

Name:

Lab time:

Lab day:

Hardy Weinberg Equilibrium

BIO 182 laboratory
Spring 2015
ASU West

The Hardy Weinberg Equilibrium model uses probability to demonstrate that the frequency of dominant and recessive alleles at the same locus in a population will remain unchanged from one generation to the next if certain conditions are met. Although no actual population can completely meet all of the conditions, the model can be used to detect if allele frequencies are changing from one generation to the next (microevolution) by comparing the frequencies of alleles, genotypes and phenotypes present in the population at the same time. The conditions necessary for allele frequencies to remain unchanged from one generation to the next (equilibrium) are:

- a sufficiently large population size
- no emigration (individuals leave the population)
- no immigration (outside individuals join the population)
- random mating
- no mutations
- recessive and dominant phenotypes contribute equally to fitness (selection does not favor one phenotype over the other)

The conditions above only apply to the alleles that determine the trait being examined. Random mating does not mean mate choice is determined by chance alone, analogous to rolling dice or cutting a deck of cards. The mating is only random in its relationship to the two forms of the phenotype. For example, presence or absence of a widow's peak does not appear to correlate to mate choice in humans. People are not more or less likely to mate with someone because they do or do not have a widow's peak. The alleles that determine this trait meet the condition of random mating.

The Hardy Weinberg equilibrium model uses allele frequencies to predict genotype frequencies for a gene in equilibrium. The symbol p is used in the equations for the frequency of the dominant allele and q for the frequency of the recessive allele. If you were studying the inheritance of albinism in humans, the frequency of the normal pigment allele (A) would be p and the frequency of the albino allele (a) would be q . Since p and q are the only alleles at this locus, the sum of their frequencies should be 100%, or 1. That is the model's first equation:

$$p + q = 100\% = 1$$

Remember from algebra that if you square both sides of any equation that is true, the resulting equation will also be true (if $a=b$, then $a^2=b^2$). So squaring both sides of the equation above yields the following:

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

$$p^2 + 2pq + q^2 = 1$$

This equation will calculate the frequency of the different genotypes and phenotypes using the allele frequencies. The frequency of homozygous dominant individuals (AA – normal pigment phenotype) in the population is p^2 , the frequency of heterozygous individuals (Aa – normal pigment phenotype, carrier) is $2pq$, and the frequency of homozygous recessive individuals (aa – albino phenotype) is q^2 . The frequency of the normal pigmentation (dominant) phenotype is $p^2 + 2pq$. The frequency of the albino (recessive) phenotype is q^2 . However, the equation is only true if the conditions of equilibrium are met-including no difference in fitness between the two phenotypes (no selection).

Two of the most common uses of the Hardy Weinberg equilibrium model are I): estimating allele frequencies from the number of individuals of each phenotype and II): determining if allele frequencies are changing from one generation to the next (microevolution) from the number of individuals of each genotype present. You will perform the necessary calculations for both of these applications of the model in the following problems.

Estimating allele frequencies from the number of individuals of each phenotype:

A population of mice contains individuals with one of two distinct coat colors, tan and mahogany (light and dark brown). The trait is determined by two alleles at the same locus, with the mahogany allele dominant to tan. A sample of the population counted 36 mahogany and 24 tan individuals. If the gene for coat color is in equilibrium, then estimate the frequency of the dominant and recessive alleles:

The sample of mice is used to calculate the frequency of the two phenotypes.

$$\text{The mahogany phenotypic frequency} = \frac{36 \text{ mahogany mice}}{60 \text{ total mice}} = 0.6$$

$$\text{The tan phenotypic frequency} = \frac{24 \text{ tan mice}}{60 \text{ total mice}} = 0.4$$

If the population is at equilibrium, then the frequency of the recessive phenotype should be q^2 , the frequency of the recessive allele squared.

To calculate q , take the square root of the recessive phenotypic frequency:

$$\sqrt{q^2} = \sqrt{0.4}$$

$$q = 0.632$$

Then calculate p as 1 minus q :

$$p = 1 - 0.632$$

$$p = 0.368$$

This estimate is assuming that the allele frequencies are constant-the gene is in equilibrium. Since you do not know how many of the dominant phenotype individuals are heterozygous and how many are homozygous dominant, you can only estimate the allele frequencies.

If you know the genotype of each individual, you can use the Hardy Weinberg equilibrium model to determine if allele frequencies are changing from one generation to the next-to detect if microevolution is occurring. This is a fundamentally different operation from estimation of allele frequency, and none of the calculations done previously will be used in the next part.

