

Short-Run Mechanistic Probability

Marshall Abrams

Department of Philosophy, University of Chicago

August 15, 2000

This paper sketches a concept of higher-level objective probability (“short-run mechanistic probability”, SRMP) inspired partly by a style of explanation of relative frequencies known as the “method of arbitrary functions”. SRMP has the potential to fill the need for a theory of objective probability which has wide application at higher levels and which gives probability causal connections to observed relative frequency (without making it equivalent to relative frequency). Though this approach provides probabilities on a space of event types, it does not provide probabilities for outcomes on particular trials. This allows SRMP to coexist with lower-level probabilities which do govern individual trials.

1. Introduction. A theory of higher-level objective probability should satisfy at least these rather loose desiderata:

1. There should be intimate relationships between probabilities and relative frequencies: Relative frequencies should usually be closer to probabilities when the number of trials is larger. The theory should be able to account for the intuition that objective probabilities explain relative frequencies. And these relationships between probability and frequency should hold for time scales relevant to everyday observations of nature, games of chance, etc. (A theory that tells us only about what happens after infinite time does little to support scientific realism.)
2. Nevertheless, the theory should allow relative frequencies to differ significantly from probabilities, perhaps quite often in cases in which only a few trials have taken place.

3. The theory shouldn't give rise to an inconsistency between higher-level and lowest-level probabilities.

As far as I know, we have no theory of probability of this sort. Hypothetical and actual frequency theories have fairly well-known problems. Single-case propensities plausibly arise only in quantum mechanics, if anywhere. Some philosophers have suggested that quantum mechanical indeterminacy routinely infects higher level processes, but I find their arguments unconvincing. Attempts to capture intuitions about objective probability by definitions in terms of subjective probability are instructive, but they can't support a strong higher-level realism.

Now, I don't claim to have a theory of probability of the sort I think we need. Instead I'll offer a sketch of a kind of probability, which I call "short-run mechanistic probability" ("SRM probability", or "SRMP"), which has the potential to satisfy these desiderata. This will in fact be a sketch of a sketch; a longer discussion can be found in my dissertation. I hope to convey how plausible the idea of SRMP is, and thereby convince you that the work required to pay its debts is worth doing.

What I'll do is describe an example of a type of argument known as "the method of arbitrary functions". I'll discuss some of the background assumptions needed to make the argument work. My suggestion then will be that the general principles that make the method of arbitrary functions work can be generalized and placed in a certain framework which will give us a theory of probability of the right sort.

2. The Method of Arbitrary Functions. Consider a simplified roulette wheel. Rather than spinning a ball around the edge of the wheel, we'll just spin the wheel and note the relative frequencies with which a fixed pointer points at this or that sector. Let the wheel have some even number n of sectors, alternately red and black, where the ratio between the sizes of any two adjacent sectors is let's say $1/2$; each red sector is $1/2$ the size of the black sector to its right. Then $1/3$ of the wheel is colored red; $2/3$ is colored black. Intuitively, the probability of red is then $1/3$ and the probability of black is $2/3$. And we'd predict that if someone spun the wheel many times, the relative frequencies of red and black would be near to these values. Why should this be the case?

A variant of a kind of argument known as the "method of arbitrary functions" (MAF), due originally to Poincaré (1912, introduction and §§90-92)

