The Reference Class Problem in Evolutionary Biology: Distinguishing Selection from Drift

Michael Strevens

1.   The Reference Class Problem

"The reference class problem is your problem too" writes Alan Hájek in a penetrating paper of the same name—regardless, he argues, of your interpretation of probability (Hájek 2007). He is right, and the problem is a particular difficulty for the foundations of evolutionary biology, as it threatens to undermine the distinction between natural selection and drift. This paper outlines the problem and presents a solution. At the paper's end, many important questions about the proper definition of drift will remain unresolved, but I hope to have provided at least a robust criterion for individuating the probabilities that so many writers invoke in attempting to give those questions answers.

In its broadest formulation, the problem of the reference class as follows: in determining the value of the physical probability of an outcome, what factors should be taken into account? And why? The problem is especially salient in those domains of scientific theory and everyday thinking in which the probabilities customarily assigned to certain outcomes take into account some causal factors relevant to the outcomes but ignore others. In determining the probability of heads on a coin toss, for example, we take into account the physical symmetry of the coin, ascribing a probability of one-half to heads on a toss of a fair coin but not on a toss of an unbalanced coin. Yet we ignore other initial conditions of the toss, such as its initial translational and angular

velocity: since these are more or less sufficient to determine the outcome, were we to take them into account we would say that the probability of heads is always either one or zero, depending on the initial conditions. (There is good reason to think that coin tosses are effectively deterministic.)

We make a distinction, then, between two kinds of physical quantities relevant to the outcome in question. Some are taken into account in determining physical probabilities. Call them *parameters*; as the parameters change, then, the probabilities change. The rest—call them *variables*—are not taken into account; differences in their values make no difference to the probability assigned to an outcome. It is as though all processes with the same values for their parameters are considered as a unified class, in spite of the differences in the values of their variables—differences that have just as big an impact on the outcome as differences in parameters. In the frequentist tradition, such classes are called *reference classes*; the problem of the reference class, then, is the problem of deciding how to individuate classes, or in other words (when dealing with causal processes), the problem of deciding whether a causal factor is to be treated as a variable or a parameter.

The physical probabilities at the heart of both classical statistical mechanics and evolutionary biology are like the probability of heads in this crucial respect: some causally relevant quantities are treated as parameters and some as variables. Some causally relevant quantities, then, are ignored when fixing probabilities. On what basis? If the question matters to you at all, then the reference class problem is your problem too.

The case of evolutionary biology is especially interesting for reference class theorists because it presents good reason to think that the distinction between parameters and variables has objective significance. It is not just a matter of what we know, or of how we have decided to organize inquiry; rather, it reflects deep differences in the way the world works. I say this because it seems, if the majority of philosophers of biology writing on the topic are correct, that the distinction between natural selection and drift is only as real

as the distinction between parameters and variables. But biologists regard selection and drift as qualitatively quite different kinds of explanation for evolutionary change. Such an attitude can be maintained only if the reference class problem can be solved—for biological probabilities at least.

Two caveats before I go on to explain the problem and to develop a solution. First, I will not supply my own characterization of drift, or seek to judge between characterizations that others have offered. Rather, I hope to provide something that (almost) all participants in these debates are in need of— something that is foundational, and therefore crucial to the debates, but that does not dictate any particular solution or support any particular side.

Among the many questions on which I will do my best to remain neutral are the following. Are selection and drift distinct kinds of processes or are they different aspects of the same process (Millstein 2002; Matthen and Ariew 2002; Brandon 2005)? If they are aspects of the same process, is that process causal (Hodge 1987; Walsh et al. 2002)? Is drift constitutive of population equilibrium, or is it a force disturbing equilibrium (Brandon 2006)? Can the contribution of drift to a change in relative frequency be quantified precisely (Beatty 1984; Millstein 2008)? Indeed, this paper does not even resolve all issues that might be grouped under the heading of "finding the right reference class" (I am thinking of the questions raised by Abrams (2009) in particular).

It is not that the problem of the reference class is completely independent of these other problems. Some arguments for the "statisticalist" position that neither selection nor drift are causal processes invoke the presumed arbitrariness of reference classes (Walsh 2007; Matthen 2009). But I will make no attempt here to untangle the often very complex connections involved.

The second caveat is that my resources for dealing with the reference class problem are drawn from the general philosophy of science: I will argue that the parameter/variable distinction can be understood in the same way in biological systems as in simple gambling devices, in the systems treated by statistical physics, and so on. This is not a paper, then, for those who say

"Philosophy of biology solutions for philosophy of biology problems" (as one anonymous referee made very clear). At same time, I am very conscious, as I said in the preceding paragraphs, that what I say here is only a small part of a complete treatment of the nature of drift—a complete treatment that most certainly will derive much of its content from what is special and distinctive about the philosophy of biology.

2. Parameters, Variables, and Drift

Many philosophers have recognized that the problem of the reference class poses a prima facie challenge to the selection/drift distinction, and thus arguably to the thesis that fitness and selection are real and explanatorily potent: Sober (1984, 129–134), Rosenberg (1994), Brandon (2005), Matthen (2009), among others. Let me begin by presenting what I take to be the standard understanding of the challenge.

Begin with a paradigm of drift: neutral evolution. Neutral evolution occurs when frequencies (of genes, traits, or whatever) change without selective pressure; although no particular gene or trait has a systematic advantage, their relative numbers change simply "by chance".

The clearest cases of neutral evolution are those in which variations in the structure of a gene have no phenotypic consequences. One variant may be reproduced at a greater rate than another, but this cannot be due to some feature of the variant that makes it more likely to be reproduced; the difference in outcomes is therefore due to "random genetic drift".

Perhaps, for example, the gene in question is wholly devoted to determining its organism's pattern of camouflage. Suppose that, because the gene's two variants are phenotypically equivalent, the organisms in which they appear have identical camouflage patterns (something of an idealization, to be sure). We may distinguish two classes of organisms, then, differing only with respect to which of the two genetic variants they possess (and ignoring the complications brought on by diploidy, gene linkage, and so on): call them

α and β. There is variation within each class, but it is the same variation. Intuitively, then, the two classes should be, on the whole, equally susceptible to predation.[1] But it might be that over some period of interest, simply by chance, αs are eaten less frequently than βs. Then the proportion of αs, and so the proportion of the underlying genetic variant, will increase. This increase is due to drift.

If, by contrast, the α allele makes a difference to the camouflage pattern, and this difference results in a lesser susceptibility to predation, the decrease in the β population would be due to natural selection. These seem to be very different explanations of the relative increase in αs. The challenge to the philosopher of biology is to spell out the qualitative difference between the two explanations—or to argue that, contrary to appearances, there is no difference.

Traditionally, drift has been distinguished from natural selection using the figure of "sampling error". Fill an urn with balls, half red and half black. Conduct one hundred draws from the urn (replacing the drawn ball each time). You will very likely draw red about one-half of the time, reflecting the one-half probability of such a draw. But most likely, you will not draw red exactly 50 times. The difference between the actual frequency of red and the probability of red is called sampling error. The idea is then generalized: sampling error occurs when the frequency of an outcome differs from its probability.

An evolutionary outcome is (on the traditional conception) attributed to drift when it is due to sampling error. That such attributions can be made

---

1. Throughout this paper, I consider only simple binary scenarios of evolutionary change, that is, scenarios in which the population is divided into two variants. Evolutionary biology is frequently concerned with scenarios in which there are many variants, even (notionally) continuously many, as in the case of traits such as height, weight, or leaf size. I am confident, though I will not make the case in these pages, that the problem of the reference class can be substantially posed and solved using only binary scenarios, although some discussion, omitted here for reasons of length, is needed to explain how to set up the treatment in sections 5 and 6 for the more complex cases.

is clearest by far in cases of neutral evolution (hence my beginning with the neutral case). If all variants have the same probability of experiencing some evolutionarily significant outcome, such as successful mating or death, then any relative change in their numbers resulting from outcomes of that sort must be due to a deviation of frequencies from probabilities.

In the scenario sketched above, the probabilities of all evolutionarily significant outcomes are the same. In particular, the probabilities of predation, the kind of event responsible for the relative increase of αs, are the same. So that increase, due entirely to the deviation of predation frequencies from predation probabilities, is attributed to drift.

