

Introduction:

Population Genetics is the study of inheritance based on alleles and their frequency within a population. Although typically applied to autosomal genes with patterns of inheritance dependent on only one loci (the physical location of a gene on a chromosome) the use of allele frequency can still shed light on sex-linked traits or polygenic traits (traits dependent on multiple genes on different loci). Studying the rates of recombination between different loci for example and their correlation to trait frequency can show the way two genes may influence a certain trait. This type of analysis is beyond the scope of this article. This article and the use of linear algebra to Population Genetics will focus primarily on autosomal traits with a dominant-recessive mode of inheritance.

For a given trait with such an autosomal dominant/recessive pattern of inheritance, the frequency of alleles is determined by the rate of their occurrence within the sexual population. This is the population restricted to the individuals capable of breeding and passing on their genes. For a given dominant-recessive gene, the frequency of alleles will not change if there is no selective pressure placed on the trait. This is what something called the Hardy-Wienberg Principle is based on which can be used to predict the equilibrium rates of given alleles as well as the frequencies resultant after one generation of mating. For example, in a given population with only dominant allele frequency A and recessive allele frequency b , after a generation of random mating, the frequency for homozygous dominant will be A^2 , homozygous recessive b^2 , and heterozygous rate $2*A*b$

Linear algebra can also be used to find the same results with more specificity. It can also be used to model systems with very high selective pressure or deleterious traits such as a deadly homozygous recessive such as in achondroplasia.

Frequency Vector:

For a given population the frequency vector is a list of rank of as many phenotypes as there are in the population giving their rates as decimals. If the trait is autosomal with one dominant (A) and one recessive (a) allele then there will be three entries. If this example population has an equal distribution for genotype then the vector, which we will call P has values $\frac{1}{3}$.

$$P_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Each of these corresponds to the genotype AA, Aa, and aa, respectively. From this we can create a frequency of the offspring for a given cross. In this case, if the cross is made with an inbred dominant AA then the resultant matrix is given by

$$AA = \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The matrix product of AA*P_0 here will produce a new vector P_1 of the first generation offspring of a homozygous dominant with random sampling from the population. The matrix AA is called the transition matrix. And the equation for the frequency is given in general by

$$P_n = AAP_{n-1} = \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{bmatrix} P_{n-1}$$

Second Generation Frequency Vector:

This is simple to translate into maple code use commands

```
P_0 := <<(1/3), (1/3), (1/3)>>;
AA := <<1, .5, 0>>
      |<0, .5, 1>
      |<0, 0, 0>>;
P_1 := AA . P_0;
```

This will generate for P_1, and use recursively to produce:

$$P_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}, P_2 = \begin{bmatrix} 3/4 \\ 1/4 \\ 0 \end{bmatrix}, P_3 = \begin{bmatrix} 7/8 \\ 1/8 \\ 0 \end{bmatrix}, P_4 = \begin{bmatrix} 0.9375 \\ 0.0625 \\ 0 \end{bmatrix}, \dots$$

Of course it is necessary if there are other genotypes within the sexual population to produce other transition matrices. For a cross with a homozygous recessive, the result is much the same, however for a heterozygous crossed with the given population, the transition matrix looks like this:

$$Aa = \begin{bmatrix} 1/2 & 1/4 & 0 \\ 1/2 & 1/2 & 1/2 \\ 0 & 1/4 & 1/2 \end{bmatrix}$$

This matrix under left multiplication with the matrix P_0 gives the frequencies after a cross with the heterozygote which are the frequencies given in the transition matrix Aa second column.

Traits with Multiple Genes of Influence

This method does not entirely break down when multiple genes come into play but it does become more tedious. A simple example of a trait with two genes of influence would be one in which a single copy of a dominant allele along two separate loci causes full expression. Take for example, in a hypothetical population exposed to radiation there is a 15/16 chance that any child conceived will die in utero, it is discovered that this is due to a mutation along either of two genetic loci. Health individuals are of genotype AABB and produce all healthy offspring with other AABB. However, within the affected population there is an even distribution of phenotypes (AABB, AABb, AAbb, AaBB, AaBb, Aabb, aaBB, aaBb, and aabb) For the cross of this population with only health individuals AABB, the following transition matrix can be formed.

$$AABB = \begin{bmatrix} 1 & 1/2 & 0 & 1/2 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1 & 0 & 1/4 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/4 & 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Each column is one of the genotypes listed, in that same order. I-th entry is then the frequency of that phenotype when crossed with AABB.

