Name: _____

Hardy-Weinberg Principle

In 1908, two scientists, Godfrey H. Hardy, an English mathematician, and Wilhelm Weinberg, a German physician, independently worked out a mathematical relationship that related genotypes to allele frequencies.

Their mathematical concept, called the Hardy-Weinberg principle, is a crucial concept in population genetics. It predicts how gene frequencies will be inherited from generation to generation given a specific set of assumptions. The Hardy-Weinberg principle states that in a large randomly breeding population, allelic frequencies will remain the same from generation to generation assuming that there is no mutation, gene migration, selection or genetic drift. This principle is important because it gives biologists a standard from which to measure changes in allele frequency in a population.

The Hardy-Weinberg principle can be illustrated mathematically with the equation:

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

Where 'p' and 'q' represent the frequencies of alleles. It is important to note that **p added to q always equals one (100%).** $p + q = 1$

To illustrate how the Hardy-Weinberg principle works, let us consider the MN blood group. Humans inherit either the M or the N antigen which is determined by two different alleles at the same gene locus. If we let the frequency of allele M=p and the frequency of the other allele N=q, then the next generation's genotypes will occur as follows:

- Frequency of homozygous MM genotype = p^2
- Frequency of heterozygous MN genotype = 2pq
- Frequency of homozygous NN genotype = q^2

We can take a sample of the population and count the number of people with each genotype. For example, a sample of 5000 from Forensic Town, USA, has:

- 1460 individuals of type MM, that is 1460/5000 or 29.2%
- 2550 of type MN, that is 2550/5000 or 51%
- 990 of type NN, that is 990/5000 or 19.8%

If we apply the Hardy-Weinberg equation ($p^2 + 2pq + q^2 = 1$) we can calculate the allele frequencies as:

- Frequency of M = p = $\sqrt{0.292}$ = 0.540
- Frequency of N = q = 1 - p = 1 - 0.540 = 0.460

We can now calculate our expected genotype frequencies:

- MM = p^2 = 0.292, or 1460 individuals in the sample
- MN = 2pq = 2 x 0.540 x 0.460 = 0.4968, or 2484 individuals
- NN = q^2 = 0.2116, or 1058

When a population meets all of the of the Hardy-Weinberg conditions, it is said to be in Hardy-Weinberg equilibrium (HWE). Human populations do not meet all of the conditions of HWE exactly, and their allele frequencies will change from one generation to the next and the population will evolve. How far a population deviates from HWE can be measured using the "goodness of fit" or chi-squared test (χ^2).

Mathematically the chi-squared test is represented:

$\chi^2 = \Sigma$ [(observed value – expected value)2 / expected value]

Applying the above data to the chi-square test gives:

- χ^2 = [(990 - 1058)2 / 1058] + [(2550 – 2484)2 / 2484] + [(1460-1460)2 / 1460]
- χ^2 = [4624 / 1058] + [4356/ 2478] + [0 / 1460]
- χ^2 = [4.371] + [1.754] + [0] = 6.125

To determine what this chi-squared value means, we must next look at a chi-squared distribution table.

Chi-Square Table

	Degrees of Freedom							
p	1	2	3	4	5	6	7	8
0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51
0.01	6.64	9.32	11.34	13.28	15.09	16.81	18.48	20.09

Since we have three genotypes, we therefore have 3 minus 1, or 2 degrees of freedom. Degrees of freedom is a complex issue, but we could look at this in simple terms: if we have frequencies for three genotypes that are truly representative of the population then, no matter what we calculate for two of them, the frequency of the third must not be significantly different for what is required to fit the population.

Looking across the distribution table for 2 degrees of freedom at 0.05 p (95% confidence), we find our chi-squared value of 6.125 contradicts the hypothesis that the differences in the Observed and Expected data did not arise by chance. Since the chi-squared value falls above the 0.05 (5%) significance cutoff, we can conclude that the Forensics Town population differs significantly from what we would expect for a Hardy-Weinberg equilibrium of the MN blood group. This means that the population is evolving due to not conforming to one of the conditions that are assumed in Hardy-Weinberg. ie: is no mutation, gene migration, selection or genetic drift

Example

Determining Hardy-Weinberg Equilibrium

Suppose that scientists are observing a population of lab-bred flies, and discover a gene controlling eye color. The *R* allele produces regular-colored eye pigment, while the *r* allele produces red pigment. Individuals that are heterozygous (*Rr*) have pink eyes. In a population of 150 flies, 15 flies have red eyes, 90 have normal eye color, and 45 have pink eyes.