Enter the problem of the reference class. There is more than one way to determine predation probability, and different ways may yield different answers to the question whether the probabilities of predation are the same or different for the two variants α and β, and so different answers to the question whether α's relative success is due to drift or selection.

To precisify the worry, let me make an assumption for purely strategic reasons: I will suppose that the processes underlying evolution are fully deterministic. The precept of determinism makes it easier to press home a serious version of the reference class problem, so increasing the challenge of constructing a response. A solution to the problem developed under these conditions is as a consequence all the stronger.

Given determinism, an event in which a specimen of α is eaten but a nearby specimen of β is not must differ in some causally relevant way from an otherwise similar event in which it is the β that is devoured and the α that escapes to live, love, and leave its genes to the next generation. Perhaps two such events differ only in that the positions of the α and the β are switched, with the α occupying the fatally conspicuous spot in the first event and the β in the second.

Conditional on these and other details, the probability that the α is captured by the predator is much higher in the first kind of scenario than in

the second, and vice versa for the β.[2] In other words, declare the details of position to be parameters rather than variables—adopt a very fine-grained individuation of predation probabilities—and the probabilities will vary a great deal from episode to episode.

In my evolutionary scenario, more βs are eaten than αs, presumably just because of unlucky positioning. This fact will be reflected in the fine-grained probabilities: the probability that a β will be eaten in any particular situation will be higher on average than the predation probability for an α.

But in that case, the relative increase in the number of αs caused by predation is not (principally) due to sampling error. It is not that αs and βs are equally likely to be eaten, but that simply by chance (that is, sampling error), more βs are eaten. Rather, βs are on the whole more likely to be eaten than αs, and this difference in probabilities is reflected in the outcomes. That is natural selection.

Use coarse-grained predation probabilities, by contrast, and as I wrote above, the predation probabilities are equal and so the success of the αs is entirely due to sampling error. That is drift.

In short, whether the αs' increase in numbers should be explained by natural selection or by drift appears to depend on which causes of predation are treated as variables and which as parameters. Treat fine-grained matters of positioning as variables and you have drift; treat them as parameters and you have selection. It follows—so you might think—that there is no objective distinction between explanations from drift and explanations from selection. Yet many biologists, and philosophers of biology, feel the distinction in their bones. As Millstein (2002) points out, some of the most important debates in evolutionary biology since the modern synthesis turn upon it. We ought to look harder to see on what foundations it might rest.

---

2. Take into account enough information, and the probabilities are all ones and zeros, but there is no need to go this far; all that is dialectically necessary is information sufficient to generate a probabilistic inequality.

Before beginning the search for an objective criterion to distinguish parameters and variables in the probabilities of evolutionary biology, however, I need to take a closer look at the existing literature on drift. Not every writer on drift characterizes it in the traditional manner reflected in my presentation above; some have proposed solutions to the reference class problem that ought to be considered; and some dispense with a probabilistic understanding of drift altogether.

3.   The Process Approach to Drift

To say, along with the textbooks and the traditionalists, that evolutionary change is due to drift rather than natural selection insofar as it constitutes sampling error, can seem rather unhelpful, even obscure. I toss a coin ten times and obtain six heads. If had obtained only five heads, the frequency of heads would have matched the probability of heads and so there would have been no sampling error. Is one of the six tosses that landed heads, then, responsible for the error? If so, which one? The question seems misconceived. There has been sampling error, but no discrete part of the process causes that error. To attribute an outcome to drift, then, cannot coherently be understood as attributing the outcome to some feature of the process that produced it—in which case drift cannot explain outcomes. What has gone wrong?

According to what I will call the process-driven account of drift, the problem is a definition of drift that focuses on outcomes—frequencies that diverge from probabilities—rather than the processes that produce the outcomes. (In discussing the process account, I do not assume that either the account or even its diagnosis of the problem is correct; I just want to cover all the bases, showing that almost every definition of drift has a reference class problem to wrangle.)

The first step in the process account is to divide the causal totality driving evolutionary change in a population into discrete causal mechanisms, each potentially producing an evolutionarily significant outcome such as death.

Some such mechanisms will be those that determine the consequences of encounters with predators; some will be those underlying other dangerous events such as electrical storms and forest fires; and so on. The division need not be completely determinate, just as the distinction between selection and drift might admit of some fuzziness. Mechanisms might run in parallel—there can be predation during storms. Some mechanisms, such as predation, will run to completion over short periods of time; some, such as gestation or drought, will unfold over months.

Once the division is made, the real business begins: classifying each mechanism as a discriminate or an indiscriminate sampler. The net evolutionary change in the population—the net change in the proportion of each variant—will be determined by the various outcomes of the many evolutionary episodes in which these mechanisms operate. Part of the the net change will be explained by episodes of indiscriminate sampling; part by episodes of discriminate sampling. The first part—the part due to indiscriminate sampling—is, according to the process-driven approach, explained by drift, while the second part is explained by natural selection. An explanatory attribution to drift just is, then, an explanatory attribution to indiscriminate sampling. (When all change is due to indiscriminate mechanisms, you have neutral evolution.)

Or at least, that is the basic idea. Many sophistications are possible. On Gildenhuys (2009) account, for example, indiscriminateness of sampling is necessary but not sufficient for a mechanism to qualify as a drift-type explainer: the mechanism must also be "non-interactive" and "non-pervasive". I hope I will be forgiven for not going into further details about the taxonomy of process-driven accounts; there is just one feature of these accounts on which I think I need to focus, namely, the definition of indiscriminate sampling.

Traditionally, an evolutionary mechanism is said to sample indiscriminately if all variants are equally likely to experience each of its possible outcomes:

A sampling process is indiscriminate if and only if each entity in

9

the pool to be sampled has an equal probability of being chosen (Brandon 2005, 156)

This is Gildenhuys's sense of indiscriminate sampling:

> A cause [e.g., predation, lightning strike] is indiscriminate if its probability of influencing the reproduction of population members, in any given generation or time slice, is independent of the type of those members (p. 540)

since probabilities that are independent of type are the same for all types, all other things being equal.

On this traditional conception of what constitutes indiscriminateness, the traditional problem of the reference class presented in section 1 plainly has its bite: probabilities that are equal when causally relevant factors such as position are treated as variables become unequal when the same factors are treated as parameters. Any notion of drift based on the equality of the probabilities of evolutionarily significant outcomes such as mating and death must provide or presume some solution, then, to the reference class problem.

There is, however, another notion of indiscriminate sampling in the literature on process accounts of drift. It is found in Beatty (1984), who writes that indiscriminate sampling occurs when

> Physical differences... between the entities in question are irrelevant to whether or not they are sampled (p. 189)

and Millstein, who defines an indiscriminate sampling process as

> A process where heritable physical differences between entities (e.g., organisms, gametes) are causally irrelevant to differences in reproductive success (Millstein 2008, 353).

I will call this the modern, as opposed to the traditional, view of indiscriminate sampling.

On the face of it, the modern view makes no appeal to probability at all. But I think that both Beatty and Millstein understand causal relevance in terms of probabilistic relevance: a factor is causally irrelevant to an outcome (equivalently, to an entity's being sampled) just in case its presence makes no (causal) difference to the outcome's probability. Does that mean that the reference class problem affects even the modern view of indiscriminateness?

Not necessarily. On one plausible understanding of difference-making, indiscriminateness in the modern sense does not require equality of probabilities, or indeed any fixed facts about probabilities at all. Consider, for example, the scenario introduced at the beginning of this paper. Variants with the α camouflage gene get eaten less often than variants with the β camouflage gene, even though the two alleles are functionally identical and so result in exactly the same camouflage pattern. Specimens of β just happened to be in the wrong place at the wrong time more often than specimens of α. On a fine-grained individuation of the probabilities, details of positioning are treated as parameters, and so on average the αs' probability of death by predation is less then the βs' probability of the same. That makes sampling discriminate on the traditional view. On the modern view, however, it is indiscriminate: the probabilities may differ, but not in virtue of physical differences (let alone heritable physical differences) between the two variants.

