

Homework Assignment #5

Date Assigned: 11 November 2010.

Date Due: 23 November 2010.

1. Suppose that we have a haploid population with two alleles, and their absolute fitnesses are $w_A = 4$ and $w_a = 2$. If the initial frequency of A is 0.001, what will it be after 20 generations?
2. In a haploid system with two alleles, A and a , with fitnesses $1 + s$ and 1, how long will it take to change the frequency of A from 0.1 to 0.2 if $s = 0.01$? How long will it take to change the frequency of A from 0.9 to 0.8 if $s = -0.01$? Explain why these numbers are or are not the same.
3. Assume a population at HWE for a diallelic locus with alleles A and a . Further, assume frequency-dependent selection occurs in such a way that the relative fitnesses of AA , Aa , and aa are $w_{11} = c/p^2$, $w_{12} = c/pq$, and $w_{22} = c/q^2$, respectively, for some constant c . What would the allele frequencies be in one generation? Explain what the expression implies.
4. The average time to fixation of new neutral mutation in an ideal diploid population is approximately $4N$. What is the average number of copies of the neutral mutation that will ever exist from the time of its origin by mutation to the time of its fixation by random genetic drift?
5. If two sequences differ by 65% of their positions and have evolved by a Jukes-Cantor model, what is the best estimate of the branch length between them?
6. Assume the Jukes-Cantor model. Let S_1 and S_2 be two sequences that differ at 10% of their sites.
 - (a) What is the branch length between S_1 and S_2 ?
 - (b) Assume S_2 evolves into S_3 by changing a completely different 10% of its sites. What is the branch length between S_2 and S_3 ?
 - (c) Compare the total branch lengths you obtained in (a) and (b) to the branch length you get when comparing S_1 and S_3 assuming they differ at 20% of their sites. Why the discrepancy?
7. Consider sequence alignments where each alignment has 4 sequences, each sequence has 5 sites, and each site exhibits one of two states (or, nucleotides): 0 and 1.
 - (a) Give an example of an alignment of sequences that evolved on a genealogy under the infinite sites model.
 - (b) Give an example of an alignment of sequences that could not have evolved on a genealogy under the infinite sites model.
 - (c) Is there a necessary and sufficient condition that if an alignment satisfies then it could have evolved on a genealogy under the infinite sites model?
8. In the infinite sites model, if $\theta = 10$, how many segregating sites are expected in a sample of size 20?
9. Assume that we have sampled one allele per locus for a large number k of selectively neutral loci from a set of 5 species.
 - (a) Describe an algorithm to construct the evolutionary history of the 5 species using the divergence estimates from the data.
 - (b) Discuss how ancestral polymorphism may affect the accuracy of the history computed by the algorithm.
 - (c) Assume that horizontal gene transfer (HGT) may occur across species boundaries, but that it has a very low rate. Discuss how this may affect the accuracy of the history computed by your algorithm, and propose a solution to the problem.