Practice problems for population genetics (and some continuous time ecology)

(1) (a) Prove: $p = D + \frac{1}{2}H$

$D = \frac{N_{AA}}{N}$, $H = \frac{N_{Aa}}{N}$, and $R = \frac{N_{aa}}{N}$ by definition of D, H, and R. The total number of alleles in the population is $2N$ and the number of A alleles contained in an AA genotype is 2 and the number in an Aa heterozygote is 1. If p is the fraction of A alleles in the population, then p is the average number of A alleles over all genotypes or,

$$p = \frac{2N_{AA} + N_{Aa}}{2N} = \frac{N_{AA}}{N} + \frac{1}{2}\frac{N_{Aa}}{N} = D + \frac{1}{2}H.$$

Prove: $q = R + \frac{1}{2}H$

By a similar argument,

$$q = \frac{2N_{aa} + N_{Aa}}{2N} = \frac{N_{aa}}{N} + \frac{1}{2}\frac{N_{Aa}}{N} = R + \frac{1}{2}H.$$

(b) Prove: $p^2 + 2pq + q^2 = 1$

$q = 1 - p$ by definition

$$p^2 + 2pq + q^2 = p^2 + 2p(1-p) + (1-p)^2 = p^2 + 2p - 2p^2 + 1 - 2p + p^2 = 1$$

(c) Random mating, large population size, Mendelian segregation and no mutation.

(2) The value of p_t for some time t is the square root of the frequency of individuals homozygous for that particular allele. Thus, if $D_t = p_t^2 = p_0^2$ for all t, then $p_t = p_0$ for all t. Since $q_t = 1 - p_t$, then it follows that $q_t = q_0$ for all t. Assuming random mating then the frequency of heterozygotes will be $2p_t q_t = 2p_0 q_0$ for all $t > 0$.

(3) The nonzero equilibrium for a locus with heterozygote advantage is $\hat{p} = \dfrac{t}{t+s}$, where the relative fitnesses of the three genotypes are $AA : 1 - s$, $Aa : 1$ and $aa : 1 - t$. So, for this problem the equilibrium frequency is $\hat{p} = \frac{0.5}{0.25 + 0.5} = 0.67$.

(4) (a) The frequency of p in the next generation after selection is given by

$$p'_{sel} = \frac{w_{AA}p^2 + w_{Aa}p(1-p)}{w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2}$$

$$= \frac{w_{AA}p + w_{Aa}(1-p)}{w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2}p.$$

The change in p from one generation to the next due to selection is $p'_{sel} - p$ or,

$$\frac{w_{AA}p + w_{Aa}(1-p)}{w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2}p - p.$$

(b) For p very small, $p^2 \approx 0$, and $(1-p) \approx 1$. Thus, the equation $w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2$ is approximately equal to $2w_{Aa}p + w_{aa}$. If we further assume that $2w_{Aa}p \ll w_{aa}$ because p is very small, then $2w_{Aa}p + w_{aa} \approx w_{aa} = 1$ (by definition of w_{aa}).

(c) If $w_{AA}p^2+w_{Aa}2p(1-p)+w_{aa}(1-p)^2 \approx 1$, $w_{AA}p \approx 0$ and $w_{Aa}(1-p) \approx w_{Aa}$, then

$$\Delta p_{sel} = \frac{w_{AA}p + w_{Aa}(1-p)}{w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2}p - p$$

$$\approx \frac{w_{Aa}}{1}p - p$$

$$= (1-s)p - p = -sp$$

(d) In this problem there is one-way mutation from $a->A$ at rate v. Another way to interpret v is that it is the fraction of a alleles that mutate to A per generation. If the fraction of a alleles in the population is $1-p$, then the total fraction of alleles that mutate from a to A is $v(1-p)$. Thus the change in p due to mutation (Δp_{mut}) is $v(1-p)$.

(e) If p is approximately 0, then $1-p \approx 1$. Thus, $v(1-p) \approx v$.

