



RADMAS- 2016

CALCULATION OF GENETIC VARIATION OF A POPULATION USING HARDY-WEINBERG EQUATION

Dr. K. MANIKYA KUMARI

Head, Department of Botany, St. Joseph's College for Women (A), Visakhapatnam-530 004, Andhra Pradesh

Email: mk.kambham@gmail.com

ABSTRACT

The Hardy-Weinberg equation is a mathematical equation that can be used to calculate the genetic variation of a population at equilibrium. The Hardy-Weinberg formulas allow determining whether or not evolution has occurred and any changes in the gene frequencies in the population over time can be detected. It involves the calculation of allele frequencies to find out percentage or proportion of the population with a particular trait in homozygous or heterozygous condition. The Hardy-Weinberg equilibrium cannot exist in real life. Some or all of these types of forces all act on living populations at various times and evolution at some level occurs in all living organisms. The Hardy-Weinberg formulas allow us to detect some allele frequencies that change from generation to generation. This paper discusses the application of Hardy-Weinberg equation to predict the phenotype and genotype expression of genetic traits and the amount of genetic variation in a population using some heritable traits of a population at equilibrium plants, animals and humans as examples.

Keywords: Population, Genetic traits, Genotype, offspring, Allele frequency, Probability, Hardy Weinberg equation

1. INTRODUCTION

Most genetics research focuses on the structure of genes on chromosomes, the function of genes, and the process of genetic transmission from parent to offspring. **Population genetics** instead focuses on the overall **gene pool** in a population of interbreeding organisms - that is, the frequency of all alleles of all genes in the population - and whether the gene pool may be changing across generations in a population. Population genetics examines the relationship among genotype frequencies in a population, allele frequencies in its gene pool, and factors that can change these frequencies over time.

The Hardy-Weinberg equation is a mathematical equation that can be used to calculate the genetic variation of a population at equilibrium. In 1908, G. H. Hardy and Wilhelm Weinberg independently described a basic principle of population genetics, which is now named the Hardy-Weinberg equation. The equation is an expression of the principle known as Hardy-Weinberg equilibrium, which states that the amount of genetic variation in a population will remain constant from one generation to the next in the absence of disturbing factors.

2. APPLICATION OF HARDY-WEINBERG EQUATION

To explore the Hardy-Weinberg equation, we can examine a simple genetic locus at which there are two alleles, A and a. The Hardy-Weinberg equation is expressed as:

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



where,

p is the frequency of the "A" allele

q is the frequency of the "a" allele in the population.

In the equation, p^2 represents the frequency of the homozygous genotype AA, q^2 represents the frequency of the homozygous genotype aa, and $2pq$ represents the frequency of the heterozygous genotype Aa. In addition, the sum of the allele frequencies for all the alleles at the locus must be 1, so $p + q = 1$. If the p and q allele frequencies are known, then the frequencies of the three genotypes may be calculated using the Hardy-Weinberg equation. In population genetics studies, the Hardy-Weinberg equation can be used to measure whether the observed genotype frequencies in a population differ from the frequencies predicted by the equation

3. DERIVATION OF THE HARDY-WEINBERG THEOREM:

Three Methods

- In a **monoecious** species, for a **single locus** with **two alleles A & a**:
 Let frequency of **A** allele **$f(A) = p$**
 Let frequency of **a** allele **$f(a) = q$**
 Let allele frequencies in males & females be equal.
 The **expected genotype frequencies** can be calculated thereafter.
- The **probabilities** of obtaining various gametic combinations
 prob. of **A** from male & **A** from female: **$p_m \times p_f = p^2$**
 prob. of **a** from male & **a** from female: **$q_m \times q_f = q^2$**
 prob. of **A** from male & **a** from female,
 or **a** from male & **A** from female: **$(p_m \times q_f) + (p_f \times q_m) = 2pq$**
- The result of the **binomial expansion** of p & q in males & females
 $(p_m + q_m) (p_f + q_f) = p^2 + 2pq + q^2$