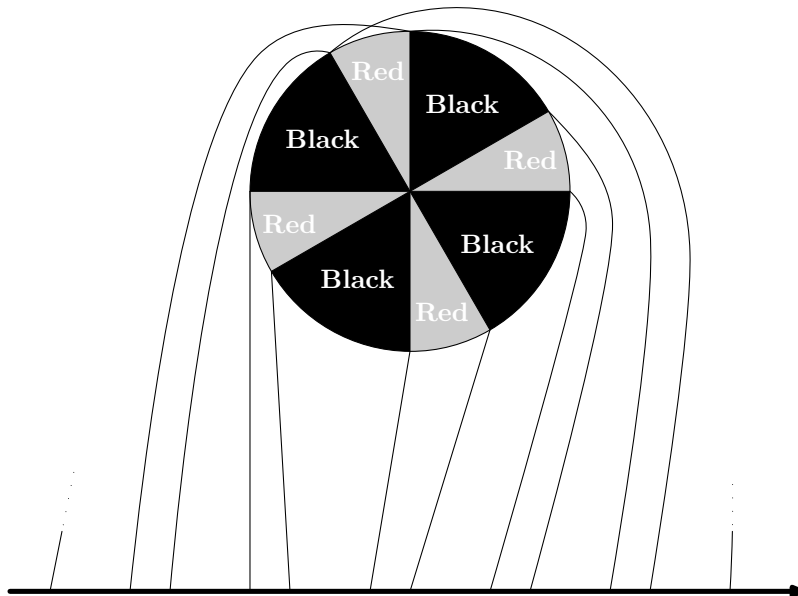


Figure 1: A portion of the map from velocity intervals to wheel sectors.

can explain the relative frequencies that such a roulette wheel would produce. Here I'll only give an intuitive argument which captures the flavor of the MAF though lacking generality and rigor. More careful arguments for the same conclusion are possible. What the argument shows is that in the case of the roulette, intuitively, “most” distributions of input spins will produce roughly the same set of relative frequencies for red and black outcomes. These relative frequencies will correspond to the relative portions of the roulette wheel covered by red and black. Poincaré applied arguments related to the one here to roulette wheels and to the distribution of asteroids. Other authors applied similar arguments to other cases.¹

I'll assume that once spun, the roulette wheel moves at a constant angular velocity; after a fixed amount of time we bring it to a dead stop. (A more realistic treatment would complicate matters without changing the basic situation, I believe.) We can divide up the initial angular velocities of wheel spins into intervals according to the color they would cause the wheel to stop on (figure 1). Two adjacent sectors of the wheel, a red one and black one,

¹See (von Plato, 1994, ch. 5) and (von Plato, 1983) for references.

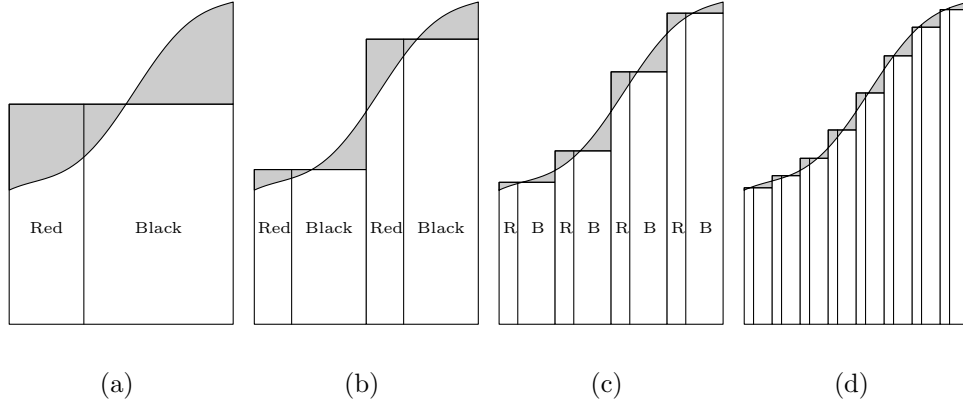


Figure 2: A portion of an approximately continuous velocity distribution.

correspond to two adjacent velocity intervals; for convenience call these “red” and “black” velocity intervals. We’ll start with a simple measure on velocity intervals, Lebesgue measure, the difference between an interval’s minimum and maximum velocities. I’ll assume that there are largest and smallest possible input velocities for a given case, so that the range of possible velocities is not infinite; we can then normalize the measure, scaling it by the total range of possible velocities.

Suppose we have a set of spins distributed between the various red and black velocity intervals. Suppose that there are very many spins, so that we have a reasonable approximation to a continuous distribution. (The curved line in figure 2(a) represents a portion of the distribution. The measure of velocity intervals is represented by width. The number of spins at a given velocity, or in a very small interval, is represented by the height of the curve. Area under the curve over a red or black interval represents the number of spins in that interval.)

