

On Equilibria of Allele Frequencies In a Gene Pool - A Simple Simulation

Alek Westover

AP Chemistry

October 14, 2019

THE EXAMPLE OF EQUILIBRIUM: ALLELE FREQUENCIES

I made a simulation of how the genotypes of a population will converge over time to an equilibrium (in the simple case that I consider) as a result of selective pressure.





In my model I consider animals (sheep) that have 2 possible phenotypes (physical characteristics) as determined by 2 alleles (their genotype).

The phenotype is displayed by the indicator of color, there are black sheep and there are blue sheep.

THE GENOTYPES AND CORRESPONDING PHENOTYPES

The phenotype is determined by a single gene.

The organisms gets an allele of the gene from their father and their mother. The black color is dominant. The phenotypes that arise from each genotype are depicted in the following table:

	A	a
A		
a		

WHAT IS THE EQUILIBRIUM THAT ARISES?

In my first simulation I give each phenotype a certain chance of dying in any given year. For the Black sheep I made this probability 0.2, and for the blue sheep I made this probability 0.1 (Note: this is a very simplistic model). Every year, each sheep that is not dead by the end of the year mates with a random sheep and they produce an offspring. The offspring gets a randomly chosen allele from each of its parents, which determines its phenotype.

After the mating I randomly kill sheep until so that the total sheep population is the same as it was the previous year. Note that this should essentially preserve the ratios of each genotype as long as the populations are big enough. This step is my way of enforcing a carrying capacity in the population (without this, sheep are expected to reproduce much more than they die off, so my simulation would quickly get really big numbers which would make it slow).

MATHEMATICALLY MODELLING THE EQUILIBRIUM

Let the total population size be n . Let $AA(t)$, $Aa(t)$, $aa(t)$ be the frequencies of the genotypes AA, Aa, aa at year t . By the end of the year, before the mating, there will be (on average)

$0.8nAA(t)$, $0.8nAa(t)$, $0.9naa(t)$ of each genotype of sheep (according to the death probabilities for each phenotype).

Let f_{AA}, f_{Aa}, f_{aa} be the frequencies of the alleles in the population after the sheep die for the year, before they mate for the year.

COMPUTING THE NEXT GENERATION

Now we compute the probability of each genotype.

- Probability of aa:

$$\begin{aligned} f_{aa} \cdot f_{aa} + \frac{1}{2}f_{aa} \cdot f_{Aa} + \frac{1}{2}f_{Aa} \cdot f_{aa} + \frac{1}{4}f_{Aa} \cdot f_{Aa} \\ = (f_{aa} + \frac{1}{2}f_{Aa})^2 \end{aligned}$$

- Probability of AA:

$$\begin{aligned} f_{AA} \cdot f_{AA} + \frac{1}{2}f_{AA} \cdot f_{Aa} + \frac{1}{2}f_{Aa} \cdot f_{AA} + \frac{1}{4}f_{Aa} \cdot f_{Aa} \\ = (f_{AA} + \frac{1}{2}f_{Aa})^2. \end{aligned}$$

- Probability of Aa: $1 - (f_{AA} + \frac{1}{2}f_{Aa})^2 - (f_{aa} + \frac{1}{2}f_{Aa})^2$

GETTING EXPLICIT FORMULAS

Let $c = 0.8 \cdot AA(t) + 0.8Aa(t) + 0.9aa(t)$. Note that cn is the population after the death, and that

$$f_{aa} = \frac{0.9aa(t)}{c}, f_{Aa} = \frac{0.8Aa(t)}{c}, f_{AA} = \frac{0.8AA(t)}{c}.$$

