary to assume that the input probability distribution assigns probabilities to regions in a way proportional to the sizes of those regions; all we need to assume is that this is approximately true within any sufficiently small sub-space of the initial-state space. In other words, we need only assume that the input probability density is sufficiently smooth: It must remain approximately constant over any sufficiently small region.

Thus, the input probability could be any old probability distribution that is sufficiently smooth. This maneuver was called ‘the method of arbitrary functions,’ by Poincaré, who also argued for the stronger conclusion that all we need to assume is that the input probability distribution is continuous.[[7]](#footnote-7) Here he seems to have overstated his case. His argument involved looking at what happens in the limit as the size of the red and black compartments approach zero (and thus, the length of the alternating red and black segments of the initial-spin speed line approach zero). In that limit, the property of remaining approximately constant over any segment that is no longer than a few of the alternating red and black segments is equivalent to the property of continuity. In the real world, however, the compartments always have some finite size, and we are interested in probabilities for things like games of chance that are far from the limit.[[8]](#footnote-8) So we require our input probabilities not only to be continuous, but also not to be ‘bumpy’ in the relevant, objectionable way. In Strevens’s terms, we must require them to be *macroperiodic*; this means that the probability distribution must be approximately constant over any sufficient small interval.[[9]](#footnote-9)

**3. One Strategy for Solving the Problem: Appeal to Frequencies**

As we have seen, the problem facing the range conception has two parts: First, to supply a metaphysical interpretation of the probability distribution over initial spin speeds, and second, to justify the assumption that that probability distribution is macroperiodic. One way to address the first problem would be to interpret the input distribution as an objective chance distribution, and rely on one of the existing theories of the metaphysics of objective chance to explain what an objective chance distribution is. Abrams, Rosenthal, and Strevens want to avoid doing anything like that, though; each of them is looking for a more metaphysically austere solution to the problem that allows the Kriesian machinery to do a bigger share of the work in our understanding of probabilities. For this reason, both Abrams and Strevens seek to understand the initial-spin-speed distribution as a frequency distribution: The idea is roughly that among actual cases in which a wheel of fortune gets spun by a human croupier – and also among other physical processes that are similar in the relevant respects – the actual frequencies of the initial spin speeds are distributed in a way that approximates a macroperiodic density.

To be a bit more precise: Abrams characterizes input probabilities in terms of an initial-condition distribution for a given ‘causal map device.’ A causal map device is a stable device (such as a wheel of fortune) that repeatedly undergoes a process in which an initial state leads to some final state within an algebra of outcomes. He defines the input probability distribution for processes involving a given such device D as the distribution of a set of ‘far-flung frequencies,’ which he defines as ‘frequencies in a large set of all and only those natural collections of inputs to actual devices D\* (similar to D at least in having the same input space) within a large spatiotemporal region around D,’ where a ‘natural collection of inputs’ with respect to D is ‘a large set of all and only those actual inputs produced by a single source device G to a particular causal map device D\* during a single interval of time T\*, where D\* has the same input space as D.’[[10]](#footnote-10) In short, to find the input probability distribution for a given wheel, what you must do is first look throughout a large spatiotemporal region surrounding your wheel and find all of the relevantly similar wheels in that region with the same range of possible initial spin speeds. Then for each such wheel, you must find every croupier (or other suitable spinning device) that initiates a large set of spins of that wheel. Then you form the collection of all the spins of all those wheels made by all those croupiers (or other spinning devices). This is a complex procedure, but the basic idea is to find a collection of actual spins that is guaranteed to be very large and not to be gerrymandered, and use the actual frequencies within this collection as your input probability distribution. Of course, this will not general determine a unique distribution, since there are multiple large spatiotemporal regions containing D. But Abrams argues convincingly that for the purpose of defining his ‘mechanistic probabilities,’ this lack of uniqueness will turn out not to matter.

Strevens does things a bit differently: He eschews all talk of a probability distribution over initial conditions, and instead makes use of frequencies, refusing to call them ‘probabilities’ at all: ‘A probability distribution over initial conditions, whether it exists or not, has no place in my metaphysics of deterministic probability.’[[11]](#footnote-11) On his view, the probabilities of outcomes associated with a certain type of process should be regarded as well-defined just in case that type of process has a microconstant dynamics and the actual initial conditions of instances of that process are have a frequency distribution that is approximately macroperiodic. When these conditions are met, the probabilities are to be identified with the strike-ratios.

