Review of Tychomancy*

Wolfgang Schwarz
03 November 2014

We humans are good at estimating probabilities. Despite the dauntingly complex dynamics involved in tossing a die, we readily see that each side has about a one in six probability of coming up. Michael Strevens's *Tychomancy* [Strevens 2013] is about this ability of ours. Strevens presents a number of inference rules that he suggests guide our probabilistic judgements. He argues that these rules are (1) innate, (2) reliable, and (3) crucial to many areas of science.

To illustrate the last point, consider evolutionary theory. It has long been recognized that fitness is a probabilistic matter. There is no logical or physical guarantee that faster wolves will catch more deer. If slow wolves frequently come across old or wounded deer while fast wolves happen to forage in areas devoid of deer, it may well be the slower wolves that prevail. But we can see that this is an unlikely scenario. We know that in normal environments, faster wolves have a greater probability of catching fast-moving deer.

Another example Strevens discusses in some detail is Maxwell's 1859 discovery of the Maxwell-Boltzmann law, according to which the velocity of gas molecules in equilibrium has a Gaussian probability distribution in each direction. As Strevens points out, Maxwell's original derivation of the distribution, from apparently a priori symmetry assumptions, is not only unsatisfactory, but doesn't even do justice to Maxwell's own evidence. Maxwell knew the probabilities, but he didn't know how he knew.

Tychomancy is a welcome philosophical contribution to an important topic. Probabilistic judgements are central to our scientific and everyday understanding of the world, but we know little about how these judgements are formed and what makes them reliable. Strevens addresses these questions in a pleasantly accessible but careful style, combining insights from child psychology, the history of science, and the mathematics of probability.

So what guides our probabilistic judgements? From examples like dice one might think they are directly based on physical symmetries, following some principle of indifference. But things are not that simple. We know that a die tossing machine is unlikely to produce the usual probabilistic patterns, despite the physical symmetries of the die. We also realize that if a die is rolled without shaking, then some outcomes are at a disadvantage. Moreover, it is not obvious how physical symmetries enter into evolutionary explanations,

* Thanks to Michael Strevens for helpful discussion.

or how indifference could directly justify Maxwell's Gaussian (and thus non-uniform) velocity distribution.

What is missing in the indifference-based picture are considerations about the dynamics of the relevant systems. According to Strevens, the central principles that drive our probabilistic judgements are not traditional principles of indifference but what he calls principles of *equidynamics*. He expresses these principles in a number of rules, tailored to different types of process and different questions. Among other things, there is a *stirring rule*, a *microdynamic rule*, an *equilibrium rule*, a *uniformity rule*, a *majority rule*, and a *microequiprobability rule*. Readers familiar with Strevens's earlier work will not be surprised that the "method of arbitrary functions" figures prominently in setting up and justifying these rules.

Consider the process of shaking a die. Each collision between the die and the container causes a reorientation of the die in which the new orientation is sensitive to fine details about the previous orientation. After a few collisions, the die's *approximate* initial orientation has become irrelevant to its new orientation: the information that the die started with, say, the six facing up entails nothing about its orientation after a few collisions; to compute which side is now facing up we would have to know the precise initial orientation, as well as precise details about the initial position etc. Slight differences to these initial conditions lead to different sides facing up after a bit of shaking. More specifically, the space of possible initial conditions can be partitioned into small, contiguous regions such that any probability distribution that is roughly uniform over each of these regions evolves, by the dynamics of the shaking, into a distribution that assigns to each side an approximately equal probability of facing up.

It is tempting to think that these observations somehow explain or "ground" the probabilities involved in shaking and tossing a die. More importantly for the present topic, they suggest a way of coming to know the probabilities. According to Strevens, the process can be divided into two steps. First, we recognize that the die's final orientation is the product of a series of collisions in each of which the post-collision state is determined by independent and essentially random details about the original state. The die's dynamics therefore approximates a Markov process, which means that after sufficient shaking, the probabilities over the die's orientation must reach a unique equilibrium. Here we apply Strevens's equilibrium rule. In a second step, we can use the uniformity rule to infer that in equilibrium each side has an equal probability of facing up. The rule says that whenever a system's short-term dynamics preserves uniformity over certain variables, then we may assume that the equilibrium distribution is uniform over these variables. (See pp.101–106 for the details.)

