University of Sri Jayewardenepura
Faculty of Graduate Studies

Department of Mathematics

Generalisation of the Hardy-Weinberg equation in the case of polygenic inheritance in population genetics and an estimation of the number of genes controlling human height

A Thesis

By

Valemurugan Balamurari

Submitted in partial fulfillment of the requirements for the

Postgraduate Diploma

in

Industrial Mathematics

Signature : ...V. Balamurari...

Name : V. Balamurari

To my ever loving parents and all my teachers

We approve the thesis titled "*Generalisation of the Hardy-Weinberg equation in the case of polygenic inheritance in population genetics and an estimation of the number of genes controlling human height*" of Mr. Valemurugan Balamurari.

Date : 2001/02/23

Signature : ....

Prof. H.G. Nandadasa
Professor of Botany and Head of the
Department of Botany,
University of Sri Jayewardenepura.

Date : 2001-02-23

Signature : ....

Mr. P. Dias
Senior lecturer in Mathematics
Department of Mathematics,
University of Sri Jayewardenepura.

Date : 2001.02.23

Signature : ....

Prof. Sunethra Weerakoon
Professor of Mathematics and the
Co-ordinator of the Pg.Dip / M.Sc. in Industrial Mathematics Programme
Department of Mathematics,
University of Sri Jayewardenepura.

Declaration

I hereby declare that the thesis titled "*Generalisation of the Hardy-Weinberg equation in the case of polygenic inheritance in population genetics and an estimation of the number of genes controlling human height*" contains my own work and that no part of this thesis has been submitted previously to this or any other university or institution.

Signature ....V. Balamurari............

Name : V. Balamurari

Date : December 2000

Abstract

Most of the characteristics found in human beings are controlled by one or more genes. There are certain characters of which we know the exact number of genes that are involved. The variation shown by a character determined by a single gene is usually easily distinguishable as opposed to the variation shown by a character determined by two or more genes. There is very little evidence of research done using statistical techniques to determine the number of genes controlling human height. As such, the aim of my research work is to obtain an approximate estimate of the number of genes involved.

The basic Hardy-Weinberg equation can be used only when we wish to calculate the allelic and genotypic frequencies for a character determined by a single gene having two alleles. Therefore it becomes necessary to have a more general equation relating allelic frequencies in order to perform calculations when we consider a polygenic trait. A generalized Hardy-Weinberg equation for polygenic inheritance too has been derived as part of my research work.

Acknowledgements

I take this opportunity first to express my gratitude to my loving mother who was instrumental in giving me the initial encouragement and support when I decided to focus my attention on an exciting but not much researched topic of Hardy-Weinberg law and its applications to polygenic inheritance. I am greatly indebted to Professor H.G. Nandadasa for having spent his valuable time in explaining to me the essentials of genetics and providing the necessary background knowledge in biology in general. I am very grateful to Mr. P. Dias for putting the project on a firm mathematical footing, without which I would have had to drop the project topic of my interest. I am indeed thankful to Professor S.Weerakoon for having given me the great opportunity to follow the program in Industrial Mathematics, of which this project is a part, and hence approving my project proposal.

I shall now thank all the people who helped me in collecting the data (heights of students) ,which is an essential component of my project, individually. Mr. S.P Senaratne, vice principal of Royal College, Colombo, for permitting me to collect the data I required, Mrs. P.N. Attygalle, senior botany teacher of the college, who took her very valuable time in providing me with as many data values as she could get amidst her commitments at the college, Mr. D.A. Pakianathan, acting warden of the St. Thomas' College, Mount Lavania, again for the permission he gave for collecting data at the college and Mr. Siri Edirisinghe, Mr. R. M Jayasena and Mr. S. Wickramaratne for their personal involvement in providing me with whatever the data they could obtain in the college.

Contents

Chapter 1

Introduction

1.1 Genetics

The theory of genetics is concerned with the study of the manner in which parental characteristics are being passed onto their offspring. Consequences of the studies carried out with regard to the mechanism of inheritance have made it possible to completely prevent certain diseases and some undesirable characters from being passed on genetically from the parents to their offspring in reproduction. These studies have also enabled geneticists to produce genetically modified plants and animals to gain economic advantage. So we, as humans, would generally like to know what characters in what proportions, relative to the population concerned, are passed on from one generation to another. The study of genetics was originally initiated by Gregor Mendel of Austria. He formulated two laws of genetics after carrying out numerous experiments on plants. The two laws are as follows.

Mendel's first law (law of segregation)

Each organism contains two factors for each trait, and the factors segregate during the formation of gametes so that each gamete contains only one factor from each pair of factors. When fertilization occurs, the new organism will have two factors for each trait, one from each parent.

Mendel's second law (law of independent assortment)

Members of one pair of factors segregate (assort) independently of members of another pair of factors. Therefore, all possible combinations of factors can occur in the gametes.

1.2 Hardy-Weinberg principle

A mathematical model to study the allelic frequencies of a gene responsible for a particular character was devised by a British mathematician G.H. Hardy and a German physician Wilhelm Weinberg. The Hardy-Weinberg principle is as follows.

The frequencies of alleles of a gene in a population will remain the same in every succeeding generation as long as the following conditions are satisfied;

1. Large population with random mating
2. No selection or mutation
3. No migration

Given below are the brief explanations of the conditions of the Hardy-Weinberg principle.

