Probability and Ecological Complexity

MARK COLYVAN
Philosophy Program
University of Queensland
Brisbane, Queensland 4072
Australia
E-mail: mcolyvan@uq.edu.au

There is something genuinely puzzling about large-scale simplicity emerging in systems that are complex at the small scale. Consider, for example, a population of hares. Clearly, the number of hares at any given time depends on hare fertility rates, the weather, the number of predators, the health of the predators, availability of hare resources, motor vehicle traffic, individual hare locations, colour of individual hares, and so on. Indeed, given the incredibly complexity of the hares' environment at the small-scale, it is amazing that *anything* can be said about hare abundances. But not only can we say something about hare abundances, we can formulate equations for hare abundance as a function of time that are remarkably accurate. But most amazing of all is that such equations have very few parameters—in the simplest cases, just the growth rate and an initial population abundance. How can this be? How can we ignore all the small-scale factors when they clearly play major roles in determining abundance? Put somewhat more grandiosely: How is population ecology possible?

Of course it's not just population ecology that exhibits such small-scale complexity and large-scale simplicity. Other important systems include the weather (in which various cyclic behaviours like El Niños and ice ages occur despite the day-to-day chaos) and gases in equilibrium (in which apparently random motions of molecules result in the gas obeying the ideal gas law). The examples can easily be multiplied. The general problem is the same: How can seemingly unpredictable and complex microbehaviour of complex systems result in predictable and simple macrobehaviour? Providing an answer to this question is the central task of Michael Strevens' excellent book *Bigger than Chaos: Understanding Complexity through Probability*.

I aim to explain why so many laws governing complex systems have only a few variables. I leave it to the individual sciences to explain why those few variables are related in the way that they are; my question is one that the individual sciences seldom if ever pose: the question as to why there should be so few variables in the laws to begin with. (p. 6)

From Micro-complexity to Macro-simplicity

Bigger than Chaos provides a detailed and systematic proposal in terms of *enion probability analysis*. In this section, I will give a very brief overview of this theory, but first I'll need to introduce some of Strevens' terminology. *Enions* are the basic units of the system under study. In population ecology the enions are the individual organisms; in thermodynamics the enions are the molecules. The level of the enions is the microlevel: a particular lynx eating a particular hare, or a particular molecule of gas colliding with the wall of a container. *Microvariables* are variables concerning the enions: the location of a particular enion, for example is a microvariable. The macrolevel abstracts away from individual enions and involves only enion statistics. And *macrovariables* are variables concerning only macrostates. Population abundance and gas pressure, for example, are macrovariables in population ecology and thermodynamics respectively.

Enion probability analysis is a method for understanding how to get from micro-complexity to macro-simplicity. It involves three steps. First, the behaviour of the system's enions is specified probabilistically—the probability of a given hare dying, say. These are the *enion probabilities*. Next, these probabilities are aggregated, giving a probability distribution describing the behaviour of macrovariables in terms of only macrovariables. Finally, a macrolevel law is derived from the macrolevel probability distribution (pp. 12–16). These laws are typically quite remarkable in that, despite the huge number of microvariables, the laws themselves are very simple in the sense that they contain very few macrovariables. It would seem that the system's macrolevel behaviour does not depend very much on the microvariables, and yet the macrolevel behaviour is fully determined by those microvariables (p. 12). This is another way of stating the problem that enion probability analysis is designed to address.

It would seem that the crucial move in all this is step two: the aggregation process. This is where microlevel information drops out of the picture. But the techniques typically employed in this aggregation process rely on the

2

enion probabilities having certain features, so step one is really the crucial step. More specifically, the enion probabilities must satisfy what Strevens calls *the probabilistic supercondition*:

> *The values of enion probabilities are unaffected by conditioning on microlevel information.* (p. 22)

For example, if this condition is satisfied for our hare population, the probability of a particular hare dying in a given time period is stochastically independent of the probability of any other hare's death (or any other microlevel event). It also ensures that the probability of death for the hare in question is a function of only macrolevel information. Most of Strevens' book (chaps. 2–3) is devoted to showing that this condition holds and explaining why it holds—at least for a large class of cases.