Determining if allele frequencies are changing from one generation to the next (microevolution) from the number of individuals of each genotype present:

The following steps are used to determine if allele frequencies are changing:

- I) Calculate Allele frequency from the number of individuals of each genotype**
- II) Calculate expected genotypic frequencies and individuals in a population from allele frequencies:**
- III) Test the goodness of fit between the data and the Hardy Weinberg equilibrium model generated expectations.**

Although the same mouse population is used here, the values for p and q calculated in the previous example problem above are never used in these equations. The two types of problems are completely different uses of the Hardy Weinberg equilibrium model.

The mice sampled had blood taken for DNA analysis. The results indicate that of the 36 mahogany individuals, 21 are heterozygous. Is the gene in equilibrium?

The genotypes are given, and are listed with the phenotypes here:

Mahogany homozygous dominant: 15 Mahogany heterozygous: 21

Tan homozygous recessive: 24

I) Calculating Allele frequency from the number of individuals of each genotype

The first step is to calculate the observed allele frequencies (p & q) from the number of each genotype. The frequency of the mahogany (dominant allele) is:

$$p = \frac{[15 \text{ homozygous dominant} \times 2 \text{ alleles}] + [21 \text{ heterozygotes} \times 1 \text{ allele}]}{60 \text{ individuals} \times 2 \text{ alleles/individual}}$$

$$p = 0.425$$

And from p you can calculate q:

$$q = 1 - p$$

$$q = 0.575$$

The next step is to use the allele frequencies to calculate the number of individuals of each genotype you would expect if the population was in equilibrium.

II) Calculating expected genotypic frequencies and individuals in a population from allele frequencies:

First calculate the expected frequency of homozygous dominant individuals expected:

$$(\text{frequency of homozygous dominant}) = p^2 = (0.425^2) = 0.181$$

To find the number of homozygous dominant individuals expected, multiply this frequency by the sample size. For the mice:

$$(\text{expected number of homozygous dominant individuals}) = 0.181 \times 60 = 10.84.$$

For heterozygotes, the expected frequency of the genotype is:

$$2pq = 2 \times 0.425 \times 0.575 = 0.489$$

And the expected number of heterozygous individuals is:

$$(\text{expected number of heterozygous individuals}) = 0.489 \times 60 = 29.33$$

For homozygous recessive, the expected frequency of the genotype is:

$$q^2 = 0.575^2 = 0.331$$

And the expected number of homozygous recessive individuals is:

$$(\text{expected number of homozygous recessive individuals}) = 0.331 \times 60 = 19.84$$

Since these values are going to be used for a goodness of fit test, retain the decimals and do not round to the nearest individual.

III) Testing the goodness of fit between the data and the Hardy Weinberg equilibrium model generated expectations.

Next from statistics a goodness of fit test is used to compare the observed numbers to the ones expected from the Hardy Weinberg equilibrium model. The test commonly used is a chi-square goodness of fit test. The test uses differences in observed and expected values to generate a test statistic. The test statistic is a measure of the total deviation the observed data has from the expected values. The formula used to calculate the test statistic (χ^2) is:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

For this problem, the following would be calculated:

$$\begin{aligned} \chi^2 &= \frac{(15 - 10.84)^2}{10.84} + \frac{(21 - 29.33)^2}{29.33} + \frac{(24 - 19.84)^2}{19.84} \\ &= 4.84 \end{aligned}$$

Compare this to a chi-square distribution with one less degree of freedom than the number of classes in the problem. Although there are three genotypes, the genotypes are determined by two alleles. So there are $2 - 1 = 1$ degrees of freedom. (Other population geneticists calculate the degrees of freedom as $3 - 2$, losing a second degree of freedom for an hypothesis intrinsic to the data.) A chi square distribution with one degree of freedom yields a value of 3.84 when evaluated at 5%. If a test statistic is greater than this, the differences between the data and expectations should occur less than 5% of the time, and are large enough to call this difference significant. If the test

statistic is less than 3.84, then the differences are expected more than 5% of the time and are not significant. In this problem, the chi-square value 4.84 is greater than the critical value 3.84, so the difference is significant. The mice population is not in equilibrium for the coat gene.

If you do the above calculations, your answers may have a slight variation. All numbers listed on this handout in the examples were rounded to three decimals for presentation only. Calculations were done using Microsoft Excel and no number was rounded by the software. It is recommended you round only the answer, not all numbers during the calculations.