Abrams and Strevens both go on to argue that actual frequencies are good enough to ground the macro-level probabilities they are interested in.[[12]](#footnote-12) Their arguments differ in many details, but they have the same upshot: If a class of deterministic processes, leading to one of a certain range of outcomes, is such that (i) its dynamics is microconstant with respect to those outcomes, and (ii) the actual frequency distribution over the initial conditions of such processes are approximately macroperiodic, then it is reasonable to expect that in the long run, the relative frequencies of the possible outcomes will be approximately equal to their strike ratios. This shows that systems behave in the way we would expect systems to behave if they were governed by probabilities equal to those strike ratios: In that case, we should expect it to be very difficult to predict the outcome of a single process, but rational to predict that in long runs, the frequencies will be about equal to the probabilities. If we identify the probabilities with the strike ratios, that is exactly what we find to be so. Thus we have a case for treating the strike ratios as probabilities, and this case does not depend on taking for granted any sort of ‘input probability’ distribution except for an actual frequency distribution.

**4. Counterfactual Robustness**

However, both Abrams and Strevens agree that this frequency distribution cannot be the whole story. For they both agree that it is not enough for the actual initial-spin speed distribution to be macroperiodic: It must also be *counterfactually robustly* macroperiodic. This is important, because otherwise, their account of probability would share one of the most important failings of the actual-frequency interpretation of probability: It would imply that the probabilities might have been completely different if circumstances had been just a little different. For example, what if I were to sneeze right now? That would shake things up a little bit. So it would make a difference to the initial conditions of all the wheel-spins in my future-directed light-cone. Maybe in the actual world, the initial speeds of all those spins are macroperiodically distributed, but if they all got shaken up by my counterfactual sneeze, then who knows how they might have been distributed? If they had not still ended up macroperiodically distributed, then on the Abrams-Strevens approach, the probability of getting Red on a spin might have failed to be well-defined. But it is very odd to suggest that the probability of getting red on a spin of a wheel of fortune counterfactually depends on whether I sneeze or not.

So Strevens and Abrams each needs to assume that it is not just true that initial-spin-speed distribution is macroperiodic; they also need to assume that this fact is counterfactually robust: the distribution would still have been macroperiodic under a wide range of counterfactual suppositions – including, but not limited to, the counterfactual supposition that I sneezed a minute ago. Both Strevens and Abrams recognize this, and each of them attempts to justify the assumption. Their attempted justifications are different, but they have something important in common. I think that, unfortunately, the thing they have in common is a weakness that undermines both of their attempts.

One thing they both say is that a counterfactual of the following form is true, where P is some ordinary counterfactual supposition, for example that I sneezed a few minutes ago:

(1) If it had been the case that P, then it would most likely still have been the case that the initial-spin-speed distribution was macroperiodic

because:

(2) By far most of the relevant nearby possible worlds where P is true are worlds where the distribution is macroperiodic.

Here is Abrams on this:

I claim that a consequence of the way FFF mechanistic probability [Abrams’s terms for the probabilities that are generated by his version of the range conception] is defined in terms of bubbliness is that there is a natural sense in which there are more ways to preserve FFF mechanistic probability than to break it. … (Abrams (2011), p. 364.)

Here is Strevens:

I have endorsed a natural extension of Lewis’s truth conditions to conditionals of the form *If A had occurred at time t, then B would likely have occurred*, true … if B holds in *nearly all* the closest possible worlds in which A occurs at t. (Strevens (2011), p. 357.)

The closest worlds in which [a set of counterfactual coin-tosses] take place are therefore *nearly all* worlds with macroperiodic sets of initial speeds. The counterfactual conditional is true, then, as for similar reasons are the many other conditionals that diagnose the robustness of macroperiodicity. (ibid.)

I think that Abrams and Strevens succeed in making it plausible that (2) is true. But unfortunately, this does not support the claim that (1) is true. The thought that it does rests on an implausible assumption about the semantics of counterfactuals. The assumption is this:

(3) The conditional ‘If it had been the case that P, then it would likely have been the case that Q’ is true whenever Q is true at most of the closest possible worlds where P is true.