Is this population in Hardy-Weinberg equilibrium?

In order for a population to be considered to be in equilibrium, it must remain the same from generation to generation. Therefore, in order to determine if this population of fruit flies is in Hardy-Weinberg equilibrium, the genetic distribution of the current generation must be compared to a prediction of the genetic distribution of the next generation, as calculated using the Hardy-Weinberg equation.

Step 1: Determine the gene frequencies of the current generation.

Phenotype	Genotype	# of Individuals
Normal Eyes	*RR*	90
Red Eyes	*rr*	15
Pink Eyes	*Rr*	45

Given this information, calculating the allele frequencies is simply a matter of counting up all of the alleles.

- Remember, each parent carries *two* alleles, so the total # of alleles is twice the population.
- Also remember that *heterozygous* individuals carry one of *each* allele.

Taking these two factors into account,

f(R) = [(90 x 2) + (45)] / (150 x 2) = 225/300 = **0.75 = p**

[(90 RR x 2 - one for each R) + (45 Rr - not multiplied by 2 since only 1 of the alleles is R)]/ 150 (total number of flies) x 2 (because each fly has two alleles) = 0.75

f(r) = [(15 x 2) + (45)] / (150 x 2) = 75/300 = **0.25 = q**

[(15 rr x 2 - one for each r) + (45 Rr - not multiplied by 2 since only one of the alleles is r)]/ 150 (total number of flies) x 2 (because each fly has two alleles) = 0.25

Step 2: Determine the expected genotype frequencies for the next generation.

Plugging the frequencies of each allele into the Hardy-Weinberg equation, we find the expected numbers of each genotype in the population:

f(RR) = p2 = f(R) x f(R) = 0.75 x 0.75 = **0.5625**

f(rr) = q2 = f(r) x f(r) = 0.25 x 0.25= **0.0625**

f(Rr) = 2pq = 2 x [f(R) x f(r)] = 2 x (0.75 x 0.25) = **0.375**

Multiplying each of these genotype frequencies with the total population number, we find that there should be:

- 0.5625 x 150 = 84 normal-eyes flies (*RR*)
- 0.0625 x 150 = 9 red-eyed flies (*rr*)
- 0.375 x 150 = 56 pink-eyed flies (*Rr*)

(Since partial individuals do not exist, the numbers are rounded off.)

Step 3: Compare the expected frequency with the original population numbers.

Comparing the expected numbers with the actual numbers of each phenotype, population geneticists can determine if populations are either in equilibrium (or very close to it) or are experiencing *disequilibrium* of some sort. In this example:

Phenotype	**Genotype**	**Expected #**	**Observed #**
Normal Eyes	*RR*	84	90
Red Eyes	*rr*	9	15
Pink Eyes	*Rr*	56	45

In this example, the population is not in equilibrium since the expected and observed values do not match. Once a population geneticist determines that a population is in disequilibrium, the reasons can be explored. Disequilibrium can be attributed to different possible mechanisms, depending on (1) the context of the population, and (2) the manner in which the population is skewed.

Step 4: Chi-Squared Test

$\chi^2 = \Sigma$ [(observed value – expected value)2 / expected value]

$X^2 = [(90\text{-}84)^2 / 84)] + [(15\text{-}9)^2 / 9)] + [(45\text{-}56)^2 / 56]$

$X^2 = [36/84] + [36/9] + [121/56]$

$X^2 = [0.429] + [4] + [2.161] = 6.59$

Chi-Square Table

	Degrees of Freedom							
p	1	2	3	4	5	6	7	8
0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51
0.01	6.64	9.32	11.34	13.28	15.09	16.81	18.48	20.09

Since we have three genotypes, we therefore have 3 minus 1, or 2 degrees of freedom. Degrees of freedom is a complex issue, but we could look at this in simple terms: if we have frequencies for three genotypes that are truly representative of the population then, no matter what we calculate for two of them, the frequency of the third must not be significantly different for what is required to fit the population.

Looking across the distribution table for 2 degrees of freedom at 0.05 p (95% confidence), we find our chi-squared value of 6.59 contradicts the hypothesis that the differences in the Observed and Expected data did not arise by chance. Since the chi-squared value falls above the 0.05 (5%) significance cutoff, we can conclude that the Fly population differs significantly from what we would expect for a Hardy-Weinberg equilibrium of the eye color. This means that the population is evolving due to not conforming to one of the conditions that are assumed in Hardy-Weinberg. ie: is no mutation, gene migration, selection or genetic drift