(f) Since $\Delta p = \Delta p_{sel} + \Delta p_{mut}$, and $\Delta p_{sel} \approx -sp$ and $\Delta p_{mut} \approx v$ (from 4c and 4d, respectively) then $\Delta p = -sp + v$.

(g) The equilibrium is now easy to solve if we remember that at equilibrium $\Delta p = 0$.

$$\Delta p = 0 = -sp^* + v$$

$$p^* = \frac{v}{s}.$$

Therefore, in the case when a deleterious allele is dominant, the equilibrium frequency is $\frac{v}{s}$, under a mutation-selection balance.

(5) In contrast to problem 4, in this problem the deleterious allele is recessive.

(a) The change in p due to selection has the same form as 4a, except now the terms w_{AA}, w_{Aa} and w_{aa} are $1-s$, 1, and 1, respectively since the allele A is recessive. Therefore, the equation for Δp_{sel} is

$$\Delta p_{sel} = \frac{w_{AA}p + w_{Aa}(1-p)}{w_{AA}p^2 + w_{Aa}2p(1-p) + w_{aa}(1-p)^2}p - p.$$

Substituting the values of w_{AA}, w_{Aa} and w_{aa} into the equation for Δp_{sel} results in the equation,

$$\Delta p_{sel} = \frac{(1-s)p + (1-p)}{(1-s)p^2 + 2p(1-p) + (1-p)^2}p - p.$$

Expanding terms results in the equation,

$$\Delta p_{sel} = \frac{-sp + 1}{-sp^2 + 1}p - p.$$

Since p is very small, $p^2 \approx 0$, and the term $-sp^2 \approx 0$ which results in the equation,

$$\Delta p_{sel} = (1-sp)p - p.$$

(b) Simplification of Δp_{sel} in 5a, yields $\Delta p_{sel} = (1-sp)p - p = p - sp^2 - p = -sp^2$. Since the change in p as a result of mutation is the same as it was in 5d and 5e, $\Delta p_{mut} = v$. The overall change in p is a consequence of selection and mutation or $\Delta p = \Delta p_{sel} + \Delta p_{mut} = -sp^2 + v$.

(c) At equilibrium $\Delta p = 0$, in which if we solve for p^* results in $p^* = \sqrt{\frac{v}{s}}$. The main assumption is p is small enough to assume $p^2 = 0$ in our derivation.

(6) Figure 1 provides a plot of the isoclines defined by the red and black scorpion competition model. We see that based on graphical analysis of these isoclines that the internal equilibrium is unstable. Therefore, if the scorpions start at the population size configuration indicated by the black dot in the figure, the Black scorpion will eventually reach its carrying capacity and the Red scorpion will eventually go extinct.

(7) For stable coexistence, the conditions: $K_1 < \dfrac{K_2}{\alpha_{21}}$ and $K_2 < \dfrac{K_1}{\alpha_{12}}$ must be met. If $K_2 = 100$ and $\alpha_{21} = 0.5$, then K_1 must be less than $\frac{100}{0.5} = 200$ according to the first condition. According to the second condition, $K_1 > \alpha_{12}K_2$, or $K_1 > 1.5(100) = 150$. Consequently for stable coexistence K_1 must be between 151 and 199. For species 1 to outcompete species 2, its carrying capacity needs to be greater than 200, based on the previous analysis.

(8) In figure 2 is a plot of the 0-isoclines for two species experiencing competition in which there is a stable internal equilibrium. By adding intraguild predation, the N_1 intercept for the 0-isocline of the prey (in this example species 2) is changed from a value of $\dfrac{K_2}{\alpha_{12}}$ to a value of $\dfrac{r_2 K_2}{\delta K_2 + \alpha_{21} r_2}$. Provided the parameter δ is sufficiently large then the new intercept will be less than K_2 resulting in the 0-isoclines not crossing each other and the predator excluding its prey.