A simple example will help. The result of a roulette wheel spin will be either red or black. Let's say we're interested in the outcome *red*. Clearly the result of the roulette spin will depend on many microlevel factors, namely the mechanics of the wheel and the (microlevel) initial conditions such as initial wheel position, the wheel's spin speed, and ball trajectory and speed. But it is well known that when determining the macrolevel statistical law for the probability of red, we can ignore all the microlevel information. The reason is that the roulette wheel spin has a very nice property which Strevens calls *microconstancy*.

The wheel is designed so that from any starting configuration of initial conditions (except some very special ones like no spin on the wheel and releasing the ball so that it rolls straight down into a red slot) the probability of red is the same as the probability of black. As everybody knows, this is achieved by making sure that the friction on the wheel is constant—there are no rough patches asymmetrically distributed—and the red and black slots alternate and are of the same size. So it doesn't matter whether the wheel turns 20 times or even 3/4 of a turn before the ball comes to rest. The ratio of red outcomes to black outcomes in any region is the same, so long as the wheel is operated under normal conditions (i.e., there are no illegitimate null spins as above). If the roulette wheel were to have all the red slots on one side and all the black slots on the other, its spins would no longer have this property of microconstancy—some sets of initial conditions would be more likely to result in red than black, while other sets of initial conditions would be more likely to result in black. But if the wheel were to have twice as many red slots as black slots, it might still yield microconstant trials, so long as the slots were arranged in a suitable way—red-red-black, say. The wheel would no longer be a fair wheel—the probability of red would be twice that of black—but it would still be microconstant, because the probability of red

would not depend on the initial conditions. (Actually, microconstancy must be relativised to classes of initial conditions. The normal roulette wheel is microconstant for almost all initial conditions—we've already seen that null spins or the like will not result in microconstant trials. And the wheel with all the red slots on one side will be microconstant for sufficiently fast spins. But let's leave such complications aside.)

Various gambling devices like roulette wheels, dice, and coin tosses are microconstant. That's what makes them statistically so tractable. But Strevens goes on to argue that many complex systems also exhibit microconstancy (or near enough) and this is what explains the satisfaction of the probabilistic supercondition, and ultimately explains the simplicity of macrolevel laws. Strevens sums up his project rather nicely in the following passage.

> At the microlevel of a complex system there is chaos—a proliferation of enions, interacting in many and various ways. Because of this chaos, the behaviour of complex systems might be thought to be unstable, impossible to describe simply, and quite unpredictable.
>
> Yet the microlevel also contains the seeds of something much bigger than chaos: microconstant probability. Microconstant probability is bigger than chaos because it is indifferent to the microlevel details that exhibit chaos. But it is bigger, too, in the way that it takes the source of chaos—sensitivity to initial conditions—and creates microdynamic independence, the independence of enion probabilities, and ultimately the macrolevel probabilities that sculpt the simple dynamic lines of our world. (p. 332)

I can't possibly do justice to all the intricacies on enion probability analysis here. Suffice to say that I found Strevens' account interesting, very rigorous, and rather plausible. There are some gaps to be filled in to be sure—and, most such gaps are acknowledged by Strevens—but *Bigger than Chaos* makes a big start on an important problem for both complex systems theory and philosophy of science. It is not an easy book though. Strevens is not shy of mathematics and he invokes quite a bit of it along the way, proving a number of formal results. This is hardly surprising though. One would expect that a book on the foundations of probability would employ quite a bit of mathematics and, by and large, all the mathematics Strevens invokes is genuinely needed. (Though, of course, if you're willing to take Strevens' word on technical matters, much of the technical material can be avoided. Indeed, the most technical material is relegated to appendices in anticipation of readers whose primary interests are not in the formal details.)

But the book is technical in another way. Strevens employs a great deal of technical terminology that is unique to his account. Terms such as *enion*, *enion probability*, *microconstancy*, *IC-evolution function*, *complex probability*, *simple probability*, *simplex probability*, *strike ratio*, *randomizing variable*, and many more, are all Strevens' inventions. The reader must keep track of all these new terms of art and as a result I found the book rather slow and heavy going. In fairness to Strevens, though, he does provide a very good glossary to help the reader in this regard, and the text itself contains many reminders about the definitions of key terms. I should also add that the terms Strevens introduces do real work—he is not just introducing jargon for the sake of it. The reason he introduces so many new technical terms is that they are needed to distinguish various concepts important for the account developed (and there are no preexisting terms for the concepts in question). At the end of the day, the additional technical terminology is needed, Strevens does his best to help the reader keep track of it all, but the result is not an easy read.