A remnant of indifference is still required. We have to assume that small variations in the initial conditions have approximately equal probability – in Strevens's terminology, that the initial conditions are *microequiprobable*. This is licensed by another rule, the

microequiprobability rule. The rule says that we may (defeasibly) assume that "standard variables" have an approximately uniform probability distribution over small intervals. Here, standard variables are "roughly, variables that measure physical quantities in proportion to the SI units" (p.69). The restriction is necessary to avoid Bertrand-type paradoxes.

So indifference assumptions are still required, but they play a much smaller role than one might have thought. We don't have to start with the assumption that each side of a die has an equal probability of facing up. Since the dynamics of shaking ensures that any reasonably smooth probability distribution over initial conditions leads to an approximately uniform distribution over final orientation, we only need to assume that small variations among the initial conditions have roughly equal probability. What's more, Strevens shows that these applications of the microequiprobability rule can often be justified by further dynamical considerations. For example, if the initial conditions are produced by processes that are subject to perturbations by noise, it is reasonable to assume that minor variations of an initial condition should have roughly equal probability. Of course, nothing comes from nothing. At some point, smoothness in the distribution of standard variables – for example, in the parameters of noise – must be taken for granted.

When we think about dice, we realize not only that each toss comes with an approximately uniform probability for the six outcomes, but also that the outcomes of different tosses are stochastically independent. For example, we judge that the probability of getting two sixes in a row is about 1/36. Assuming a deterministic dynamics, this requires that the initial conditions for the two tosses are suitably uncorrelated. It is not actually required that the probabilities over initial conditions are independent, which would be implausible. Rather, it is enough that the *joint* distribution over pairs of initial conditions is sufficiently smooth. Once again this assumption can be justified by further dynamical considerations, as Strevens argued in detail in ch.3 of *Bigger than Chaos* [Strevens 2003].

In *Tychomancy*, the required assumption about joint distributions is folded into the microequiprobability rule, by stipulating that a probability measure qualifies as microequiprobable only if it is approximately uniform over small, contiguous regions in the compound space of initial conditions for several trials. It would have been more perspicuous, I think, to give a separate name to the independence assumption. As it stands, the assumption has a tendency to slip under the radar – at least for the reader, but arguably also for Strevens himself. One sign of this is that Strevens's own use of the term 'microequiprobable' seems to fluctuate between the stronger sense in which it entails smoothness over joint distributions and the weaker sense in which it doesn't. A more substantive problem arising from the neglect of independence will be mentioned below.

Like most of the literature on arbitrary functions, I have focused on comparatively simple dynamical systems like dice. Strevens deserves credit for pushing the method a lot

further, showing how the kinds of considerations that apply to dice and roulette wheels might carry over to the kinetic theory of gases, to population dynamics, or to weather models. There is a lot to learn from these discussions. Nonetheless, I want to use the remainder of this review to focus on two general aspects of Strevens's treatment where I think it might be useful to explore alternatives.

The first concerns the status of Strevens's inference rules. Strevens does not pretend that his rules are complete or deductively valid. (They are clearly neither.) Given these concessions, more could have been said about what role the rules are meant to play. Are they supposed to be heuristics that we apply to arrive at probabilistic judgements? According to Strevens, even 8-month old infants use equidynamic reasoning, so the rules are presumably not meant to encode explicit steps of reasoning. Are we meant to be actively guided by the rules, or is the hypothesis merely that whenever we have evidence for the premises, then (*ceteris paribus*) we tend to come to believe the conclusion?

