Indian Institute of Technology, Ropar
Department of Mathematics

Genotype Distribution

Anubothula Hema Priya
2020MCB1228

MA614: Applied Linear Algebra
Dr.G.Raju

October 9, 2022

Contents

- Introduction
- Genotype Distribution
- Assumption
- Objective
- Application of Linear Algebra

1. Methodology
   - Problem Statement
   - Discussion
   - Result

2. Plan for next part
   - Part 2

3. References
   - References

Introduction

- What is Genetics?
  - Genetics is the branch of biology that studies organisms' heredity and variation. Throughout the study of Genetics, there are many factors that take part by which organism looks and acts the way that they do, both in their physical shape and also in their genes. In the study of Genetics an organism genes change to fit their surroundings either through "natural selection" or to adapt to their environment. Both are major causes for these adaptations in which the dominant traits are inherited and recessive traits are "lost" throughout their existence.
- Population Genetics
  - It is a branch of genetics that studies the genetic structure of a certain population and seeks to explain how transmission of genes changes from one generation to another.

Introduction

- What is Genes?
  - A unit of heredity which is transferred from a parent to offspring and is held to determine some characteristic of the offspring. It governs the inheritance of traits like sex, colour of eyes, hair(for plants and animals), petal colour(for plants).
- Genotype
  - Each individual in a population carries a pair of genes and these pair of genes are called the individual's genotype.
- Autosomal Inheritance
  - There are several types of inheritance, but we are mainly concerned about autosomal inheritance where a heritable trait is assumed to be governed by a single gene (generally called allele).

Genotype Distribution

- This is the method which allow for estimation of the expected genotype count or fractions for one or two individuals based on genotype likelihoods. This can be very useful for a number of population genetic statistics.

Assumption

- According to Autosomal Inheritance model
- First, we consider two different forms of genes. Let them be A and a.
- Therefore, we obtain three possible genotypes for each inheritable trait: AA, Aa and aa (Aa and aA are genetically same).
- Consider each of the genotype is equally likely to occur.

Objective

- To find the probability of trait produced based on the genotypes of offspring after undergoing fertilization between these three different pair of genotypes of parents and the probability of three genotypes in the nth generation of trait.
- To predict what trait will occur when three different genotypes of parents are pairing each other.

Application of Linear Algebra

- In genetics, Linear Algebra is used to determine the genotype distribution of a certain trait in a population after any number of generations based on the starting population's genotype distribution.
- In this case, we use the application of eigen values, eigen vectors and diagonalisation.

The genotype distribution of a particular trait in a population in the n-th generation can be represented by a genotype vector $\mathbf{x}_n = \begin{matrix} a\_n \\ b\_n \\ c\_n \end{matrix}$ where a_n, b_n, and c_n denote the portion of the population with genotype AA, Aa, aa, respectively in the n-th generation. Since the genotype distribution changes over time, we can represent the succession of genotype distributions from one generation to the next in the form of a difference equation,

$$\mathbf{x}_n = M\mathbf{x}_{n-1} \qquad (1)$$

Application of Linear Algebra

- For a suitable matrix M and n = 1,2,3,....We seek an explicit description of xn whose formula for each xn does not depend on A or on the preceding terms in the sequence other than the initial term x0 (the initial genotype distribution in the population). From equation (1), we have,

$$\mathbf{x}_n = \mathbf{M}\mathbf{x}_{n-1} = \mathbf{M}^2\mathbf{x}_{n-2} = \ldots\ldots = \mathbf{M}^n\mathbf{x}_0$$

  Therefore, if we can find an explicit expression for Mn, we can use equation (1) to obtain an explicit expression for xn in terms of x0. We will proceed to diagonalize M (that is, we will seek to find an invertible matrix P and a diagonal matrix D such that M = PDP-1). Since $M^n = PD^nP^-1$

- This diagonalization will enable us to compute Mn quickly for large values of n. M is diagonalizable, then the diagonal entries of D are the eigenvalues of M, and the columns of P are n linearly independent eigenvectors corresponding, respectively, to each of the eigenvalues of M. Once P and D have been specified, we will have

Application of Linear Algebra

$$\mathbf{x}_n = PD^nP^{-1}\mathbf{x}_0$$

This is an explicit description of xn that is relatively easy to compute, and is based upon x0.

Problem Statement

- A farmer notices that most of people would like to eat red apples more than green apples. He wants to find the way how to produce more red apples than green apples for selling to the buyer. Assume he has a huge population of apples with some distribution of each the three genotypes RR,Rr and rr to start a breeding programme in which each apple in the population is fertilised with a RR apple and subsequently replaced by one of its offspring. After any number of generations, we want to compute an expression for the distribution of three genotypes in the population.
- Let R be the dominant allele and r be the recessive allele:
  - A plant with genotype RR will have red apples.
  - A plant with genotype Rr will have red apples.
  - A plant with genotype rr will have green apples.

Problem Statement

In table 1, we list the probabilities of the possible genotypes of the offspring for three possible combinations of the genotypes of the parents which is RR.