But this is not true in general. For example, suppose that there is a big machine at CERN, hooked up to a powerful particle accelerator; there is a red button on this machine, and if you push the button, then the machine will turn on, and it will flip a coin in a biased manner, such that there is a 99% chance that it will land heads and a 1% chance that it will land tails. If the coin lands heads, then the machine will then initiate a process in the particle accelerator, which will create very dangerous conditions that have a very high chance of causing all of the matter and energy in the universe to disappear. (So, please, nobody push the button.) Now, suppose that, counterfactually, somebody had pushed the button. What would have happened? Well, perhaps the universe would have been annihilated, but perhaps we would have gotten lucky and life would have gone on. Which outcome would be more likely? It seems obvious that there are vastly many more ways in which life could continue as usual then there are ways in which the universe could be quickly annihilated. After all, once everything is annihilated, there is really only one way for things to go on from there. But if the universe sticks around, there are uncountably many ways it might go on from there. So principle (3) implies that this is true:

(4) If the button had been pushed, then likely, the universe would not have been annihilated, and life would have continued as usual.

But (4) is clearly false. Whether a certain outcome would likely have happened does not depend on whether there are more ways in which it could have happened than there are ways in which it could have failed to happen; it also depends on how likely or unlikely those ways are. (Thus fact is the heart of my objection; nothing of importance turns on the details of the annihilation-device example, which is merely a brutally dramatic illustrative example.) But that means that (3) is false. One might reply that this is only because we are counting possible worlds in the wrong way: We should use a measure that gives more weight to worlds where more probable things happen than to worlds where less probable things happen. But to make this move is to appeal to a probability measure over possible worlds, which will require its own interpretation; this is just the kind of move that the range conception of probability and the method of arbitrary functions are intended to enable us to avoid having to make.

So Abrams and Strevens need another way of justifying their assumption that (1) is true.

**5. A Second Strategy[[13]](#footnote-13)**

Abrams might have another way: He goes on to suggest that for the case of the fortune-wheel, the counterfactual ways things could have been that would have messed up the frequencies in a way that undermined his mechanistic probabilities are more special than the counterfactual ways things could have been that would not havemessed up those frequencies. So there is a sense in which the former require ‘larger miracles’ than the latter:

Thus breaking FFF mechanistic probability via transfers of inputs requires very special combinations of such transfers, which are reasonably considered ‘larger miracles’ (cf. Lewis 1979). (Abrams (2012), p. 365.)

According to this idea, what makes the macroperiodicity of the input distribution counterfactually robust is not that so few nearby possible worlds are ones where the distribution is not macroperiodic, but rather that a bigger counterfactual perturbation (a ‘larger miracle’) is required to jostle the distribution in a way that makes it no longer macroperioidc than is required to jostle it in a way that leaves it macroperiodic.

Here is one crude way to implement Abrams’s thought here (not Abrams’s own way): Suppose that you adopt a god’s perspective on the world, and you are considering ways of tinkering with the frequency distribution for the initial-spin speeds of fortune-wheels. In the actual world, this distribution is approximately macroperiodic. What this means is that if you break up the whole initial-spin-speed line into suitably small intervals – or ‘pockets’ as I will call them here[[14]](#footnote-14) – the number of spins with an initial speed in one pocket will be approximately equal to the numbers of spins with speeds found in the adjacent pockets. Now suppose that you want to change the frequency distribution in a way that destroys this feature. This will require doing more than changing the initial speeds of some of the spins; the changes must be at least large enough to move a spin from one pocket into one the adjoining ones, or else the number of spins found in each pocket will remain the same, so the distribution will still be approximately macroperiodic. Furthermore, it will not be enough to move some of the spins from one pocket to an adjoining pocket, for if a relatively small number of such moves are made, the distribution will still be approximately macroperiodic. So it takes a counterfactual intervention that moves a much larger number of spins from one pocket to another – a ‘larger miracle’ – to wreck the approximate macroperiodicity of the frequency distribution that it does simply to change that distribution.