In light of the technical nature of *Bigger than Chaos*, it runs the risk of only being read by aficionados of complex systems theory, philosophers of probability theory, and the like. That would be a shame. This book really does deserve the attention of a wider audience. Though it is most certainly not a book about population ecology, population ecologists and philosophers of ecology stand to gain a great deal from this book. Indeed, *Bigger than Chaos* rather nicely complements some of the recent work on philosophy of ecology (e.g., Cooper 2003) and some of the recent mathematical treatments of population ecology (e.g., Turchin 2003). For one thing, this book might just explain why theoretical population ecology is possible.

In the remainder of this article I will focus on a couple of the philosophical issues that arise in relation to one of Strevens' primary examples of a complex system: population ecology. More specifically, enion probability analysis has some very interesting consequences for the issue of the status of laws in population ecology and the nature of explanation in ecology. These issues are touched upon by Strevens but both warrant further attention.

Ecological Laws

Strevens makes it very clear from the outset (p. 5) that he is not in the business of deriving laws for ecology or for statistical mechanics (his two ongoing, examples in the book); he merely wants to demonstrate how it is possible to derive simple macrolevel laws for disciplines such as these. And true to his word, he never even states an example of such a macro law for

5

ecology, though at one point (p. 15) he alludes to the logistic equation,

$$\frac{1}{N}\frac{dN}{dt} = r(1 - \frac{N}{K}),$$

(where N is the population abundance, r is the growth rate, t is time, and K is the carrying capacity). In any case, the logistic equation is certainly a paradigmatic example of what Strevens has in mind. The logistic equation has only one variable, time, and three parameters: initial population, carrying capacity, and growth rate. All three parameters concern macrolevel information. So it is worth keeping the logistic equation in mind as an example in what follows.

Now let's turn our attention to the *status* of equations like the logistic equation. In particular, we might ask if the logistic equation is a law. In recent times there has been a very lively debate in population ecology on the issue of whether the discipline is law governed at all, let alone whether there are *simple* ecological laws (Cooper 2003; Colyvan and Ginzburg 2003; Lawton 1999; Murray 1999; Turchin 2001). The problem is that the logistic equation is not only simple, it is *simplistic*. It ignores age structure of populations, for instance. And despite being capable of producing incredibly complex behaviour (May 1974), the standard Lotka-Volterra predator–prey equations are also unrealistically simple. (See Ginzburg 1986 for criticism of such models and Turchin 2003 for details of more sophisticated predator–prey models.) While I think there is good reason to accept equations like Malthus's equation or the logistic equation as ecological laws, the fact remains that few ecologists are content with such simple equations. Ecologists want more realistic mathematical models that result in more accurate predictions. Introducing age structure is a very common move in this regard.

Strevens considers complications such as age and health structure of the population (pp. 287–290). On the face of it, at least, the introduction of age and health structures of populations presents a serious problem for Strevens. It amounts to a violation of the probabilistic supercondition, since the probability of a given hare's survival will depend, in part, on microlevel information, namely, the hare's health and age. Strevens has some ingenious responses to such cases. For example, we can introduce a small number of age structures so that we can talk of a hare's probability of survival, given that it belongs to a particular age class. A minor complication and, moreover, such age-structure models are well known in ecology. But surely the probability of a hare's survival also depends on the health of others—the health of its predators, for instance. Here Strevens suggests that since any individual hare is no more likely to encounter any particular predator than any other, we can average over the predator population and consider only av-

erage predator health. Again, a minor complication and again the microlevel information drops out.

