Problem 1

In Project 2, we constructed a model for the change in a single gene in a population over time. If the gene has two variants, \( r \) and \( R \), then three distinct types of individuals may exist in the population: \( RR \), \( Rr \), and \( rr \). \( A(t) \) is the number of individuals with genotype \( RR \). \( B(t) \) is the number of individuals with genotype \( Rr \). \( C(t) \) is the number of individuals with genotype \( rr \). The model is then

\[
\frac{dA}{dt} = \frac{(A + B/2)^2}{A + B + C}(m - n(A + B + C)) - aA \\
\frac{dB}{dt} = \frac{2(C + B/2)(A + B/2)}{A + B + C}(m - n(A + B + C)) - bB \\
\frac{dC}{dt} = \frac{(C + B/2)^2}{A + B + C}(m - n(A + B + C)) - cC
\]

To analyze this model, we will use a computer to numerically approximate the solution to this system of ODE’s. Use fourth order Runge-Kutta to approximate the ODE. You may use any programming language you like. Use this implementation to explore the following by running simulations with a variety of initial conditions and parameters. In each case below, determine what affect the parameters have on the long term behavior of the population. Does the fraction of \( R \) allele, \( g(t) = \frac{A(t) + B(t)}{A(t) + B(t) + C(t)} \), change over time? Does it approach some fixed value (such as 0 (\( R \) dies off), \( \frac{1}{2} \) (equality), or 1 (\( r \) dies off))? How does this depend on the initial populations? For each problem, turn in plots for at least four sets of parameters and initial conditions which together tell a convincing story of what the population is doing. In each plot, show \( A(t) \), \( B(t) \), and \( C(t) \) over time.

(a) \( a = b = c \).
(b) \( a = c \) but \( b < a \). (heterozygote advantage).
(c) \( a = c \) but \( b > a \). (heterozygote disadvantage).
(d) \( a = b \) but \( c < a \). (\( rr \) offers some advantage).
(e) \( a = b \) but \( c > a \). (\( rr \) offers some disadvantage).
(f) \( c > a > b \). This can occur with recessive traits, where the recessive trait \( rr \) is harmful, but yet the trait offers heterozygote advantage. How do the results differ qualitatively from those of case (e)?

(a) To make parameter selection easier, note that scaling \( m \) and \( n \) is equivalent to scaling time and population. Scaling population is not necessarily a problem; it is okay to measure populations in hundreds or thousands of individuals rather than in units of individuals. Thus, I can fix \( m = n = 1 \).

In the model presented above \( g \) is observed to be constant, though the ratios \( f_A \), \( f_B \), and \( f_C \), will change.
(b) The allele fraction tends to $\frac{1}{2}$. That is, equilibrium is reached in which $A = C$. For my model, the population will tend to an equilibrium where $A = C = \frac{b}{2a}B$. (Unless of course one allele is initially extinct, in which case it stays that way.) Heterozygote advantage tends to equalize the population.
In this case, $A = C$ is unstable. For any other initial condition, the allele that is initially less prevalent in the population goes extinct. Even a very slight disadvantage in the heterozygote population will cause one of the alleles to die off. The dynamics in (a), (b), and (c) are quite different, even if the changes in $b$ are very slight.
(d) The $R$ allele dies out.
In this case, the populations $A$, $B$, and $C$ reach some equilibrium value with both alleles surviving in the population. The $r$ allele will not go extinct, even if the homozygote condition $rr$ is absolutely fatal ($c = 1$). This is in sharp contrast to (e) and (f), where even a slight disadvantage leads to extinction of the allele.
Problem 2

Sickle-cell and Cystic fibrosis are two fairly common genetic diseases. Both are recessive and, without significant medical intervention, both are effectively fatal. Sickle-cell is known to confer heterozygote advantage (such individuals fare better when infected with malaria). Cystic fibrosis has no known heterozygote advantage. Yet even though no such advantage has been found, it is nevertheless hypothesized that such an advantage must exist. Why?

Without a heterozygote advantage, one would predict that the Cystic fibrosis allele would have vanished from the population long ago. The existence of a heterozygote advantage would explain why the disease has not disappeared from the population.
Problem 3

The model proposed in these solutions has a major problem when the population is over its carrying capacity and one of the populations is zero. What is this problem? Suggest a modification to the model that would correct this.

In that case, negative populations can be produced. This is because the growth rate should not be negative, but nothing in the model prevents that from occurring. One possible modification to prevent this is to let $R = \max(m - nT, 0)$, which instead causes all three populations to decay exponentially under these conditions.

Equilibria

The equilibria of this system can be studied, albeit with significant difficulty. Assuming all of the parameters are strictly positive, any equilibrium must fall into one of four cases:

- General case (only if $d \neq 0$ and $e \neq 0$):
  
  \[ d = b^2 - ac \quad e = (a+c)b - 2ac \quad z = \frac{dn - be}{ned^2} \quad A = abz(b-c)^2 \quad C = cbz(a-b)^2 \quad B = 2acz(b-c)(b-a) \]

- Irrelevant allele: $a = b = c$, which is discussed above

- Allele $r$ vanishes: $B = C = 0$

- Allele $R$ vanishes: $B = A = 0$

The general case is interesting and merits some addition analysis. We require $A > 0$, $B > 0$, and $C > 0$. $A > 0$ implies $z > 0$. Then, $B > 0$ implies $(b > a$ and $b > c)$ or $(b < a$ and $b < c)$. That is, the heterozygote condition must either be the most favorable or least favorable. The general case only applies to cases (b), (c), and (g). Note that if $b$ is largest, then $d > 0$ and $e > 0$, but if $b$ is smallest, then $d < 0$ and $e < 0$.

Determining the stability of the general case is quite difficult. There will be three eigenvalues, which cannot be solved for in any simple way. However, it turns out that $\lambda_1 \lambda_2 \lambda_3 = Bdn/2$. If the system is stable then (i) all three roots are negative, in which case $\lambda_1 \lambda_2 \lambda_3 < 0$, or (ii) one root ($\lambda_1$) is negative and the other two are complex conjugates, in which case $\lambda_1 \lambda_2 \lambda_3 = \lambda_1 |\lambda_2|^2 < 0$. Either way, $Bdn/2 < 0$, so $d < 0$. This corresponds to $b$ being smallest. Even with this primitive analysis, we can already say that heterozygote advantage is necessary for this equilibrium to be stable.