More generally, however the probabilities are individuated—however you draw the boundary between variables and parameters—neither the genetic difference, nor (arguably) any other intrinsic physical difference between the variants, will affect the probability of death. So even without there being a fact of the matter about the probabilities, there can be a fact of the matter about whether physical differences affect the probabilities (by supervaluation, if you like). In this case, they have no such effect, and so the increase in αs can be attributed, objectively, to drift.

Other cases, however, pose greater problems for the modern view. Many organisms appear to make foraging decisions at random, as though they

conduct a kind of internal mental coin toss to decide where to seek out food. Imagine a population of organisms that function in just this way: when confronted with the sort of choice that would defeat Buridan's ass, they toss some neurological equivalent of a coin, and then act accordingly: heads they go left; tails they go right. Or at least, some of them—the αs—work this way. In the others, the βs, it is the other way around: heads they go right; tails they go left. (Although this is a ridiculously simple toy example, it seems plausible that there are many real-life, though of course more complicated, analogs.)

In some situations, if you go left you get a nice lunch, whereas if you go right you are what's for lunch; or in other words, the outcome of the mental coin toss can make a difference between life and death. Over a period of time, suppose, many of the βs get unlucky in just these circumstances and so get eaten: consulting their mental coin toss, they happen to go in precisely the wrong direction. The αs are on the whole more fortunate, and so the proportion of αs in the population increases. Selection or drift? On the process-driven approach, the question becomes: discriminate or indiscriminate sampling?

If we are allowed to construct probabilities in a fine-grained way, then we can take both the position of the hungry predator—whether it is on the left or on the right—and the outcome of an organism's mental coin toss to be parameters. Such probabilities will vary from encounter to encounter, but in general, βs will have a higher probability of being eaten than αs. Now the modern test for indiscriminateness: does the physical difference between the two variants make a difference to the probabilities? Clearly it does: had the βs followed the α rule, they would (usually) have fared much better.[3]

Relative to the fine-grained probabilities, then, the sampling counts as discriminate. But relative to more coarse-grained probabilities that treat

---

3. It is important for this conclusion that, as specified above, the βs' bad luck consists largely in their choosing more often than not the direction in which the predator lies, and not in their running into such situations more often in the first place, or in their more often getting into situations where there are predators on both sides.

predator position and so on as variables, the difference between the variants will make no difference to the probability of death (which will be equal for both). So the sampling will count as indiscriminate. In the absence of a solution to the reference class problem, the modern view therefore fails to deliver a determinate judgment about this scenario. But it is clearly a case of drift.

This is not an argument against the modern view of indiscriminate sampling or more generally against the process-driven account of drift. It is rather an argument that these approaches to thinking about drift, like all others that depend on probabilities, require some solution to the reference class problem, some criterion for distinguishing variables from parameters—a criterion that is as robust and objective as the distinction between selection and drift itself.

* * *

Do we need probabilities to draw a line between selection and drift? As intimated above, the modern notion of indiscriminateness might be rendered non-probabilistic if it were supplied with a suitable non-probabilistic criterion for causal relevance—a criterion that sometimes counts even decisive difference-makers like position relative to a predator as irrelevant. But I know of no such criterion.

I do know of a characterization of episodes of drift that plainly makes no appeal to probabilities, that suggested by Godfrey-Smith (2009). It is not clear to me, however, that Godfrey-Smith's criterion succeeds in distinguishing episodes of drift from episodes of selection. His criterion for a drift-type episode—roughly, that it involve external causes and that the outcome of the episode be sensitive to small changes in initial conditions—seems to apply equally well to certain selection-type episodes, such as those involving predation in which small differences in position make a difference to whether the prey is or is not spotted.

These brief and far from decisive considerations hardly close the door, however, on non-probabilistic approaches; they should be regarded only as

giving my own reasons for not pursuing this path.

4.   Solving the Reference Class Problem

If drift and the reference class cannot be separated, then what to do about the reference class?

You might opt a brute force or trivializing solution: declare that every factor that plays a role in causing an outcome should be treated as a parameter for the purposes of determining the outcome's probability. In a deterministic world, then, all probabilities are zero or one. This is, in effect, Rosenberg's (1994) solution; he concludes that drift (or the concept thereof) is more an instrumental aid to theorizing than an objectively valid explanatory construct.

Another option is to relativize the distinction between variables and parameters, and so (possibly) between drift and selection. Matthen (2009) cogently explores two forms of relativization.

The first form is epistemic. Perhaps information about parameters is easy to come by, whereas information about variables is often unknown and expensive to acquire. (Certainly, that is true of the coin toss.) Reichenbach's solution to the reference class problem is to "use the narrowest… class for which reliable statistics are available" (Reichenbach 1949); what makes a causal factor a parameter, on this approach, is that we are in a position to learn its value and then to use that value to reliably determine a probability (in Reichenbach's frequentist interpretation of probability, by consulting statistics). The resulting doctrine on drift implies, like Rosenberg's, that the concept of drift is useful only because we are imperfect knowers of causes.

Matthen's second form of relativization is institutional: on this approach, the distinction between variables and parameters is dictated by a theoretical framework, with different theories or domains prescribing different distinctions. It is part of the rulebook for doing statistical physics, for example, that you take into account the temperature of a gas but not the velocities of particular gas molecules. It is part of the rulebook for doing the science of

games of chance that you take into account a coin's physical symmetry but not its initial spin speed. And so on.

One rationale for establishing demarcations of this sort is epistemic; the epistemic and institutional forms of relativization are not, then, mutually exclusive. Another rationale, which presents a genuine alternative to the epistemic approach, is the efficient organization of research: just as a factory assembly line assigns different squads of workers to different tasks, so science tells evolutionary biologists to pay attention to some variables but not to others—with responsibility for the others being assigned to lower-level sciences such as physiology, molecular biology, and physics. What count as variables in one science, then, might be perfectly knowable, but pursuit of such knowledge is for reasons of economy assigned to a different, complementary branch of inquiry. Whereas on Reichenbach's epistemic approach, the parameter/variable distinction will dance around as our epistemic situation changes, on the economic approach it is sociologically fixed: not immovable, then, but slow to change.

Matthen argues that physical probabilities built on a parameter/variable distinction dictated by a socially mandated division of labor are genuinely explanatory.[4] Thus drift is genuinely explanatory—though this explanatoriness must surely be relativized to the sanctioning institutional structure.

The importance and (relativized) explanatory significance of divisions of cognitive labor are important themes in my own work—see Strevens (2008) and especially Strevens (2016). Nevertheless, I will not take the institutional approach in the present paper. Something better is possible: an entirely non-relativistic conception of drift's explanatory power.[5]

---

4. Note that on Hempel's epistemic approach to statistical event explanation, the inductive-statistical model, probabilities built on a parameter/variable distinction dictated by epistemic limits are also genuinely explanatory (Hempel 1965).

5. A third explanation of institutional parameter/variable distinctions, also not to be pursued in this paper, puts them down to historical contingencies in the development of the various disciplines that have no single explanation and perhaps no rational justification at all.

* * *

The objective explanatory validity of drift must rest, I believe, on an unrelativized, objective criterion for marking the boundary between parameters and variables. Seeking to provide an objective foundation for the probabilities in his statistical-relevance account of explanation, Salmon (1970) developed just such a criterion. More recently, Brandon (2005) has suggested that Salmon's criterion can be put to use to ground the objective explanatory status of drift.

Salmon's proposal turns on what he called the homogeneity of reference classes. Formulated in terms of the parameter/variable distinction, the homogeneity criterion stipulates that a causal factor should be counted as a variable rather than a parameter just in case conditionalizing on the value of the variable does not affect the probability of the outcome. Mechanisms that differ only with respect to quantities satisfying this condition form what Salmon calls a homogeneous class (with respect to the outcome in question).[6]

The homogeneity criterion is suitable for some purposes, but it does not draw the line between parameters and variables in the places needed to make sense of biologists' judgments about what is selection and what is drift. The problem is that any causal factor that makes a difference to whether an outcome occurs will count, according to the homogeneity criterion, as a parameter for the purposes of determining the probability of that outcome. But biologists treat many such causal difference-makers as variables.

