Bio 417 Homework 4

(for the weeks of Mar 11th and 18th, due April 1st)

1. (Selection vs. drift) Consider the two-allele haploid model in a finite population, as simulated in the Jupyter notebook.
a) Set $w_A = 1.01$, and $w_B = 1$. Run 100 replicate populations for 1000 generations, starting with an initial allele frequency for A $p = 0.01$. Run simulations for population sizes $N = 10, 100, 500, 1000, 5000, 10000$ and 100000. For each population size, record the average final frequency of A in the 100 replicate populations. Plot this final average frequency against population size (use a log-scale for population size). What do you observe? What does that mean for natural selection?

You should observe that at small population sizes, the A allele most often will fail to go fixation despite having a fitness advantage. This is because when it is low frequency, the A allele is susceptible to loss by drift (bad luck, essentially). As population size increases, drift becomes weaker (the effects of luck average out more), and therefore the allele with the fitness advantage becomes more likely to prevail. At large population sizes, the fixation probability approaches 1, so the favored allele fixes deterministically. This means that selection is strongest in large populations, whereas drift is strongest in small.

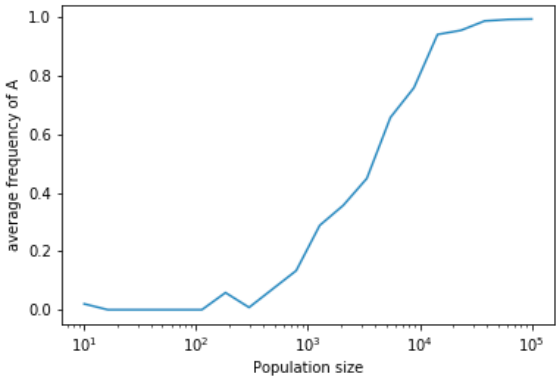

Figure 1: Population size vs. average final frequency of allele A.

b) Now set $w_A = 1.1$, leaving $w_B = 1$, and do the same as above. What do you observe?

With stronger fitness advantage to A, selection overwhelms drift at lower population sizes.

2. (Marginal fitness) Show that you can write the expression for new allele frequency p' in the diploid case (eq 4 in the notes) in an analogous way to the haploid case (eq. 1 in the notes), by defining a composite fitness measure w_{A*}. What is this fitness measure, and what is its interpretation?

We can write $w_{A*} = pw_{AA} + (1-p)w_{AB}$. Factoring p out from equation 4 in the notes, we immediately see that

$$p' = \frac{w_{A*}p}{\bar{w}} \, .$$

w_{A*} is sometimes called the marginal fitness of an allele; it is the expected fitness of a randomly chosen A allele, taking into account that the allele will find itself in the same body as another A allele p of the time, and as a B allele $(1-p)$ of the time.
3. (Long-term evolution in the iterated PD) Assume that inequality 22 in the notes holds (and $s < p < r < t$ as in the notes).
a) Consider a population composed of ALLC at frequency q and TFT at frequency $(1-q)$. Suppose a mutant arises that plays ALLD. Under which conditions, if any, will it be able to increase?

The fitness of TFT and ALLC against each other is exactly the same, and $\frac{1}{1-\delta}r$. But ALLD gets $\frac{1}{1-\delta}t$ against ALLC and $t+\frac{\delta}{1-\delta}p$ against TFT. So, a rare ALLD mutant has expected payoff of:

$$w_{ALLD} = q\frac{1}{1-\delta}t + (1-q)\left(t + \frac{\delta}{1-\delta}p\right) \, , \tag{1}$$

which has to be greater than $\frac{1}{1-\delta}r$ for ALLD to be able to invade. That condition reduces to:

$$\frac{\delta(t-p) - (t-r)}{\delta(t-p)} < q \, ,$$

which you can verify will be less than 1, given $s < p < r < t$. Thus, if ALLC is at sufficiently high frequency, ALLD can invade the population, because it effectively exploits ALLCs.
b) Is the TFT strategy evolutionarily stable? Prove your answer.

No. The proof is that when TFT is fixed, ALLC does equally well against TFT, so satisfies the first part of the ESS condition only with equality. For the second part, TFT has to do against ALLC strictly better than ALLC against TFT, but again, TFT and ALLC have exactly the same payoff against each other. So, TFT is not an ESS. In the long run, ALLC will be able drift up against TFT (b/c it's not getting

selected against), and once ALLC reaches high enough frequency, ALLD will be able to invade.
c) If you have a finite but large population composed of all TFT initially, but ALLC and ALLD can spontaneously arise through mutation, what will happen in the long run?

ALLC will increase by drift, because it is neutral. Eventually, it will cross the threshold q derived above, at which point a mutant ALLD will be able to invade and take over the population. As such, ALL D is the only long-term stable outcome.
4. (Fisher's fundamental theorem) Calculate the "transmission bias" ($E[w_i\delta_i]$) term in the Price equation for fitness in the (one shot, i.e., not iterated) Prisoner's Dilemma game model (set $r = b - c$, $t = b$, $p = 0$, and $s = -c$, as in the donation game, with $b > c$), and show that it is always negative and bigger than the covariance term, so that fitness always goes down.

Let p denote the frequency of Cooperators at time 1. The fitness of cooperators is then $w_C = pb - c$ while that of the defectors $w_D = pb$. At time 2, the frequency of cooperators will be

$$p' = \frac{w_C}{pw_C + (1-p)w_D}p = \frac{pb-c}{pb-pc}p = \frac{pb-c}{b-c}\,.$$

Thus, the fitness of Cooperators at time 2 will be $w'_C = \frac{pb-c}{b-c}b - c$ and $w'_D = \frac{pb-c}{b-c}b$. We can then write:

$$\begin{aligned}
\mathrm{E}\left[w_i\delta_i\right] &= pw_C(w'_C - w_C) + (1-p)w_D(w'_D - w_D) \\
&= p(pb-c)\left(\frac{pb-c}{b-c}b - pb\right) + (1-p)pb\left(\frac{pb-c}{b-c}b - pb\right) \\
&= \left(\frac{pb-c}{b-c}b - pb\right)p(b-c) \\
&= pb\left[(pb-c) - p(b-c)\right] = -pb(1-p)c < 0
\end{aligned}$$

The covariance term:

$$\begin{aligned}
\mathrm{cov}(w_i, w_i) = \mathrm{var}(w_i) &= \mathrm{E}\left[w_i^2\right] - \bar{w}^2 = p(pb-c)^2 + (1-p)pb^2 - (pb-pc)^2 \\
&= pc^2 - p^2c^2 = p(1-p)c^2 > 0
\end{aligned}$$

Now, adding these terms:

$$\mathrm{cov}(w_i, w_i) + \mathrm{E}\left[w_i\delta_i\right] = p(1-p)c^2 - pb(1-p)c = p(1-p)c(c-b) < 0\,,$$

because $b > c$.