 On that way of understanding Abrams’s thought, it is subject to an important objection. When we think that the probabilities associated with macro-level devices such as fortune-wheels should be counterfactually stable, we mean that at the very least, they should remain the same under counterfactual interventions that involve making relatively small changes to the actual course of events – for example, having me sneeze on an occasion when I actually do not. But changes that are ‘relatively small’ in this sense are not the same things as changes that are small in the sense that they result in moving only at most a small number of initial spin speeds from one pocket to another. A counterfactual disturbance introduced at one spacetime location, however small it is, can be expected to have influences that propagate throughout the entire future-directed light cone of that location; even if my counterfactual sneeze is small, it can be expected to have at least some impact on the initial speeds of vastly many spins. In fact, if you wanted to intervene in the course of events in a way that changed the initial speed of one spin while leaving the initial speeds of all spins in its causal future unchanged, that would require an extremely delicate operation – arguably a much ‘larger miracle’ than the simple introduction of a counterfactual sneeze.

 However, there is another way to understand Abrams’s argument here that avoids this difficulty. Consider a typical counterfactual intervention that produces changes in the initial speeds of very many spins that are big enough to move them from one pocket to another. Again imagine the spin-speed line divided up into suitably small pockets, and this time imagine the pockets to be sequentially numbered. Presumably, there is no reason to expect that whether the speed of a given spin gets moved to a different pocket or not, or which direction it gets moved in, or how many ‘pocket-lengths’ it moves, to be correlated with whether it is actually in an even-numbered or odd-numbered pocket. (If I were to sneeze, that would shake up all the spins randomly.) So, on average, we should expect the even-numbered pockets to lose individual spins to their left-side neighbors and to their right-side neighbors in about equal numbers, and to expect the neighboring odd-numbered pockets to lose individual spins to their left- and right-side neighbors in about the same numbers.[[15]](#footnote-15) (This is simply a point about combinatorics: By far, most of the possible ways of rearranging a large number of spins by moving them from their actual pocket to an adjacent one will be as just described – just as most of the possible ways of writing a sequence of 10,000 1s and 0s have about equal numbers of 1s and 0s.) Similarly, on average, we should expect the pockets to lose individual spins to their neighbors two pockets down on either side at about the same rate, and at about the same rates as the other very nearby pockets do. And so on. So on average, the number of spins located within each pocket should change very little. Therefore, if the actual frequency distribution is approximately macroperiodic, we should expect that typical counterfactual distributions resulting from small interventions will also be approximately macroperiodic. There are possible counterfactual disturbances which would make the distribution become non-macroperiodic. For example, a counterfactual bump could have the effect of moving vast numbers of spins whose initial speeds are actually located in odd-numbered pockets into even-numbered pockets while moving few or no spins from even-numbered pockets into odd-numbered pockets. Such a bump would turn an approximately flat distribution into a very bumpy one. But such ways of rearranging the initial spin speeds are very rare within the space of all such rearrangements.

 This avoids the problem mentioned above, but it does so only by collapsing into a version of the strategy we considered in the preceding section. For it is not the case that there are no nearby possible worlds where I sneeze and as a result the frequency distribution is no longer approximately macroperiodic; it is just that such worlds are very rare within the space of nearby possible worlds. In other words, most of the nearby possible worlds where I sneeze are ones where the frequency distribution is still approximately macroperiodic. One could say that the unusual worlds here are not merely unusual but also extremely improbable in some sense. But that would require one to introduce a probability distribution over the nearby possible worlds, and we would need to interpret this probability distribution and motivate the assumptions one would have to make about it – the same difficulties that the range conception seeks to avoid by appealing to actual frequency distributions in the first place.

**6.**  **A New Way Forward**

The situation we confront can be summarized as follows. Abrams and Strevens argue that if, for systems of kind K, the dynamics are microconstant (so that the strike-ratios of the possible outcomes are well-defined), and the actual initial-condition distribution for systems of type K is approximately macroperiodic, then the strike-ratios of the outcomes will behave in the way we expect probabilities to behave: Though it will be very difficult to predict the outcome of any particular K-type process, it will be a very safe bet that in any large collection of such processes, the frequencies of the various possible outcomes will approximately equal their strike-ratios. This, they argue, justifies thinking of the strike-ratios in systems with microconstant dynamics with approximately macroperiodically distributed initial conditions as the macro-level probabilities governing those processes. Their argument is very plausible, but it runs into one problem: If it makes sense at all to speak of probabilities governing these processes, then the long-run frequencies for such processes should not be a fragile fluke; they should be counterfactually robust across a broad range of possible local modifications to the actual course of events. So the range conception of probability is acceptable only if it gives us some reason to think that the macroperiodicity of the initial-condition distribution is itself a counterfactually robust feature of the world. It has proven difficult to guarantee that this is so without either introducing a separate probability distribution over initial conditions (which effectively re-introduces the problem of interpreting the probabilities and justifying our assumptions about them – problems that the range conception was itself intended to solve for us) or else making implausible assumptions about the semantics of counterfactuals (such as that ‘Were A the case, C would be the case’ is true whenever C is true at most of the nearby A-worlds).