(9) To analyze the model we first find the equilibrium population sizes and then determine their stability. The equilibrium population sizes occur when $\frac{dN}{dt} = 0$:

$$rN^*(1 - \frac{N^*}{K}) - \frac{cN^*}{d + N^*} = 0$$

$r(1 - \dfrac{N^*}{K}) = \dfrac{c}{d + N^*}$ obtained by bringing the term $\dfrac{cN^*}{d + N^*}$ to the right side and then dividing through by N^*

$(1 - \dfrac{N^*}{K})(d + N^*) = \frac{c}{r}$, by multiplying each side by $(d + N^*)$ and dividing each side by r

$$-\frac{N^{*2}}{K} + (1 - \frac{d}{K})N^* + d - \frac{c}{r} = 0$$

The solutions to this equation are $N^* = \dfrac{-(1 - \frac{d}{K}) \pm \sqrt{4\frac{(d - \frac{c}{r})}{K} + (1 - \frac{d}{K})^2}}{-2\frac{1}{K}}$.

Another solution that satisfies $\frac{dN}{dt} = 0$, is $N^* = 0$. Thus, there are three equilibrium population sizes, namely,

$$N^* = \{0, \frac{-(1 - \frac{d}{K}) + \sqrt{4\frac{(d - \frac{c}{r})}{K} + (1 - \frac{d}{K})^2}}{-2\frac{1}{K}}, \frac{-(1 - \frac{d}{k}) - \sqrt{4\frac{(d - \frac{c}{r})}{K} + (1 - \frac{d}{K})^2}}{-2\frac{1}{K}}\}.$$