Let $f'_{aa}, f'_{AA}, f'_{Aa}$ be the frequencies of the alleles after the death and mating. Then

$$f'_{aa} = \left(\frac{0.9aa(t) + \frac{1}{2}0.8Aa(t)}{c} \right)^2$$

$$f'_{AA} = \left(\frac{0.8AA(t) + \frac{1}{2}0.8Aa(t)}{c} \right)^2$$

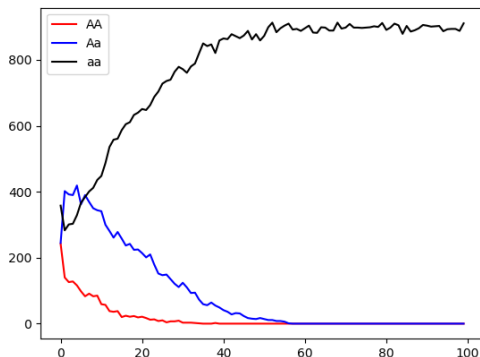
$$f'_{Aa} = 1 - f'_{aa} - f'_{AA}$$

SO WHAT?

It is not obvious at surface inspection that these formulas imply that an equilibrium will ensue. I will create a simulation to get reasonable values for these frequencies, and then come back to these equations to see if they are stable at the equilibrium point suggested by my simulation.

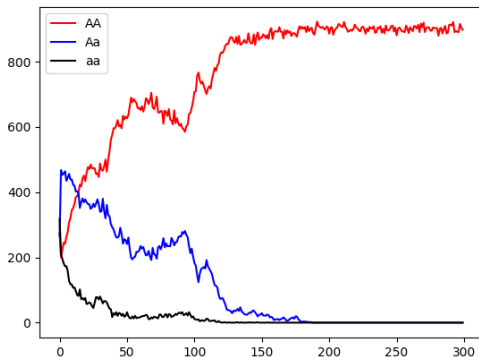
A SIMULATION

Not surprisingly, aa quickly dominates and the equilibrium is where all sheep are aa. Note that our formula predicts that this is an equilibrium (plug in $aa(t) = 1$, $Aa(t) = 0$, $AA(t) = 0$ and the ratios remain constant)



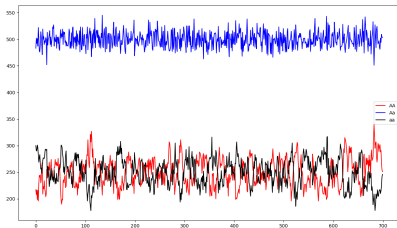
CHANGING THE PROBABILITIES

If we change the situation so that the black sheep have a 90% chance of surviving in a year and the blue sheep have a 80% chance of surviving in the year then the equilibrium becomes all black sheep (specifically AA).



A MORE INTERESTING EQUILIBRIUM

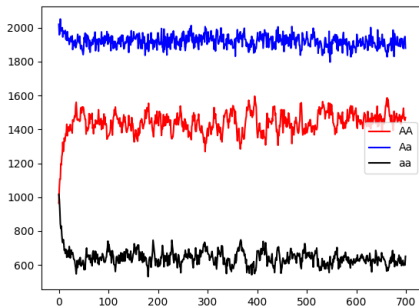
By making the probability of heterozygotes surviving higher (this could be reasonable because heterozygotes are more flexible) we get the following graph:



Note that while the populations of AA and aa are in constant flux they stabilize out to being centered around some value, which I call the equilibrium. If the population size were bigger then I imagine that this oscillation would be harder to see.

OSCILLATION AT EQUILIBRIUM IS LESS VISIBLE WHEN THE POPULATION SIZE IS LARGER

Here is a graph of what could happen when the probability of AA dying is 30%, the probability of Aa dying is 10% and the probability of aa dying is 40% (in a given year):



WHY DOES THE EQUILIBRIUM OCCUR? - MATHEMATICAL ANSWER

In the situation described in the previous slide the long run averages of the frequencies of AA, Aa, aa are 0.368836, 0.47636, 0.154804.

Let's plug these into the obvious extension of our formula for $f'_{aa}, f'_{Aa}, f'_{AA}$ First note that

$$c = 0.7AA(t) + 0.9Aa(t) + 0.6aa(t) = 0.7798$$

$$f'_{aa} = \left(\frac{0.6aa(t) + \frac{1}{2}0.9Aa(t)}{c} \right)^2 = \frac{0.6 \cdot 0.155 + 0.45 \cdot 0.476}{0.7798} = 0.155.$$

Note that f'_{aa} is very close to the long run average of $aa(t)$. We would find similar things if we compared f'_{Aa} and the long run average of Aa in the simulation, and the same goes for AA.

