Homework Set #2 Key

See Nielsen & Slatkin for solutions to problems 3, 5, 7, 8, & 10

1. (a) The only chain is the common ancestor A. So \( f = (1 + f_A)/2 = (1+0)/2 = \frac{1}{2} \). (Note that this result can be argued directly by applying Mendel’s principle of segregation to a selfer.)

(b) Two chains: CAD and CBD. \( f = (\frac{1}{2})(1 + f_A)/2(\frac{1}{2}) + (\frac{1}{2})(1 + f_B)/2(\frac{1}{2}) = (\frac{1}{2})(1 + 0)/2(\frac{1}{2}) + (\frac{1}{2})(1 + 0)/2(\frac{1}{2}) = (\frac{1}{2})^3 + (\frac{1}{2})^3 = \frac{1}{4} \)

(c) Two chains: ECADF and ECBDF. \( f = (\frac{1}{2})(1 + f_A)/2(\frac{1}{2}) + (\frac{1}{2})(1 + f_B)/2(\frac{1}{2}) = (\frac{1}{2})(1 + 0)/2(\frac{1}{2}) + (\frac{1}{2})(1 + 0)/2(\frac{1}{2}) = (\frac{1}{2})^4(1+0)/2 + (\frac{1}{2})^4(1+0)/2 = (\frac{1}{2})^5 + (\frac{1}{2})^5 = \frac{1}{16} \)

2. Disease frequency = 0.001 = \( q^2 + 0.005pq = q^2 + 0.005(1-q)q = 0.995q^2 + 0.005q \). Solving for \( q \) (using the quadratic formula) gives \( q = \frac{-0.005 + \sqrt{0.005^2 - 4(0.995)(-0.005)}}{2(0.995)} = 0.029 \). In a random mating population, the frequency of the disease would be \( q^2 = (0.029)^2 = 0.00084 \). Inbreeding increases the prevalence of the disease in this population by a factor of \( 0.001/0.00084 = 1.2 \). Although the level inbreeding is very low (an order of magnitude smaller than 1st cousins), this is a fairly substantial increase in disease prevalence.

4. Nielsen & Slatkin, p. 74 #4.2
   a. \( p_1 = (2*20 + 20)/(2*60) = 0.5 \)
   \( p_2 = (2*15 + 15)/(2*60) = 0.375 \)
   \( p_3 = (2*20 + 25)/(2*60) = 0.375 \)
   \( \bar{p} = \frac{0.5 + 0.375 + 0.54}{3} \approx 0.472 \)
   b. \( H_S = \frac{2(0.5)(0.528) + 2(0.375)(0.625) + 2(0.54)(0.46)}{3} \approx 0.488 \)
   \( H_T = 2(0.472)(0.528) \approx 0.498 \)
   \( F_{ST} = \frac{H_T - H_S}{H_T} = \frac{0.498 - 0.488}{0.498} \approx 0.02 \)
6. Nielsen & Slatkin, p. 33 #2.2
   Let $q_t$ be the expected frequency of the mutation in generation $t$. If $q_t = 0.1$ then
   
   $$q_{t+1} = (1 - u)q_t + v(1 - q_t) = (1 - 10^{-6})(0.1) + (5 \times 10^{-6})(1 - 0.1) = 0.1000044$$

9. Nielsen & Slatkin, p. 58 #3.6
   \[
   \theta (\text{Watterson}) = \frac{S}{\Sigma_{k=1}^{n} \frac{1}{k!}} = \frac{5}{1 + \frac{1}{2} + \frac{1}{3}} = 2.73
   \]
   \[
   \theta (\text{Tajima}) = \pi = 3
   \]

11. Nielsen & Slatkin, p. 58 #3.8
   Site frequency spectrum: $E(\text{proportion of singletons}) = E(f_1) = \frac{1}{\Sigma_{k=1}^{n-1} \frac{1}{k!}} = \frac{1}{1 + \frac{1}{2} + \frac{1}{3}} = \frac{6}{11} = 0.545$.

   Assuming A is ancestral, the observed value of $f_1 = 0$, which is much less than the expected value of 0.545.

   Folded SFS: $E(f_1^*) = E(f_1 + f_{n-1}) = E(f_1 + f_3) = E(f_1) + E(f_3) = \frac{1}{\Sigma_{k=1}^{n-1} \frac{1}{k!}} + \frac{1/3}{\Sigma_{k=1}^{n-1} \frac{1}{k!}} = \frac{8}{11} = 0.727$. The observed proportion of folded singletons is $f_1^* = f_1 + f_3 = 0 + 2/5 = 0.4$, which is also less than the expected value of 0.727.