Next we would like to know which of these equilibrium population sizes is stable and which is unstable. Clearly the equilibrium solutions are complicated so graphical analysis may be a better approach than an analytical analysis.
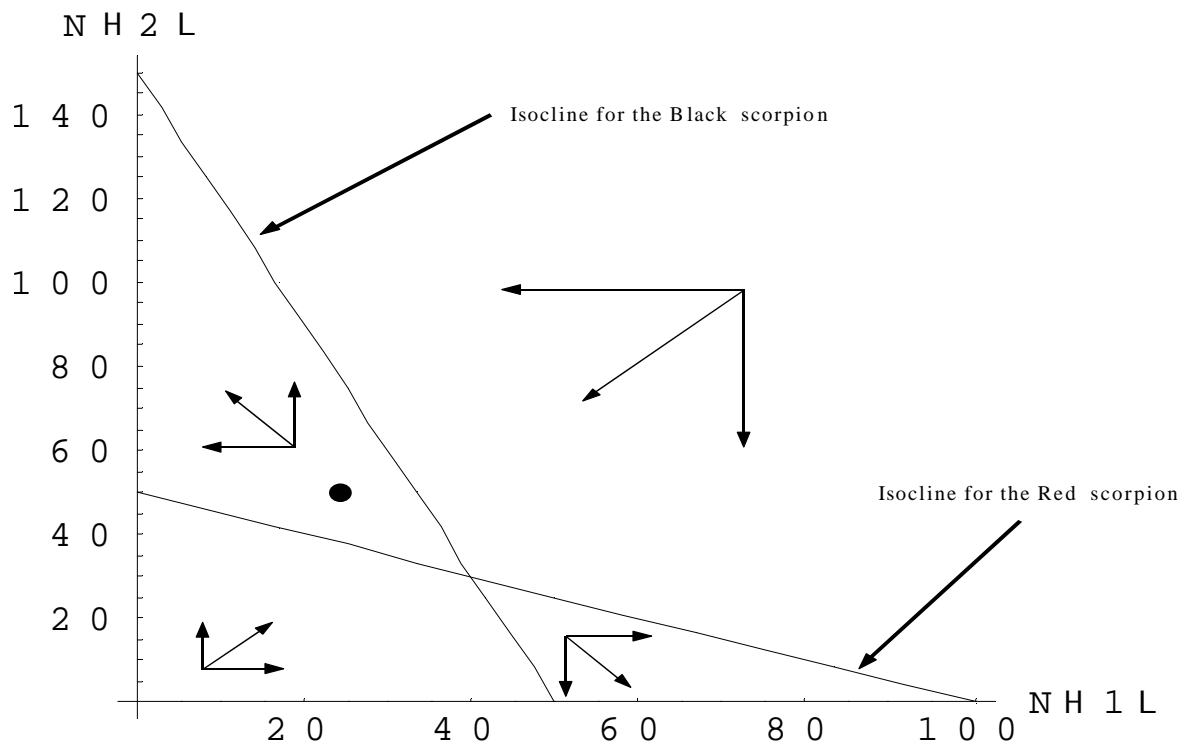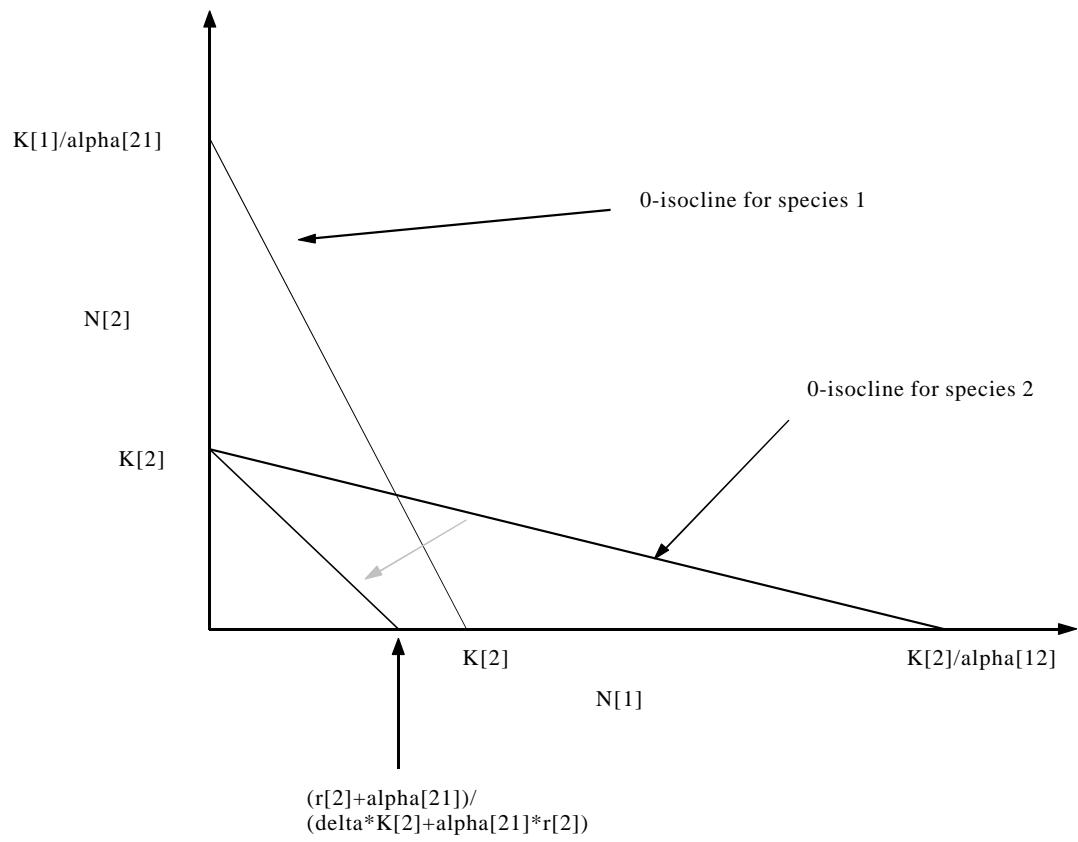
3

NH2L

Isocline for the Black scorpion

Isocline for the Red scorpion

NH1L

Figure 1:

K[1]/alpha[21]

0-isocline for species 1

N[2]

0-isocline for species 2

K[2]

K[2]