 There is a very simple and straightforward solution available. What we need is some way to guarantee that the fact that the initial-condition distribution over the systems for which range-conception probabilities exist is approximately macroperiodic is not a fragile fluke, but is rather counterfactually robust. So why not say that this fact is not merely a fact, but a law of nature? That is, why not say that in order for range-conception probabilities over a set of outcomes for processes of kind K to be well-defined, it must be nomologically necessary that the initial-condition distribution is approximately macroperiodic? This would give Abrams and Strevens exactly what they need: It would guarantee the counterfactual robustness of the facts about frequencies that ground the probabilities, and it would do so in a way that does not involve introducing a new probability distribution over initial conditions.

 There are three skeptical questions that this suggestion immediately provokes, however. First, how could there be laws of nature governing initial-condition frequency distributions? Second, even if there could be such laws, would it not be bizarre for there to be specific laws governing the initial condition distributions of, for example, spins of fortune-wheels? And finally, even if we agree that it is possible for there to be such laws, how could we be justified in believing that there are? This last question is pressing because, on the proposal under consideration, the kind of macro-level probabilities Abrams and Strevens are interested in are well-defined only if there exist such laws, so unless we could be justified in believing that there are such laws, we could not be justified in believing that there is a well-defined probability for getting a Red outcome in a spin of a fortune-wheel, which would be surprising. I will offer answers to these three questions in the following three sections.

**7. Could There Be Laws that Govern Frequencies?**

 To the first skeptical question, the proper reply is that it is of course still not clear what it is for there to be laws of nature at all, but there is no special problem about laws concerning frequency distributions. It might be a fact that:

(5) Every electron has charge –q.

This fact might also be a law of nature. Similarly, it might be fact that:

(6) The frequency with which electrons have x-spin up lies in the interval

(½ – e, ½ + e).

Might not this fact also be a law of nature? (If so, it would be a non-local law, since it would constrain the behaviors of all electrons in the universe as a class, instead of constraining what they do on particular occasions one by one. But by now, we are quite used to the idea of non-local laws.)

To answer this question, start by considering a different one: What makes the difference between (5)’s being merely true and it’s being a law of nature? Different theorists of laws give different answers. On a non-reductive realist account[[16]](#footnote-16), the difference lies only in the former’s having a certain elevated modal status, which results in its having great counterfactual robustness and explanatory power. There is no deeper account of what it is about the fact that gives it these special features; that is just the way laws are. In that case, it is not easy to see why there should be any barrier to (6)’s enjoying the special modal status as well.

On a contingent-necessitation account[[17]](#footnote-17), (5) is a law just in case there exists a higher-order relation of necessitation linking the property of electronhood to the property of having a charge of –q, such that it is metaphysically necessary that when two properties stand in this relation, whatever instantiates the first property instantiates the second one as well. If there might exist such higher-level necessitation relations among properties, then why could there not also exist a higher-order relation such that it is metaphysically necessary that when two properties stand in it, the frequency with which instances of the first property are also instances of the second one lies within some range?[[18]](#footnote-18) This second sort of necessitation relation is only minimally less simple than the first; in fact, the first is just a special case of the second. There is no obvious reason why the first sort of necessitation relation should be a metaphysical possibility whereas the second one is not. So on this approach to laws, it is also possible for facts like (6) to be laws.

On Lewis’s (1994) best-system account, what it means for (5) to be not merely true but also a law is that it finds a place in the system of true regularities that achieves an optimal combination of information content and simplicity. But there is no reason why a proposition of the form of (6) could not find a place in such an optimal system. Information about relative frequencies, no less than information about exceptionless regularities, can be very informative and at the same time far simpler than a complete catalogue of the worldly facts would be.

