Sickle Cell Anemia and Malaria

Information presented is adapted from material developed by Jim Sandefur and Rosalie Dance.

www.georgetown.edu/projects/ handsonmath/downloads/pdf/scel-s.pdf

Sickle cell anemia is caused by a mutant form of the gene coding for a subunit of the hemoglobin protein. We will use N to represent the dominant (normal) hemoglobin allele and S to represent the recessive (sickle cell) allele. Persons with sickle cell anemia are born a with recessive pair, SS, and do not generally live to a reproductive age. Persons with a dominant pair, NN, are not affected by sickle cell, nor are people carrying the pairs NS or SN. But people who are NS or SN do carry the lethal recessive allele and can potentially pass it on to their children. As it turns out, the defective allele that causes sickle cell anemia also helps protect it's carriers from malaria. Having the sickle cell trait (that is, the pairing NS or SN) appears to indicate a partial resistance to malaria.

At any point in time, the gene pool contains a proportion (*n*) of the normal allele N and the rest are sickle cell alleles S. The proportions of each type of allele may change from one generation to the next. We will assume that people who die from malaria will not pass on genes to the next generation and neither will people who have sickle cell anemia.

To summarize: Persons who have two N genes (NN) are susceptible to malaria. Persons who have one N gene and one S gene (NS or SN) are not susceptible to malaria. Persons who have two S genes (SS) will have sickle cell anemia.

Example 1: Suppose 70% of the genes in the gene pool are N and that 80% of the people susceptible to malaria survive it. Fill in the table to show the composition of the gene pool in a population of 1000 people.

		NN	NS	SS	Total
First generation	# of people				
	# of genes				
	# of N genes				
Next generation	# of people				
	# of genes				
	# of N genes				

What is the survival rate (what proportion of the first generation survives to reproduce)?
What is the proportion of N genes in the next generation?

Example 2: Suppose 90% of the genes in the gene pool are N and that 60% of the people susceptible to malaria survive it. Fill out the table and find the survival rate and the proportion of N genes in the next generation.

		NN	NS	SS	Total
First generation	# of people				
	# of genes				
	# of N genes				
Next generation	# of people				
	# of genes				
	# of N genes				

Example 3: Fill out a table for an arbitrary *n* and *k*.

		NN	NS	SS	Total
First generation	# of people				
	# of genes				
	# of N genes				
Next generation	# of people				
	# of genes				
	# of N genes				

What proportion of individuals survives to reproduce in the next generation?

$$\frac{1000n\left(2-\left(2-k\right)n\right)}{1000}=n\left(2-\left(2-k\right)n\right).$$

Note that for a fixed value of k, this proportion is a quadratic function of n.

$$proportion = n\left(2-(2-k)n\right) = 2n-(2-k)n^2$$

The proportion surviving will be zero when $n=0$ and when $\left(2-(2-k)n\right)=0$, or $n=\dfrac{2}{2-k}$. The maximum proportion occurs midway between these zeros at $n=\dfrac{1}{2-k}$.

Now we want to focus on how *n* changes from generation to generation if *k* is fixed.

Refer back to the table we already created for an arbitrary *n* and *k* in a population of 1000 people. The proportion of N genes in the first generation is represented by n_0.

		NN	NS	SS	Total
First generation	# of people	$1000n_0^2$	$2000n_0(1-n_0)$	$1000(1-n_0)^2$	1000
	# of genes	$2000n_0^2$	$4000n_0(1-n_0)$	$2000(1-n_0)^2$	2000
	# of N genes	$2000n_0^2$	$2000n_0(1-n_0)$	0	$2000n_0$
Next generation	# of people	$1000kn_0^2$	$2000n_0(1-n_0)$	0	$1000n_0(2-(2-k)n_0)$
	# of genes	$2000kn_0^2$	$4000n_0(1-n_0)$	0	$2000n_0(2-(2-k)n_0)$
	# of N genes	$2000kn_0^2$	$2000n_0(1-n_0)$	0	$2000n_0(1-(1-k)n_0)$

In the first generation, the proportion of N genes was $\dfrac{2000n_0}{2000}=n_0$. In the next generation the proportion of N genes was $n_1=\dfrac{2000n_0(1-(1-k)n_0)}{2000n_0(2-(2-k)n_0)}=\dfrac{1-(1-k)n_0}{2-(2-k)n_0}$. The proportion of N genes in two generations will be $n_2=\dfrac{2000n_1(1-(1-k)n_1)}{2000n_1(2-(2-k)n_1)}=\dfrac{1-(1-k)n_1}{2-(2-k)n_1}$. We can see that the proportion of N genes varies from one generation to the next according to the recursive equation $n_{i+1}=\dfrac{1-(1-k)n_i}{2-(2-k)n_i}$.

How do the proportions $n_0, n_1, n_2, n_3 \ldots$ behave? We can use some particular values for n_0 and *k* to investigate numerically.

Example 4: Suppose $n_0=0.7$ and $k=0.8$. Find n_1, n_2, n_3 and n_4. After many generations, to what value does *n* eventually stabilize?

Example 5: Suppose $n_0 = 0.9$ and $k = 0.75$. Find n_1, n_2, n_3 and n_4. After many generations, to what value does *n* eventually stabilize?

It is worth finding out whether the recursive equation $n_{i+1} = \dfrac{1-(1-k)n_i}{2-(2-k)n_i}$ has an attracting fixed point. To do this, we can try to find the solution of the equation $n_{i+1} = n_i$.

Example 6: Solve $n_i = \dfrac{1-(1-k)n_i}{2-(2-k)n_i}$ to find n_i in terms of k. Have you seen the solution before? Are you surprised?