Consider, for example, the scenario from section 2 in which variants α and β have identical camouflage patterns, but in which β, over some period of interest, suffers relative to α because its specimens more often, simply by

---

6. Salmon, at this stage of his thinking, avoids all mention of causation. Thus he does not restrict his potential probability-determiners—the factors that might count as parameters—to factors that play some causal role in producing the outcome. Rather, as Brandon explains, he restricts the potential probability-determiners to factors that are not characterized in terms that depend on the outcome of interest. This paper has presumed from the very beginning that potential probability-determiners must be involved in causally producing the outcome; my formulation of Salmon's criterion imports that presumption—not altering Salmon's view, but rather applying it within the paper's framework.

chance, find themselves in the "locus of doom", where they will be picked out and picked off by a predator. We want to count this as a case of drift, which means—on the traditional approach—assigning to both variants equal probabilities of death by predation.[7] The homogeneity criterion does not oblige. It requires that a specimen's being located within the locus of doom, since it makes a big difference to the probability of death, be treated as a parameter. The probability of death must take position relative to predators into account, then, and so will vary from specimen to specimen (and from moment to moment) and will be higher on average for βs. In short, the homogeneity criterion, applied to this scenario, judges that the proportion of βs decreases not because of drift but because of selection against bad timing and unhappy placement.

Indeed, in a deterministic world, the probabilities delivered by the homogeneity criterion are invariably equal to zero or one; it finds stochasticity only where there is indeterminism. A parameter/variable distinction based on homogeneity, then, is none other than Rosenberg's trivializing distinction. Some other approach to the reference class problem is required.

We want a parameter/variable distinction that provides probabilities suitable for objectively valid explanation. I propose that we work backward from explanatory concerns: what distinguishes parameters and variables, I suggest, is that parameters play an explanatory role that variables do not, somehow connected to their counting as probabilistically relevant to evolutionary outcomes in a way that variables are not. This explanationist approach promises not only to draw a line between parameters and variables, but also to tell us why we ought to care about that line: it demarcates factors that are explanatorily relevant to evolution, neutral or selective, from those that are not.

What is explanatory relevance? According to Salmon (1984), Woodward

---

7. I omit here the parallel argument that the homogeneity criterion will not cooperate with the a process-driven conception of drift driven by a "modern" definition of indiscriminate sampling (section 3).

(2003) and many other writers, it is causal relevance. On such a view, the explanationist approach to the parameter/variable distinction seems to be entirely unhelpful. Initial conditions such as position, which I have taken to be paradigms of the sort of quantities that we want to count as variables rather than parameters, are as I have emphasized throughout, highly causally relevant to the outcomes in question: an organism's position may play as important a causal role as any fact about it and its situation in determining whether or not it gets eaten. Surely, then, position is as good an explainer as camouflage or predator sensory prowess or any other paradigmatic parameter?

If it is individual outcomes that are to be explained, this is true: a concern with the explanation of some particular death provides no reason to favor, say, facts about camouflage patterns over facts about position. But what matters in evolutionary explanation are the frequencies of outcomes, not the outcomes themselves. The evolutionary biologist cares, and attempts to explain, why the relative frequency of αs has increased at the expense of the βs (or why the frequencies of some larger range of variants have changed—see note 1), but does not care at all which particular αs and βs reproduced and which did not.

Some initial conditions that are causally relevant to individual outcomes are explanatorily irrelevant to frequencies because they do not make a difference to frequencies, I will argue in the next section—and it is precisely what biologists typically take to be the variables that are irrelevant and the parameters that are relevant. The probabilities of evolutionary models are constructed, then, to take into consideration just what matters in explaining frequencies, and so just what matters in accounting for the standard explananda of evolutionary theory—namely, relative changes in the numbers of genes, traits, organisms, and other biological protagonists.[8] The distinction between selection and drift mirrors this explanatory divide.

---

8. The view, then, is the reverse of Matthen's, which is also in some sense explanationist: Matthen's institutional standards prescribe (relativized) explanatory standards, whereas my (objective) explanatory standards prescribe institutional standards.

5. The Probabilities that Explain Frequencies

5.1   A Wheel of Fortune   Simple stochastic models such as the drawing of balls from an urn have often served as toy examples for thinking about drift. Let me follow suit.

Begin with an especially simple probabilistic setup that has long been used to understand physical probabilities in deterministic systems, the wheel of fortune (Poincaré 1896). The wheel is a disc with alternately red and black sections, mounted on an axis around which it turns freely. In addition to the wheel itself there is a fixed pointer. To conduct a trial on the device, the wheel is given a vigorous spin. Eventually it comes to rest; the outcome of the trial, *red* or *black*, is then determined by the color of the section indicated by the pointer.

Such a wheel might have 100 sections, 50 red and 50 black, all of equal size. Or it might have the same 100 sections but with each of the black sections slightly wider than the red sections, so that the wheel is only 40% red.

Spin the wheel many times, and as everyone knows, the frequency of *red* in the resulting sequence of outcomes will tend to equal the proportion of the wheel that is painted red: you will get around 50% *red* from the first wheel described above and about 40% *red* from the second wheel. Why? A schematic answer: First, because of the wheel's construction, the physical probability of obtaining *red* on any given spin is equal to the proportion of red. Second, because of the law of large numbers, the frequency of *red* will, over many spins, tend to equal the probability of *red*. What explains the frequency, then, is whatever explains the probability.

Now intuitively, what explains the probability are properties of the wheel such as its characteristic paint scheme, and in particular the ratio of the width of the red sections to the width of the black sections; the fact that the wheel rotates smoothly around its axis; and the fact that the wheel's spin speeds are sufficiently fast and variable. As explainers of the probability, these are also explainers of the frequency.

Certain facts that are causally relevant to individual outcomes—facts that explain why some particular spin yielded *red* rather than *black*—do not help to explain the probability: the speed of any particular spin, the position of the pointer (provided that it has some position or other), the diameter of the wheel. The position of the pointer, for example, makes a difference to whether any particular spin gives you *red* or *black*, but makes no difference to the probability of *red*: put the pointer anywhere you like (or even vary its position between spins) and the probability of *red* remains the same, equal to the ratio of red to black. Because these things do not explain the probability of *red*, they do not explain the frequency of *red*—in spite of the fact that they help to cause and therefore to explain the individual occurrences of *red* and *black* that jointly determine the frequency.

The same line of thought applies to the biological case. Imagine two variants α and β, differing perhaps only in their camouflage, with α being slightly less conspicuous to predators. Their population dynamics may be modeled as follows. Each variant has its own wheel of fortune. Once a month each organism spins the applicable wheel. If the outcome is *black*, it survives predation for that month. If the outcome is *red*, death. Because α has the more effective camouflage scheme it has a little more black and a little less red on its wheel than β. Over time, then, it will tend to suffer less from predation and—all other things being equal—it will more likely than not take over the population. (Ignore drift for now.)

Suppose that this is indeed what happens: α goes to fixation at the expense of β. The fixation of α, and the extinction of β, is of course to be explained by selection, which means that it is to be explained by a difference in the frequency of (in this case) predation which is explained in turn by a corresponding difference in the probabilities of predation: β's more bloody wheel renders it more likely to be picked off. What explains selection, then, are those differences between α and β that contribute to the probability difference. In the model, the relevant difference is the proportion of red on the

wheel; pointer position and the magnitudes of individual spins are by contrast probabilistically irrelevant. In the real biological world, the relevant difference is something about the camouflage scheme; the locations of particular specimens during predator/prey encounters and various fixed features of the environment (the analogs of pointer position) are probabilistically irrelevant.

5.2   *The Skeptical Response*   I have elicited your intuitions about what is relevant to the probability, and so to the matching frequency, of *red* on the wheel of fortune, but I have not given those intuitions an objective vindication. Counting the red to black ratio but not the pointer position as relevant to the probability of *red* seems right, seems natural—but so what? Skeptics such as Rosenberg and relativists such as Matthen will put this down to nothing more than prejudice, context, convention.

Let me run the argument undermining the parameter/variable distinction in more detail. Looking at the wheel of fortune, the skeptic says: What we want to explain is the fact that the frequency of *red* outcomes on some particular long series of spins is close to one-half (supposing in this case that the wheel in question is half red and half black). What determines the facts about the frequency? The outcomes of individual spins, of course. What determines the outcomes of individual spins? A number of things: the paint scheme, the pointer position, the initial speed of particular spins, and so on. All these quantities are bona fide causes of, and explainers of, outcomes, thus they are causes and explainers of any facts determined by outcomes—including the frequency of *red*. If we are more inclined to attribute the frequency to the paint scheme than to the pointer position, it is only because the role of the paint scheme is more salient, both in its spatial extent and because of the striking mathematical correspondence between the proportion of red paint and the frequency of *red*.