Problems:

Estimating allele frequencies from numbers of individuals or each phenotype:

1.1 In peas, flower color is controlled by genes at a single locus, with the purple allele dominant to the white allele. A population of peas has 15 purple flowers and 24 white flowers. What are the allele frequencies if this gene is in equilibrium?

p = _____

q = _____

1.2 In humans, the hitchhiker's thumb trait is controlled by genes at a single locus, with the allele that does not have the trait dominant to the allele that does. A population of humans has 37 people that do not have the thumb and 45 people with the hitchhiker's thumb. What are the allele frequencies if this gene is in equilibrium?

p = _____

q = _____

1.3 In peas, seed color is controlled by genes at a single locus, with the yellow kernel allele dominant to the green kernel allele. A population of peas has 114 plants with yellow seeds and 75 plants with green seeds. What are the allele frequencies if this gene is in equilibrium?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

1.4 In humans, the ability to taste phenylthiocarbamide (PTC) is controlled by genes at a single locus, with the allele that gives PTC a bitter taste dominant to the allele that gives PTC no taste. A population of humans has 9 people that taste PTC and 7 people that do not taste PTC. What are the allele frequencies if this gene is in equilibrium?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

1.5 In fruit flies, body color is controlled by genes at a single locus, with the gray allele dominant to the black allele. A population of fruit flies has 415 gray fruit flies and 397 black fruit flies. What are the allele frequencies if this gene is in equilibrium?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

1.6 In fruit flies, eye color is controlled by genes at a single locus, with the red allele dominant to the white allele. A population of fruit flies has 1420 red eyed fruit flies and 1659 white eyed fruit flies. What are the allele frequencies if this gene is in equilibrium?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

Determining if allele frequencies are changing from one generation to the next (microevolution) from the number of individuals of each genotype present:

The following steps are used to determine if allele frequencies are changing:

I) Calculate Allele frequency from the number of individuals of each genotype

II) Calculate expected genotypic frequencies and individuals in a population from allele frequencies:

III) Test the goodness of fit between the data and the Hardy Weinberg equilibrium model generated expectations.

The following problems are the calculations used to determine if allele frequencies are changing-if the gene is not in equilibrium. The first two sets will have you calculate the steps used to detect allele changes. The final set of problems will use all of the steps to determine if allele frequencies are changing in a population.

I) Calculating Allele frequency from the number of individuals of each genotype

2.1 In a species of lizard, individuals have either thirteen or fourteen brown and white horizontal stripes on their back. This trait is controlled by genes at a single locus, with the allele for thirteen stripes dominant to the allele for fourteen. A population of lizards is sampled; 87 with thirteen stripes and 43 with fourteen stripes are recorded. Of the 87 with thirteen stripes, 42 are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

p = _____

q = _____

2.2 In a species of bird, individuals have either red or yellow feathers. This trait is controlled by genes at a single locus, with the red allele for dominant to the yellow allele. A population of birds is sampled; 13 red and 10 yellow are recorded. Of the 13 red ones, 5 are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

p = _____

q = _____

2.3 In humans, individuals have either a dimpled chin or a chin without a dimple. This trait is controlled by genes at a single locus, with the allele for a dimple dominant to the allele to not have a dimple. A population of humans is sampled; 17 have dimples and 12 do not. 10 of the people with dimples are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

2.4 In a species of pine tree, individuals have either narrow or wide cones. This trait is controlled by genes at a single locus, with the narrow cone allele dominant to the wide cone allele. A forest is sampled; 107 trees with narrow cones and 136 trees with wide cones are counted. Of the narrow cone trees, 68 are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

2.5 In humans, individuals may or may not have a widow's peak. This trait is controlled by genes at a single locus, with the allele for having a widow's peak dominant to the allele to not have a widow's peak. A population of humans is sampled; 59 have a widow's peak and 44 do not. 25 of the people with a widow's peak are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

2.6 In a species of corn, individuals have either narrow or wide corn cobs. This trait is controlled by genes at a single locus, with the narrow cob allele dominant to the wide cob allele. A farmer counts 74 narrow cob and 113 wide cob plants in a field. Of the narrow cob plants, 28 are homozygous dominant. What are the frequencies of the dominant and recessive alleles observed in this sample?

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

II) Calculating expected genotypic frequencies and individuals in a population from allele frequencies:

3.1 In a population of peas, the frequency of the dominant allele for a purple flower is 0.61 and the frequency of the recessive allele for a white flower is 0.39. What would the genotypic frequencies be if the population is in equilibrium?

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

How many individuals would you expect of each genotype in a population of 500 peas?

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

3.2 In mice, a mahogany coat color allele is dominant over the tan coat color allele. In a population of mice, the frequency of the allele for mahogany coat color is 0.19 and the frequency of the allele for tan coat color is 0.81. What would the genotypic frequencies be if the population is in equilibrium?

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

How many individuals would you expect of each genotype in a population of 100 mice?

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

3.3 In humans, the free earlobe allele is dominant over the attached ear lobe allele. In a population of humans, the frequency of the allele for free earlobes is 0.42 and the frequency of the allele for attached earlobes is 0.58. What would the genotypic frequencies be if the population is in equilibrium?

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

How many individuals would you expect of each genotype in a population of 25 humans?

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

3.4 In a population of peas, the frequency of the dominant allele for a yellow seed is 0.29 and the frequency of the recessive allele for a green seed is 0.71. What would the genotypic frequencies be if the population is in equilibrium?

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

How many individuals would you expect of each genotype in a population of 350 peas?