Genotype of offspring	Genotype of parents		
	$RR - RR$	$RR - Rr$	$RR - rr$
RR	1	$\frac{1}{2}$	0
Rr	0	$\frac{1}{2}$	1
rr	0	0	0

Table 1

Discussion

- From Table 1, we create a matrix A with row as Genotype of Offspring and x0 (Initial generation) is column matrix with probabilities of genotype.(P(RR)= P(Rr) = P(rr) = 1/3)

$$A = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$x_0 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

Discussion

First of all, we must find the eigenvalues and eigenvectors for matrix A.

$$A - \lambda I = \begin{bmatrix} 1-\lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}-\lambda & 1 \\ 0 & 0 & 0-\lambda \end{bmatrix}$$

$$= \begin{vmatrix} 1-\lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}-\lambda & 1 \\ 0 & 0 & 0-\lambda \end{vmatrix}$$

$$= (0-\lambda)\left[(1-\lambda)\left(\frac{1}{2}-\lambda\right) - 0\right]$$

$$= -\lambda^3 + \frac{3}{2}\lambda^2 - \frac{1}{2}\lambda$$

$$\therefore \lambda = 1, \frac{1}{2}, 0$$

$$\lambda_1 = 1;\ \lambda_2 = \frac{1}{2};\ \lambda_3 = 0$$

Discussion

We already found the eigenvalues for matrix A. Now, we need to find the corresponding eigenvectors for all eigenvalues.

$$\lambda = 1$$

$$\begin{bmatrix} 1-1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}-1 & 1 \\ 0 & 0 & 0-1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1 = 1$$

$$\frac{1}{2}x_2 = 0$$

$$-x_3 = 0$$

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda = \frac{1}{2}$$

$$\begin{bmatrix} 1-\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}-\frac{1}{2} & 1 \\ 0 & 0 & 0-\frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{1}{2}x_1 + \frac{1}{2}x_2 = 0 \rightarrow x_1 + x_2 = 0 \rightarrow x_2 = -x_1$$

$$x_1 = 1$$

$$x_2 = -1$$

$$-\frac{1}{2}x_3 = 0 \rightarrow x_3 = 0$$

$$v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Discussion

$$\lambda = 0$$

$$\begin{bmatrix} 1-0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2}-0 & 1 \\ 0 & 0 & 0-0 \end{bmatrix}\begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$x_1 + \frac{1}{2}x_2 = 0 \to \frac{1}{2}x_2 = -x_1$$

$$x_1 = 1$$

$$\frac{1}{2}x_2 = -1 \to x_2 = -2$$

$$\frac{1}{2}x_2 + x_3 = 0 \to x_3 = -\frac{1}{2}x_2$$

$$x_3 = -\frac{1}{2}\left(-2\right) \to x_3 = 1$$

$$v_3 = \begin{bmatrix}1\\-2\\1\end{bmatrix}$$

After we get all the corresponding eigenvectors of all the eigenvalues, we can determine the diagonalization for matrix A.

$$D^n = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1}{2}\right)^n & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

After that we must find the corresponding matrix obtained by the eigenvectors.

$$P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Discussion

All the values for formula $x_n = PD^nP^{-1}x_0$ have we found out. So we fill the formula with the corresponding values.

$$\begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1}{2}\right)^n & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \left(\frac{1}{2}\right)^n & 0 \\ 0 & -\left(\frac{1}{2}\right)^n & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1-\left(\frac{1}{2}\right)^n & 1-\left(\frac{1}{2}\right)^{n-1} \\ 0 & \left(\frac{1}{2}\right)^n & \left(\frac{1}{2}\right)^{n-1} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

The last step, we need to find the value of a_n, b_n and c_n

$$a_n = (1)\left(\frac{1}{3}\right) + \left[1-\left(\frac{1}{2}\right)^n\right]\left(\frac{1}{3}\right) + \left[1-\left(\frac{1}{2}\right)^{n-1}\right]\left(\frac{1}{3}\right)$$

$$a_n = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$$

$$b_n = \left[\left(\frac{1}{2}\right)^n\right]\left(\frac{1}{3}\right) + \left[\left(\frac{1}{2}\right)^{n-1}\right]\left(\frac{1}{3}\right)$$

$$b_n = 0 + 0 = 0$$

$$c_n = 0$$

$$\begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Result

- These are the calculations obtained for the distribution of three genotypes in the population by using diagnolisation method, $xn = PD^nP^-1x0$

$$RR = 1$$
$$Rr = 0$$
$$rr = 0$$

This means that the limit of all plants in the population will be produce genotype RR, which is red apples. Based on the last result, the apples from genotype RR will produce the red apples as long as the fertilisation conditions remain the same during their lifetime.

Part 2

- We considered equally likely probability for three different forms of genotypes in this part, next we would consider different probabilities for each genotype.
- Inbreeding program
- Implementation

References

- application linear algebra to genetics. (-, - -). Retrieved from yumpu: https://www.yumpu.com/en/document/read/24378049/application-of-linear-algebra-to-genetics-living-things-
- Howard Anton, C. R. (2014). Elementary Linear Algebra, 11th edition . In C. R. Howard Anton, Elementary Linear Algebra, 11th edition (p. 678). Canada: Laurie Rosatone.
- John Gallego, S. K. (-, - -). Linear Algebra Genetics. Retrieved from http://web.csulb.edu/: http://web.csulb.edu/ jchang9/m247/m247$_s$p10$_D$anial$_J$ohn$_S$unny.pdf

- Mehebub, M.S.(2017, September 30).Linear algebra to solve autosomal inheritance. Retrieved from slideshare: https://www.slideshare.net/Shababmehebub/linear-algebra-to-solve-autosomal-inheritance