The areas of the four boxes in a **Punnet Square**

	p	q
p	p^2	pq
q	pq	q^2

By all three methods: **$f(AA) = p^2$ $f(Aa) = 2pq$ $f(aa) = q^2$**

4. POPULATION GENETICS AND THE HARDY-WEINBERG LAW

The Hardy-Weinberg formulas allow scientists to determine whether evolution has occurred. Any changes in the gene frequencies in the population over time can be detected. The law essentially states that if no evolution is occurring, then equilibrium of allele frequencies will remain in effect in each succeeding generation of sexually reproducing individuals. In order for equilibrium to remain in effect (i.e. that no evolution is occurring) then the following five conditions must be met:

1. No mutations must occur so that new alleles do not enter the population.
2. No gene flow can occur (i.e. no migration of individuals into, or out of, the population).



3. Random mating must occur (i.e. individuals must pair by chance)
4. The population must be large so that no genetic drift (random chance) can cause the allele frequencies to change.
5. No selection can occur so that certain alleles are not selected for, or against.

Obviously, the Hardy-Weinberg equilibrium cannot exist in real life. Some or all of these types of forces all act on living populations at various times and evolution at some level occurs in all living organisms. The Hardy-Weinberg formulas allow us to detect some allele frequencies that change from generation to generation, thus allowing a simplified method of determining that evolution is occurring. There are two formulas that must be memorized:

$$p^2 + 2pq + q^2 = 1 \text{ and } p + q = 1$$

Where,

p = frequency of the dominant allele in the population

q = frequency of the recessive allele in the population

p^2 = percentage of homozygous dominant individuals

q^2 = percentage of homozygous recessive individuals

2pq = percentage of heterozygous individuals

4. 1. Population Genotypes and Alleles

The Hardy-Weinberg principle applies to individual genes with two alleles, a dominant allele and a recessive allele. A population with such a gene can be described in terms of its **genotype numbers** - the number of individuals with each of the three resulting genotypes - or in terms of the three **genotype frequencies**. The frequency of each genotype is the number of individuals in the population with that genotype divided by the total number of individuals in the population, as displayed in Table 1. (Note that total genotype frequencies sum to 1.0.)

Table 1. Genotype numbers and genotype frequencies in a hypothetical population.

Genotypes	Number with these Genotypes	Genotype Frequencies
AA	180	0.36
Aa	240	0.48
aa	80	0.16
Total	500	1.0

4. 2. Calculating Allele Frequencies

We can also describe a population, somewhat more abstractly, in terms of its **allele frequencies**. The frequency of an allele is defined as the total number of copies of that allele in the population divided by the total number of copies of all alleles of the gene. We can calculate population allele frequencies from genotype numbers.

The total number of dominant **A** alleles in our population equals 600, which is the sum of:

- the number of AA individuals times 2 (the number of **A** alleles per individual) = $180 \times 2 = 360$
 - the number of Aa individuals (times 1, the number of **A** alleles per individual) + 240
- 600



RADMAS– 2016

The total number of all alleles of the gene equals 1000, which is 2 times the number of individuals in the population (since the individuals are diploid).

By definition, the frequency of the dominant **A** alleles in our population equals to:

$$600/1000 = 0.60.$$

We can calculate the total number of **a** alleles in the population, and divide by the total number of alleles:

$$(2 \times 80 + 240)/1000 = 400/1000 = 0.4$$

Or, since the total of all the allele frequencies sums to 1.0 and since there are only two alleles, **A** and **a**, we can calculate the allele frequency by subtracting the **A** allele frequency from 1.0:

$$1.0 - 0.60 = 0.40$$

4. 3. Testing for Genetic Equilibrium

We can check if a population is in genetic equilibrium by testing if the Hardy-Weinberg principle applies, as follows:

Given the population genotype numbers,

(1) calculate the allele frequencies from the observed population genotype numbers.

(2) calculate the genotype frequencies from the observed genotype numbers.

(3) apply the Hardy-Weinberg principle to calculate the **expected** genotype frequencies from the allele frequencies in the population.

(4) If the population is in Hardy-Weinberg equilibrium the observed genotype frequencies in step 2 will be (roughly) the same as the expected frequencies in step 3. (A Chi-Square test is used to determine if the observed and expected genotype frequencies are significantly different from each other or not.)