Consider any pair of adjacent regions, a red one and a black one. Their widths have the ratio $1/2$, but the ratio between their areas will not generally be $1/2$. In figure 2(a), for example, the curve representing density of spins goes a little higher over the black region, giving it a larger area compared to the red region. It will be informative to compare this situation to one in which the combined number of spins for a pair of red and black regions is similar, but where the density curve is constant across both regions. I’ve put this new curve (a horizontal line) at about the average height of the original

curve over the combined region, though all that we really need is that the new line's height be in the same general range as the original curve over the two intervals. The new, flat curve is similar to the original curve in overall height, but it makes the ratio between red and black *areas* (the rectangular regions) the same as the ratio between red and black *widths*. Then the shaded regions in figure 2(a) show how much the alternative, rectangular red or black region's area differs from that of the corresponding original curve-topped region. In other words, the shaded regions show something like the difference between actual relative frequency and what relative frequency would be if it were equal to relative measure. Or to put it another way, since the ratio between the areas of the two rectangular regions is equal to the ratio between their widths, we can think of the shaded regions as what accounts for the difference between the ratio of the two original regions' areas and that of their widths. The shaded regions in effect represent the difference between on the one hand, the ratio between the actual relative frequencies for the two intervals, and on the other the ratio between their measures.

Consider what happens if we increase the number of sectors on the wheel, making each sector narrower. This has the effect of making the corresponding velocity intervals smaller as well. But the density curve just shows how often spins cause the wheel to stop at this or that point on the wheel's circumference, so the shape of the curve doesn't change as the widths of the regions get smaller (figures 2(b)–2(d)). Notice how the shaded regions which account for the difference between the width and area ratios get smaller as we increase the number of sectors on the wheel. And how a greater portion of the original regions is included in the rectangular regions which bear the 1:2 ratio. If we make the sectors on the wheel small enough, always keeping the 1:2 red/black width ratio constant, the shaded regions will be very tiny. Then the ratio of the area of each red region to that of its neighboring black region will be nearly 1:2. And since the ratio between the red and black areas for each pair of neighbors will be close to $1/2$, the ratio between the area of all red regions combined and the area of all black regions combined will also be close to $1/2$. Thus, when the sectors on the roulette wheel are small, making velocity intervals small in relation to the range of velocities, relative frequencies of red and black outcomes will be near to the proportions of red and black on the wheel. This conclusion is largely independent of the particular distribution of spin velocities. It will hold as long as the density curve representing numbers of inputs in small velocity intervals does not become too steep at any point. In other words, within very broad limits,

it doesn't matter if the person spinning tends to give faster spins or slower spins; the end result will be approximately the same.

3. What makes the MAF work? What is it about the roulette wheel and other mechanisms that makes the method of arbitrary functions applicable? What makes it so likely that relative frequencies come out the same way in most large sets of spins of the wheel? First, the roulette wheel exhibits instability: it often maps similar inputs to dissimilar outputs. For example, two input velocities can be very near to each other, yet one results in red while the other results in black. Second, the wheel often maps dissimilar inputs to similar outputs. Two very different input velocities can both lead to black, for example. When these two conditions are met, causal chains from inputs will be “shuffled” on the way to their corresponding outputs, and the causes of a set of similar outputs can be drawn roughly equally from all parts of the input space.

Let us say that a “cell” is a set of similar inputs which lead to similar outputs (for example, a red interval or a black interval in the roulette example). Then the third condition needed for the MAF to work is that nearby cells mapping to dissimilar outputs should receive comparable numbers of inputs. This last condition has the consequence that dissimilar outputs (e.g. red and black) drawn from the same multi-cell input region (a pair of adjacent intervals) will appear in roughly the same numbers. Since this is true for each such input region, it's true in general.²

The next section discusses a subtlety related to the third requirement.