These questions become even more pressing given that Strevens's rules are not sufficient to account even for the main examples in his study. As we saw above, application of the uniformity rule presupposes that we can figure out a system's short-run dynamics, which is not given to us by any of Strevens's rules. Similarly, to make use of the equilibrium rule we need to know how long a die or an urn must be shaken in order for each outcome to have an equal probability. In both cases, Strevens appeals to a largely unexplained skill of *probabilistic dynamics*. Probabilistic dynamics must also be invoked to generate probabilistic judgements about outcomes if a shaking or shuffling has not gone on for long enough to apply the equilibrium and uniformity rules. For actual drawings from an urn or a deck of cards, this is the rule rather than the exception. Yet another appeal to probabilistic dynamics is needed to account for Maxwell's realization that correlations between the velocity components of gas molecules become disassociated in the coarse of several collisions. Given all these uses of probabilistic dynamics, one would rather like to know more about this ability, and one wonders how much work it leaves for Strevens's inference rules.

Strevens hints at the idea that what drives probabilistic dynamics is a kind of mental simulation or modelling of the relevant process. For example, to apply the equilibrium and uniformity rule, we might represent the shaking die as a Markov process. The idea is appealing, and might be worth exploring further. If our probabilistic judgements are largely based on the use of stochastic models, new questions would come into focus that presently play little role in Strevens's account. Most obviously, there would be questions about model selection and parameter estimation. The answer to these might supersede many of Strevens's rules: if such-and-such evidence makes it reasonable to represent a process by a certain stochastic model, and that model attributes a given probability to an outcome, then the evidence also serves as the basis for the probability judgement. Closer to the level of implementation, we can ask how a cognitive system can realistically

employ stochastic models. In each case, one might hope to draw on developments in machine learning, where the use of "generative models" has become a *de facto* standard.

I am also not convinced that the relevant models have to be dynamical – in other words, that the assumption of a shuffling dynamics is as central to our probability judgements as Strevens suggests. Imagine you find a well-used (standard) deck of cards in an abandoned building in a drawer. What's the probability that the 10th card from the top is the ace of spades? About 1/52, I would say, even if I know little about the process that gave the stack its order. Or consider the usual explanation of why quantitative features tend to have a Gaussian frequency distribution. The explanation is that the relevant features are determined as the sum or average of a large number of independent factors, and summing or averaging independent distributions tends to produce a Gaussian distribution. This strikes me as a good explanation, and as a useful guide for estimating probabilities, but it is not dynamical.

The second point where I think it might be worth investigating alternatives to Strevens's treatment concerns the interpretation of probability – a topic on which Strevens tries to maintain an "ecumenical agnosticism" (p.55). At first glance, this may seem puzzling. One might think that before we ask what guides judgements of a certain kind, or what makes them reliable, we ought to get clear about their content. How can we tell whether a rule generally leads from true premises to true conclusions if we don't know what it would mean for the probabilistic conclusions to be true? The relevant probability statements are clearly not statements about "objective chance" in the metaphysician's sense in which non-trivial chance requires indeterminacy in the basic laws of physics. On the other hand, Strevens insists that he is not just talking about degrees of belief. The probabilities in question are meant to be "physical", as witnessed by the fact that they are empirically discoverable, arise from physical features of dynamical processes, and are closely related to relative frequencies. There is indeed a growing consensus that we need such a conception of probability, even though none of the usual "interpretations of probability" seem to fit the job. On this background, Strevens's procedure to investigate probabilistic inferences without first settling on an interpretation actually makes sense. We know better what kinds of observations support probability judgements than what these judgements mean. Indeed, we might hope to get a firmer grip on the interpretation of probabilistic judgements by looking at what sorts of observations count as evidence for or against them.

Towards the end of *Tychomancy*, in chapter 12, Strevens drops his agnosticism and puts forward his own interpretation of physical probability, also outlined in his recent paper [Strevens 2011]. The starting point is the above-mentioned observation that the dynamics of certain physical processes – which Strevens calls *microconstant* – carries any smooth distribution P_I over initial conditions to approximately the same distribution P_O over outcomes. P_O can then be identified with the physical probability of the outcomes

provided that the actual frequencies of initial conditions approximate a microequiprobable distribution, and that they would continue to do so under a variety of counterfactual circumstances.