4. 4. Examples:

PROBLEM 1: Sickle-cell anemia is an interesting genetic disease. Normal homozygous individuals (**SS**) have normal blood cells that are easily infected with the malarial parasite. Thus, many of these individuals become very ill from the parasite and many die. Individuals homozygous for the sickle-cell trait (**ss**) have red blood cells that readily collapse when deoxygenated. Although malaria cannot grow in these red blood cells, individuals often die because of the genetic defect. However, individuals with the heterozygous condition (**Ss**) have some sickling of red blood cells, but generally not enough to cause mortality. In addition, malaria cannot survive well within these "partially defective" red blood cells. Thus, heterozygotes tend to survive better than either of the homozygous conditions. If 9% of an African population is born with a severe form of sickle-cell anemia (**ss**), what percentage of the population will be more resistant to malaria because they are heterozygous (**Ss**) for the sickle-cell gene?

Answer:

$$9\% = .09 = ss = q^2.$$

To find **q**, simply take the square root of 0.09 to get 0.3.

Since $p = 1 - 0.3$,

Then, **p** must equal 0.7. $2pq = 2(0.7 \times 0.3) = 0.42 = 42\%$

i. e., 42% of the population are heterozygotes (carriers).



RADMAS- 2016

PROBLEM 2: Cystic fibrosis is a recessive condition that affects about 1 in 2,500 babies in the Caucasian population of the United States. Please calculate the following:

- A. The frequency of the recessive allele in the population.
- B. The frequency of the dominant allele in the population.
- C. The percentage of heterozygous individuals (carriers) in the population.

Answer:

Since $2pq$ equals the frequency of heterozygotes or carriers, then the equation will be as follows:

$$2pq = (2)(.98)(.02) = 0.04 \text{ or } 1 \text{ in } 25 \text{ are carriers.}$$

PROBLEM 3. In a given population, only the "A" and "B" alleles are present in the ABO system; there are no individuals with type "O" blood or with O alleles in this particular population. If 200 people have type A blood, 75 have type AB blood, and 25 have type B blood, what are the allelic frequencies of this population (i.e., what are p and q)?

Answer:

To calculate the allele frequencies for A and B, we need to remember that the individuals with type A blood are homozygous AA, individuals with type AB blood are heterozygous AB, and individuals with type B blood are homozygous BB.

The frequency of A equals the following:

$$2 \times (\text{number of AA}) + (\text{number of AB}) \text{ divided by } 2 \times (\text{total number of individuals}).$$

$$\text{Thus } 2 \times (200) + (75) \text{ divided by } 2 (200 + 75 + 25).$$

$$\text{This is } 475/600 = 0.792 = p.$$

$$\text{Since } q \text{ is simply } 1 - p, \text{ then } q = 1 - 0.792 \text{ or } 0.208$$

5. Conclusions: Hardy-Weinberg equation states that the amount of genetic variation in a population will remain constant from one generation to the next in the absence of disturbing factors. We can check if a population is in genetic equilibrium by using the Hardy-Weinberg principle. We can also describe a population, somewhat more abstractly, in terms of its allele frequencies. The Hardy-Weinberg principle applies to individual genes with two alleles, a dominant allele and a recessive allele. A population with such a gene can be described in terms of its genotype numbers - the number of individuals with each of the three resulting genotypes - or in terms of the three genotype frequencies. Any changes in the gene frequencies in the population over time can be detected to know whether evolution occurred or not. It is highly useful for finding genetic variation in a population.

6. References

- [1]. Artin, E. (1964) The Gamma Function. Holt, Rinehart and Winston, New York, Chicago, San Francisco, Toronto, London.
- [2]. Derman, C., Gleser, L. J. and Olkin, I. (1973) A Guide to Probability and Application. Holt, Rinehart and Winston, Inc., New York, Chicago, San Francisco, Toronto, London. *52 Stochastic Processes in Genetics and Evolution*
- [3]. Feller, W. (1968) An Introduction to Probability Theory and Its Applications. Volume I, Third Edition. John Wiley and Sons, New York, London, Sydney.
- [4]. Snustad, D. P. and Simmons, M. J. (2006) Principles of Genetics. John Wiley & Sons, Inc., New York, London.