There are other theories of lawhood to consider, but enough has been said to show that it is plausible that whatever lawhood turns out to be, there is no principled reason why propositions like (5) can have it but propositions like (6) cannot.[[19]](#footnote-19) What really concerns us, though is laws of a more general type. Consider (7):

(7) For any charge-value q, every particle with q in an electric field **E** suffers a force of –**E**q

This illustrates a more general form that laws can take:

(8) For each property P of type T, every object that instantiates P also instantiates F(P)

where F is some function mapping properties to properties. Armstrong (1983) argues that the contingent-necessitation approach to laws can accommodate such ’functional’ laws by treating them as still-higher order relations of necessitation; T in (8) and ‘charge-value’ in (7) are both kinds of (first-order) properties, which get linked to other properties of properties by higher-order laws, yielding functional laws such as (7). Here is another possible form for such higher-order laws to take:

(9) For each property P of type T, the frequency with which instances of Q are instances of P is related in way R to the frequency with which instances of Q are instances of F(P)

where again F is some function mapping properties to properties. If there can be higher-order laws of the form (8), as well as laws about frequencies of the form of (6), there is no apparent reason why there could not also be ones of form (9): The lawhood of (9), like that of (8), amounts to there being a necessitation relation between two higher-order properties, which constrains the first-order necessitation relations that properties that are instances of those higher-order properties can enter into; thus, they are essentially the same kind of law. One instance of (9) is:

(10) For each r-sized interval of initial-spin-speed values J, the frequency with which spins start out with an initial speed in range J differs by no more than d from the frequency with which spins start out with an initial speed in an adjacent r-sized interval.

Thus, Armstrong’s treatment of functional laws such as (7) illustrates a pattern that can be used again to give a treatment of laws like (10) – which states the approximate macroperiodicity of a frequency distribution.

 On the other philosophical theories of lawhood we have looked at – the non-reductive realist account and the best-system analysis – it is far more obvious that it could be a law that a certain frequency distribution is macroperiodic. On a non-reductive realist account, any generalization might have the elevated modal status of lawhood. (Perhaps generalizations involving grue-like predicates need not apply, but no such predicates occur in the propositions about frequencies that we are interested in.) As the theorems of Abrams and Strevens show, the assumption of approximate macroperiodicity in a frequency distribution can have greatly informative implications about the actual course of events, and they are relatively simple, so there is no reason to think they could not find a place in the best system. In short, not only should we acknowledge the possibility of laws about frequencies, but we should also recognize the possibility of laws to the effect that certain frequency distributions are approximately macroperiodic.

**8. Could There Be Laws about Such Specific Kinds of System?**

This brings us to the second skeptical question: Even if there could be such laws, why should we think such laws might pertain specifically to such narrowly-defined sorts of systems as spins of a fortune-wheel? The answer is that we need not think that. There might be more general laws governing initial-condition frequency distributions in macroscopic systems of which these more specific laws are consequences. For example, perhaps it is a law that the frequency with which systems of particles are found (at any time) within a region of size R in their phase space differs from the frequency with which such systems are found in any of the adjacent R-sized regions by no more than a certain small fraction. Things need not be that simple; there are many more possibilities to consider. But the important point here is that the law that guarantees that the initial spin-speed distribution for wheels of fortune is approximately macroperiodic need not be a fundamental law; it might be a consequence of a more general law governing a more general class of frequency distributions. Moreover, in order to be justified in believing the less fundamental law governing frequency distributions, we do not need to know the form of the more fundamental law from which it derives – just as in general, we do not need to know the true underlying fundamental laws in order to be justified in believing that we have discovered a phenomenological law.

**9. Why Should We Believe in Such Laws?**

This brings us to the third skeptical question. It is one thing for it to be metaphysically possible for there to laws of nature governing initial-condition distributions, constraining them to be approximately macroperiodic; it is another thing for there to be a good reason to think that there really are such laws in the actual universe. What could possibly be a good reason to believe in such bizarre laws? The question is pressing because, if there is no good reason to believe in such laws, then there is no good reason to believe that there are any such things as macro-level probabilities, as they are construed by the range conception supplemented with my suggestion about how to solve the problem of the input probabilities. The proposal would then turn out to be nothing more than a far-fetched speculation.