- **Large population with random mating**

Large population is to ensure that small changes in the allelic frequencies will not significantly affect genetic composition of the population from one generation to the next. Mating among the individuals of a population is said to be random if every individual of the population has an equal chance to reproduce with any other individual in the population.

Random mating among the individuals of a population does not mean haphazard mating that is found in many animals and birds. It is the selection of a mate for reproduction without regard to any genotype.

In other words, the selection of a mate for reproduction should not have any genetic basis, for the existence of a population with random mating.

Consider a random sample of people for a study of the genetic basis of their heights. If the height of an individual is not used as a criterion in selecting a partner for marriage, then the children of such parents can be considered to have come as a result of random mating with respect to height. If the marriages of the parents of the individuals in the random sample were not influenced by their heights, the random sample can be considered to have come from a population with random mating.

The height of an individual is generally not used as an important characteristic in selecting a partner for marriage, although it may be the case in certain extreme cases. Race, religion, language and the ethnic group of an individual are the most, among many others, widely used criteria in selecting a partner for marriage. Therefore, it is a reasonably good assumption to say that the height of an individual does not play a vital role in selecting a partner for marriage.

- **No selection or mutation**

No individual should have a greater chance of reproductive success over another. Selection is said to take place when the selection of mate for reproduction has some genetic basis. We may have a sample of individuals, chosen on the basis of a character which is genetically determined, with random mating among the individuals. Selection is used in plant and animal breeding. Selection, in general, does not exist in human populations.

As the selection of a partner for marriage is not generally affected by the height of the partner, we may conveniently assume that a sample of randomly selected individuals have come from a population where there is no selection.

Mutation is a change in the genetic material (DNA). Mutation may occur spontaneously due to natural causes or man-made radiation and chemical pollutants of the environment. Mutations are very infrequent and hence they alone will not alter the allele frequencies in a population radically. As such, mutation can be assumed to have no influence on any particular character being studied.

- **No migration**

Migration is the movement of individuals into and out of a population. Immigration or emigration may alter the genetic composition of the population. Migration will have an effect on the allele frequencies only if a comparably large number of individuals of a population with different allele frequencies move into the receiving population. Movement of individuals, in relatively small numbers, into and out of a population will not alter the allele frequencies to a greater extent. Effect of migration in the study of a particular character of individuals chosen randomly from a population can be assumed to be minimum if all the individuals chosen have lived in that population for a major part of their life time.

1.3 Hardy-Weinberg (H-W) equation

A mathematical equation, involving the allelic frequencies of a gene, was developed by G.H. Hardy and W. Weinberg for calculating the genotype frequencies when the frequency of one of the alleles of the gene is known. It should be noted that the number of alleles involved in the development of the equation has been assumed to be two although a gene may have more than two alleles. The following are the steps involved in the derivation of the Hardy-Weinberg equation.

Suppose that a gene has alleles **A** (dominant) and **a** (recessive). Let the frequencies of the alleles in a given population be as follows.

A – p (p is the frequency of the allele **A**)

a – q (q is the frequency of the allele **a**), where $p + q = 1$.

	A (p)	**a** (q)
A (p)	AA (p^2)	Aa (pq)
a (q)	Aa (pq)	aa (q^2)

Using the fact that sum of the frequencies of all the genotypes is one, we get the equation,

$$\mathbf{p^2 + 2pq + q^2 = 1}, \text{ where } p + q = 1. \qquad (1.1)$$

This is known as the Hardy-Weinberg equation.

1.4 Applications of H-W equation

Example 1 *Tongue-rolling in humans. The ability to roll the tongue is caused by a single gene. Presence of the dominant allele is responsible for the ability to roll one's tongue into a "U" shape. Suppose that in a certain population 84% of the people can roll their tongues whereas the remaining 16% cannot. Find the percentages of the people having different genotypes.*

Solution

Let R and r denote the dominant and recessive alleles of the genes responsible for one's ability to roll his/her tongue.
Individuals having the genotypes RR and Rr are tongue rollers and those having the genotype rr are non-tongue rollers.

It is clear from the information given that 16% of the individuals have the genotype rr.

$$\therefore \quad q^2 = 0.16 \Rightarrow q = 0.4$$

$$\therefore \quad p = 0.6 \text{ since } p + q = 1$$

Now the percentage of individuals having the genotype RR $= 100p^2\% = 36\%$

The percentage of individuals having the genotype Rr $= 100(2pq)\% = 48\%$

Example 2 *ABO blood group in humans are determined by a gene having* ***three*** *alleles. The four blood groups are generally known as* ***A, B, AB*** *and* ***O****. Let* I^A, I^B *and i be the three alleles of the gene.* I^A *and* I^B *are the dominant alleles and i is the recessive allele. The possible genotypes are* I^AI^A, I^BI^B, I^Ai, I^Bi, I^AI^B *and ii. The following table gives the type of blood group and the genotype associated with it.*

Genotype	***Blood group***
I^AI^A, I^Ai	*A*
I^BI^B, I^Bi	*B*
I^AI^B	*AB*
ii	*O*

Let the frequencies of I^A, I^B *and i be p, q and r respectively, where* $p + q + r = 1$*. If the percentage of the individuals having at least two different types of blood group are known from the population, then all the allelic and genotypic frequencies can be calculated.*