How to answer the skeptic? Intuitions about probabilistic relevance are not enough—they can be explained away. Some physical and philosophical

foundation for the intuitions must be found.

In what follows, I examine the physical structure of the wheel of fortune to find a difference between what we instinctively take to be the parameters of the probability of *red* and what we take to be mere variables. My principal move is, however, philosophical rather than physical: on a conception of explanation as difference-making, I will argue, what explains the low-level facts that determine a high-level fact such as frequency may not explain that high-level fact.

5.3 Microconstancy and Smoothness Take a closer look at the wheel of fortune. Suppose that the outcome of a spin on the wheel is entirely determined by the initial spin speed v and the fixed properties of the setup (thus, fluctuations in atmospheric pressure, vibrations from passing trucks, and quantum probabilities are assumed to have no impact on the wheel's final state). Then a function can be defined specifying the outcome obtained for any value of the spin speed—a function mapping values of v to either one, representing the outcome *red*, or zero, representing *black*. I call this the wheel's *evolution function*. (Here "evolution" refers to the time evolution of dynamic systems in general, not to biological evolution.)

The form of the evolution function (for a plausible though idealized physics of the wheel's operation) is shown, for a wheel with equal amounts of red and black, in figure 1. The gray regions mark values of v that produce *red*
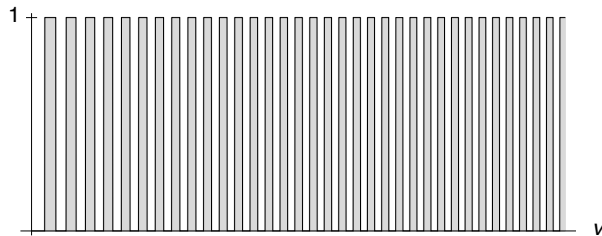

Figure 1: Evolution function for the idealized wheel of fortune

outcomes; the white regions values that produce *black* outcomes.

Since physical probability muscled its way into the sciences in the later nineteenth century, philosophers, mathematicians, and scientists have perceived in the evolution function of the idealized wheel of fortune a clue to the existence (or apparent existence) of probability in deterministic systems (von Kries 1886; Poincaré 1896; Reichenbach 2008; Hopf 1934; Strevens 2003; Abrams 2012).[9] What is striking about the evolution function is that, given almost any probability distribution over the spin speed, the probability of *red* will be the same, equal to the ratio of red to black paint on the wheel: in this case, one-half.

As an informal visual demonstration of this claim, consider figure 2. The
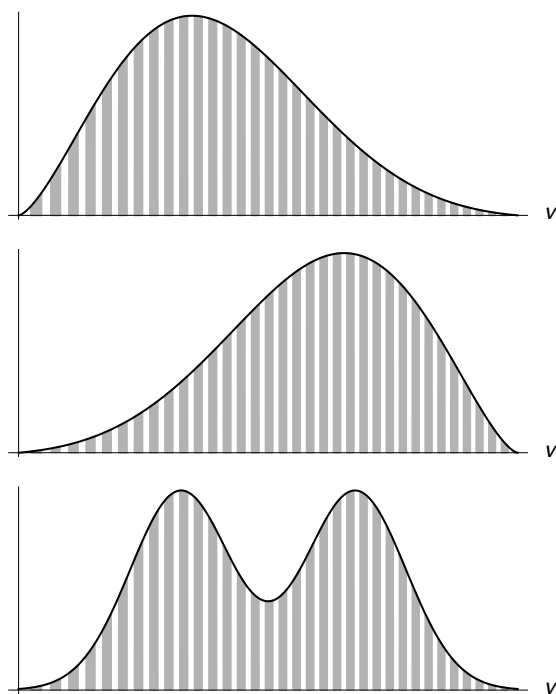
Figure 2: Almost any probability distribution over initial spin speed *v* results in a one-half probability for *red*

9. The early history of what has been called "the method of arbitrary functions" is surveyed by von Plato (1983).

figure shows three possible probability densities for spin speed superimposed on the evolution function. The probability of *red* is equal to the proportion of the density that coincides with gray parts of the evolution function. Mash, stretch, knead the density almost any way you like, and the shaded area will persist in filling about half of the total, implying a probability for *red* of about one-half.

Almost any way, but not any way at all: there are some probability distributions over spin speed that will induce a probability for *red* other than one-half. The distribution in figure 3, for example, results in a probability for red that is much greater than one-half (since considerably more than one-half of the density is shaded, implying that values of the spin speed v that result in *red* are considerably more probable than values resulting in *black*). A
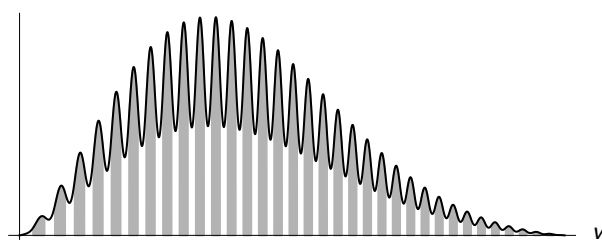

Figure 3: A rapidly oscillating probability density induces a probability for *red* much greater than one-half

sufficient condition for the probability's being one-half is that the probability distribution is relatively smooth, meaning that the corresponding density changes only slowly over small intervals—ruling out the rapid oscillations shown in figure 3. From here on, the term *smooth* should be understood in this technical sense.[10]

Let me state a more general result. The distinctive property of the wheel of fortune's evolution function I call *microconstancy*. Thus, a deterministic

---

10. In more formal work I have called smoothness *macroperiodicity* (Strevens 2003) or *microequiprobability* (Strevens 2013).

system is microconstant with respect to a given outcome if its space of initial conditions can be divided into many small ("micro-sized") contiguous regions, in each of which the proportion of initial conditions producing the outcome is the same. I call this proportion the *strike ratio* for the outcome. As you can see by inspecting the wheel of fortune's evolution function for *red* (figure 1), the spin speed v can be divided into many small intervals in each of which the proportion of speeds producing *red* is one-half, so the wheel of fortune is microconstant with respect to *red* with a strike ratio of one-half.

The following is true: if a system is microconstant with respect to a given outcome with a strike ratio of p, and its initial condition distribution is smooth in the sense characterized above (with both microconstancy and smoothness assessed using the same standard of what is micro-sized), then the probability of the outcome is approximately equal to p.[11]

Some writers have attempted to build a metaphysics of physical probability in deterministic systems upon this result, arguing that there is something inherently stochastic about a microconstant dynamics (Reichenbach 2008; Rosenthal 2010; Strevens 2011; Abrams 2012). In the remainder of this section, by contrast, I will not argue that the microconstancy-based probability has a special ontological status, but rather that it has a special explanatory status: it is the probability that best explains an approximately one-half frequency of *red*.

5.4 *Explanation as Difference-Making*   A long series of spins on the wheel of fortune yields *red* with a frequency of approximately one-half; how should that frequency be explained? Intuitively, it is the microconstancy of the wheel's dynamics, and in particular its strike ratio for *red* of one-half, together with the smoothness of the initial condition distribution, that provides the best explanation of the frequency. Citing less information (omitting the value of the strike ratio, for example) would leave the frequency a mystery, while

---

11. A formal treatment is given in Strevens (2003).

citing more information—say, the position of the pointer, or the diameter of the wheel, or the exact speed of the nineteenth spin—would add nothing explanatorily enlightening.

Against this line of thought the skeptic argues in this way. Pointer position, wheel diameter, and the speeds of particular spins are uncontroversially both causes and explainers of individual outcomes. Frequencies are nothing over and above patterns in individual outcomes. Thus, pointer position, wheel diameter, and individual spin speeds are causes and explainers of frequencies.

Where has the skeptic gone wrong? The principle of explanatory transitivity implicit in the skeptical argument does not, I submit, hold. Often, what explains a low-level state of affairs is irrelevant to the explanation of a high-level state of affairs that is realized in part by the low-level state, because the explainer is responsible for some aspect of the low-level state of affairs that makes no difference to the obtaining or otherwise of the high-level state of affairs. The seam on a baseball that breaks a window may explain why some particular shard of glass spins away during the breaking in the way that it does, but it is irrelevant to the breaking itself: with or without the seam, the window would break.