 I believe that this question is not as difficult to answer as it might first appear. The hypothesis that the initial-condition distributions are approximately macroperiodic is relatively simple and straightforward, and it[[20]](#footnote-20) predicts something we actually observe: namely, that in systems like wheels of fortunes, in which there are symmetries that we can exploit to find out about strike ratios, the frequencies in long series of trials tend to be approximately equal to those strike ratios. The fact that the hypothesis makes that successful prediction is good evidence for it.

But you might press further, and ask: ‘Okay, but we could derive this prediction from the hypothesis that the processes *are as a matter of fact* distributed in a way that is approximately macroperiodic. Why should we think there is evidence that this is a law of nature?’ The answer to this is very simple and straightforward, too: If it is a fact that the initial conditions of the processes in question are distributed in a way that is approximately macroperiodic, that is a very broad and relatively simple generalization about physical systems, universally true as far as we can tell, that is very rich in consequences for what happens in the physical world, and which we have discovered on the basis of inductive reasoning. Those are exactly the conditions under which we usually think we are justified in thinking we have discovered a law. Why should this case be special?

**10. Conclusion**

Familiar examples of putative laws of nature are universal regularities – facts to the effect that every system of a certain kind has a certain property or satisfies a certain equation. But there appears to be no reason not to allow that there might be laws of other forms, such as laws that constrain the values of relative frequencies or laws that require certain frequency distributions to have global properties, such as macroperiodicity. In fact, thanks to arguments due to Abrams and Strevens, we can see that the hypothesis that there are such laws elegantly predicts and explains many familiar classes of phenomena. In particular, it predicts and explains why so many kinds of large-scale system appear to behave as if relatively simple probabilistic laws governed them. So we have good reasons to believe there are such laws.[[21]](#footnote-21)

 This hypothesis enables the range conception of macro-level probabilities to solve the problem of the input probabilities in a very simple and straightforward way. We have well-defined macro-level probabilities exactly when we have a class of systems with a common state-space, whose dynamics are microconstant with respect to a set of possible outcomes, and whose initial conditions are constrained by natural laws to be approximately macroperiodic. This predicts that such systems will behave in the way we expect systems governed by probabilistic laws to behave. It also implies that this behavior is counterfactually robust, and explains why it occurs, at a minimal ontological cost (requiring only that we posit that there are such things as laws of nature), without taking for granted the existence of an input probability distribution. Thus, just as Abrams, Rosenthal and Strevens hoped, the Kriesian machinery does all of the work in accounting for the macro-level probabilities: No additional, more fundamental probability distribution at the micro-level need be called upon.[[22]](#footnote-22)

**References:**

Abrams, Marshall (2012). Mechanistic Probability. *Synthese*, 187(2), 343-375.

Armstrong, D. M. (1983). *What is a Law of Nature?* Cambridge, Cambridge University Press.

Bird, Alexander (2007). *Nature’s Metaphysics: Laws and Properties*. Oxford, Oxford University Press.

Carroll, John (1994). *Laws of Nature*. Cambridge, Cambridge University Press.

Dretske, Fred I. (1977). Laws of Nature. *Philosophy of Science*, 44, 248-268.

Lewis, David (1979). Counterfactual Dependence and Time’s Arrow. *Nous* 13(4). 455-476.

Lewis, David (1994). Humean Supervenience Debugged. *Mind* 103, 473-490.

Maudlin, Tim (2007). *The Metaphysics within Physics*. Oxford, Oxford University Press.

Poincaré, Henri (1952). *Science and Hypothesis.* New York, Dover.

Roberts, John T. (ms). Laws about Frequencies. Unublished manuscript. Available online at http://philosophy.unc.edu/people/faculty/john-t.-roberts

Rosenthal, Jacob (2012). Probabilities as Ratios of Ranges in Initial-State Spaces. *Journal of Logic, Language, and Information* 21(2), 217-236.

Strevens, Michael (2003). *Bigger than Chaos.* Cambridge, MA., Harvard University Press.

Strevens, Michael (2011). Probability Out of Determinism. In Claus Beisbart and Stephan Hartmann (eds.), *Probabilities in Physics*, Oxford, Oxford University Press, 339-364.

Tooley, Michael (1977). The Nature of Laws. *Canadian Journal of Philosophy* 7, 677-698.

von Kries, Johannes (1886). *Die Principien der Wahrscheinlichkeitsrechnung.* Tübingen, Mohr Siebeck.

