

Population Genetics Homework - Week 8 (Project 2)

Christian Polania

1. *Why will independent samples from the same population taken at different times differ more, on average, than independent samples from the same population taken at the same time?*

Samples from the same biological population taken at different times are not from the same population, statistically. Since population allele frequencies change over time, sample allele frequencies will also change.

2. *What might explain the drop at year 4 and the increase from there?*

If one set up a population where every single breeder was born in the year before they have offspring, the allele frequency will change over time due to drift. If one were to somehow introduce individuals from year 1 to the breeding population in year 10, they would “pull” the resulting allele frequency back closer to what it originally was. You could even think of them as migrants, pulling the frequency in the direction of their homeland (year 1). Since the salmon simulation picks breeders from 4 years ago, a drop in allele frequency difference at 3/4 years (when the first offspring begin to contribute to breeding) might be from those year 1 offspring pulling the population closer to the original state. I’m not quite sure where the simulation is sourcing breeders from before the 3/4 year mark, but if they’re from the same-year population it would make sense that once the offspring of the first year started breeding that it would pull the frequency back temporarily. After that initial introduction, the allele frequency would begin to drift away from the original again (as it began to do before the first returned).

3. *Why does the fraction of significant tests become so large when the sample size is large relative to the effective number of breeders, i.e., when S/Nb is large?*

By taking a sample that’s larger in relation to the effective number of breeders (as it is postulated the hatcheries did), you’re artificially magnifying the difference between populations (here, different temporal populations). A demonstration: Set A (1,2,3) and Set B (8,9,10) will appear less different than Set C (1,1,2,2,3,3) and Set D (8,8,9,9,10,10) do, even though the difference in mean between A and B is the same as the difference in mean between C and D.

4. *Given what you know about drift and selection, why do you think there are fewer significant changes in allele frequency detected when populations are large than when they are small?*

With a larger sample, the sample mean approaches the population mean. If you consider the number of breeding individuals as a sample and the allele frequency as the population mean, a larger “sample” will mean a resultant allele frequency closer to the original. Iterate this process many times (many generations), and a case where sample size is larger will be closer to the original than a case where sample size is smaller. In other words, larger populations aren’t as susceptible to stochastic processes.

5. *If you were extending this simulation study, do you think it would be more useful to (a) calculate the fraction of allele frequency comparisons that show statistically significant differences between time periods or (b) study how the mean and variance of allele frequency differences changes over time?*

Since the fraction of statistically different comparisons was shown to be so difficult to interpret and so easily inflated by sample size in relation to Nb , I’d have to imagine means and variances would be easier to work with while remembering what will influence them.