Here the neglected independence condition turns out to complicate the picture. According to Strevens, microequiprobability over initial conditions requires a smooth joint distribution for selections of two or three or ... trials, which on the proposed interpretation of probability requires an approximately smooth joint frequency distribution for such selections. That is, if we map out the frequencies of initial conditions in all suitable selections of, say, 100 die tosses, we should find no correlations between slight variations of the 100 initial conditions. The problem is that this requires a very large overall number of die tosses – arguably more than there are in the history of the world. Evidently the problem only gets worse if we consider not tosses of ordinary dice, but tosses of sixteen-sided dice or other uncommon probabilistic setups. Here Strevens hopes to get help from counterfactuals about scenarios in which there are many more instances of the relevant setup, but that threatens to bring on board many of the problems plaguing hypothetical frequentism.

Strevens's interpretation also has surprising consequences for young infants. Recall that according to Strevens, 8-month-old children perform equidynamic reasoning. The evidence is, for example, that when balls are drawn from an urn, 8-month-olds are more surprised by non-representative samples. What's more, their expectations about the outcome takes into account the process of sampling. Some form of dynamical reasoning therefore seems to take place. But what is the end product of that reasoning? Strevens speaks of "the conception of probability found in infants' equidynamic judgments". But do 8-month-olds really have a full-blooded Strevensian concept of physical probability? Do they really come to believe – if only in some implicit, non-conceptual manner – that the sampling process is microconstant and that the actual and counterfactual frequency distribution over initial conditions is approximately microequiprobable? These claims certainly go far beyond the evidence from psychology. One might argue that all the evidence actually supports is that the infants acquire precise and stable degrees of belief, reflected in their expectations about frequencies.

The difficulty of defining a suitable concept of physical probability, together with the issues raised by attributing any such concept to young infants, may justify another look at the projectivist interpretation of probability defended by de Finetti [1937], Skyrms [1984] and Jeffrey [2004]. Crudely speaking, on this account the function of physical probability statements is not to represent special probabilistic facts in the world; rather, it is to express (and perhaps recommend) certain stable degrees of belief. The account does not require de Finetti's radical subjectivism, on which all coherent degrees of belief are equally good. We can accept that observations about dynamics, frequencies, and other physical facts tend to make some attitudes rational and others irrational. The

projectivist interpretation goes nicely with the above idea that we use stochastic models to think about physical processes, as endorsing a model gives rise to precise and resilient degrees of belief.

From this perspective, justifying our probabilistic inferences means to explain how observations about symmetries and dynamics can warrant representing the relevant process by a certain stochastic dynamical model. As before, the change in perspective might bring into focus some interesting questions currently not addressed by Strevens. For example, we may ask why physical probability judgements are useful in the first place. Why model the dynamics of dice or gas molecules as a stochastic process, if the underlying dynamics is deterministic? The obvious answer is that going probabilistic is a means to achieve simplicity and generality, abstracting away from complex lower-level details. This is useful not only because the microphysical details are often hard to know, but also because they would obscure important higher-level regularities. Our ability to form probabilistic judgements might therefore be seen as part of a more general ability to employ abstract, high-level models.

The central issues discussed in *Tychomancy* still arise: How do we construct probabilistic models? Why are these models so successful? What kinds of physical facts make it reasonable to use a given model? We are still far from a full answer to these questions. But Strevens takes us a lot closer than we were before.

References

Bruno de Finetti [1937]: "La Prévision: ses lois logiques, ses sources subjectives". *Annales de l'Institute Henri Poincaré*, 7: 1–68

Richard Jeffrey [2004]: *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press

Brian Skyrms [1984]: *Pragmatics and Empiricism*. Yale: Yale University Press

Michael Strevens [2003]: *Bigger than Chaos*. Cambridge (Mass.): Harvard University Press

— [2011]: "Probability Out Of Determinism". In Claus Beisbart (Ed.) *Probabilities in Physics,* Oxford: Oxford University Press, 339–364

— [2013]: *Tychomancy: Inferring Probability from Causal Structure*. Cambridge (Mass.): Harvard University Press