1. Strevens (2003), p. 52; (2011), p. 346. [↑](#footnote-ref-1)
2. This term is used in this way by Strevens (2011), p. 346; see also Strevens (2003), p. 53. Abrams (2012) uses the term in a similar way, though his definition, instead of quantifying over all suitably small patches of the state space, quantifies only over ones that contain points leading to all possible outcomes. He calls such patches ‘bubbles,’ and he makes his ’mechanistic probabilities’ well-defined only in cases where the input-state space is ‘bubbly,’ i.e. it can be partitioned into a large number of bubbles. These details are important to the main theorem that Abrams proves, but they will not matter much for our purposes, so I will neglect them henceforth. [↑](#footnote-ref-2)
3. See also Rosenthal (2012), p. 23. This will suffice to determine these probabilities only approximately, of course. But that should not worry us. We are talking here about probabilities of macro-level outcomes in macroscopic processes; perhaps these are defined only to certain degree of precision, as many other mathematical features of the macro-world are (e.g. the length of a coastline). [↑](#footnote-ref-3)
4. Rosenthal (2012), p. 23. [↑](#footnote-ref-4)
5. Poincaré (1952), pp. 201-203. [↑](#footnote-ref-5)
6. Strevens (2003), pp. 29-32. [↑](#footnote-ref-6)
7. Poincaré (1952), pp. 201-203. [↑](#footnote-ref-7)
8. See also Strevens (2011), 348-349. [↑](#footnote-ref-8)
9. Strevens (2003), pp. 54-55; (2011), p. 348. [↑](#footnote-ref-9)
10. Abrams (2012), p. 358. [↑](#footnote-ref-10)
11. Strevens (2011), p. 350. [↑](#footnote-ref-11)
12. Strevens gives an informal argument in his (2011), 352-353 and a more technical argument in his (2003). Abrams proves a theorem in the appendix to his (2012) in which he shows that there are strict limits on the difference between the actual long-run relative frequencies and the mechanistic probabilities. [↑](#footnote-ref-12)
13. I am grateful to Marshall Abrams for very helpful discussion of the issue addressed in this section. [↑](#footnote-ref-13)
14. These are no larger than any of the little alternating Red- and Black-producing intervals; this must be so if it is to be meaningful to speak of the distribution being approximately constant across any give pair of such intervals. [↑](#footnote-ref-14)
15. This is consistent with the possibility that spins with relatively slow initial speeds are more likely to get bumped into the next pocket than spins with higher initial speeds are, or vice versa; the only assumption required is that on average, approximately equal numbers will get bumped from pockets nearby to one another. [↑](#footnote-ref-15)
16. Such as Carroll (1994) or Maudlin (2007). [↑](#footnote-ref-16)
17. Such as Dretske (1977), Tooley (1977), or Armstrong (1983). [↑](#footnote-ref-17)
18. This suggestion is not unprecedented: Tooley (1977) admits the possibility of multiple necessitation relations, yielding laws of nature of different logical forms; Armstrong (1983), in Chapter 9, proposes that the necessitation relation comes in different quantitative grades in order to accommodate probabilistic laws. [↑](#footnote-ref-18)
19. One view whose influence has been growing lately is the dispositional essentialist view of laws defended by Bird (2007) among others. The most important difference between this view and the contingent-necessitation relation discussed in the text is that it says that the necessitation relations among properties are metaphysically necessary, rather than contingent, because they belong to the essences of the properties related. Nothing in the argument of the text depends on the modal status of the necessitation relations, so it is hopeful that the same considerations show that laws about frequencies are possible on an essentialist view of laws as well. [↑](#footnote-ref-19)
20. In conjunction with some well-confirmed auxiliary hypotheses, such as that the dynamics of the system is microperiodic. [↑](#footnote-ref-20)
21. I argue in my (ms) that the hypothesis that there are laws about frequencies can also be used to develop an interesting interpretation of objective probabilities in general. [↑](#footnote-ref-21)
22. For helpful feedback on earlier versions of this material, I am grateful to Marshall Abrams, Claus Beisbart, Jacob Rosenthal, and the other participants at the workshop “Johannes von Kries’ Conception of Probability, its Roots and Impact” at the University of Bonn in September, 2012. [↑](#footnote-ref-22)