There are some interesting issues here though. First, a minor point. It is not clear that hares are just as likely to encounter a healthy predator (lynxes in the classic example) as an unhealthy one. It is highly plausible that healthy lynxes are more mobile and so hares are more likely to encounter the healthy ones. Second, and more interesting, is a problem that Strevens raises: the averaging strategy assumes that the health of the predators is not changing over time. That is, the strategy assumes that the predator population is in equilibrium, with respect to health. Strevens comments that the averaging strategy "can reasonably be applied [...] in situations where equilibrium can reasonably be expected" (p. 290). He does not pursue the matter further. He obviously takes this as a satisfactory resolution of the problem, but there are a number of questions left hanging. What are the situations in which predator health is in equilibrium? And what about situations where it is known that predator health is not in equilibrium? Consider, for example, a hare population when some epidemic impacts on the health of the lynx population. Whether such non-equilibrium cases present serious problems for Strevens is not clear. Perhaps when predator health is not in equilibrium, the standard population laws break down as well. Enion probability analysis would not be able to explain such laws if there were any, but this is not a problem if there are no such laws. In such cases, enion probability analysis might even be thought to be able to shed light on the underlying assumptions and limitations of population laws. But all this needs further investigation.

But there are more troubling cases for Strevens. I have in mind cases where a particular hare's probability of survival depends on other hares or lynxes, and where the encounters between individuals is not random. Strevens suggests that such cases are rare (p. 289), but it is not clear to me that they are. Consider, for instance, herding behaviour. This amounts to highly co-ordinated enion interactions where an individual's probability of survival depends on others in the herd. (If others in the herd direct the herd to a region densely populated with predators, for instance, the probability of survival of the individual in question decreases.) In populations, like humans, there are social structures and these too amount to non-random enion interactions that impact on the probability of survival. (As an academic, I am much more inclined to have interactions with members of my own faculty than with, say, ruthless drug lords. As much as I dislike faculty meetings, they are rarely life threatening, so I take it that my chances of survival are higher as a result of being an Australian academic rather than a Columbian drug dealer.) It is also worth noting explicitly that in regard to both social structures and herding behaviour we are not inclined to think that the

relevant population laws break down. In fact, a major application of population modelling is in making predictions about future human population abundances, both nationally and globally.

Consider another kind of non-random interaction that impacts on an individual's probability of survival: the maternal effect. According to the maternal effect hypothesis, well-nourished mothers (or more generally, parents) produce not only more offspring, but also healthier offspring (see Mousseau, and Fox 1998 for details). It can be thought of as the inheritance of (non-genetic) quality. If this hypothesis is correct, then *every* individual's probability of survival depends on other individuals (namely their parents and their parents' environment) and, moreover, the parent–offspring interaction is not random. For a start, the probability that an individual has the parents she does is surely maximal and so is no random event. Though this reading of the probability in question might be thought to be the wrong way to approach the issue. Instead, it might be argued that the probability of interest is the probability that an individual has some particular individual as their parent. But in this sense too, the encounter between individuals and parents is not (obviously) random. After all, healthy parents produce more offspring, so in any given population, the chance of an individual coming from a healthy parent is a function of both the number of healthy parents and the differential reproductive rates between healthy and unhealthy parents.

Whichever way you look at it, it seems that the maternal effect, if correct, results in widespread violation of the probabilistic supercondition. Moreover, this spells trouble for enion probability analysis because there *are* simple models of population growth based on the maternal effect (Ginzburg and Taneyhill 1994). It's just that enion probability analysis does not (at present) have the resources to explain the simplicity of such models. (Strevens suggests various strategies for dealing with such cases (pp. 290–292), such as renormalisation, but he leaves these strategies for possible future extensions to enion probability analysis.) What is interesting about the maternal effect example is not that it provides any insurmountable problem for the method of enion probability analysis. Rather, the maternal effect suggests that whether the method of enion probability analysis is applicable to a given complex system may depend on empirical details of the system in question—whether the maternal effect or herding behaviour are true of the population in question. This result is perhaps not so surprising but it does run against Strevens' hope of providing a tool-kit of methods for explaining simple macrolevel laws wherever they occur. That, however, may have been hoping for just a bit too much.

Although Strevens takes for granted that there are simple macrolevel laws of ecology, he does not presuppose anything about the details of such laws.

His aim is to explain how such laws are possible. But, as I've already mentioned, there are some who would deny the existence (or even the possibility) of such laws. Strevens' project, however, has an interesting and unexpected consequence for this debate. Strevens, in effect, provides us with a conditional existence claim about ecological laws: if the conditions required for enion probability theory are satisfied, then simple macrolevel laws are possible. This conditional existence claim is interesting for a couple of reasons. First, it serves to highlight the importance of determining whether the probabilistic supercondition holds in ecology. If the supercondition holds, the sceptics about ecological laws are wrong—such laws exist, though there remains the substantial issue of discovering what those laws are. Second, in light of this, the debates over the maternal effect and the like take on new significance. For the maternal effect and other such complications strike me as the most serious challenge to the satisfaction of the probabilistic supercondition in ecology. And the satisfactory resolution of such complications would seem to depend on the extensions of enion probability analysis that Strevens gestures toward in section 4.6.4 (pp. 290–292). If what I've suggested above is correct, these extensions are needed sooner rather than later.

