Probability Models, Homework 5, due May 2

1. Consider the Wright–Fisher model with $2N = 6$ with 4 copies of A and 2 copies of a in generation 0.

(a) Find the probability that A becomes fixed in the first generation.

(b) Find the probability that a becomes fixed in the first generation.

(c) If the population in **(a)** doubles between the 0th and first generation instead of stying constant, what are the average population size and the effective population size?

2. The smallest possible (haploid) population size in the Wright-Fisher model is 2. Let n be any nonnegative integer and show that the harmonic mean of 2 and n is always less than 4.

3. Recall Problem 2 on HW 3 where we let the individuals $1, 2, ..., 2N$ have numbers of offspring $X_1, X_2, ...X_{2N}$ where $X_1 + X_2 + \cdots + X_{2N} = 2N$. Also recall that in the Wright–Fisher model, each X_k has a certain binomial distribution (and the X_k are not independent). Thus, random sampling with replacement corresponds to the X_k having this binomial distribution. More generally, we can let the X_k be any nonnegative integer-valued random variables with the same distribution as long as they are such that $X_1 + X_2 + \cdots + X_{2N} = 2N$ (in order to keep a constant population size and note that this also means that we must have $E[X_k] = 1$). We are no longer necessarily sampling with replacement; for example, we can put bounds on how many offspring an individual may have. As we have seen, in the Wright–Fisher model, the time T until the most recent common ancestor equals $E[T] \approx 4N$ (time unit: generations) if the sample k is reasonably large. In this more general model, it can be shown that

$$E[T] \approx \frac{4N}{\sigma^2}$$

where σ^2 is the limit of the variance in the offspring distribution as $N \to \infty$, $\sigma^2 = \lim_{N\to\infty} \mathrm{Var}[X_k]$ (from a practical point of view, σ^2 is the approximate variance in the offspring distribution assuming N is large).

(a) Find σ^2 in the Wright–Fisher model and show that the general formula for $E[T]$ applies.

(b) Suppose we choose N individuals randomly and give them 2 offspring each. The remaining individuals get no offspring. Find σ^2 and $E[T]$. Note: For a given individual k, you need to find the probability that it is included in the sample of N individuals. This is the probability that $X_k = 2$. To find the variance, use the formula $\mathrm{Var}[X] = E[X^2] - (E[X])^2$.

(c) If σ^2 increases, $E[T]$ decreases and vice versa. Give an intuitive argument why this makes sense.

(d) If $\sigma^2 = 0$, we get $E[T] = \infty$. What situation does this describe? Think about what it means for a random variable to have variance equal to 0 and recall what we said about $E[X_k]$ above. Also, what can you say about T itself?