What is driving these judgments of relevance is the notion that, in order to explain an event or state of affairs, a causal factor must make a difference to whether or not the event occurs or the state of affairs obtains. How can something be a causal factor but not a difference-maker? As follows: its causal contribution to the high-level explanandum event is to make a difference to *how* the event occurs without making a difference to the fact *that* it occurs—it makes a difference to how the event is realized without making a difference to the fact that it is realized. The baseball's seam, for example, causally influences the window's breaking by helping to determine the way that the breaking is realized, namely, in part by some particular shard's spinning in some particular way. Difference-making in the "how" sense is quite consistent with the seam's not being a difference-maker in the "that" sense.

That "that" difference-making is necessary for explanatory relevance is an old idea in the philosophy of explanation; I will not defend it here (Salmon 1970; Garfinkel 1981; Lewis 1986; Salmon 1997; Woodward 2003). Nor will I advance a specific account of difference-making, though elsewhere I have argued for my own "kairetic" account over the counterfactual and probabilistic approaches offered by other writers (Strevens 2008). For the purposes of this paper, the choice does not matter greatly. In what follows, I will rely on a simple counterfactual criterion for difference-making: a causal factor makes a difference to an explanandum just in case, had the factor been absent, the explanandum would almost surely not have occurred. Although the counterfactual criterion is universally acknowledged to have limited validity, it delivers reliable judgments for principled reasons in the cases I will discuss (Strevens 2008, §§2.4, 3.81).[12]

Consider, then, pointer position and wheel diameter. Had the pointer position been moved prior to spinning the wheel, the sequence of outcomes would almost surely have been different, but the frequency of *red* would almost surely have been the same—around one-half. Likewise, had the wheel been slightly larger or slightly smaller (while retaining the same symmetry and paint scheme), particular outcomes would have differed but the frequency would have been roughly the same. Thus, though position and diameter make a difference to how the one-half frequency of *red* is realized, they do not make a difference to the fact that that approximate frequency, and not some other, is realized.[13]

---

12. The best-known failure of the counterfactual criterion occurs in scenarios where there is some backup to the actual difference-maker: because of the backup, even if the actual difference-maker had not been present, the explanandum would have occurred anyway.

13. Apparently crucial to the truth of these observations is the "almost surely" qualification in the counterfactual criterion for difference-making, concerning which, briefly, two points.

First, the qualification bridges, in my own story about difference-making, a considerably more sophisticated story, developed in Strevens (2008), Part Four.

Second, since it is the objective validity of probabilities that is at stake in this paper, you might wonder which probabilities are used to assess the "almost surely". If there is an objective probability distribution over the relevant initial conditions, then that one, of course. But I

What about spin speeds? They are not entirely irrelevant: they must be smoothly distributed. But given smoothness, particular values of actual spin speeds are irrelevant: had the nineteenth spin had a speed of 7.8 rather than 8.7 radians per second, for example, the frequency of *red* would have been much the same.

To make sense of these claims, it must be borne in mind that it is not things but the properties of things that are difference-makers (although in some cases, no property of a thing makes a difference, in which case there is no harm in saying that the thing itself is not a difference-maker). It is not quite correct, then, to say that the spin speeds were not difference-makers for the frequency. A certain property of the speeds—their smooth distribution—was a difference-maker, and certain other properties—their exact values—were not.

5.5   *The Reference Class Problem Solved*   The story so far: certain features of a wheel of fortune's physics and paint scheme and the smoothness of its initial condition distribution are relevant to explaining why the frequency of *red* outcomes in some long sequence of spins is approximately one-half; the precise values of the spin speeds and the position of the pointer are, by contrast, explanatorily irrelevant.

This shows why a one-half frequency of *red* can be explained by the microconstancy-based one-half probability of *red*: the facts that determine, and so constitute, that probability are identical to the explanatorily relevant facts.

---

have assumed, for the sake of the argument, that the world is a deterministic place. In that case, what is the source of the initial condition distribution function?

One approach to the question puts aside any need for probabilities over initial conditions. A sense is defined in which a large but finite set of actual spins can be smoothly distributed. This smoothness of the actual speeds implies, given microconstancy, a frequency approximately equal to the strike ratio (Strevens 2003, §2.33, §2.72). Further, holding this smoothness fixed, changes to pointer position, wheel size, and so on determinately have no effect on the frequency. Thus, the "almost surely" qualification to the counterfactual criterion for difference-making can be dropped.

It also shows why the frequency cannot be explained by a set of more specific probabilities that conditionalize on precise spin speeds and such matters as the pointer position. Suppose for the sake of the argument that such fine-grained probabilities exist, say, probabilities of nearly one for *red* on about half of the spins and probabilities of nearly zero on the rest. Such probabilities together make it very likely that the frequency of *red* will be approximately one-half. But they are not properly explanatory, because they are constituted in part by information that is irrelevant to the explanation of the frequency. The problem with an explanation that deploys the probabilities, then, is not that it leaves something out but that it contains *too much* information, thus explanatorily irrelevant information.

In short, for explanatory purposes you should cite a probability for *red* that takes into account the structure of the wheel's paint scheme but not its pointer position or any of the speeds with which it is initially spun, which is to say that when individuating explanatory probabilities, you should use a reference class constructed using the former but not the latter information. In other words, you should treat the paint scheme as a parameter and the pointer position and spin speeds as variables (in my technical sense, of course—the pointer position does not actually vary).

We have a solution, then, to the reference class problem. Treat as parameters those features of a system that determine microconstancy, strike ratio, and smoothness of the initial conditions; treat all other causally relevant features as variables.

This prescription is limited in two ways. First, the correctness of a reference class is relative to an explanatory task—to an explanandum. I have not said that there is a uniquely correct, objective probability for *red*, only that there is a uniquely correct probability of *red* for the purpose of explaining frequencies of *red* on long series of spins that approximately match *red*'s strike ratio. Second, the solution as framed requires a microconstant dynamics (although as Strevens (2008, §10.3) shows, there is some scope for generalization).

6. Biological Probability and Drift

6.1 Individuating Probability for Selection   The ecological dynamics that produce evolutionarily significant outcomes such as feeding and dying are microconstant with respect to those outcomes. Or at least, they are microconstant when the processes run for reasonably long times: the evolution function for a snail's being eaten by a song thrush over the course of a month is microconstant, though the evolution function for its being eaten over the course of the next three minutes most likely is not.

Further, the same is true for less significant outcomes such as being in such and such a physical state or even having such and such a position in the habitat. Thus, provided that your waiting period is long enough (a month rather than three minutes), and given a smooth distribution over initial conditions, there is a probability of a snail's being in some particular place rather than some other place that is entirely determined by the factors that determine the strike ratio of the microconstant evolution function for the outcome of *being in position x at the end of time period t*. Or as I will say, there is a microconstant probability distribution over position at the end of t—for any sufficiently long t. (Whereas in the case of the wheel of fortune, there is just one varying initial condition, spin speed, in the ecological case there will be a huge number of such conditions.)

The justification of these claims about pervasive ecological microconstancy is, as you might expect, a complicated matter; I make my best attempt in Strevens (2003). In this paper, I will simply assume microconstancy without further argument.

Biological outcomes that have microconstant probabilities will tend to occur with frequencies equal to strike ratios. When explaining such frequencies—not just the frequencies of evolutionarily significant outcomes such as births and deaths, but even the frequencies with which snails are found in this position, as opposed to that position, on the old log at the back of the garden—you have a rationale for making the distinction between parameters and variables:

what is explanatorily relevant to the frequencies, aside from the smoothness of the initial conditions, is just the set of properties that determine the strike ratio for the outcome in question; these are the properties, then, that should be used to individuate the probabilities used to explain the frequencies, solving the reference class problem.

The properties relevant to the strike ratio are more or less the properties you would expect: a snail's camouflage is relevant to death by predator, and the distribution of nutritious moss is relevant to the snail's location on the log, but the snail's exact starting position at the beginning of the month (or other long time period) is, like initial spin speed on a wheel of fortune, relevant to neither.