4. Measures and input distributions. The measure on the velocity space plays a crucial role in the MAF argument above. The argument's conclusion was that if the alternating red/black intervals in the velocity space had small measure relative to the slope of the input distribution, then the relative frequencies of red and black would be close to the ratio between measures of red and black intervals. Lebesgue measure must be pretty special, if facts about measure imply facts about what relative frequencies there will be. Why not some other measure? Suppose, for example, we arbitrarily choose

²According to von Plato (1994, p. 172), three conditions for the applicability of the MAF similar to the ones given here were first stated by von Smoluchowsky. Hopf 1935, p. 5 gives a requirement similar to the third: that in small regions inputs be roughly uniformly distributed.

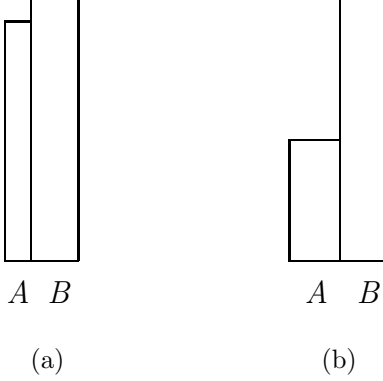


Figure 3: Width: measure; height: average trials; corresponding areas equal.

a measure that happens to weight red regions more than Lebesgue measure does? Could we then argue that red would have a higher relative frequency?

Consider figure 3. Here it will be simpler if we don't pretend that the input distribution is continuous. Again, velocity intervals' measures are represented by width; the number of inputs having a value in a given interval is represented by the area of the rectangle over an interval. The height of a rectangle then represents the average number of inputs for that interval. Let's say that in figure 3(a), width represents Lebesgue measure defined in terms of velocity as above. Now suppose that without changing the actual set of inputs, we use a different measure, shown in 3(b), which gives the set A twice the measure that it has in 3(a) while keeping the measure of B the same. The number of inputs in A is the same in both cases, so we have to keep the area of the corresponding rectangle the same, thus making it half as tall. What we saw above was that relative frequencies will be near to measure given that inputs are distributed to adjoining velocity intervals in such a way that the slope of the distribution curve is not too steep. The analogous requirement here is that the height difference between adjoining rectangles not be large. What figure 3 shows is that the heights of adjoining intervals can be similar relative to one measure but very different for another. Given the measure in figure 3(a), the MAF may apply because the "slope" from one interval to the next is not large. But given the measure in figure 3(b), the intervals A and B are not small enough to make relative frequency close to relative measure. On the other hand, if there were significantly more

inputs in A , then this second measure might allow the MAF to go through while the first measure might not.

So there isn't necessarily an obvious measure on which to base an MAF argument. An appropriate measure has to reflect actual relative frequencies to a certain extent; it has to "match" the facts about certain sets of inputs in our world. If Lebesgue measure works for MAF-arguments concerning roulette wheels, that means that among certain common sets of roulette spins, we will find mostly shallow "slopes" relative to Lebesgue measure. Nevertheless, this fit between a measure and sets of inputs allows a wide range of variation in those sets.

There do have to be restrictions on what can count as a relevant set of inputs in determining the appropriate measure for application of the MAF. Aside from the need for the input sets to be large, they can't just be arbitrary collections of actual inputs. At a minimum, I suggest that a relevant set of inputs must be one which contains all and only events which are the product of a single token mechanism during a continuous interval of time, and which are events in the input space of a mechanism, such as the roulette wheel, which is under consideration. For example, viewing people working at casinos as mechanisms, all of the roulette spins by a single croupier in a given month might count (whether at the same roulette wheel or not). But a set containing only those of the croupier's spins which lead to red could not count. I discuss this issue in more detail my dissertation, where I call the sets of inputs that should count in determining the appropriate measure "natural" sets of inputs.

5. Mechanistic Probability. Versions of the method of arbitrary functions have been applied to a handful of idealized mechanisms, all relatively simple. Related ideas can be found in an area of mathematics known as ergodic theory, though generally only applied to systems in which a mechanism repeatedly operates on its own outputs.³ And ergodic theory is typically concerned with behavior in the infinite limit, which needn't have any implications for relative frequencies we observe. But if we could generalize some of the ideas from the method of arbitrary functions and ergodic theory, we could base a theory of higher-level objective probability on them. In particular, we'd need to generalize, in some reasonably precise manner, the idea of inputs getting "shuffled" in the above sense on the way to the output space.