N[1]

K[2]/alpha[12]

(r[2]+alpha[21])/
(delta*K[2]+alpha[21]*r[2])

Figure 2:

A graphical analysis shows that for properly chosen r, K, c and d, the plot of $\frac{dN}{dt}$ versus N looks like figure 3, in which there are two equilibria with values greater than or equal to zero in which the largest equilibrium population size is stable. There is an internal equilibrium (figure 4) for large values of c relative to d and for r sufficiently small. The internal equilibrium is unstable and lies between the two stable ones (figure 4). When c is too large $\frac{dN}{dt}$ is always negative resulting in the equilibrium $N^* = 0$ being stable (figure 5).

(10) For this problem we need to return to 8th grade algebra. If there are 10x as many MN genotypes as NN genotypes and $2pq$ is the frequency of MN genotypes and q^2 is the frequency of NN genotypes, then $2pq = 10q^2$. Remembering that $p = 1 - q$, then the relation $2pq = 10q^2$ can be written as $2(1 - q)q = 10q^2$, from which we need to solve for q.

$$2(1 - q)q = 10q^2$$
$$2q - 2q^2 = 10q^2$$
$$12q^2 = 2q$$
$$q = \tfrac{1}{6}$$

(11) (a) If MN is at a frequency 0.42, then $2p(1 - p) = 0.42$; which is a quadratic equation with two solutions, $p = \{0.3, 0.7\}$, so you can't determine a specific p from the information because there are two possible solutions.

(b) If you are provided information that p is the more common allele, then p is the largest solution of $p^2 - 2p + 0.42 = 0$ which is $p = 0.7$.

(12) If one allele is dominant and you cannot assume that the population is in Hardy-Weinberg equilibrium, then given only the observed phenotypes it is impossible to estimate accurately allele frequencies by taking the square root of the frequency of observed individuals with phenotype M because heterozygotes are indistinguishable from homozygotes for the dominant allele.

(13) Let p be the frequency of the horn allele and assume genotype frequencies are in Hardy-Weinberg equilibrium. Based on the laws of probability, $p^2 + 2pq + q^2 = 1$. Or by rearranging terms: $p^2 + 2pq = 1 - q^2$. Since the horn allele is dominant in males, then both homozygous and heterozygous males have horns and there frequency is equal to $p^2 + 2pq$. By substituting 0.96 for $p^2 + 2pq$ in the relation $p^2 + 2pq = 1 - q^2$, it is possible to solve for q.

$$0.96 = 1 - q^2$$
$$q = 0.2$$

Since $q = 0.2$, then $p = 0.8$ and the frequency of females that have horns is $p^2 = 0.64$.

(14) (a) If the populations are of equal size then upon fusion of the two populations the frequency of the a allele is $\dfrac{q_1 + q_2}{2}$. Given the new frequency of the a allele, the frequency of aa individuals in the new population is $\left(\dfrac{q_1 + q_2}{2}\right)^2$.