The facts about the microconstancy of biological dynamics are reflected, then, in our intuitions about probabilistic relevance in biology. Why are we so expert in these matters? This is a question I answer in Strevens (2013), but not here.

To summarize the story so far: microconstancy provides an objective individuation of probabilities for an array of outcomes—ranging from position to personal extinction—of relatively long-term biological processes, provided that those probabilities are used to explain corresponding strike ratio–matching frequencies.

When do frequencies that match probabilities appear in evolutionary explanation? Most saliently, they appear in explanations that attribute changes in the makeup of a population to selection. Suppose, for example, that variant α has replaced β in the population because of its superior camouflage scheme. In virtue of the difference in camouflage (compare to the difference in the wheels' paint schemes in section 5.1), α's probability of going for a month without being eaten is somewhat higher than β's probability of the same. Typically, the frequencies will roughly match the probabilities: α will evade predation more often than β. All other things being equal, the relative frequency of α will therefore continue to increase, month after month, until β is crowded

out of the habitat. Because such an explanation turns entirely on probability-matching frequencies, it should be conducted entirely in terms of information relevant to explaining such frequencies. Probabilities individuated as I have recommended are the right tools for the job.

So much for explanation by natural selection. How should the parameter/variable distinction be drawn for probabilities invoked in explanation by drift? We should use the same strategy as in the case of selection, counting as parameters just those factors that affect evolutionarily significant outcomes' strike ratios—so I claim.

It looks to be an easy claim to sustain on the traditional view of drift as sampling error, since precisely the same sorts of probabilities feature in both cases: the probability of being eaten in the course of month and so on. But that is too quick: not only the probabilities but the explananda must be of the same type. You might wonder whether the explananda targeted by drift explanations are frequencies. If not, there is trouble, since it is only with respect to frequencies that microconstancy draws an explanatory distinction between variables and parameters. This issue is explored in the next section, section 6.2.

What about alternatives to the traditional view of drift? I will consider only the process-driven view (section 3). The probabilities that matter in the process view are those invoked in the characterization of indiscriminate sampling. On both the traditional and modern notions of indiscriminateness, many of these probabilities are attached to processes unfolding over relatively brief periods of time, such as the time taken by a typical predator/prey encounter. The dynamics mapping the initial conditions of such processes to evolutionarily significant outcomes are likely not microconstant, so the process view's drift-defining probabilities are likely not microconstant—in which case, the individuation criterion proposed above does not apply. This objection is confronted in section 6.3.

6.2 Drift Explanation Is Frequency Explanation   Two identical twins are strolling along a ridge top; one is killed by lightning but the other survives. Why? You might be tempted to reply "Drift". In this case, drift explains a single evolutionarily significant event, rather than a frequency with which such events occur. The explanation is equivalent, in effect, to saying that the outcome occurred "by chance".

The great majority of serious evolutionary explanations citing drift are not, however, mere attributions of arbitrary deaths and wonky statistics to "chance". They rather use mathematical models of evolutionary processes to make predictions about differential reproduction.

Consider, for example, the explanation of homozygosity due to a population bottleneck. A study of northern elephant seals, to take a superb example of the genre, found population-wide homozygosity at a large number of genetic loci. This is explained by the seals' having been hunted almost to extinction in the 1890s (with perhaps only 20 specimens at one point surviving). As the population recovered from this low ebb, it was highly likely that any particular allele would drift to extinction, even given modest fitness advantages to heterozygotes. The vast majority of alleles of the genes in question evidently did go extinct, leaving only one survivor at each locus (Bonnell and Selander 1974).

The probabilistic reasoning here is as follows. Any allele can go extinct if has a run of bad luck (if, for example, its sole possessors, through no fault of the allele, fail to contribute offspring to the next generation—quite probable in elephant seals where a few males do all the reproduction—or if the offspring they contribute by chance contain only its rival allele). In extremely small populations, this run need not be very long; if the population remains small for a while, then, so that there are many chances for any given allele to suffer such a run, there is a high probability that most alleles go to extinction. Hence wholesale homozygosity.

Consider the external ecological outcomes that play a role in allele extinc-

tion (putting aside, then, the vicissitudes of meiosis, the fertilization process, and so on). The relevant outcomes are those that enable or frustrate an allele's transmission, thus the same outcomes by which selection occurs, such as birth and death. Compared to explanation by selection, there is a complication: we are interested in explaining not frequencies of individual births and deaths, but rather frequencies of "bad runs"—the evolutionary equivalent of explaining the frequency with which you see, on a wheel of fortune, ten *red* outcomes in a row. The same probabilities that explain frequencies of individual outcomes, however, explain frequencies of runs. A probability for *red* of one-half, for example, explains both why the number of *red* outcomes in 1000 trials on a wheel of fortune is about 500, and also why the number of sequences of five or more *red* outcomes in a row in the same number of trials is about 16.[14]

The probabilities that explain evolution by drift, then—as opposed to one-off incidents—have the same sort of explananda as the probabilities that explain evolution by natural selection: frequencies of evolutionarily significant events.

6.3   *Short-Period Probabilities*   To formulate the problem of probabilities attached to short-period ecological processes that lack microconstancy, consider again an encounter between a song thrush and a snail. Divide the initial conditions for the encounter into two sets, central and peripheral. In the central set, the snail is near the center of the thrush's visual field, while in the peripheral set, it is nearer the edges. Consequently, more initial conditions within the central set than the peripheral set result in snail capture and death. Define a binary variable, *centrality*, to represent whether the initial conditions of a particular thrush/snail encounter are central or peripheral. (Perhaps centrality takes the value one when the conditions are central, zero otherwise.)

---

14. Not counting sub-sequences.

Now consider the short-period probability of a snail's surviving a thrush encounter—the sort of probability that figures explicitly in the process approach's traditional definition of indiscriminate sampling and that is also, I have argued, presumed by the modern definition. For the purpose of applying these definitions, relative to what reference class ought the probability to be defined? Should we distinguish two "narrow probabilities", the probability of death given that centrality is equal to one and the probability given that centrality is zero? Or should we use a single "broad probability", a combination of these two probabilities weighted by the probability distribution over centrality? In short: should we treat centrality as a parameter or as a variable?

If the dynamics underlying each of the narrower probabilities were microconstant, with a higher strike ratio for death when centrality is equal to one, a straightforward answer could be given: centrality affects the strike ratios, so it should count as a parameter. Suppose, however, that the dynamics is not microconstant. Can any objective, substantive parameter/variable distinction be sustained? Yes—and in seeing how, you will see also that the "straightforward answer" is in fact incorrect.

Consider a wheel of fortune analogy. You generate a series of outcomes, *red* and *black*, using the following procedure: first, you spin a wheel that is half red and half black—call it the *even wheel*. Then, depending on the outcome, *red* or *black*, you spin one of two other wheels; call them the red and black wheels. The red wheel is 75% red and so produces *red* with a three-quarters probability; the black wheel is 75% black and so produces *red* with a one-quarter probability. (Think of these two wheels as representing the two kinds of thrush/snail encounter, and think of the outcome that determines which of the two wheels to spin as representing the value of centrality. In the wheel of fortune case, the red and black wheels have a microconstant dynamics, whereas in the biological case by assumption the analogous dynamics is non-microconstant, but I promise not to appeal to microconstancy; this will enable me to move faster by not having to describe a new, non-microconstant

probabilistic setup.)

Now suppose that you conduct 100 trials, generating 100 outcomes. (You will make 200 spins, then, since each trial consists of two spins: a spin on the even wheel to determine whether to use the red or the black wheel, and then a spin on the indicated wheel to determine the final outcome.) About one-half of the outcomes are *red*. How to explain this probability-matching frequency?

Here are two strategies. First, you could note that the probability of *red* on any particular trial (consisting of two spins) is one-half, and that the frequency reflects this probability. Second, you could record which of the red and black wheels was spun on each trial, noting that (as I will presume) in about one-half of the trials the red wheel was spun and in the other half the black, and that the frequency of *red* resulting from the spins of the red wheel was about three-quarters and that the frequency of *red* resulting from the spins of the black wheel was about one-quarter, for a net frequency of about one-half.