WHY DOES THE EQUILIBRIUM OCCUR? - INTUITIVE ANSWER

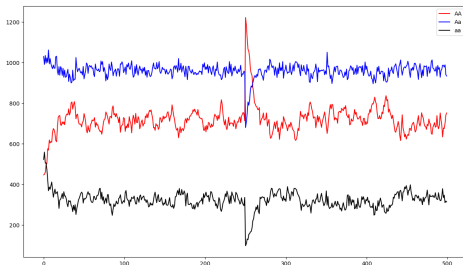
There are opposing processes that balance each other out at equilibrium:

- ▶ the aa and the AA are much more likely to die than the Aa
- ▶ But the Aa are fairly likely (depending on the makeup of the population) to have children that are aa or AA

This is why, unlike in the cases we considered where a homozygous individual had the selective advantage, the equilibrium happens at a nonzero frequency for all populations.

WILL EQUILIBRIUM BE RESTORED?

Equilibrium will be restored if we disrupt the populations by. The graph below is an example of this. I simulate suddenly having a large number of AA immigrate into the population (this is the spike in the curves). Note that the equilibrium was not changed.



WHY IS EQUILIBRIUM RESTORED?

Equilibrium is restored because the AA are more likely to die than the Aa, so when the frequency of AA is very high the rate at which AA die is much higher than the rate at which AA are created by mating. As the frequency of AA goes down the rate at which AA die and are produced by mating equal each other.

AN EQUILIBRIUM CONSTANT

If I were to describe my situation in terms of reactions it would look like:



The analogy is somewhat flawed, but what I mean is that there is some probability of an aa and a AA combining to make an Aa, and some probability of an Aa and an Aa combining to make either a AA or a aa. Then the "equilibrium constant" would be

$$\frac{[Aa]}{[AA][aa]}.$$

"Concentrations" should naturally be interpreted as frequencies. The fact that given the death probabilities of each genotype this is a constant means that it makes sense as an equilibrium constant.

SOME INTUITION FOR THE MEANING OF THE EQUILIBRIUM CONSTANT

- ▶ If this "equilibrium constant" is large then the forward reaction is favored. What this means here is that if Aa have a lower probability of dying then they will become prominent in the population.
- ▶ The analogy isn't perfect, because even if Aa are sure to survive while AA, aa are sure to die each year, then from the mating process about $\frac{1}{4}$ of the offspring from the all Aa population will be AA and $\frac{1}{4}$ will be aa.
- ▶ If this "equilibrium constant" is small, then either AA or aa will be very large. From my simulations this seems to mean that it is very likely that over time all other genotypes will die out eventually. This could be interpreted as the reaction going to completion.

THE END

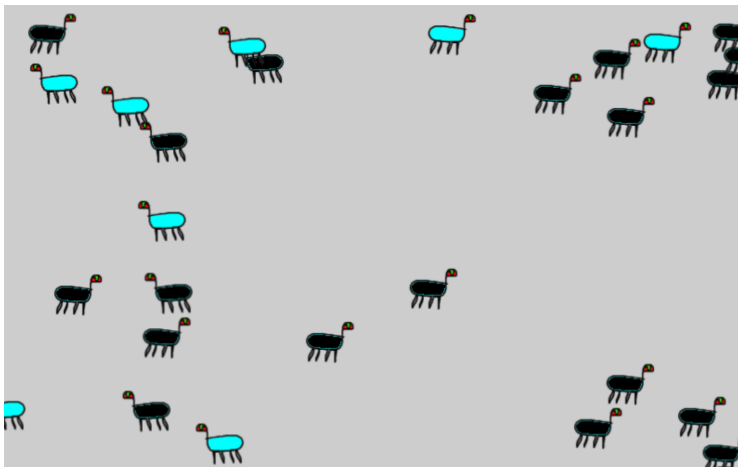


Figure: The game version of the simulation