One final point on enion probability analysis and laws. Lets suppose we have two competing and equally simple macrolevel laws, L_1 and L_2. Further suppose that both laws fare equally well with respect to empirical adequacy and other theoretical virtues. Might enion probability analysis help us decide which of L_1 or L_2 we ought to believe? Maybe. Suppose that L_1 but not L_2 is based on microlevel assumptions that are seen to satisfy the probabilistic supercondition. In such cases, the simplicity of L_1 is explicable but the simplicity of L_2 is not. In the name of mystery reduction, we ought to opt for the less mysterious law, L_1. It would seem that enion probability analysis might well have consequences for theory choice. (Strevens makes a related point in the final section of the book (section 5.6) when he suggests that enion probability analysis gives us some reason to believe that quantum probabilities are not fundamental.) This amounts to the suggestion that we should not only seek simple laws, but we should also seek laws underwritten by microconstancy.

The Explanatory Role of Mathematics in Ecological Laws

Strevens points out (pp. 8–9) that in the history of probability theory there have been at least three quite distinct attitudes towards the explanatory power of statistical laws. The first view is that the statistical regularities expressed in statistical laws are explained by the law of large numbers; here,

probability theory plays the starring role. This view was argued for by Poisson and later by Maxwell and Boltzmann. Another, perhaps more popular, view is that the statistical regularities are due to some non-probabilistic causes. This view was held by Quetelet and Buckle. On this view, probability theory plays only a supporting role; the law of large numbers is invoked to argue that short-term probabilistic disorder will, in the long run, cancel out, leaving non-probabilistic order. The third view is frequentism, according to which probabilities are identified with relative frequencies. On this view the law of large numbers plays no explanatory role at all—it's a trivial consequence of the definition of 'probability'. Probabilities cannot explain statistical regularities because probabilities just *are* those regularities.

Strevens points out that enion probability analysis shares a great deal with the first view—the view that probability plays the primary explanatory role in statistical laws. Though, again as Strevens points out (p. 8), the case for the first view is made stronger by enion probability analysis, because enion probability analysis emphasises the explanatory significance of the physical properties that underwrite the legitimacy of the application of the law of large numbers (namely, stochastic independence). Without this, the Poisson view would seem to rely on the mathematics itself for the explanation of the statistical regularities. In this section I'd like to explore this issue a little further.

There may be a fundamental difference in the way physicists treat mathematical explanation and the way ecologists do (Ginzburg and Colyvan 2004, pp. 30–33). Fundamental physics, at least, is very close to scientific bedrock and so causal explanation (if not all explanation) can reach an end in physics. In the place of causal explanation we may find mathematical explanations, appeals to various unifying principles, and sometimes no explanations at all— just appeals to brute fact. But ecology is a long way from scientific bedrock. Ecology is presumably underwritten by biology, physiology, theories of animal behaviour (even animal psychology), and ultimately biochemistry and physics. Mathematics plays an important role in ecology, of course, but it doesn't seem appropriate for mathematics to play the explanatory role it often does in physics. (See Colyvan 2001, pp. 49–51 for more on the explanatory role mathematics can play in science.)

Consider the difference between Strevens' two main examples in the book: population ecology and statistical mechanics. In ecology we are not inclined to assume any brute facts about genuinely stochastic behaviour (as we do in statistical mechanics). Although various animal foraging decisions, for instance, may *look* stochastic, they are nevertheless underwritten by some, perhaps unknown, neuro-physiological story. Or consider the various body-size allometries of macroecology (Calder 1996). The Generation Time allom-

etry (Bonner 1965) is the relationship between body weight and maturation time: maturation time is proportional to a 1/4 power of body weight. Although to date there is no generally accepted explanation for the Generation Time and other allometries, no-one would be happy with accepting them as simple brute ecological facts. They cry out for explanation and in this regard the allometries seem quite different from, say, the inverse square law of Newtonian gravitational theory. (Though the allometries are not so different from Kepler's laws (Ginzburg and Colyvan 2004).) The allometries must be underwritten by something, presumably something metabolic and hence bio-chemical.

