Chance, Determinism and Biology: Introduction

Note: a wonderful, readable book that covers some material from this part of the course in more detail than we have time to is: *Chance in Biology: Using Probability to Explore Nature* by Mark Denny and Steven Gaines, Princeton University Press, 2000.

This section of our course deals with unpredictability, a factor which occurs in every area of the life sciences, and of course in every area of our lives (weather, traffic conditions, catching the flu, the sex and genetics of our offspring, when we die). It might be useful in some circumstances to ignore the aspects of biology that we can't predict, which means that we take a purely "deterministic" view. In the deterministic view we assume that there is no unpredictability in our measurements, or in our capacity for determining the future course of events. We saw one example of this in the Leslie matrix models for the age structure of a population. Given the initial age structure of the population as a vectr x_0, and a Leslie matrix P, then at the next time period the age structure will be P x_0. Here, given the present population structure, we can precisely predict the population structure one time period later, and in fact at any future time n as we saw by calculating P^n x_0. In using this, we are deciding to ignore uncertainty about many aspects of the future of the population, for example, because we assume the survivorships and fertilities in the matrix P are fixed and don't vary through time.

The Leslie matrix model also assumes that all individuals in the population reproduce and survive exactly consistently according to the elements of the Leslie matrix, and this doesn't vary in any way. In one sense, this implies that every individual in the population produces exactly the same number of offspring as every other individual of the same age. Clearly this assumption would not hold in reality, where not only do individuals differ in the capacity to survive and reproduce (the individual differences which, if heritable, are the basis for the process of natural selection), but these may vary through time and space as environmental conditions change.

The area of mathematics that provides us with methods to account for unpredictability is called "probability". This is closely related to the area of statistics, which applies probability to questions such as how likely it is that the outcomes of two experiments will differ or be the same (e.g. do the outcomes of the experiments differ "significantly") and how we can best design experiments to evaluate some hypothesis.

Although in this course we will only cover basic probability and models, in addition to the field of statistics there are many applications of probability to sub-disciplines of biology that are absolutely essential. A few examples:

Genomics - this is the application of probability to analyze genetic sequences, determine differences between sequences, and compare sequences between different individuals/species. The techniques to do this utilize computational methods that are part of the general area of "bioinformatics".

Population genetics - this uses probability to analyze the genetic structure of a population (just as we have analyzed the age structure using the Leslie model). The reason probability matters here is the process by which mating and assortment of genetic material occurs in many populations - it is not possible to determine exactly what egg and sperm cells will combine and therefore there is an unpredictable component as to what the next generation will look like. Think of the simple situation of two birds landing on an island, mating and founding a population on the island. If by chance a genetically determined characteristic of the parents is not passed on to the offspring, then that characteristic will be completely lost from future generations unless there is a mutation which returns it or an immigrant arrives with that characteristic and interbreeds. The jargon name for this is the "founder effect". The limited genetic material in the small number of founders of a new population acts to constrain the future genetic composition of the population.

Disease spread - the entire field of epidemiology deals with how diseases spread within and between populations. The initial phase of this is unpredictable - think of the 2003 case of hepatitis A in Knoxville, which was spread quite unpredictably among some individuals who ate at a restaurant and not to others. Similarly, harmful E. coli infections might affect some but not all individuals who eat improperly handled meat – determining why some individuals are affected and others are not is a primary question in epidemiology. This is also associated with the issue of vaccination. Under what circumstances is it likely that without vaccination a disease will spread and how do you trade off the costs of this with the side effects and other costs associated with the vaccination.

Vision - certain organisms have the capacity to perceive images and movement extremely well under very low light conditions. In this situation, there are very few photons hitting the cells (within our retinas) that respond and pass a neural signal on to the brain. So photons hit the retina in an unpredictable manner. How then does an organism perceive an image under low light, since a decision has to made as to whether that part of the image is really dark, or that it is not dark but that a photon from that part of the image by chance has not hit the retinal cell? This is part of the problem of signal detection - determining what a signal is in a system with "noise" – a central problem in sensory perception.