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

3.5 In humans, the widow's peak allele is dominant over the straight hairline allele. In a population of humans, the frequency of the allele for a widow's peak is 0.61 and the frequency of the allele for a straight hairline is 0.39. What would the genotypic frequencies be if the population is in equilibrium?

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

How many individuals would you expect of each genotype in a population of 75 humans?

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

III) Testing the goodness of fit between the data and the Hardy Weinberg equilibrium model generated expectations.

4.1 In a species of bird, feather color is controlled by genes at a single locus, with the red feather allele dominant to the yellow feather allele. A population has 19 red and 16 yellow birds, with 7 of the red birds having a homozygous dominant genotype. Is this population in equilibrium?

Calculate p and q from the number of individuals of each genotype:

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

Calculate the expected frequency of each genotype if the population is in equilibrium:

$\underline{\hspace{2cm}}$ = Frequency of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Frequency of heterozygous individuals

$\underline{\hspace{2cm}}$ = Frequency of homozygous recessive individuals

Calculate the expected number of individuals of each genotype in a population of 35 birds if the gene is in equilibrium:

$\underline{\hspace{2cm}}$ = Number of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Number of heterozygous individuals

$\underline{\hspace{2cm}}$ = Number of homozygous recessive individuals

Test how well your data fits the expected values from the equilibrium model:

$\underline{\hspace{2cm}}$ = Chi-square test statistic

$\underline{\hspace{2cm}}$ (y/n) in equilibrium?

4.2 In a species of mouse, tail length is controlled by genes at a single locus, with the long tail allele dominant to the short tail allele. A population has 44 long tail and 36 short tail mice, with 27 of the long tail mice having a homozygous dominant genotype. Is this population in equilibrium?

Calculate p and q from the number of individuals of each genotype:

$p =$ _____

$q =$ _____

Calculate the expected frequency of each genotype if the population is in equilibrium:

_____ = Frequency of homozygous dominant individuals

_____ = Frequency of heterozygous individuals

_____ = Frequency of homozygous recessive individuals

Calculate the expected number of individuals of each genotype in a population of 80 mice if the gene is in equilibrium:

_____ = Number of homozygous dominant individuals

_____ = Number of heterozygous individuals

_____ = Number of homozygous recessive individuals

Test how well your data fits the expected values from the equilibrium model:

_____ Chi-square test statistic

_____ (y/n) in equilibrium?

4.3 In humans, the hitchhiker's thumb trait is controlled by genes at a single locus, with the non-hitchhiker's thumb allele dominant to the hitchhiker's thumb allele. A population has 127 people that do not have the hitchhiker's thumb and 75 that do. Of the humans without a hitchhiker's thumb, 42 have a homozygous dominant genotype. Is this population in equilibrium?

Calculate p and q from the number of individuals of each genotype:

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

Calculate the expected frequency of each genotype if the population is in equilibrium:

$\underline{\hspace{2cm}}$ = Frequency of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Frequency of heterozygous individuals

$\underline{\hspace{2cm}}$ = Frequency of homozygous recessive individuals

Calculate the expected number of individuals of each genotype in a population of 202 humans if the gene is in equilibrium:

$\underline{\hspace{2cm}}$ = Number of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Number of heterozygous individuals

$\underline{\hspace{2cm}}$ = Number of homozygous recessive individuals

Test how well your data fits the expected values from the equilibrium model:

$\underline{\hspace{2cm}}$ Chi-square test statistic

$\underline{\hspace{2cm}}$ (y/n) in equilibrium?

4.4 In a certain species of prickly pear, having straight or curved spines is a trait controlled by genes at a single locus, with the straight spine allele dominant to the curved spine allele. A population of prickly pears has 37 individuals with straight spines and 42 individuals with curved spines. Of the prickly pears with straight spines, 26 have a homozygous dominant genotype. Is this population in equilibrium?

Calculate p and q from the number of individuals of each genotype:

$$p = \underline{\hspace{2cm}}$$

$$q = \underline{\hspace{2cm}}$$

Calculate the expected frequency of each genotype if the population is in equilibrium:

$\underline{\hspace{2cm}}$ = Frequency of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Frequency of heterozygous individuals

$\underline{\hspace{2cm}}$ = Frequency of homozygous recessive individuals

Calculate the expected number of individuals of each genotype in a population of 79 prickly pear cacti if the gene is in equilibrium:

$\underline{\hspace{2cm}}$ = Number of homozygous dominant individuals

$\underline{\hspace{2cm}}$ = Number of heterozygous individuals

$\underline{\hspace{2cm}}$ = Number of homozygous recessive individuals

Test how well your data fits the expected values from the equilibrium model:

$\underline{\hspace{2cm}}$ Chi-square test statistic

$\underline{\hspace{2cm}}$ (y/n) in equilibrium?