The bottom line is that statistical regularities in ecology require explanation. Enion probability analysis takes this intuition very seriously. Despite Strevens' claim that enion probability analysis has a great deal in common with the Poisson approach to probability, it also has a great deal in common with the second view outlined at the beginning of this section—the Quetelet and Buckle view. Indeed, enion probability analysis might be seen to be providing an account that takes the best of both Poisson's and Quetelet's attitude towards the explanatory power of statistical laws. From the first approach Strevens takes the mathematics very seriously—probability theory is a major player in the explanation of macrolevel laws. From the second account he takes the idea that non-probabilistic factors underwrite the legitimacy of the statistical aggregation techniques. It is this combination that gives enion probability analysis so much of its appeal, particularly for population ecology.

Conclusion

Population ecology is a highly mathematical science. From its inception, in the 1930s, it has been dominated by differential and difference equations. More recently ecological theory has been a driving force behind investigations of chaotic behaviour arising from non-linear and initial-condition-sensitive equations (May 1974). Yet the mathematisation of ecology is treated with considerable suspicion in some quarters. In the wrong hands, the mathematics can obscure the underlying biological mechanisms or, worse still, take on a life of its own and ignore the underlying biology. Indeed, the fact that the familiar equations of population ecology—Malthus's equation, the logistic equation, the Lotka-Volterra equations, and others—seem to ignore all the underlying biologically significant factors, lends support to this concern. The (microlevel) biologically relevant factors completely determine the population abundance at a given time, and yet these factors are missing from

the (macrolevel) mathematical laws. But this is precisely the puzzle we began with, and the puzzle which Strevens so expertly tackles in *Bigger than Chaos*. Strevens provides a very plausible explanation for the absence of microlevel information from the macro laws. Moreover, he shows how the rather simple macrolevel laws can arise out of the properties of the complex microlevel behaviour. In doing so, he shows how, in the right hands at least, the mathematisation of population ecology does not need to obscure or ignore the underlying biology. Rather, the mathematics can be seen to represent the underlying biology in a systematic, simple, and natural way.

Acknowledgements

References

Bonner, J.T.: 1965, *Size and Cycle*, Princeton University Press, Princeton.

Calder, W.A. III: 1996, *Size, Function, and Life History*, Dover, Mineola, NY.

Colyvan, M.: 2001, *The Indispensability of Mathematics*, Oxford University Press, New York.

Colyvan, M. and Ginzburg, L.R.: 2003, 'Laws of Nature and Laws of Ecology', *Oikos* **101**, 649–653.

Cooper, G.J.: 2003, *The Science of the Struggle for Existence: On the Foundations of Ecology*, Cambridge University Press, Cambridge.

Ginzburg, L.R.: 1986, 'The Theory of Population Dynamics I: Back to First Principles', *Journal of Theoretical Biology* **122**, 385–399.

Ginzburg, L.R. and Colyvan, M.: 2004, *Ecological Orbits: How Planets Move and Populations Grow*, Oxford University Press, Oxford.

Ginzburg, L.R. and Taneyhill, D.E.: 1994, 'Population Cycles of Forest Lepidoptera: A Maternal Effect Hypothesis', *Journal of Animal Ecology* **63**, 79–92.

Lawton, J.H.: 1999, 'Are there General Laws in Biology?', *Oikos* **84**, 177–192.

May, R.M.: 1974, *Stability and Complexity in Model Ecosystems*, second edition, Princeton University Press, Princeton.

Mousseau, T.A. and Fox, C.W. (eds.): 1998, *Maternal Effects as Adaptions*, Oxford University Press, Oxford.

Murray, B.M. Jr.: 1999, 'Is Theoretical Ecology a Science?', *Oikos* **87**, 594–600.

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

Turchin, P.: 2001, 'Does Population Ecology Have General Laws?', *Oikos*, **94**, 17–26.

Turchin, P.: 2003, *Complex Population Dynamics: A Theoretical/Empirical Synthesis*, Princeton University Press, Princeton.