(b) If population 1 is four times larger then population 2, when the two populations fuse, four out of every five alleles in the new population will have
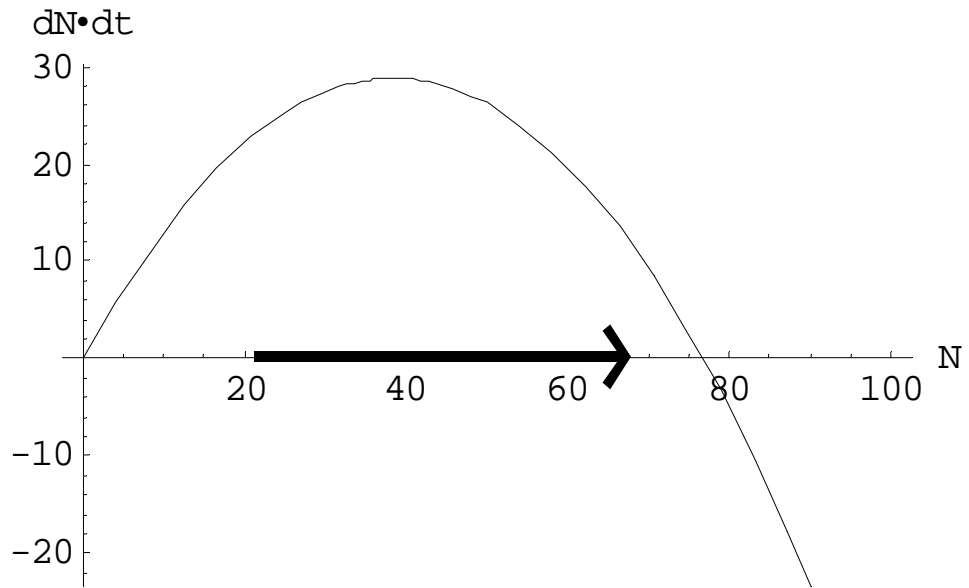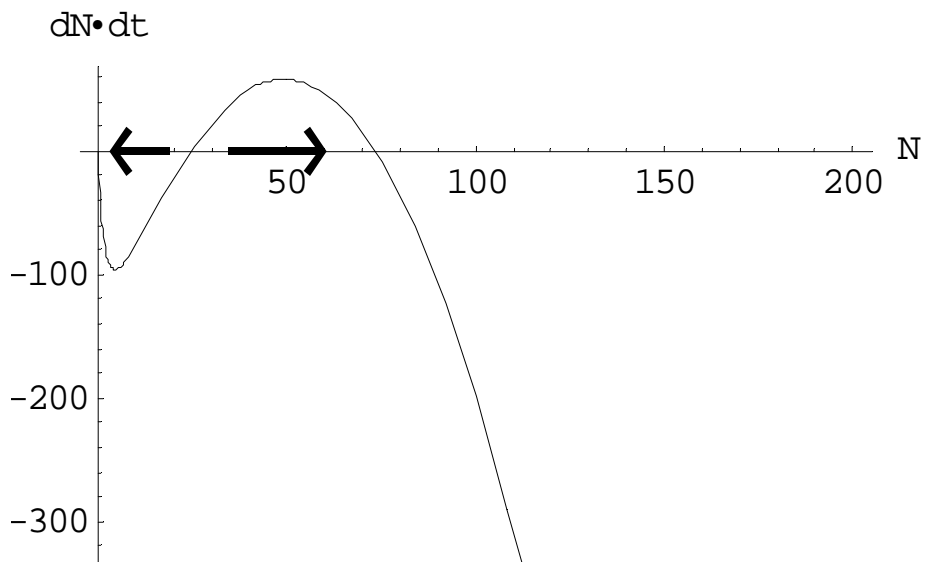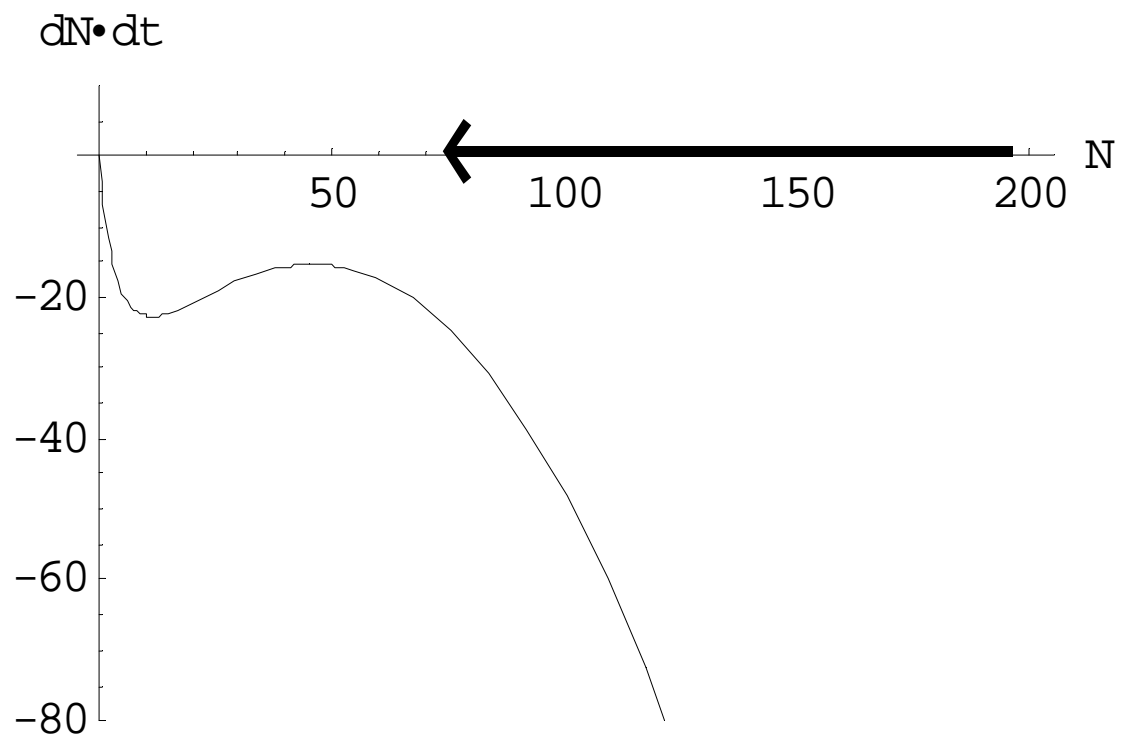
Figure 3:

Figure 4:

Figure 5:

been from population 1, and the frequency of the a allele will be $\dfrac{4q_1 + q_2}{5}$. As a result, the frequency of aa individuals in the new population is $\left(\dfrac{4q_1 + q_2}{5}\right)^2$.

(15) At an overdominant locus we would expect a polymorphic equilibrium. To solve this problem we need to determine the equation that defines the change in the frequency of one of the alleles and then find the allele frequencies such that the change is equal to zero. Let p be the frequency of the allele with a fitness of 1 and q be the frequency of the other allele.

Let p' be frequency p in the next generation after selection and reproduction:

$p' = \dfrac{w_A(t)}{\bar{w}(t)}$, where $w_A(t)$ is the marginal fitness of allele A and $\bar{w}(t)$ is the mean fitness of the population. Based on the relative fitnesses of the three genotypes, $w_A(t) = p + q(1 + hs)$ and $\bar{w}(t) = p^2 + 2pq(1 + hs) + q^2(1 - s)$. Substituting these values into the equation for p' yields,

$p' = \dfrac{p + q(1 + hs)}{p^2 + 2pq + q^2 + 2pqhs - sq^2}p = \dfrac{p + q(1 + hs)}{1 + sq(2ph - q)}p.$

Then $\Delta p = p' - p$ is equal to

$\dfrac{p + q(1 + hs)}{1 + sq(2ph - q)}p - p.$

We seek the equilibrium values of p that satisfy $\Delta p = 0$. Let p^* be the equilibrium values of p, and likewise q^* are the equilibrium values of q.

$\dfrac{p^* + q^*(1 + hs)}{1 + sq^*(2p^*h - q^*)}p^* - p^* = 0$

$\dfrac{p^* + q^*(1 + hs)}{1 + sq^*(2p^*h - q^*)}p^* = p^*$

$\dfrac{p^* + q^*(1 + hs)}{1 + sq^*(2p^*h - q^*)} = 1$, by dividing each side by p^*

$p^* + q^*(1 + hs) = 1 + sq(2p^*h - q^*)$, by cross multiplication

$p^* + q^* + q^*hs = 1 + 2p^*q^*hs - sq^{*2}$, by expansion of terms

$1 + q^*hs = 1 + 2p^*q^*hs - sq^{*2}$, noting that $p^* + q^* = 1$

$q^*hs = 2p^*q^*hs - sq^{*2}$

$h = 2p^*h - q^*$

$h = 2p^*h - 1 + p^*$, by substituting $q^* = 1 - p^*$

$p^* = \dfrac{1 + h}{1 + 2h}$

The other, more trivial equilibrium values for p are 0 and 1.