³(Friedman, 1970) and (Nadkarni, 1998) are good introductions to ergodic theory.

The ROF requirement below is intended as a stand-in for such a future generalization. In addition, we’d need to generalize the idea that there must be some sort of loose “match” between a measure on the mechanism’s input and certain kinds of common sets of inputs, what I referred to above as “natural” sets of inputs. The MID requirement below is intended as a stand-in for that kind of generalization. What I’ll do now is explain how one can construct a theory of objective probability given the right kind of generalizations.

First, from an instance of an input state, a deterministic mechanism produces an instance of an output state. More abstractly, such a mechanism implements a map T from an input space of event types to an output space of event types. For present purposes, we can treat event types as sets of possible events. (Here “event type” corresponds roughly to what mathematicians call “event”; a mathematician might speak of a particular outcome of a trial where I use “event”.) Since we’re assuming that the mechanism is deterministic, if we have some kind of objective probability for a set of inputs to the mechanism, this determines a probability for a corresponding output set. Or to put it a different way, we can define the probability $\mathbf{P}_{\mathcal{O}}$ of any set of outputs A in terms of the probability $\mathbf{P}_{\mathcal{I}}$ of the set of inputs which might cause it:

$$\mathbf{P}_{\mathcal{O}}(A) = \mathbf{P}_{\mathcal{I}}(T^{-1}(A)). \quad (1)$$

Now suppose that a mechanism T satisfies the following two-part condition MP in our world.

MP: MID Existence of a Measure Matching Input Distributions:

There’s a measure $\mathbf{P}_{\mathcal{I}}$ on T ’s input space which “matches” most actual natural sets of events in T ’s input space. These sets of inputs must in some sense be very diverse. If $\mathbf{P}_{\mathcal{I}}$ is not unique, then any such measure must generate similar values for $\mathbf{P}_{\mathcal{O}}(A)$ for any A in the relevant algebra $\mathcal{B}_{\mathcal{O}}$ of subsets of T ’s output space. Thus, the actual set of natural input sets must determine a unique, nonempty near-equivalence class of measures $\mathbf{P}_{\mathcal{I}j}$ which match most of those sets.

ROF Robustness of Output Frequencies:

T is such that relative frequencies among outputs will be robust with respect to variations among natural sets of inputs. Specifically, natural input distributions which match $\mathbf{P}_{\mathcal{I}}$ cause output distributions with relative frequencies near to the measures assigned by $\mathbf{P}_{\mathcal{O}}$.

MID and ROF clearly need to be fleshed out in several respects with more precise mathematical and metaphysical criteria. Suppose that they have been. Then the *mechanistic probability* of A will be defined to be $\mathbf{P}_O(A)$. Call this “short-run mechanistic probability” (“SRM probability”, or “SRMP”) if the relative frequencies mentioned in ROF include relative frequencies in finite sets of outputs, including sets which are small enough that these frequencies might be of concern to us. Otherwise call it “medium-run” or “long-run mechanistic probability”, the latter for the case where ROF mentions only limiting frequencies.

It’s essential to MID that the matching relation allow a wide range of input distributions to match the same input measure. But ROF guarantees that this range will be wider, in some sense, than the range of output distributions which the matching input distributions cause. As long as a set of inputs is in the right general ballpark, the structure of the mechanism satisfying ROF should guarantee that output relative frequencies will fall within narrow ranges of values.

It’s plausible that the shuffling causal structure which exemplifies ROF is common among games of chance and many other aspects of the world. For example, many ways of flipping a coin lead to heads. Yet it seems likely that a small difference of initial velocity could result in a different outcome. And although the details are complex, I believe that SRMP can ground the probabilities underlying popular definitions of biological fitness. Consider the fact that placing the same organism in diverse conditions can result in it having the same number of offspring, while a small shift in the wind or in the location of another organism can lead to an early death or to the failure to eat or mate, thus changing the number of offspring.