The first of these explanatory strategies appeals only to the coarse-grained or broad probability of *red*, a composite of the probabilities attached to each of the three wheels which is the same in each trial; the second appeals to the fine-grained or narrow probability of *red*, which depends on whether the red or the black wheel is spun and which therefore varies from trial to trial. Both strategies predict what is to be explained, the one-half frequency of *red*, by citing factors causally involved in the production of that frequency. But the first explanation does so with strictly less information—it does not specify which of the red and black wheels was used for each trial, and so it does not specify the narrow probabilities at work in each trial. That the prediction is possible using broad probabilities alone shows that the additional information needed to deploy narrow probabilities does not make a difference to the frequency of *red*, any more than does the exact sequence of *red* and *black* outcomes that determines the frequency.

I want to say the same thing about the explanatory relevance of information about centrality in the thrush/snail encounters, namely, that centrality

is explanatorily irrelevant to the frequency with which thrushes eat snails. If it is these frequencies you are trying to explain, then—as is the case when invoking the mathematics of drift to explain why, say, the ratio of two equally effective snail camouflage schemes in a small population exhibits a certain pattern of variation over time—the snails' centrality or otherwise in particular encounters should not be mentioned.

To individuate short-period probabilities, then, you should as a general rule put probability distributions over variables such as centrality, citing for explanatory purposes broad probabilities, rather than conditioning on specific values and thereby delineating and citing narrow probabilities. Of course, centrality is a contrived variable, an unsubtle coarse-graining of predator and prey position. The probability distribution should go over the predator and prey positions themselves, and over all other initial conditions causally relevant to the outcome: orientation, hunger, physical condition, and so on.

Where do these initial condition probabilities come from? With respect to what reference class should they be individuated? They are the long-term microconstant probabilities introduced at the beginning of section 6.1, individuated in the same way as all microconstant probabilities. To find the probability distribution over centrality, then, look to the microconstant probability distribution over the positions of snails and thrushes induced by the dynamics of the previous month (or any other appropriately long period—it will be the same distribution for any sensible choice).

The individuation rule for short-period probabilities advocated here—do not condition on actual values of initial conditions of short-period encounters, but rather put a long-term probability distribution over absolutely all of them—creates a pressure toward breadth in probability, that is, a pressure toward using the widest feasible reference class. Other principles of explanation will prevent things from going too far. Consider, for example, the following worry. Some snails are eaten by thrushes; others are struck by lightning. In treating these events, either to define drift using short-period probabilities or

for any other purposes, you want to distinguish the probability of being eaten (when in close proximity to a thrush) from the probability of being struck by lightning (when caught in the open during an electrical storm). But you might think that the impetus toward breadth forces you to amalgamate the two, as follows. Consider as a "scenario type" a snail's either running into a thrush or being caught in a storm. (Both thrush and storm encounters fall under this disjunctive type, then.) Define the variable thrush-not-storm as follows: it has the value one for thrush encounters and zero for storm entrapments. Like the centrality variable, then, it divides a broader scenario into two narrower scenarios. Intuitively, the division is in this case a good one; it is a mistake to amalgamate for explanatory purposes the dynamics of predation and lightning into a single broad probability. But can you not put a probability distribution over the thrush-not-storm variable to define a broad probability of death-by-thrush-or-lightning-strike for snails that find themselves in a thrush-or-storm scenario? If such a distribution exists, which I will not question, you can. Why not use the broad probability then?

You should not use it because explanations should cite a single causal mechanism, not a disjunction of causal mechanisms (a consideration understood very well by process theorists of drift). When explaining a death, then, you should specify the particular mechanism at work: bird or bolt. Thus, you should not cite probabilities that fail to distinguish which of these mechanisms is at work.[15]

Let me finish there. Microconstancy, I have argued, provides the foundation for an objective individuation of the probabilities used by process theorists to characterize indiscriminate sampling, and so for an objective process-theoretic account of the distinction between selection and drift. This may not comprise a complete solution to evolutionary biology's reference class problems, but it is a good start.

---

15. On the individuation of mechanisms for causal explanation, see Strevens (2008, §3.6).

Acknowledgments

References

Abrams, M. (2009). What determines biological fitness? The problem of the reference environment. *Synthese* 166:21–40.

———. (2012). Mechanistic probability. *Synthese* 187:343–375.

Beatty, J. (1984). Chance and natural selection. *Philosophy of Science* 51:183–211.

Bonnell, M. L. and R. K. Selander. (1974). Elephant seals: Genetic variation and near extinction. *Science* 24:908–909.

Brandon, R. (2005). The difference between selection and drift: A reply to Millstein. *Biology and Philosophy* 20:153–170.

———. (2006). The principle of drift: Biology's first law. *Journal of Philosophy* 103:319–335.

Garfinkel, A. (1981). *Forms of Explanation*. Yale University Press, New Haven, CT.

Gildenhuys, P. (2009). An explication of the causal dimension of drift. *British Journal for the Philosophy of Science* 60:521–555.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford University Press, Oxford.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese* 156:563–585.

Hempel, C. G. (1965). Aspects of scientific explanation. In *Aspects of Scientific Explanation*, chap. 12, pp. 331–496. Free Press, New York.

Hodge, M. J. S. (1987). Natural selection as a causal, empirical and probabilistic theory. In L. Krüger, G. Gigerenzer, and M. S. Morgan (eds.), *The Probabilistic Revolution*, volume 2. MIT Press, Cambridge, MA.

Hopf, E. (1934). On causality, statistics and probability. *Journal of Mathematics and Physics* 13:51–102.

von Kries, J. (1886). *Die Principien der Wahrscheinlichkeitsrechnung*. Mohr, Freiburg.

Lewis, D. (1986). Causal explanation. In *Philosophical Papers*, volume 2, pp. 214–240. Oxford University Press, Oxford.

Matthen, M. (2009). Statistically abstractive explanation. *Philosophy of Science* 76:464–487.

Matthen, M. and A. Ariew. (2002). Two ways of thinking about fitness and natural selection. *Journal of Philosophy* 49:55–83.

Millstein, R. L. (2002). Are random drift and natural selection conceptually distinct? *Biology and Philosophy* 17:33–53.

———. (2008). Distinguishing drift and selection empirically: The great snail debate of the 1950s. *Journal of the History of Biology* 41:339–367.

von Plato, J. (1983). The method of arbitrary functions. *British Journal for the Philosophy of Science* 34:37–47.

Poincaré, H. (1896). *Calcul des Probabilités*. First edition. Gauthier-Villars, Paris.

Reichenbach, H. (1949). *The Theory of Probability*. University of California Press, Berkeley, CA.

———. (2008). *The Concept of Probability in the Mathematical Representation of Reality*. Translated and edited by F. Eberhardt and C. Glymour. Open Court, Chicago. Reichenbach's doctoral dissertation, originally published in 1916.

Rosenberg, A. (1994). *Instrumental Biology or The Disunity of Science*. University of Chicago Press, Chicago.

Rosenthal, J. (2010). The natural-range conception of probability. In G. Ernst and A. Hüttemann (eds.), *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*. Cambridge University Press, Cambridge.

Salmon, W. C. (1970). Statistical explanation. In R. G. Colodny (ed.), *The Nature and Function of Scientific Theories*, pp. 173–231. University of Pittsburgh Press, Pittsburgh.

———. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press, Princeton, NJ.

———. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science* 64:461–477.

Sober, E. (1984). *The Nature of Selection*. MIT Press, Cambridge, MA.

Strevens, M. (2003). *Bigger than Chaos: Understanding Complexity through Probability*. Harvard University Press, Cambridge, MA.

———. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.

———. (2011). Probability out of determinism. In C. Beisbart and S. Hartmann (eds.), *Probabilities in Physics*, pp. 339–364. Oxford University Press, Oxford.

———. (2013). *Tychomancy: Inferring Probability from Causal Structure*. Harvard University Press, Cambridge, MA.

———. (2016). Special-science autonomy and the division of labor. In M. Couch and J. Pfeifer (eds.), *The Philosophy of Philip Kitcher*. Oxford University Press, Oxford.

Walsh, D. M. (2007). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science* 74:281–303.

Walsh, D. M., T. Lewens, and A. Ariew. (2002). The trials of life: Natural selection and random drift. *Philosophy of Science* 69:452–473.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.