(16) (a) From lecture and problem 5, the mutation-selection equilibrium frequency of a recessive deleterious allele is $p^* = \sqrt{\dfrac{v}{s}}$. If $s = 1$, then $p^* = \sqrt{v}$, and solving for v yields $v = p^{*2}$. Thus the necessary mutation rate is 0.0004 if the incidence of cystic fibrosis is 0.0004, i.e., the frequency of individuals with cystic fibrosis (p^2) is 0.0004.

(b) A mutation rate of 0.0004 per generation per locus is very high. Normal mutation rates at single sites in coding regions of DNA is 10^{-9}, with per locus rates about $10^{-8} - 10^{-6}$. From problem 15, the equilibrium frequency of the allele with a relative fitness of $1 - s$ is $q^* = 1 - p^* = \dfrac{h}{1 + 2h}$. If $q^{*2} = 0.0004$, then $q^* = 0.02 = \dfrac{h}{1 + 2h}$. Solving for h yields $h = 0.021$. So, an alternative explanation for the frequency of cystic fibrosis is that there is a slight heterozygote advantage for having the "cystic fibrosis" allele.

(17) A way to solve this problem is to determine w_{AA}, w_{Aa}, and w_{aa} such that $\Delta p_{sel} = 0$ for any p. The equation that describes the change in p after selection is $\Delta p_{sel} = \dfrac{w_A(t)}{\bar{w}(t)} p - p$. From this equation it is apparent that for Δp_{sel} to be equal to zero, then $\dfrac{w_A(t)}{\bar{w}(t)} = 1$. Since $\dfrac{w_A(t)}{\bar{w}(t)} = \dfrac{w_{AA}p + w_{Aa}(1 - p)}{w_{AA}p^2 + 2w_{Aa}p(1 - p) + w_{aa}(1 - p)^2}$, then the numerator and denominator must be equal $w_{AA}p + w_{Aa}(1 - p) = w_{AA}p^2 + 2w_{Aa}p(1 - p) + w_{aa}(1 - p)^2$. We can choose any w_{AA} and w_{Aa} and then simply solve for a w_{aa} that would satisfy the previous relation.

Let $w_{AA} = a$ and $w_{Aa} = b$, then w_{aa} satisfies

$ap + b(1 - p) = ap^2 + 2bp(1 - p) + w_{aa}(1 - p)^2$

$w_{aa} = \frac{ap + b(1-p) - ap^2 - 2bp(1-p)}{(1-p)^2}$.

This would be an example of frequency dependent selection such that the selective coefficient of aa individuals is a function of the frequency of the alleles in the population.

(18) In this problem I will use the same techniques as was done in problem 15 to find the equilibrium and then perform a stability analysis by comparing the fitness of the heterozygote with both homozygotes. For the fitnesses defined in this problem, $\Delta p = \dfrac{spq(q + h(p - q))}{1 - sq(q + 2ph)}$, where p measures the frequency of the A' allele. The polymorphic equilibrium solution is $p^* = \frac{h-1}{2h-1}$.

(a) For any $0 < h \quad 1$, the polymorphic equilibrium for the allele A' fails to exists because $p^* \quad 0$.

(b) For $h > 1$, p^* is positive, suggesting a polymorphic equilibrium exists. But for $h > 1$, the fitness of the heterozygote is less than both homozygotes, which means that for $h > 1$, the polymorphic equilibrium is unstable.

(c) For $h < 0$, p^* is positive, again suggesting a polymorphic equilibrium exists. Also notice that for $h < 0$, the fitness of the heterozygote is greater than either homozygote, which implies that the polymorphic equilibrium is stable when $h < 0$.