It may be that a fully-developed version of SRMP would support output spaces with countably additive probability measures, even though relative frequencies in finite spaces are only finitely additive. This seems possible because MP as stated is consistent with the output probability space being more finely structured than a finite output relative frequency space can be. On the other hand, it could turn out that finite additivity is all that SRMP will provide. Given the looseness of MP as stated here, it could even turn out that SRMP would not even provide a probability measure with determinate values. It might only provide probability in some looser sense, as in axiom systems for interval-valued probabilities. But if the intervals were small, they might be enough like numeric probabilities for most practical purposes.

Assuming that the details of the definition of SRMP can be fleshed out

in such a way that it does not depend on any non-objective elements, SRMP clearly satisfies the first two of the desiderata I listed at the beginning of the paper. It would bear a causal relationship, via the ROF property of a mechanism, to relative frequency, usually resulting in relative frequencies near probabilities when the numbers of trials is large. Nevertheless, that a mechanism's outputs had certain SRM probabilities would not necessarily guarantee that relative frequencies would come out near to those probabilities—after all, MP doesn't require that there be no input distributions which give rise to divergent frequencies. And by definition, if SRMP is definable at all, SRMP will be able to do all of this in terms of finite frequencies.

A significant feature of SRMP is that while satisfaction of MP has implications for relative frequencies, MP says nothing at all about outcomes on individual trials. There's no reason to think that a fully developed theory of SRMP will provide objective probabilities concerning outcomes on particular trials. It's true that there's a sense in which mechanistic probabilities must be present on each execution of a mechanism. The mechanism has the structure it has, and the distribution of natural sets of inputs is what it is, and all this is true at the moment of a given trial if it is ever true. Moreover, the mechanistic probabilities in question are probabilities of event types in the output space of the mechanism—just what probabilities of outcomes of a trial should be. All this might make one want to claim that mechanistic probabilities must apply to the outcomes of particular trials. But though the probabilities do exist in this sense at the time of a given trial, they need not be probabilities concerning what will happen on that trial; they need not be probabilities which in any sense govern individual causal interactions. SRMP is objective, since it's determined by facts about inputs in the world and about the structure of a mechanism. And it's probability, as long as it satisfies a set of probability axioms. Moreover, the space on which SRMP is a measure consists of event types. But SRM probabilities need not be probabilities of possible events on *particular* trials.⁴ This, in fact, is one of SRMP's greatest virtues. Since SRM probabilities needn't be probabilities of outcomes on particular trials, an outcome event type on a particular trial can still have an objective probability of 0 or 1 (given determinism), consistent

⁴This shows that SRMP cannot be identified with single-case propensity. I argue in my dissertation that it would also be a mistake to identify SRMP with hypothetical frequency. Shimony (1975, pp. 390f) and von Plato (1982, pp. 62f) each seem to suggest that propensities can in some sense be based on ergodic theory in a way that's reminiscent of SRMP.

with that event type having any value for its SRM probability. (If determinism is false, as is likely, a complete theory of SRMP would need to be revised to deal with the fact that most mechanisms are only nearly deterministic. Then the point would be that SRMP would not constrain higher-level probabilities to be near to 0 or 1 in typical cases.) Thus SRMP appears to be able to satisfy our third desideratum as well.

References

- Friedman, N. A. (1970). *Introduction to Ergodic Theory*. Van Nostrand Reinhold.
- Hopf, E. (1935). Remarks on causality and probability. *Journal of Mathematics and Physics* 14, 4–9.
- Nadkarni, M. G. (1998). *Basic Ergodic Theory* (2nd ed.). Birkhauser Verlag.
- Poincaré, H. (1912). *Calcul des Probabilités*. Gauthier-Villars.
- Shimony, A. (1975). Carnap on entropy. In J. Hintikka (Ed.), *Rudolph Carnap, Logical Empiricist*, pp. 381–395. D. Reidel.
- von Plato, J. (1982). Probability and determinism. *Philosophy of Science* 49, 51–66.
- von Plato, J. (1983). The method of arbitrary functions. *British Journal for the Philosophy of Science* 34, 37–47.
- von Plato, J. (1994). *Creating Modern Probability*. Cambridge University Press.